Perception of Degraded Speech in Primary Progressive Aphasia and Alzheimer's Disease

A thesis submitted to University College London for the degree of Doctor of Philosophy

Jessica Jiang

Dementia Research Centre

Institute of Neurology

University College London

2023

I, Jessica Jiang, confirm that the work presented in this thesis is my own. Where information is derived from other sources, I confirm that this has been indicated in the thesis.

1 ABSTRACT

Accurate and flexible understanding of speech in daily life depends critically on our brains' capacity to respond efficiently and adaptively to diverse auditory inputs in multiple contexts and environments. As major dementias strike the brain's auditory and language processing networks relatively selectively, early, and saliently, comprehension of speech under challenging listening conditions is a significant clinical issue in neurodegenerative diseases, particularly Alzheimer's disease (AD) and primary progressive aphasia (PPA). In this thesis, I designed new measures to probe and assess degraded speech perception in AD and PPA, in comparison to healthy older listeners. I investigated how both verbal and nonverbal signals associated with speech are affected in these diseases under degraded listening conditions (simulating those that occur in everyday life) and delineated 'phenotypes' of degraded speech processing accompanying particular diseases.

In Chapter 3, I used phonemic restoration to simulate everyday listening conditions where speech signals are interrupted by background noises. In Chapter 4 and 5, I used noise-vocoding to simulate listening scenarios where verbal and nonverbal emotional signals are of suboptimal quality. In Chapter 6, I used sinewave-transformed accents to assess whether paralinguistic features can convey nonverbal semantic information about speakers even in highly degraded acoustic environments. After taking account of peripheral hearing and general cognitive limitations, AD and PPA syndromes had distinct and separable profiles of impairment across experiments. My findings have implications for understanding the association between hearing impairment and cognitive decline, and for the design of novel diagnostic tests, markers, and interventions addressing real-world communication in major dementias.

2 IMPACT STATEMENT

Currently, 900,000 people are living with dementia in the UK and this number is projected to double every 20 years. Despite such high numbers of those living with dementia, the mechanisms associated with the diseases remain poorly understood and there is yet to be a viable treatment.

Hearing loss has recently been (and increasingly) implicated to be related to dementia, and deficits in auditory processing, such as difficulties with speech perception and communication in typically non-ideal listening environments, have been documented in patients. However, there is currently a lack of a framework to interpret and anticipate the deficits seen and to understand the difficulties that each disease may have with daily communication functions. What is needed are measures to help assess and probe these functionalities, as well as relate these measures to real-world applications of listening, communicating, and social interactions.

This thesis establishes new measures to assess perception of degraded verbal components of speech in Alzheimer's disease (AD) and primary progressive aphasia (PPA, the language-led dementias). It shows different auditory profiles seen in each of the different diseases, speaking to the underlying pathophysiology. These measures could potentially be used as neurophysiological 'stress' tests and proxies for daily life hearing and communication.

This thesis also presents the first experiments to study perception of degraded nonverbal components of speech (e.g., paralinguistic cues and emotional prosody) in AD and PPA. The findings give new insight into how paralinguistic

patterns formulate semantic information of speakers, as well as provide implications for the role of social cognition within daily communication.

The work has many possible applications clinically. Firstly, the measures designed could be presented as rapid readouts on therapeutic effects on neural circuit function, crucial for tracking and marking progress in interventions and clinical trials. Secondly, the measures could facilitate earlier diagnosis as it is likely that central auditory perception, such as degraded speech perception, is affected early in neurodegenerative diseases. Thirdly, the measures designed and findings in this thesis speak to real-world symptoms that patients experience, particularly the difficulty with communicating in our naturally occurring non-ideal listening conditions. Therefore, future clinical care should address the day-to-day troubles that patients and their families experience.

3 ACKNOWLEDGEMENTS

First and foremost, this thesis would not have been possible without the patients, their families and friends, and the healthy volunteers that participated in my research. From the start of this PhD, every single patient and their study partner that I have met has completely inspired me. Even amid a pandemic, the positivity that they have shown me, and their willingness to help with research is just completely and utterly amazing. Thank you for all the chats, sweetest emails, and buckets full of advice on music and gardening – you have all been so kind and inspirational, and this thesis is completely for you.

A very *very* large thank you to both Jason and Chris. Saying that I've been so fortunate is an understatement, as not only did I have such a supportive supervisor throughout my PhD, but I got two *incredibly* supportive supervisors! Thank you both for being so kind and generous with your time (always) in answering many questions, explaining multiple concepts, allowing me to waffle about my thoughts, and appeasing my anxieties. Thank you both for trusting me and always providing me with grace in running research projects too. I have learned so much – thank you thank you thank you.

I am incredibly grateful to the DRC for being such a welcoming working environment and community from when I first started, as a desk-hopping MSc student on the 5th floor, to a PhD student who somewhat knows how to get around the imaging room. Thank you to the loveliest, coolest, and best lab members, the BBG: Jeremy, Elia, Harri, Anthipa, Ben, Lucy, and Jochum. To Mai-Carmen, thank you immensely: you made doing PhD work during lockdowns go from impossible to possible. Thank you to Emilie for your friendship that spun out of chatting through anything, from wedding attendance to statistical analyses. Thank you to Claire, Floey, and Lucianne for all the recruitment meetings, referrals, and for making Thursday afternoons filled with smiles. Thank you Anna for your infectious positivity and passion. Thank you Anne, Suzie, and Faidon for always helping and supporting me with all things DRC-related, allowing my research work to be easier. Thank you Nick for providing me with much advice on academia career timelines and sharing your Sustrans and biking stories. Thank you Jon for your support, listening to me rehearse talks, and referring research participants to us. Thank you to a fantastic thesis committee: Rimona, Stuart, and Doris – I have learned so much from each of you, and thank you for making all our meetings so comfortable and enjoyable.

The biggest hugs and kisses go to my amazing friends (who really are family), both here in the UK and afar. To Steph, Josh, Brandon, Paul, Noah, KC, Monica, Cassie, Nate, Ben, and Lewis – you all prove that distance is horrible, but texts and calls make it…sort of better. To Yuki, MingHao, Allegra, Julie, Harry, Bailey, MJ, and Silvia – you make London always feel like home. To Katie – couldn't have imagined doing our MSc without your friendship, and the same goes for this PhD – thankful for our tea/coffee mornings, afternoon walks, dinners, and many voice notes.

To Lauren and Bert, you both are a very large reason why I decided to do a PhD. I'm not quite sure what you saw in a naïve undergraduate that made you think

‘she might do well with research’, but somehow it stuck and now I’m here. Thank you for your never-ending support and grateful for your friendship and loving families who took me in as their own.

Finally, my beautiful family: 爸爸妈妈、非常感谢您们一直以来的支持和爱。我知道你们会说这是无条件的、但这对我来说意味着整个世界。我们相隔很远、但我希望你们也能感觉到、至少早我们的心目中、我们是亲密的。To Jo and Mike, thank you for being my chilled-out sister and brother-in-law who without a doubt, regardless of any more urgent personal matters (like of course, an adorable child like Evelyn), always has time for me (really, you’ve never not taken my phone calls and always patiently wait for them to end with awkward goodbyes). To the loving and caring Greening family: Mark, Jo, Rosemary, Fiona, and Bill – thank you for welcoming me so whole-heartedly, and immensely appreciative of how supportive you’ve all been with my endeavours. To the cool neighbours five mins away, Rosie and Harry: thank you for your constant positivity, your generosity, and hilarious natures. And last but not least, Sam – I know we said that we’d *never* put significant others on acknowledgments in theses/dissertations (oops), but that’s just plainly not possible. You have been the most supportive partner, start to finish, on this PhD, and for lack of a better phrase: you’re simply the best.

This PhD is dedicated to you all!

UCL Research Paper Declaration Form
referencing the doctoral candidate's own published work(s)

Please use this form to declare if parts of your thesis are already available in another format, e.g. if data, text, or figures:

- *have been uploaded to a preprint server*
- *are in submission to a peer-reviewed publication*
- *have been published in a peer-reviewed publication, e.g. journal, textbook.*

This form should be completed as many times as necessary. For instance, if you have seven thesis chapters, two of which containing material that has already been published, you would complete this form twice.

1. For a research manuscript that has already been published (if not yet published, please skip to section 2)

a) **What is the title of the manuscript?**

Processing of Degraded Speech in Brain Disorders

b) **Please include a link to or doi for the work**

https://doi.org/10.3390/brainsci11030394

c) **Where was the work published?**

Brain Sciences

d) **Who published the work?** (e.g. OUP)

MDPI

e) **When was the work published?**

2021

f) **List the manuscript's authors in the order they appear on the publication**

Jessica Jiang, Elia Benhamou, Sheena Waters, Jeremy CS Johnson, Anna Volkmer, Rimona S Weil, Charles R Marshall, Jason D Warren, Chris JD Hardy

g) **Was the work peer reviewed?**

Yes

h) **Have you retained the copyright?**

Yes – Open Access

i) **Was an earlier form of the manuscript uploaded to a preprint server?** (e.g. medRxiv). If 'Yes', please give a link or doi)

No

If 'No', please seek permission from the relevant publisher and check the box next to the below statement:

☒

*I acknowledge permission of the publisher named under **1d** to include in this thesis portions of the publication named as included in **1c**.*

2. **For a research manuscript prepared for publication but that has not yet been published** (if already published, please skip to section 3)

   a) **What is the current title of the manuscript?**

   Click or tap here to enter text.

   b) **Has the manuscript been uploaded to a preprint server?** (e.g. medRxiv; if 'Yes', please give a link or doi)

   Click or tap here to enter text.

   c) **Where is the work intended to be published?** (e.g. journal names)

   Click or tap here to enter text.

   d) **List the manuscript's authors in the intended authorship order**

   Click or tap here to enter text.

   e) **Stage of publication** (e.g. in submission)

   Click or tap here to enter text.

3. **For multi-authored work, please give a statement of contribution covering all authors** (if single-author, please skip to section 4)

   Conceptualisation: JJ, EB, SW, JCSJ, AV, RSW, CRM, JDW, CJDH. Writing – Original draft: JJ, JDW, CJDH. Writing – Review & Editing: JJ, EB, SW, JCSJ, AV, RSW, CRM, JDW, CJDH. Supervision: RSW, JDW, CJDH

4. **In which chapter(s) of your thesis can this material be found?**

   Chapter 1 (Introduction) and Chapter 7 (Discussion)

5. **e-Signatures confirming that the information above is accurate** (this form should be co-signed by the supervisor/ senior author unless this is not appropriate, e.g. if the paper was a single-author work)

   Candidate

   Jessica Jiang

   Date:

   06/01/2023

   Supervisor/ Senior Author (where appropriate)

   Chris Hardy

   Date

   06/01/2023

UCL Research Paper Declaration Form
referencing the doctoral candidate's own published work(s)

Please use this form to declare if parts of your thesis are already available in another format, e.g. if data, text, or figures:

- *have been uploaded to a preprint server*
- *are in submission to a peer-reviewed publication*
- *have been published in a peer-reviewed publication, e.g. journal, textbook.*

This form should be completed as many times as necessary. For instance, if you have seven thesis chapters, two of which containing material that has already been published, you would complete this form twice.

1. **For a research manuscript that has already been published** (if not yet published, please skip to section 2)

a) **What is the title of the manuscript?**

Remote versus face-to-face neuropsychological testing for dementia research: a comparative study in people with Alzheimer's disease, frontotemporal dementia and healthy older individuals

b) **Please include a link to or doi for the work**

http://dx.doi.org/10.1136/bmjopen-2022-064576

c) **Where was the work published?**

BMJ Open

d) **Who published the work?** (e.g. OUP)

BMJ

e) **When was the work published?**

2022

f) **List the manuscript's authors in the order they appear on the publication**

Mai-Carmen Requena-Komuro, Jessica Jiang, Lucianne Dobson, Elia Benhamou, Lucy Russell, Rebecca L Bond, Emilie V Brotherhood, Caroline Greaves, Suzie Barker, Jonathan D Rohrer, Sebastian J Crutch, Jason D Warren, Chris JD Hardy

g) **Was the work peer reviewed?**

Yes

h) **Have you retained the copyright?**

Yes – Open Access

i) **Was an earlier form of the manuscript uploaded to a preprint server?** (e.g. medRxiv). If 'Yes', please give a link or doi)

https://doi.org/10.1101/2022.04.28.22274370

If 'No', please seek permission from the relevant publisher and check the box next to the below statement:

☒

*I acknowledge permission of the publisher named under **1d** to include in this thesis portions of the publication named as included in **1c**.*

2. **For a research manuscript prepared for publication but that has not yet been published** (if already published, please skip to section 3)

   a) **What is the current title of the manuscript?**

   Click or tap here to enter text.

   b) **Has the manuscript been uploaded to a preprint server?** (e.g. medRxiv; if 'Yes', please give a link or doi)

   Click or tap here to enter text.

   c) **Where is the work intended to be published?** (e.g. journal names)

   Click or tap here to enter text.

   d) **List the manuscript's authors in the intended authorship order**

   Click or tap here to enter text.

   e) **Stage of publication** (e.g. in submission)

   Click or tap here to enter text.

3. **For multi-authored work, please give a statement of contribution covering all authors** (if single-author, please skip to section 4)

   Experimental design: MCRK, JJ, JDW, CJH. Construction of methodology: MCRK, JJ, SB, SJC, JDW, CJD. Data collection: MCRK, JJ, LD, EB, LR, RLB, EVB, CG, CJD. Data analysis: MCRK, JJ. Writing – Original draft: MCRK, JJ, JDW, CJDH. Writing – Review & Editing: MCRK, JJ, LD, EB, LR, RLB, EVB, CG, SB, JDR, SJC, JDW, CJD. Supervision: JDW, CJDH.

4. **In which chapter(s) of your thesis can this material be found?**

   Chapter 2 (General Methods)

5. **e-Signatures confirming that the information above is accurate** (this form should be co-signed by the supervisor/ senior author unless this is not appropriate, e.g. if the paper was a single-author work)

   Candidate

   Jessica Jiang

   Date:

   06/01/2023

   Supervisor/ Senior Author (where appropriate)

   Chris Hardy

   Date

   06/01/2023

UCL Research Paper Declaration Form
referencing the doctoral candidate's own published work(s)

Please use this form to declare if parts of your thesis are already available in another format, e.g. if data, text, or figures:

- *have been uploaded to a preprint server*
- *are in submission to a peer-reviewed publication*
- *have been published in a peer-reviewed publication, e.g. journal, textbook.*

This form should be completed as many times as necessary. For instance, if you have seven thesis chapters, two of which containing material that has already been published, you would complete this form twice.

1. **For a research manuscript that has already been published** (if not yet published, please skip to section 2)

a) **What is the title of the manuscript?**

Phonemic restoration in Alzheimer's disease and semantic dementia: a preliminary investigation

b) **Please include a link to or doi for the work**

https://doi.org/10.1093/braincomms/fcac118

c) **Where was the work published?**

Brain Communications

d) **Who published the work?** (e.g. OUP)

Oxford University Press

e) **When was the work published?**

2022

f) **List the manuscript's authors in the order they appear on the publication**

Jessica Jiang, Jeremy CS Johnson, Mai-Carmen Requena-Komuro, Elia Benhamou, Harri Sivasathiaseelan, Damion L Sheppard, Anna Volkmer, Sebastian J Crutch, Chris JD Hardy, Jason D Warren

g) **Was the work peer reviewed?**

Yes

h) **Have you retained the copyright?**

Yes – Open Access

i) **Was an earlier form of the manuscript uploaded to a preprint server?** (e.g. medRxiv). If 'Yes', please give a link or doi)

Click or tap here to enter text.

If 'No', please seek permission from the relevant publisher and check the box next to the below statement:

☒

*I acknowledge permission of the publisher named under **1d** to include in this thesis portions of the publication named as included in **1c**.*

2. **For a research manuscript prepared for publication but that has not yet been published** (if already published, please skip to section 3)

   a) **What is the current title of the manuscript?**

   Click or tap here to enter text.

   b) **Has the manuscript been uploaded to a preprint server?** (e.g. medRxiv; if 'Yes', please give a link or doi)

   Click or tap here to enter text.

   c) **Where is the work intended to be published?** (e.g. journal names)

   Click or tap here to enter text.

   d) **List the manuscript's authors in the intended authorship order**

   Click or tap here to enter text.

   e) **Stage of publication** (e.g. in submission)

   Click or tap here to enter text.

3. **For multi-authored work, please give a statement of contribution covering all authors** (if single-author, please skip to section 4)

   Experimental design: JJ, CJDH, JDW. Construction of tests: JJ, DLS, CJDH. Data collection: JJ, JCSJ, MCRK, EB, HS. Data analysis: JJ, CJDH. Writing – Original draft: JJ, CJDH, JDW. Writing – Review & Editing: JJ, JCSJ, MCRK, EB, HS, AV, SJC, CJH, JDW. Supervision: JDW, CJDH.

4. **In which chapter(s) of your thesis can this material be found?**

   Chapter 3 (Experiment 1 – Phonemic Restoration)

5. **e-Signatures confirming that the information above is accurate** (this form should be co-signed by the supervisor/ senior author unless this is not appropriate, e.g. if the paper was a single-author work)

   Candidate

   Jessica Jiang

   Date:

   06/01/2023

   Supervisor/ Senior Author (where appropriate)

   Jason D Warren

   Date

   06/01/2023

UCL Research Paper Declaration Form

referencing the doctoral candidate's own published work(s)

Please use this form to declare if parts of your thesis are already available in another format, e.g. if data, text, or figures:

- *have been uploaded to a preprint server*
- *are in submission to a peer-reviewed publication*
- *have been published in a peer-reviewed publication, e.g. journal, textbook.*

This form should be completed as many times as necessary. For instance, if you have seven thesis chapters, two of which containing material that has already been published, you would complete this form twice.

1. For a research manuscript that has already been published (if not yet published, please skip to section 2)

a) **What is the title of the manuscript?**

Click or tap here to enter text.

b) **Please include a link to or doi for the work**

Click or tap here to enter text.

c) **Where was the work published?**

Click or tap here to enter text.

d) **Who published the work?** (e.g. OUP)

Click or tap here to enter text.

e) **When was the work published?**

Click or tap here to enter text.

f) **List the manuscript's authors in the order they appear on the publication**

Click or tap here to enter text.

g) **Was the work peer reviewed?**

Click or tap here to enter text.

h) **Have you retained the copyright?**

Click or tap here to enter text.

i) **Was an earlier form of the manuscript uploaded to a preprint server?** (e.g. medRxiv). If 'Yes', please give a link or doi)

Click or tap here to enter text.

If 'No', please seek permission from the relevant publisher and check the box next to the below statement:

☐

*I acknowledge permission of the publisher named under **1d** to include in this thesis portions of the publication named as included in **1c**.*

2. **For a research manuscript prepared for publication but that has not yet been published** (if already published, please skip to section 3)

   a) **What is the current title of the manuscript?**

   Comprehension of acoustically degraded speech in Alzheimer's disease and primary progressive aphasia

   b) **Has the manuscript been uploaded to a preprint server?** (e.g. medRxiv; if 'Yes', please give a link or doi)

   Yes: https://doi.org/10.1101/2022.12.05.22283108

   c) **Where is the work intended to be published?** (e.g. journal names)

   Brain

   d) **List the manuscript's authors in the intended authorship order**

   Jessica Jiang, Jeremy CS Johnson, Mai-Carmen Requena-Komuro, Elia Benhamou, Harri Sivasathiaseelan, Anthipa Chokesuwattanaskul, Annabel Nelson, Ross Nortley, Rimona S Weil, Anna Volkmer, Charles R Marshall, Doris-Eva Bamiou, Jason D Warren, Chris JD Hardy

   e) **Stage of publication** (e.g. in submission)

   Under peer-review

3. **For multi-authored work, please give a statement of contribution covering all authors** (if single-author, please skip to section 4)

   Experimental design: JJ, CJDH, JDW. Construction of tests: JJ, CJDH. Data collection: JJ, JCSJ, MCRK, EB, HS, AC. Data analysis: JJ. Writing – Original draft: JJ, CJDH, JDW. Writing – Review & Editing: JJ, JCSJ, MCRK, EB, HS, AC, AV, RN, RSW, AV, CRM, DEB, CJH, JDW. Supervision: JDW, CJDH.

4. **In which chapter(s) of your thesis can this material be found?**

   Chapter 4 (Experiment 2 – Noise-Vocoded Verbal Messages)

5. **e-Signatures confirming that the information above is accurate** (this form should be co-signed by the supervisor/ senior author unless this is not appropriate, e.g. if the paper was a single-author work)

   Candidate

   Jessica Jiang

   Date:

   06/01/2023

   Supervisor/ Senior Author (where appropriate)

   Chris Hardy

   Date

   06/01/2023

4 FUNDING

I was supported for the duration of this work by a Frontotemporal Dementia Research Studentship in Memory of David Blechner (funded through The National Brain Appeal). The Dementia Research Centre is supported by Alzheimer's Research UK, Brain Research Trust, and The Wolfson Foundation. The work is also supported by the Alzheimer's Society (grant AS-PG-16-007 to Prof Jason Warren), the Royal National Institute for Deaf People, Alzheimer's Research UK, and the National Institute for Health Research University College London Hospitals Biomedical Research Centre. This work was undertaken at University College London Hospital / University College London, which receives a proportion of its funding from the Department of Health's NIHR Biomedical Research Centre.

5 TABLE OF CONTENTS

6 TABLE OF FIGURES

7 TABLE OF TABLES

8 ABBREVIATIONS

A	ambidextrous
AAF	altered auditory feedback
AD	Alzheimer's disease
AG	angular gyrus
APP	amyloid precursor protein
ASA	auditory scene analysis
aTL	anterior temporal lobe
AUC	area under curve
BKB	Bamford-Kowal-Bench
BNT	Boston Naming Test
BPSD	behavioural and psychological symptoms
BPVS	British Picture Vocabulary Scale
bvFTD	behavioural variant frontotemporal dementia
C9orf72	chromosome 9 open reading frame 72
CBD	cortical basal degeneration
COVID-19	coronavirus disease 2019
DAF	delayed auditory feedback
DARTEL	diffeomorphic anatomical registration using exponentiated lie algebra
dB	decibel
EEG	electro-encephalography
E-KEFS	Delis Kaplan Executive System
F	female
f0	fundamental frequency
FAF	frequency-altered feedback
fMRI	functional magnetic resonance imaging
FTD	frontotemporal dementia
FTLD	frontotemporal lobar degeneration
FWE	family-wise error
FWHM	full width at half-maximum
DAW	digital audio workstation
DRC	Dementia Research Centre

GDA	Graded Difficulty Arithmetic
GNT	Graded Naming Test
GRN	progranulin
HG	Heschl's gyrus
Hz	Hertz
IFG	inferior frontal gyrus
IPL	inferior parietal lobule
L	left
lvPPA	logopenic variant primary progressive aphasia
M	male
MAF	masking auditory feedback
mAIAD	modified Amsterdam Inventory of Auditory Disability and Handicap
MAPT	microtubule associated protein tau
MEG	magnetoencephalography
mIRI	modified Interpersonal Reactivity Index
MMSE	Mini Mental State Examination
MNI	Montreal Neurological Institute
MPRAGE	magnetisation prepared rapid gradient echo
MRI	magnetic resonance imaging
ms	milliseconds
nfvPPA	nonfluent variant primary progressive aphasia
NA	not applicable
NART	National Adult Reading Test
NHNN	National Hospital for Neurology and Neurosurgery
PALPA	Psycholinguistic Assessment of Language Processing in Aphasia
PCA	posterior cortical atrophy
PET	positron emission tomography
PPA	primary progressive aphasia
PSEN	presenilin
PSP	progressive supranuclear palsy
R	right
RMS	root mean square

RMT	Recognition Memory Test
ROC	receiver operating characteristic
RSMS	Revised Self-Monitoring Scale (RSMS)
s	seconds
SD	standard deviation
SPECT	single-photon emission computed tomography
STG	superior temporal gyrus
STS	superior temporal sulcus
svPPA	semantic variant primary progressive aphasia
TDP-43	transactive response DNA binding protein 43
TPJ	temporoparietal junction
VBM	voxel-based morphometry
VOSP	Visual Object and Space Perception
WAIS-R	Wechsler Adult Intelligence Scale-Revised
WASI	Wechsler Abbreviated Scale of Intelligence
WMS-R	Wechsler Memory Scale-Revised
y	years

1 INTRODUCTION

1.1 HEARING IN THE HEALTHY BRAIN

From intentionally listening to a piece of music, to conversing with a friend, to noticing a small cough from someone across the room, our brains are constantly processing the sounds occurring around us, allowing us to fully engage with our surroundings.

The auditory system is highly complex, both in its anatomy and physiology, and has not been as thoroughly studied in comparison to its visual counterpart. When processing a single sound, the outer ear collects and transforms the sounds into vibratory (mechanical) energy, sending it through the middle ear to the inner ear, specifically the cochlea (the organ of hearing). Within the cochlea, the basilar membrane vibrates and the endolymph moves the inner hair cells, transforming the vibrations into a neural representation of the acoustic signal to be relayed to the auditory nerve and auditory cortex in our brain (Musiek & Baran, 2020). As auditory processing occurs in successive stages within the auditory pathway, once the auditory information has reached the cortex, the original auditory information has already been distilled and processed to a certain degree (Cope et al., 2015).

The neuropsychology of auditory information processing has been broadly separated into four stages: early perceptual processing, auditory scene analysis (ASA), apperceptive processing, and semantic/associative processing (Johanna C. Goll, Sebastian J. Crutch, & Jason D. Warren, 2010). Firstly, early auditory perception includes the auditory cognitive operations of feature detection, feature analysis, and scene analysis. Secondly, ASA consists of the ability to parse the

auditory environment into its constituent auditory objects (Bregman, 1990). A classic example of ASA is the ‘cocktail party effect’ (Cherry, 1953), where an individual can “spot” their name and focus on a single speaker in a busy acoustic environment. An auditory object might be defined neuropsychologically as a collection of acoustic data bound in a common perceptual representation and disambiguated from the auditory scene (Goll et al., 2011). Thirdly, the apperceptive processing stage achieves a stable representation of the source or message that has been parsed from the background. Finally, the semantic/associative processing stage is when the auditory object is successfully parsed and matched to meaning.

1.1.1 ‘Peripheral’ and ‘Central’ Hearing

Given the complexity of auditory processing, damage to any level within the auditory pathway can easily affect our hearing. To begin, there is a broad distinction that needs to be addressed when considering the damage to our hearing process: our peripheral (e.g., induced at the cochlear/auditory nerve level) and/or central (i.e., dysfunction in the cortical mechanisms involved) hearing (Hardy et al., 2016). Especially within the population age group studied in this thesis, presbycusis (age-related hearing loss) needs to be considered (Gates & Mills, 2005). Presbycusis is most commonly caused by cochlear dysfunction, but central auditory involvement is highly relevant and most likely understudied (Panza et al., 2015). However, the distinction previously emphasised between peripheral and central hearing is likely not as clear-cut (see **Figure 1.1**) (Johnson et al., 2021). In this thesis, the main focus will be on auditory cognition.

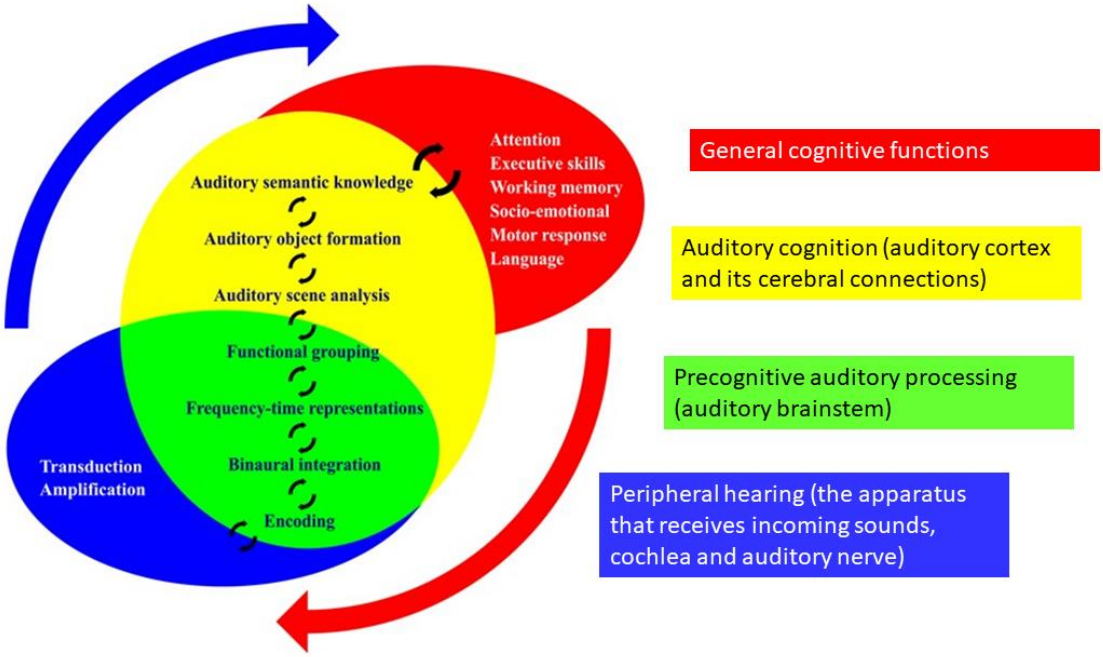


Figure 1.1. Processes and interactions in 'peripheral' and 'central' hearing.
The functional organisation of the auditory processing hierarchy (including the neuropsychological auditory information processing stages mentioned above) and the interplay of hearing with more general cognitive functions (see **Section 1.3.3**). Ellipses indicate the broad domains with colour representations detailed on the right side. Black arrows indicate the reciprocal connections between successive processing stages. External red and blue arrows signify general mechanisms in which hearing dysfunction of any cause may promote cognitive decline and vice versa. Adapted from (Johnson et al., 2021).

1.1.2 Hearing Loss and Dementia

Hearing impairment has recently been identified as a major risk factor for dementia and a driver of cognitive decline and disability (Griffiths et al., 2020; Lin et al., 2011; Livingston et al., 2017). While most studies addressing this linkage have focused on peripheral hearing function measured using the detection of pure tones (Lin et al., 2011; Loughrey et al., 2018; Powell et al., 2021), mounting evidence suggests that measures of central hearing may be more pertinent (Gates et al., 2008; Gates et al., 2011; Johnson et al., 2021; Stevenson et al., 2022). Yet the particular mechanisms involved with the relationship have yet to

be fully understood (see Griffiths et al. (2020); Johnson et al. (2021) for reviews of possible mechanisms).

The auditory system has evolved to allow for quick responses to highly dynamic auditory environments (Pickles, 2015), and one of the key auditory inputs that we encounter daily that require such a response of the auditory system is speech. Speech provides itself not only as a good paradigm to study central hearing (auditory brain) function within hearing populations, but also extends our understanding of the clinical implications in dementia (e.g., impaired speech processing is a key issue in dementia, and speech could present itself as a neural computational stress test for early detection). Parts of the next few sections, which review both speech and degraded speech perception in healthy and diseased brains, have been published in a review in *Brain Sciences* (https://doi.org/10.3390/brainsci11030394).

1.2 Speech Perception in the Healthy Brain

Speech is the core of communication between humans. It is also arguably the most complex of all sensory signals, despite our healthy brain processing it with an apparent ease that belies the complexities of its neurobiological and computational underpinnings. Speech signals arrive at the ears with widely varying acoustic characteristics, reflecting factors such as speech rate, morphology, and in particular, the presence of competing sounds (Mattys et al., 2012). They also include both crucial verbal and nonverbal (paralinguistic) information for successful communication.

Neuroanatomically, it has been well-established that the representation and analysis of intelligible speech occur chiefly in a processing network surrounding the primary auditory cortex in the Heschl's gyrus (HG), with processing 'streams'

projecting ventrally along superior temporal gyrus and sulcus (STG/STS), and dorsally to inferior frontal gyrus (IFG) in the left (dominant) cerebral hemisphere (Alain et al., 2018; Di Liberto et al., 2018; Hickok & Poeppel, 2007).

Medial temporal lobe structures in the dominant hemisphere encode and retain verbal information (Di Liberto et al., 2018; Johnsrude, 2002; Strange et al., 2002) and anterior temporal polar cortex may constitute a 'semantic hub' (Binney et al., 2010; Lambon Ralph & Patterson, 2008; Pobric et al., 2007). The reciprocal connections between auditory regions and prefrontal cortical areas, in particular the IFG (Awad et al., 2007; Peelle, 2010; Rodd et al., 2005), are essential for the top-down disambiguation of speech signals (Erb et al., 2013; Hagoort, 2005; Wild et al., 2012).

In terms of hemispheric asymmetries, there have been differences found between the left and right temporal lobes in response to speech and sounds (McGettigan & Scott, 2012; Scott & McGettigan, 2013). However, while there is bilateral activation in response to acoustic modulations, the left hemisphere does largely show dominance for intelligible speech, reflecting linguistic mechanisms (McGettigan et al., 2012).

1.3 Degraded Speech Perception in the Healthy Brain

The 'clear' speech stimuli played to participants in quiet, controlled laboratory settings are very different from the speech we encounter daily, where it is usually 'degraded' in some form. Under natural listening conditions, speech typically competes with other sounds, as well as occurs in an auditory environment that frequently changes over time. Hence, degraded speech processing is inherently dynamic.

The processing of degraded speech entails the extraction of the intelligible message despite the suboptimal natural listening conditions that adversely affect the quality. These 'conditions' can either be the external environmental factors (e.g., background sounds) (Anderson et al., 2013), individualised vocal characteristics of different speakers (e.g., unfamiliar accents) (Adank et al., 2012), or even feedback relating to one's vocal productions (Chon et al., 2013).

Several different models have been proposed to explain how a speech signal is normally and efficiently disambiguated from auditory 'noise' and the extraction of specific acoustic features, phonemes, words, syntax, and meaning (Davis & Sohoglu, 2019; Hickok & Poeppel, 2007; Okada et al., 2010; Peelle, 2010; Scott et al., 2000). Common to these models is that accurate speech decoding depends on the integration of 'bottom-up' processing of incoming auditory sensory information with 'top-down' prior knowledge and contextual information. The generic bottom-up processes that are used to parse out degraded speech signals are also engaged by other complex acoustic environments during 'auditory scene analysis' (Bregman, 1990), and the high predictability of speech signals recruits top-down processes that are relatively speech-specific: these processes normally interact dynamically and reciprocally to achieve accurate speech recognition (Davis & Johnsrude, 2007).

Broadly similar anatomical regions have been consistently identified in neuroimaging studies on degraded speech processing (using different degraded speech manipulations – see **Section 1.3.2**). The STS/STG has been implicated in accent processing (Adank et al., 2015), altered auditory feedback (Hashimoto & Sakai, 2003), dichotic listening (Hirnstein et al., 2013), noise-vocoded speech (Davis & Johnsrude, 2003; Hervais-Adelman et al., 2012; Scott et al., 2006), perceptual restoration (Sunami et al., 2013), sinewave speech (Khoshkhoo et al.,

2018; Möttönen et al., 2006), speech-in-noise (Hwang et al., 2007), and time-compressed speech (Adank & Devlin, 2010).

Another frequently implicated region is the IFG, where focus is placed on speech signals, but particularly engaged for combinatorial processing of the signals with patterns (e.g., grammar, syntax) and sequences, alongside motoric response. The IFG is activated during accent processing (Adank et al., 2015), noise-vocoded speech (Davis & Johnsrude, 2003; Hervais-Adelman et al., 2012; Scott et al., 2006), perceptual restoration (Sunami et al., 2013), and sinewave speech (Khoshkhoo et al., 2018) in the dominant hemisphere.

Additional temporo-parietal brain regions, such as the (left) angular gyrus (AG), are also engaged under challenging listening conditions (Hartwigsen et al., 2015; Hirnstein et al., 2013; Shahin et al., 2009). Therefore, a distributed fronto-temporo-parietal network consolidates information across multiple levels (acoustic, lexical, syntactic, semantic) to facilitate the comprehension of a degraded speech signal (Guediche et al., 2014).

1.3.1 Predictive Coding Model of Degraded Speech Perception

A potential framework to consider degraded speech perception is using predictive coding. Predictive coding theory postulates that the brain makes iterative inferences about the world at large by continually making predictions about sensory traffic, otherwise known as ‘priors’, and updating those ‘priors’ to minimise the mismatch between prediction and experience (Friston, 2005). Predictive coding assumes that perception is achieved through ‘perceptual inference’, whereby the perceiving brain takes in sensory input and fits it to neural representations (templates) that are optimised by reducing prediction error (Ainley et al., 2016). Here, ‘error’ is defined as the difference between the neural

representations at each processing level and the predictions constructed from the level above (Arnal et al., 2011), and these errors are minimised as priors become increasingly accurate (Heilbron & Chait, 2018), unless contradictory information is received (Ainley et al., 2016).

According to this framework, degraded speech perception depends on hierarchical reciprocal processing, in which each stage (e.g. acoustic, phonemic, lexical, semantic) passes down predictions, and prediction errors (i.e. the difference between expected and heard speech) are passed up the hierarchy (Davis & Sohoglu, 2019; Kocagoncu et al., 2020). Our ability to accurately perceive degraded speech is enhanced by computing the probability of various possible incoming messages according to context (Başkent et al., 2010; Davis & Sohoglu, 2019; Kashino, 2006; Sohoglu et al., 2014). This theory has been explored in some recent work (Donhauser & Baillet, 2020; Sohoglu & Davis, 2020).

The macro-anatomical and functional organisation of the language network suggests how predictive coding mechanisms might operate in the processing of degraded speech (see **Figure 1.2**). Cortical regions involved in 'early' analysis of the speech signal, such as the STG/STS, communicate with 'higher' regions, such as the IFG, that could instantiate predictions about degraded sensory signals. By engaging in both 'bottom-up' perception and 'top-down' processing at every stage within the hierarchy, updated templates on the auditory environment are produced, alongside a continual generation of prediction errors when the auditory input fails to match the template (Leonard et al., 2016; Sohoglu & Davis, 2020).

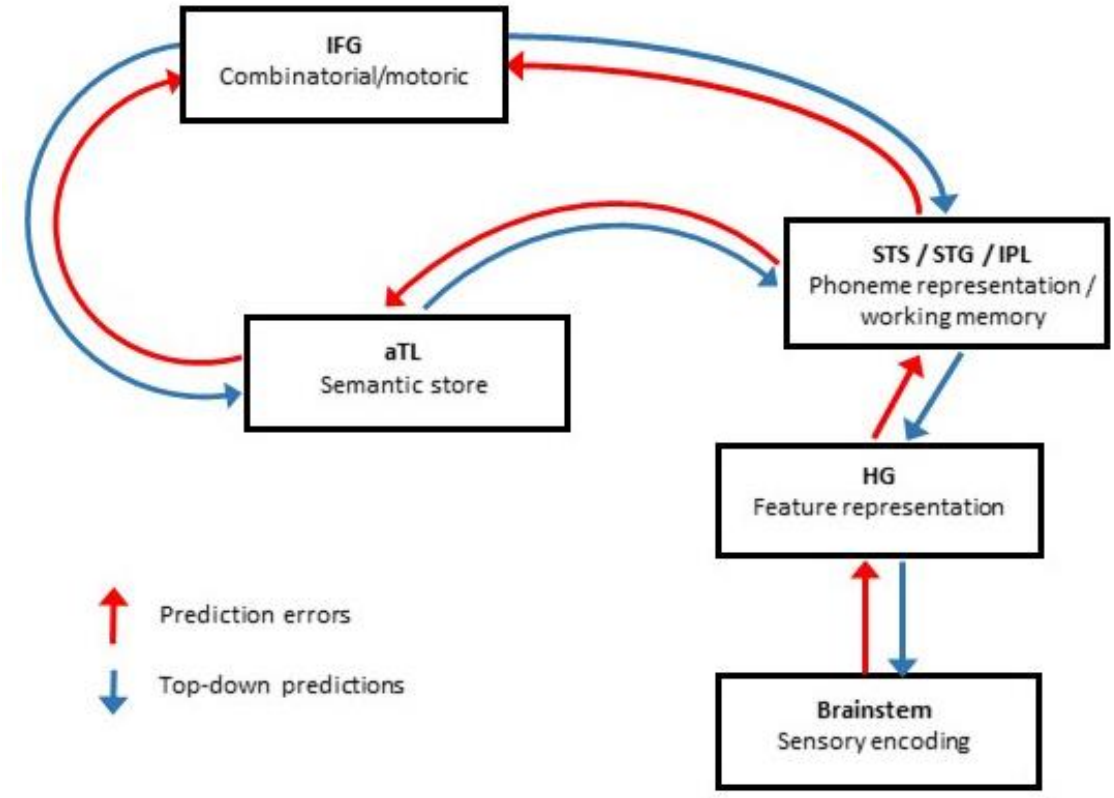


Figure 1.2. A predictive coding model of degraded speech processing with major anatomical loci for core speech decoding operations and their connections, informed by evidence in the healthy adult brain.
Different kinds of degraded speech manipulation are likely to engage these cognitive operations and connections differentially (see **Section 1.3.2**). Incoming sensory information undergoes "bottom-up" perceptual analysis in early auditory areas, while higher-level brain regions generate predictions about the content of the speech signal. Boxes indicate processors that instantiate core functions, however, the processing "levels" are not strictly confined to higher-order predictions or early sensory input as interactions occur at each level. Arrows indicate connections between levels, including reciprocal information, mediating modulatory influences, and dynamic updating/perceptual learning of degraded speech signals. This figure is an over-simplification – cortical areas that are likely to have separable functional roles are grouped for clarity of representation, and while they are not shown in this figure, intra-areal recurrences and inhibitions alongside other local circuit effects may also be operating within these regions. aTL, anterior temporal lobe; HG, Heschl's gyrus; IFG, inferior frontal gyrus; IPL, inferior parietal lobule; STG, superior temporal gyrus; STS, superior temporal sulcus.

Techniques such as electro-encephalography (EEG) and magneto-encephalography (MEG) have revealed dynamic, oscillatory activity that synchronises neural circuits and large-scale networks (Becker et al., 2013).

By delineating feedforward and feedback influences, as well as the rapid changes that attend to deviant, incongruous, or ambiguous stimuli, such techniques are well suited to predictive coding applications. Indeed, evidence from MEG paradigms suggests that induced activity in particular frequency bands may constitute signatures of underlying neural operations during the predictive decoding of speech and other sensory signals (Arnal et al., 2015; Hovsepyan et al., 2020; Sedley et al., 2016).

1.3.2 A Taxonomy of Degraded Speech Manipulations on Verbal Content

To study degraded speech intelligibility, different types of speech degradation methods can be used to address target experimental questions (see **Figure 1.3** for examples of different manipulations). The following list is not exhaustive and mainly focuses on manipulations used in spoken word recognition paradigms (for full reviews, see Cooke et al. (2014); Guediche et al. (2014); Mattys et al. (2012)). While our experience of degraded speech in daily life is infrequently as extreme as some of these manipulations, the processing of naturally degraded speech is continuous and adaptive (regardless of the amount of degradation). Thus, even if some of the manipulations are artificial, they can still elucidate the mechanisms involved in real-world degraded speech perception (e.g., the strategic use of top-down semantic knowledge will occur regardless of the extremity of the degraded speech). It is important to note that depending on the type of manipulation, the intelligibility of the speech will differ, as well as interactional factors (e.g., language background, working memory) associated with speech perception (see **Section 1.3.3**) (Paulus et al., 2020).

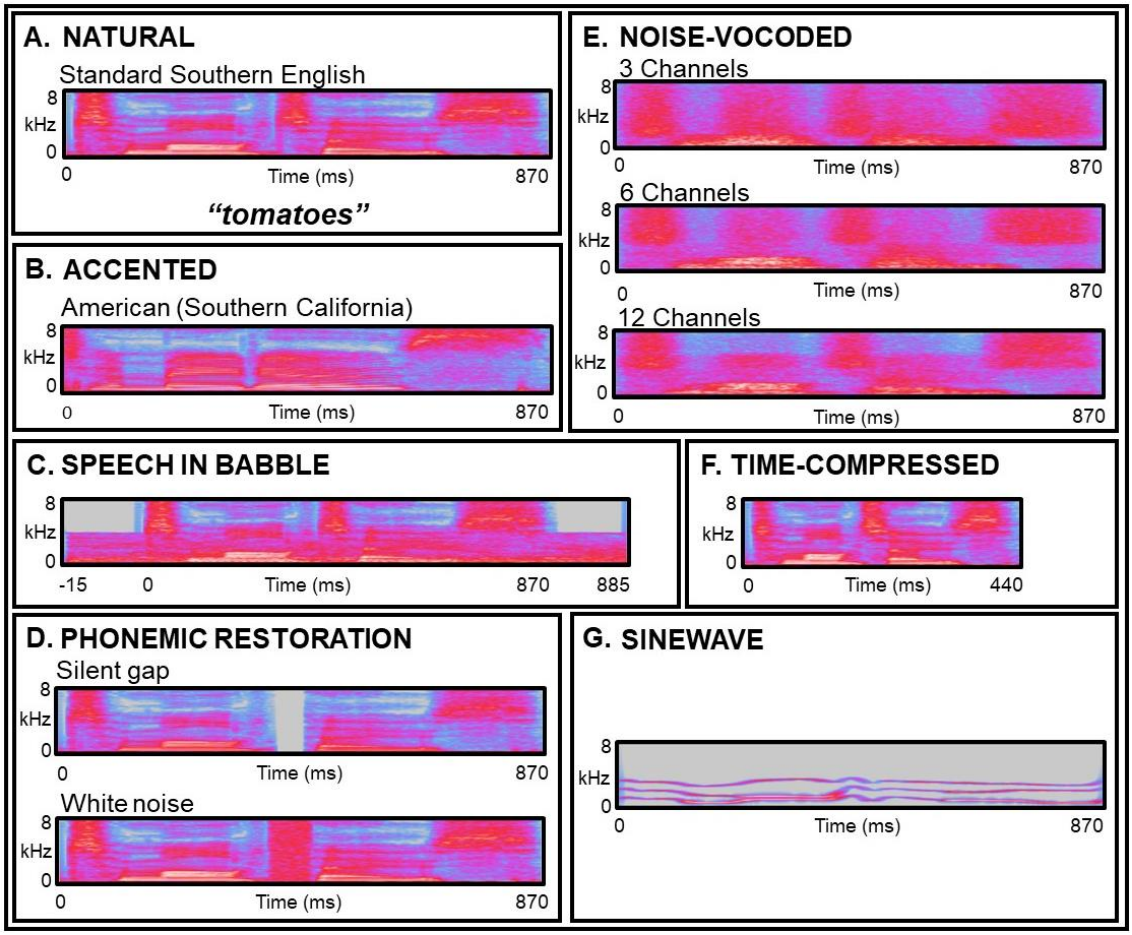


Figure 1.3. Examples of degraded speech manipulations used experimentally and their acoustic effects on the speech signal.
Broadband time-frequency spectrograms of the same speech token ("tomatoes"), were subjected to different forms of speech degradation (all samples apart from 2B were recorded by a speaker with a Standard Southern English accent). (**A**) Natural speech token. (**B**) Same speech token spoken with a Californian accent. (**C**) Speech in multi-talker babble. (**D**) Phonemic restoration. (**E**) Noise-vocoded speech. (**F**) Time-compressed speech. (**G**) Sinewave speech.

1.3.2.1 Accented Speech

An accent is a meta-linguistic feature of speech that reveals information about the speaker's geographical or socio-cultural background (Fletcher et al., 2013). As a degraded speech manipulation, it presents noncanonical variations of the verbal message (e.g., spoken phonemes) in comparison to the listeners' native accent. It also presents patterns of paralinguistic features that convey nonverbal semantic information about the speaker themself.

Understanding speakers with less familiar accents is a common challenge and occurs more frequently in our international society. Different accents modify the acoustic properties of spoken phonemes in different ways and interact with individual vocal characteristics and prosody (Fletcher et al., 2013; Hailstone et al., 2012), and thus, challenges the 'bottom-up' processing with mapping unfamiliar auditory input. The perception of accents causes a delay in word identification, regardless of the accent being regional or foreign to the listener (Floccia et al., 2009). Consistent with a predictive coding framework (see **Figure 1.2**), research in healthy individuals suggests that listeners make predictions about speakers' accents, with accurate predictions facilitating faster accent processing (Clarke & Garrett, 2004; Floccia et al., 2009). Further, the identification of accents of talkers helps us attribute semantic meaning/understanding to words (Cai et al., 2017). Therefore, 'top-down' predictions (and likely semantic grounded) is needed to full comprehend accented speech. In **Chapter 6** (Experiment 4), native British participants will be listening to and identifying international accents.

1.3.2.2 Phonemic Restoration

Phonemic (or perceptual) restoration is a phenomenon put on record by Richard Warren (1970). In the original experiment, a key phoneme was artificially excised from a given sentence. Listeners found it difficult to identify the missing phoneme when it was "filled-in" with an acoustically similar noise (e.g., cough), but could easily and accurately identify the location of the missing phoneme if there was no replacement sound (i.e., silent gap). In other words, participants misperceived the excised phoneme as being present when "filled-in" with an acoustically similar noise (Warren & Obusek, 1971; Warren & Sherman, 1974). This phenomenon is encountered fairly commonly in daily life (e.g., when a door closes in proximity to

someone talking, obliterating part of the speaker's acoustic signal). It likely reflects an interaction between 'bottom-up' early perceptual properties (e.g. the interrupting noise burst has acoustic characteristics sufficient to mask the speech (Kashino, 2006)) and 'top-down' contextual integration (Başkent et al., 2010). This mechanism will be explored in **Chapter 3** (Experiment 1).

1.3.2.3 Altered Auditory Feedback

Altered auditory feedback (AAF) is where the speech signal is electronically altered so that the speaker perceives their voice differently from normal (Lincoln et al., 2006). The interest in this methodology specifically arose from AAF paradoxically improving fluency of speech output in stutterers (Foundas et al., 2013; Lincoln et al., 2006). AAF is a collective term for many manipulations, including masking auditory feedback (MAF), delayed auditory feedback (DAF), and frequency-altered feedback (FAF).

MAF, where the vocal output is played back to the speaker embedded in some form of noise disruption, has been shown to temporarily reduce disfluencies in people who stutter due to suppression of their auditory perception system (Andrews et al., 1982; Cherry et al., 1955). DAF, where the vocal output is played back to the speaker with a delay typically between 100 and 200ms, has been shown to slow speech output rate and elicit speech errors in individuals (Chon et al., 2013; Lee, 1950; Maruta et al., 2014; Stuart et al., 2002; Yates, 1963). FAF, where the vocal output is played back to the speaker with a pitch shift, is compensated for by individuals through typically opposing adjustments to their speech acoustics (Burnett et al., 1998; Chang et al., 2013; Houde & Jordan, 1998; Jones & Munhall, 2000; Kort et al., 2016; Kort et al., 2014; Purcell & Munhall, 2006). AAF's effect on speech output is mediated by regions that comprise the dorsal language network which links auditory vocal representation with

articulatory mechanisms, engaging sensorimotor retuning during speech production (Hashimoto & Sakai, 2003; Hirano et al., 1997; McGuire et al., 1996; Zheng et al., 2013).

1.3.2.4 Speech-in-Noise

Speech-in-noise tests are widely used by audiologists to measure a person's ability to hear spoken words set against an auditory background. The signal-to-noise ratio can be adaptively adjusted to find the 'threshold' point at which speech switches from intelligible to unintelligible (Taylor, 2003). The background 'noise' used in these tests typically falls into one of two categories: 'energetic' masking (e.g. steady-state white noise) or 'informational' masking (e.g. multi-talker babble) (Lidestam et al., 2014). The latter is intuitively more ecologically valid, as exemplified in the well-known 'cocktail party effect' (Cherry, 1953) and ASA (Bregman, 1990) (see **Section 1.1**). Broadly, successful speech-in-noise processing is achieved via a combination of bottom-up and top-down processing to predict and interpret incoming sensory input within a noisy background (Golestani et al., 2013; Golestani et al., 2009).

1.3.2.5 Dichotic Listening

Another way to target ASA is with dichotic listening. Dichotic listening consists of participants being played two different auditory stimuli, delivered simultaneously, to different ears. Therefore, as in speech-in-noise, it is investigating the processing of spoken information with competing verbal material. Past studies have found that as participants focus on a message being played in one ear, they typically do not recognise the content of the unattended message (Moray, 1959). This implicates the role of 'top-down' processing to ascertain which of the compteting 'bottom-up' auditory signals to attend to. However, certain factors

such as semantic similarity of the messages can affect reaction time (Lewis, 1970), suggesting that despite active and conscious attention to one signal, the brain is processing aspects of the other message as well (Ding & Simon, 2012).

1.3.2.6 Time-Compressed Speech

Time-compressed speech, originally instantiated to simulate mental reading speeds for blind people (Foulke et al., 1962), is created by artificially increasing the rate at which a recorded speech stimulus is presented. This is created through systematic sampling and discarding segments of the signal without distorting the frequency of the signal (Musiek & Chermak, 2015). A consistent observation is that the intelligibility of time-compressed speech in healthy listeners decreases as the speech compression rate increases (Dupoux & Green, 1997; Fausto et al., 2018; Foulke & Sticht, 1969). However, listeners are capable of understanding sentences that are compressed down to 35% of their original duration (Dupoux & Green, 1997), and make continuous adjustments to increase the intelligibility of the time-compressed speech (from its decreasing rate) once 275 words per minute is reached (Foulke & Sticht, 1969; Peelle et al., 2004; Poldrack et al., 2001). This builds on the needed role of 'top-down' processing to continually update templates constituted from all facets of auditory cues for speech perception, including syllabic rhythm (which is heavily implicated in time-compressed speech).

1.3.2.7 Noise-Vocoded Speech

Noise-vocoding divides the speech signal digitally into frequency bands ('channels'), each filled with white noise and modulated by the amplitude envelope of the original signal (Shannon et al., 1995). This procedure removes fine structure and spectral detail from speech, whilst preserving temporal cues

(Davis & Johnsrude, 2007; Davis et al., 2005; Shannon et al., 1995). The intelligibility of noise-vocoded speech is related to the number of 'channels' and therefore the intelligibility of the speech signal can be controlled: fewer channels (i.e., less spectral detail) lead to less intelligible speech, and more channels lead to more intelligible speech. The interactions of 'top-down' vs 'bottom-up' processing can be seen as through manipulations of channels, with less channels, and thus, less spectral details, causing a higher reliance on 'top-down' processing due to the increased difficulty in understanding the degraded speech. Three channels are the minimum needed for consistent high-recognition performance in healthy listeners (Shannon *et al.*, 1995), although this can be manipulated based on the task, stimuli, and the precise noise-vocoding parameters. Noise-vocoding has been widely studied and its behavioural and neuroanatomical correlates in the healthy brain are fairly well established (Griffiths & Warren, 2002; Obleser et al., 2007; Warren et al., 2006), as well as frequently used to simulate cochlear implants due to acoustic similarities (McGettigan et al., 2014). Noise-vocoding will be used in **Chapters 4** and **5** (Experiments 2 and 3).

1.3.2.8 Sinewave Speech

Sinewave speech is a drastic reconstruction of the original speech signal that reduces speech by tracking and replacing the formant contours with sinewaves (Remez et al., 1981). The sinewave-reconstructed version of speech signals are acoustically-sparse, lacking the fundamental frequencies, harmonic structure, and short-term spectral cues that are present in a natural spoken sentence, resulting in 'whistled' tones (Davis & Johnsrude, 2007; Remez et al., 1981). Therefore, sinewave-reconstructed speech is highly unnatural and largely uninterpretable as speech initially, but exposure over time facilitates perceptual learning and allows for intelligibility (Barker & Cooke, 1999). Further, this rapid

perceptual learning is extended to not just the verbal content, but nonverbal information as well (e.g., individual speaker identification (Fellowes et al., 1997; Remez et al., 1997; Sheffert et al., 2002)). Thus, sinewave speech is a prime example of the influence of 'top-down' processing for speech perception, due to the difficulty in perceiving the speech with strictly 'bottom-up' auditory cues. This artificial manipulation is used in **Chapter 6** (Experiment 4).

1.3.3 Factors Affecting Degraded Speech Perception

It is crucial to consider the role of general cognitive factors that can influence degraded speech perception (see **Figure 1.1**). The list here is not exhaustive either, but factors that are particularly significant in the population studied in this thesis are the predominant focus.

1.3.3.1 Cognitive factors

The auditory system is dynamic and highly integrated with a vast array of other cognitive functions (Anderson et al., 2013; Arlinger et al., 2009). When engaging in everyday listening, the challenges occurring as a result of non-ideal listening conditions impact the amount of cognitive energy we exert for comprehension (Peelle, 2018; White & Langdon, 2021). Attention modulates the intelligibility of degraded speech, and functional magnetic resonance imaging (fMRI) research suggests that additional frontal cortical regions are recruited when listeners attend to degraded speech signals (Wild et al., 2012). Auditory working memory is also integral to degraded speech processing. Listeners with poorer auditory working memory capacity have more difficulty understanding speech-in-noise, even after accounting for age differences and peripheral hearing loss (Akeroyd, 2008; Anderson et al., 2013; Souza & Arehart, 2015), although this has been debated due to inconsistent results (Dryden et al., 2017; Füllgrabe & Rosen,

2016). Further, after controlling for similar factors, auditory working memory has been shown to significantly predict the perception of speech under adverse listening conditions in older listeners (Kim et al., 2020). Thus, auditory working memory should be taken into account when considering adaptation to degraded speech (Erb et al., 2012).

1.3.3.2 Speech Production

A complementary view to the speech perception model presented in **Section 1.2** is the motor theory of speech perception, postulating that the perception of speech "gestures" is fundamental to the understanding of speech (Galantucci et al., 2006; Liberman et al., 1967; Liberman & Mattingly, 1985).

Studies have found that the motor system, such as the primary motor cortex, the premotor cortex, and the supplementary motor area (but excluding Broca's area), responds to speech (Wise et al., 1999). Disruption of premotor cortex activity with transcranial magnetic stimulation has been shown to impair the discrimination of consonants in syllables that were masked with white noise (Meister et al., 2007), and stimulation of associated regions could enhance other syllable discrimination (D'Ausilio et al., 2009). Other functional studies have shown motor cortex activity at a whole-brain level of analysis (Wilson et al., 2004), however, there have been issues with task structure and how they affect the involvement in the motor cortex (Scott et al., 2009).

Speech production itself also relies on feedback and feedforward control (Hickok, 2012), and automatically compensates for altering auditory feedback (Tourville et al., 2008) (see **Section 1.3.2.3**). Functional neuroimaging studies show that when auditory feedback is altered, there is an increase in activation in the superior temporal cortex, extending into posterior-medial auditory areas

(Hashimoto & Sakai, 2003; Takaso et al., 2010). This corroborates other work suggesting that this region has a prominent role in sensorimotor integration and error detection (Meekings *et al.*, 2016; Sohoglu and Davis, 2016). However, some clinical evidence (e.g., nonfluent aphasic patients; (Hickok et al., 2011; Stasenko et al., 2015)) suggests a dissociation of impairment in speech production from speech perception.

It remains unclear what the role of the motor system could play, whether motor processes are essential for speech perception, or less central and more important in other linguistic and non-linguistic computations (see McGettigan and Tremblay (2018); Scott et al. (2009) for reviews).

1.3.3.3 Perceptual Learning

Improved accuracy of degraded speech processing is often associated with sustained exposure to the stimulus (Eisner et al., 2010; Floccia et al., 2009; Hervais-Adelman et al., 2008): this reflects perceptual learning (Goldstone, 1998). Perceptual learning allows listeners to learn to understand speech that has deviated from expectations (Samuel & Kraljic, 2009), and typically occurs automatically and within short periods of time (Eisner & McQueen, 2006; Norris, 2003; Sohoglu & Davis, 2016). It is likely to reflect synaptic plasticity at different levels of perceptual analysis (Petrov et al., 2005; Tsodyks & Gilbert, 2004), and (in predictive coding terms) reflects iterative fine-tuning of the internal model with increased exposure to the stimulus, leading to error minimisation and improved accuracy of future predictions (**Figure 1.2**) (Kocagoncu et al., 2020). These rapid adaptations are disrupted with transcranial direct stimulation to the left STG fields (Choi & Perrachione, 2019).

Although perceptual learning of degraded speech is strongest and most consistent if trained and tested with the same single speaker (Bradlow & Bent, 2008; Eisner & McQueen, 2005; Nygaard & Pisoni, 1998), with exposure to many speakers embodying a similar particular characteristic (e.g., similar accent), the enhanced processing of that characteristic generalises to different speakers (Clopper & Pisoni, 2004; Gordon-Salant et al., 2010; Sidaras et al., 2009; Stacey & Summerfield, 2007). Longer training (i.e. more exposure to the stimulus) also leads to more stable learning and generalization (Banai & Lavner, 2014).

Listener factors also affect perceptual learning (Perrachione et al., 2011), including language background (Francis et al., 2008; Perrachione et al., 2011), age (Peelle & Wingfield, 2005), attentional set (Huyck & Johnsrude, 2012), and recruitment of language processes in higher-level brain regions and connectivity (Eisner et al., 2010). Further, differences in perceptual learning can result from the type of feedback (e.g., explicit or implicit) (Lehet et al., 2020) or training paradigms as well (Perrachione et al., 2011). The results from past studies on auditory perceptual learning suggest that it arises from dynamic interactions between different levels of the auditory processing hierarchy (Kraljic & Samuel, 2005).

1.3.3.4 Musical Factors

The accumulated experience of speech signals and auditory environments throughout a lifetime leads to the development and refinement of internal models that direct predictions about auditory input, facilitating faster neural encoding and integration (Donhauser & Baillet, 2020). Certain experiential factors, such as musical training, can affect the processing of degraded speech (Alain et al., 2014; Anderson et al., 2013).

Both music and language are forms of human communication, rely on auditory learning, are hierarchically organised (e.g., from sounds/phonemes to melodies/sentences), and share auditory pathways (Neves et al., 2022; Peretz et al., 2015; Zatorre et al., 2002). Therefore, it follows that as auditory skills are critical for music and musical training requires high precision in the processing of acoustic differences, musical training should, and has been shown, to improve a range of basic auditory skills (Bidelman & Krishnan, 2010; Hyde et al., 2009; Koelsch et al., 1999; Kraus et al., 2009).

Many studies have found a transfer of musical abilities to improve linguistic abilities, including phonological awareness (Vidal et al., 2020), linguistic pitch pattern processing (Magne et al., 2006; Moreno et al., 2008; Schön et al., 2004), speech-in-noise perception (Başkent & Gaudrain, 2016; Hennessy et al., 2022; Merten et al., 2021; Parbery-Clark et al., 2009; Swaminathan et al., 2015; Yoo & Bidelman, 2019; Zendel & Alain, 2009), and identification of noise-vocoded words (Fuller et al., 2014). While aspects of the influence and extent are debated, there are suggested theories, such as the OPERA hypothesis. It postulates that music training induces plasticity in speech and language networks when five conditions are met: music engages sensory and cognitive networks that Overlap with those engaged by speech (e.g., auditory working memory); music places higher demand on these networks than speech, requiring more Precision of processing; and musical activities occur in a context that involves positive Emotion, extensive Repetition, and focused Attention (Patel, 2011).

1.3.3.5 Language Factors

Speech perception is likely to be highly influenced and differentiated based on an individual listener's native language background. It is likely that perception of the basic units of speech, such as phonemes, is dependent on native language

experience (Cutler et al., 1986; Cutler & Otake, 1994). Listeners are also much less accurate at discriminating against talkers in a language that they do not speak (Perrachione et al., 2011), and have no additional benefit on perceiving speech spoken by a native speaker or non-native speaker in a non-native language (Bent & Bradlow, 2003).

Bilingual speakers have more difficulty perceiving speech-in-noise in their second language (Blanco-Elorrieta et al., 2020; Jin & Liu, 2012; Lucks Mendel & Widner, 2016). This may be due to over-reliance on bottom-up processing with reduced integration of semantic and contextual knowledge for the second language (Hervais-Adelman et al., 2014; Kousaie et al., 2019; Skoe & Karayanidi, 2019) relative to more efficient top-down integration in their first language (Rammell et al., 2019). Recent studies have suggested that factors such as a bilinguals' language proficiency and age of language acquisition are not the only factors that affect the perception of degraded speech, but other factors (e.g., duration of exposure to languages; co-activation between languages) can affect the perceptual performance of bilinguals under speech-in-noise and noise-vocoded speech (Bsharat-Maalouf & Karawani, 2022).

More studies need to be conducted on speech perception in other languages that are not similar to English in terms of linguistic components (e.g., tonal languages). Thus, speech perception models can be made to be more inclusive of linguistic features present in other languages (e.g., the role of pitch/tone in the speech processing network). For example, while STG is sensitive to phonological features in speech and thus frequently features as a hub for phonological processing (Hickok & Poeppel, 2007), the STG is also sensitive to pitch and lexical tone, key for tonal language speakers (Bhaya-Grossman & Chang, 2021; Liang & Du, 2018; Zatorre & Gandour, 2008).

1.3.3.6 Healthy Ageing

Healthy ageing is associated with changes affecting multiple stages of auditory processing: from cochlea (Roth, 2015), to brainstem (Bidelman & Howell, 2016), then cortex (Henry et al., 2017). Older adults are more likely to experience functionally significant peripheral hearing loss (see **Section 1.1.1**) (Hardy et al., 2016), but there are also various age-related changes in the ascending auditory pathways and their cortical connections (Anderson et al., 2013; Frisina & Frisina, 1997; Gordon-Salant & Fitzgibbons, 1993; Pichora-Fuller et al., 1995; Ross et al., 2020). Thus, the reduced efficiency of processing degraded speech with normal ageing is likely to reflect the interaction of peripheral and central factors (Gates & Mills, 2005).

Age-related decline in cognitive functions relevant to degraded speech perception is also well-documented, encompassing domains such as episodic memory, working memory, and attention (Anderson et al., 2013; Cabeza et al., 2018; Gates & Mills, 2005; Humes et al., 2012). There is evidence to suggest that older listeners rely more heavily on 'top-down' cognitive mechanisms than younger listeners, compensating for the reduced fidelity of 'bottom-up' auditory signal analysis (Henry et al., 2017; Pichora-Fuller, 2008; Saija et al., 2014; Wolpe et al., 2016).

1.4 NONVERBAL AUDITORY PROCESSING IN THE HEALTHY BRAIN

Human speech is typically rarely spoken aloud alone, rather it occurs more typically in social engagement. In conversations, the perception of speech includes the linguistic content (e.g., verbal messages) of what is being communicated and the extraction of non-verbal information that is often crucial to understanding a speaker's intended meaning (Wilson & Sperber, 2002). In other

words,

to understand *what* is being said, it is also important to note *how* the speaker is are saying it. Therefore, in situations where the verbal 'direct' message is unclear, the paralinguistic features (e.g., prosody) provide context and insight into the message (Liebenthal et al., 2016). These features include patterned acoustic characteristics such as pitch, volume, tempo, and rhythm. As mentioned earlier in **Section 1.3.2.1**, accents consist of patterns of paralinguistic features that convey nonverbal semantic information about the speaker.

Prosody is a complex nonverbal feature associated with speech that considers individual speech sounds, pitch, intonation, stress, duration, and intensity, to convey multidimensional information and functions, including distinguishing word meanings in tone languages, disambiguating syntax, highlighting or emphasising elements in a sentence, and signalling emotion (Zatorre & Baum, 2012). Broadly, the functions of prosody can be considered linguistically (to distinguish whether a statement is declarative or interrogative) or applied in terms of decoding a speaker's emotional state (i.e. emotional prosody) (Jonathan D. Rohrer, Disa Sauter, et al., 2010).

Historically, studies on prosody processing in the brain have focused largely on lateralisation. This was initially due to patients with lesions in the right hemisphere, who were unable to express or understand prosody ('aprosodia' (Ross, 1981)). This led to many early studies arguing that the processing of prosody was predominantly in the right hemisphere, while the processing of linguistic information was predominantly in the left hemisphere (Buchanan et al., 2000; George et al., 1996; Pihans et al., 1997). Recent studies have found that the right hemisphere shows more involvement associated with the paralinguistic features of human communication (Kyong et al., 2014; Sammler et al., 2015; Zatorre et

al., 2002), but, more likely, the divide is not strict (Paulmann, 2016) as mechanisms are shared (e.g., fundamental auditory grouping) between speech perception and non-linguistic auditory content as well (Holmes et al., 2020). Similar to speech perception, emotional prosody processing has been suggested to be impaired with age (Cannon & Chatterjee, 2022; Mill et al., 2009) and hearing loss (Christensen et al., 2019).

In models of emotional prosody, acoustic information (as in speech perception) is first extracted in HG, then representation of meaningful suprasegmental features of the auditory stream (e.g., stress, tone) is processed within a wider band of the STS and temporal lobe more broadly (Wildgruber et al., 2006). Evaluation of emotional prosody is then largely mediated by the bilateral inferior frontal cortex (Alba-Ferrara et al., 2011; Ethofer et al., 2006; Frühholz et al., 2011). In addition to these auditory processing areas, other regions involved in social information processing (e.g. medical prefrontal cortex and temporo-parietal junction (TPJ)) likely play a role in decoding emotional intent in others' voices (Morningstar et al., 2022). As speech is the dominant mode for social interactions across human cultures, more research needs to be conducted on relating spoken communication to larger social and emotional contexts (Scott, 2019).

The nonverbal auditory information (e.g., paralinguistic cues, prosody) can also, like verbal messages, be degraded by adverse listening conditions. A few studies have been conducted on degraded emotional prosody perception in children and individuals with cochlear implants (Chatterjee et al., 2019), but this field of research is largely missing.

1.5 DEMENTIA

Dementia is a massive health and social care challenge, with an estimated 47 million people living with dementia in 2015 globally (Livingston et al., 2017). It encompasses many diseases, most neurodegenerative, and is characterised by progressive deterioration in cognitive abilities that impacts daily living and social functioning (Prince et al., 2013). This next section will largely review major dementias that are focused on in this thesis, as they have been documented to have a significant problem with communication and degraded speech processing.

1.5.1 Alzheimer's Disease

Alzheimer's disease (AD), the most common form of dementia, is typically considered to be an amnestic clinical syndrome underpinned by degeneration of posterior hippocampus, entorhinal cortex, posterior cingulate, medial and lateral parietal regions within the so-called 'default mode network' (Agosta et al., 2012; Zhou et al., 2010).

Advancing age is the greatest risk factor for AD, and typically, those with AD occurring after the age of 65 are sporadic cases. AD is also a common cause of young-onset dementia (under 65 years old). Familial AD (caused by mutations in one of the three major genes: PSEN1, PSEN2, and APP) accounts for a very small proportion of all cases. The patients with AD studied in this thesis were all sporadic cases.

Patients with AD typically present with episodic memory impairment, and subsequently develop parietal cortical impairments (e.g., navigation, arithmetic), attention deficits, executive dysfunction, and linguistic impairment (Dubois et al., 2014; Weintraub et al., 2012) (see **Table 1.1**). Variants of AD include posterior cortical atrophy (PCA), which is led by visuo-perceptual and visuo-spatial

Table 1.1. Diagnostic criteria for typical Alzheimer's disease used in this thesis

A	**Specific clinical phenotype**
	• Presence of an early and significant episodic memory impairment (isolated or associated with other cognitive or behavioural changes that are suggestive of a mild cognitive impairment or of a dementia syndrome) that includes the following features: o Gradual and progressive change in memory function reported by patient or informant over more than 6 months o Objective evidence of an amnestic syndrome of the hippocampal type, * based on significant impaired performance on an episodic memory test with established specificity for AD, such as cued recall with control of encoding test
B	**In-vivo evidence of Alzheimer's pathology (one of the following)**
	• Decrease Aβ1-42 together with increased T-tau or P-tau in CSF • Increased tracer retention on amyloid PET • AD autosomal dominant mutation present (in PSEN1, PSEN2, or APP)

Adapted from the IWG-2 criteria for typical Alzheimer's disease (A plus B at any stage) from Dubois and colleagues (2014).

impairments, and frontal variant AD, characterised by personality change and a dysexecutive syndrome (Warren et al., 2012). The language variant, logopenic variant primary progressive aphasia (lvPPA), will be discussed in **Section 1.5.5**. However, even within the 'typical' amnestic AD group, substantial variations in phenotypic profiles is evident (Snowden et al., 2007).Behavioural and psychological symptoms (BPSD) are common in AD, particularly depression, irritability, and delusions (Fernández et al., 2010).

The cardinal pathologies of AD are amyloid plaques and neurofibrillary tau tangles, which accumulate years prior to any cognitive decline (Jack et al., 2010). While both pathological factors are likely to be synergistically responsible for the progression of AD, the specifics remain to be fully understood (Busche & Hyman, 2020). Currently, a definitive diagnosis of AD can only be obtained with genetic confirmation of a known autosomal dominant mutation (Bateman et al., 2011) or histopathological evidence of characteristic protein aggregates in the brain post-mortem (Braak et al., 2006).

1.5.2 Primary Progressive Aphasia

Primary progressive aphasia (PPA) refers to a group of clinical neurodegenerative syndromes with heterogeneous pathological causes. Most patients have underlying frontotemporal lobar degeneration (FTLD) or AD pathology.

FTLD is a clinically and pathologically diverse group of diseases, characterised by progressive decline in behaviour and/or language associated with degeneration of the frontal and temporal lobes (Rabinovici & Miller, 2010). Frontotemporal dementia (FTD), the second major form of young onset dementia, includes three canonical syndromic groups: behavioural variant FTD (bvFTD), nonfluent variant PPA (nfvPPA), and semantic variant PPA (svPPA) (Sivasathiaseelan et al., 2019; Warren, Rohrer, & Rossor, 2013).

Speech and language problems are leading features of PPA. These 'language-led dementias' constitute a heterogeneous group of disorders, comprising three cardinal clinico-anatomical syndromic variants: nfvPPA, svPPA, and lvPPA (Gorno-Tempini et al., 2011; Marshall et al., 2018). However, this current classification, proposed in 2011, still allows considerable overlap with a general tendency for the syndromes to converge at later stages of the disease and for the PPA syndromes with underlying FTLD pathology (particularly nfvPPA) to merge with the atypical parkinsonism spectrum (progressive supranuclear palsy (PSP) and cortical basal degeneration (CBD)) (de Pablo-Fernández et al., 2021).

1.5.3 Nonfluent Variant Primary Progressive Aphasia

The nfvPPA is characterised by disrupted speech and connected language production, due to selective degeneration of a peri-Sylvian network centred on

inferior frontal cortex and insula (Gorno-Tempini *et al.*, 2011; Marshall *et al.*, 2018) (see **Table 1.2**).

The atrophy profiles vary extensively between individual patients, both in severity and extension along and around the STG. The speech presented in patients with nfvPPA is normally effortful, hesitant, and malformed (Cordella et al., 2017; Gunawardena et al., 2010).

Agrammatism is also common in their language (both in speech and in writing), however, while apraxia of speech and agrammatism are typically seen to be linked, recent research suggests these may dissociate (Josephs et al., 2013) as with aprosody as well (Utianski et al., 2018).

Therefore, the clinical spectrum of nfvPPA is diverse with a number of variant sub-syndromes (Josephs et al., 2013; Utianski et al., 2018), and a proportion of patients also further develop symptoms of Parkinsonism, overlapping with PSP and CBD (Doherty et al., 2013; Graff-Radford et al., 2012; Kremen et al., 2011). Patients with underlying CBS-PSP pathology tend to have prominent verbal

Table 1.2. Diagnostic criteria for nonfluent variant primary progressive aphasia used in this thesis

I. Clinical diagnosis of nonfluent/agrammatic variant PPA
At least one of the following core features must be present: 1. Agrammatism in language production 2. Effortful, halting speech with inconsistent of speech sound errors and distortions (apraxia of speech)
At least 2 of 3 of the following other features must be present: 1. Impaired comprehension of syntactically complex sentences 2. Spared single-word comprehension 3. Spared object knowledge
II. Imaging-supported nfvPPA diagnosis
Must show at least one of the following results: 1. Predominant left posterior fronto-insular atrophy on MRI 2. Predominant left posterior fronto-insular hypoperfusion or hypometabolism on SPECT or PET

Adapted from the diagnostic criteria from Gorno-Tempini and colleagues (2011).

adynamia (‘dynamic aphasia’) with significantly reduced spontaneous speech output disproportionate to the level of motor speech impairment (Magdalinou et al., 2018). Many patients also develop orofacial apraxia (Botha et al., 2014; Marshall et al., 2018).

Pathologically heterogeneous, a majority of patients have underlying tau pathology, but a substantial minority have TDP-43 or Alzheimer’s pathology (Rohrer & Schott, 2011; Spinelli et al., 2017). Around 10% of patients have a family history and mutations in all major genes (GRN, MAPT, and C9orf72), with some forms presenting a distinct clinical phenotype (e.g. GRN mutations present with severe agrammatism and semantic impairment without apraxia of speech) (J. D. Rohrer et al., 2010; Snowden et al., 2006).

1.5.4 Semantic Variant Primary Progressive Aphasia

The svPPA is considered the most coherent of all dementia syndromes, exhibiting highly uniform and characteristic clinical, neuroanatomical, and pathological features. It is characterised by the erosion of semantic memory due to selective degeneration of the semantic appraisal network in the antero-mesial (and particularly, the dominant) temporal lobe (Garrard & Carroll, 2006; Gorno-Tempini et al., 2011; Knibb & Hodges, 2005; Marshall et al., 2018) (see **Table 1.3**). As the other PPA syndrome that falls under FTLD, a key distinction from nfvPPA is that patients with svPPA tend to have “fluent” speech, albeit largely circumlocutory and empty, as fine-grained content is replaced by superordinate categories (e.g., ‘sparrow’ becomes ‘bird’).

The hallmark neuroanatomical profile in svPPA is ‘knife-blade’ atrophy of the anterior temporal lobe, spreading to more posterior temporal regions and homologous gyri in the right temporal lobe, as well as the bilateral orbitofrontal

Table 1.3. Diagnostic criteria for semantic variant primary progressive aphasia used in this thesis

I. Clinical diagnosis of nonfluent/agrammatic variant PPA
Both of the following core features must be present: 1. Impaired confrontation naming 2. Impaired single-word comprehension
At least 3 of the following other features must be present: 1. Impaired object knowledge, particularly for low-frequency or low-familiarity items 2. Surface dyslexia or dysgraphia 3. Spared repetition 4. Spared speech production (grammar and motor speech)
II. Imaging-supported nfvPPA diagnosis
Must show at least one of the following results: 1. Predominant anterior temporal lobe atrophy 2. Predominant anterior temporal hypoperfusion or hypometabolism on SPECT or PET

Adapted from the diagnostic criteria from Gorno-Tempini and colleagues (2011).

cortex (Rohrer et al., 2008). Pathologically, most cases are sporadic, with post-mortem analysis revealing TDP-43 type C pathology (Bocchetta et al., 2020).

While the usual leading feature of svPPA is the loss of semantic verbal knowledge, deficits encompass other domains of semantic knowledge, becoming panmodal as the disease evolves (e.g., faces, nonverbal sounds) (Bozeat et al., 2000; Hodges et al., 2000; Snowden et al., 2004).

Behavioural symptoms such as absent or misplaced empathy, social disinhibition and faux pas, a blunter sense of humour, and pathological sweet tooth are common. Other behavioural features such as exaggerated reactions to pain, rigidity with clock-watching, and obsessional interest in numbers, puzzles, and music, also seem particularly prevalent in svPPA (Rosen et al., 2006; Van't Hooft et al., 2021).

1.5.5 Logopenic Variant Primary Progressive Aphasia

The lvPPA is characterised by anomia and impaired phonological working memory, due to degeneration of dominant temporo-parietal circuitry overlapping in what is targeted in other AD variants (Gorno-Tempini *et al.*, 2011; Marshall *et al.*, 2018) (see **Table 1.4**). It shares many overlapping features with AD, such as impaired episodic memory (Mendez et al., 2019), comparably impaired visuospatial awareness (Watson et al., 2018), and deficits in attention and arithmetic skills (Kamath et al., 2020).

Brain atrophy can also overlap, with lvPPA showing early and extensive involvement in speech processing regions (Jonathan D. Rohrer, Gerard R. Ridgway, et al., 2010). The degeneration of temporal/inferior parietal cortices and connected regions underpin the multidimensional cognitive deficits, beyond language (Ramanan et al., 2022).

Pathologically, lvPPA is most likely to have amyloid plaques and tau tangles, and the cerebrospinal fluid profiles are also typically consistent with Alzheimer's pathology (Ikeda et al., 2014; Rohrer & Schott, 2011).

Table 1.4. Diagnostic criteria for logopenic variant primary progressive aphasia used in this thesis.

I. Clinical diagnosis of logopenic variant PPA
Both of the following core features must be present: 1. Impaired single-word retrieval in spontaneous speech and naming 2. Impaired repetition of sentences and phrases
At least 3 of the following other features must be present: 1. Speech (phonologic) errors in spontaneous speech and naming 2. Spared single-word comprehension and object knowledge 3. Spared motor speech 4. Absence of frank agrammatism
II. Imaging-supported lvPPA diagnosis
Must show at least one of the following results: 1. Predominant left posterior perisylvian or parietal atrophy on MRI 2. Predominant left posterior perisylvian or parietal hypoperfusion or hypometabolism on SPECT or PET

Adapted from the diagnostic criteria from Gorno-Tempini and colleagues (2011).

1.5.6 Degraded Speech Perception in Alzheimer's Disease and Primary Progressive Aphasia

As mentioned previously, an accurate and flexible understanding of speech depends critically on the capacity of speech processing circuitry (and linked executive and attentional mechanisms) to respond efficiently, dynamically, and adaptively to diverse auditory inputs in multiple contexts and environments (Samuel, 2011). Degraded speech processing is therefore likely to be highly vulnerable to brain diseases that affect these networks, in particular the primary neurodegenerative 'nexopathies' that cause dementia (Warren, Rohrer, Schott, et al., 2013).

Major dementias strike central auditory and language processing networks relatively selectively, early, and saliently (see Hardy et al., 2016 for a review). Difficulties with auditory and language processing inhibit multiple aspects of daily living for people with these disorders (e.g., difficulty engaging in conversation in busy acoustic environments), decreasing their quality of life. However, the processing of degraded speech in these neurodegenerative diseases remains poorly understood and we presently lack a framework for interpreting and anticipating deficits.

People with AD have particular difficulty with dichotic digit identification tasks (Bouma & Gootjes, 2011; Gates et al., 2008; Idrizbegovic et al., 2013; Utoomprurkporn et al., 2020). This is likely to reflect a more fundamental impairment of auditory scene analysis that includes difficulty with speech-in-noise and speech-in-babble perception (Hannah L. Golden, Jennifer M. Nicholas, et al., 2015; Goll et al., 2012; Wang et al., 2022). During perception of one's name in background babble (the classic 'cocktail party effect'), patients with AD were shown to have abnormally enhanced activation relative to healthy older controls

in the right supramarginal gyrus (H. L. Golden et al., 2015). Deficits in auditory scene analysis are most striking in PCA (the visuo-perceptual and visuo-spatial form of AD), suggesting that posterior cortical regions within the core temporo-parietal network targeted by AD pathology play a critical pathophysiological role (Hardy et al., 2020). Speech-in-noise processing deficits have also been found to precede the onset of other symptoms in AD and have been identified as a harbinger of dementia (Gates et al., 2011; Gates et al., 2010; Pronk et al., 2019; Stevenson et al., 2022).

Mild to moderate AD show enhanced compensatory responses to FAF compared to age-matched controls (Ranasinghe et al., 2017): this has been linked to reduced prefrontal activation and enhanced recruitment of right temporal cortices (Ranasinghe et al., 2019). Further, those with AD have shown to have difficulty with recognising non-native accents (Burda et al., 2004; Hailstone et al., 2012) and comprehending sinewave speech (Hardy, Marshall, et al., 2018) relative to healthy older individuals, and this has been linked to grey matter loss in left superior temporal cortex using voxel-based morphometry (VBM). Considered together with impairments of auditory scene analysis in AD, these findings could be interpreted to signify a fundamental lesion of the neural mechanisms that map degraded speech signals onto stored neural ‘templates’ representing canonical auditory objects, such as phonemes. Encouragingly, perceptual learning of sinewave speech is seen to be intact in AD (Hardy, Marshall, et al., 2018), and comprehension of sinewave speech improves following administration of an acetylcholinesterase inhibitor (C. J. D. Hardy, Y. T. Hwang, et al., 2017).

All three major PPA syndromes have been shown to have clinically significant impairments of central auditory processing affecting speech comprehension

(J. C. Goll et al., 2010; Goll et al., 2011; Grube et al., 2016; Hardy et al., 2016; C. J. D. Hardy, J. L. Agustus, et al., 2017; Hardy et al., 2019; Johnson et al., 2020; Johnson et al., 2021; Ruksenaite et al., 2021): together, these disorders constitute a paradigm for selective language network vulnerability and impaired processing of degraded speech.

In those with nfvPPA, there is a general difficulty with early auditory perceptual processing (Goll et al., 2011; Grube et al., 2016; C. J. D. Hardy, J. L. Agustus, et al., 2017), and transpires to difficulties with speech perception, in both clear and degraded form. In comparison to AD, they show a more pervasive pattern of impairment affecting more and less familiar accents at the level of single words (Hailstone et al., 2012). They also show impaired understanding of sinewave speech relative to healthy controls and svPPA patients (Hardy, Marshall, et al., 2018), as well as some evidence that at least some may be particularly susceptible to the effects of DAF (Hardy, Bond, et al., 2018). In a MEG paradigm in which noise-vocoded words were presented to participants alongside written text that either matched or mismatched the degraded words, Cope and colleagues (2017) found that atrophy of the left inferior frontal cortex in nfvPPA was associated with inflexible and delayed neural resolution of top-down predictions about incoming degraded speech signals.

In patients with svPPA, it is likely that general perceptual encoding of speech signals is accurate, and the deficits seen in speech perception are dependent on top-down mechanisms (especially semantic predictability) that are engaged. For example, in the same sinewave speech paradigm as above, patients with svPPA show a significant identification advantage for more predictable (numbers) than less predictable (geographical place name) verbal signals, highlighting the

important role of ‘top-down’ contextual integration in degraded speech perception (Hardy, Marshall, et al., 2018).

The core deficit in lvPPA is impaired activation and/or transcoding of phonemic ‘templates’, impacting phonemic representations (C. J. D. Hardy, J. L. Agustus, et al., 2017; Johnson et al., 2020), and further, the template matching needed for incoming degraded speech signals. In joint voxel-based morphometric and functional MRI studies of a PPA cohort (Chris J. D. Hardy et al., 2017; C. J. D. Hardy, J. L. Agustus, et al., 2017), a substrate for impaired decoding of spectrally-degraded phonemes in the left supramarginal gyrus and posterior superior temporal cortex is identified, most strikingly in lvPPA relative to healthy older individuals. This extends to the difficulties that lvPPA have with degraded speech perception, shown by an impaired understanding of sinewave speech relative to healthy controls and people with svPPA (Hardy, Marshall, et al., 2018).

1.5.7 Nonverbal Auditory Processing in Alzheimer’s disease and Primary Progressive Aphasia

Past studies have shown impairments in non-verbal sound recognition in AD, nfvPPA, and svPPA patients (Bozeat et al., 2000; Hannah L. Golden, Laura E. Downey, et al., 2015; J. C. Goll et al., 2010; Grube et al., 2016; Hsieh et al., 2011; Omar et al., 2010). Patients with AD showed impairments in non-verbal auditory scene analysis (Goll et al., 2012), and svPPA patients show impaired processing in semantic and emotional congruity and reduced affective integration in auditory scene analysis (Clark et al., 2017).

In svPPA, the primary deficit is semantic representation, in any modality. However, in nonverbal auditory perception, the semantic deficits may be fractionated, initially affecting some kinds of auditory information more than

others, and hierarchically, affecting knowledge of higher order before more generic attributes of sounds (Hailstone et al., 2009; Mole et al., 2019; Muhammed et al., 2018; Omar et al., 2010; Weinstein et al., 2011). Whereas, in nfvPPA, the deficits in nonverbal auditory recognition (e.g., environmental) are auditory specific, and a consequence of early apperceptive stages (Johanna C. Goll, Sebastian J. Crutch, Jenny H. Y. Loo, et al., 2010), and patients with AD showed deficits in nonverbal auditory objects (Omar et al., 2010).

In a study looking at general non-verbal basic acoustic processing, linguistic and emotional prosody, impairments were seen in lvPPA and nfvPPA patients (Jonathan D. Rohrer, Disa Sauter, et al., 2010). Specific analysis showed a particular vulnerability for longer-range prosodic structure. AD patients also show impaired emotional prosody processing (Horley et al., 2010). While auditory emotional processing research in dementia has also been conducted using different acoustic forms (e.g., music) (Agustus et al., 2015; Hsieh et al., 2012; Omar et al., 2011), there is yet to be research conducted on the perception of acoustically degraded paralinguistic and prosody perception.

1.6 Hypotheses and Experiment Outlines

Degraded speech processing is a significant clinical issue in major dementias, particularly AD and PPA. The reduced ability for auditory processing is a major contributor to the decrease in quality of life for people with dementia, posing significant challenges for the care and management of these patients (since conventional hearing aids based on amplification are unlikely to help much).

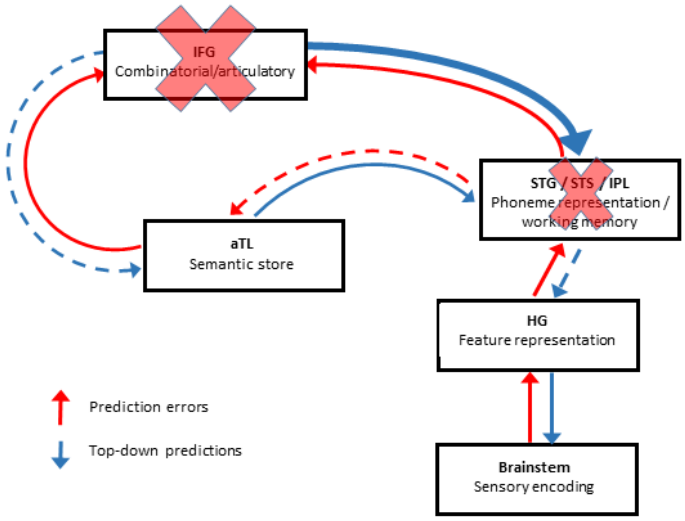
Accurate and flexible understanding of speech, verbal and nonverbal components, depends critically on the capacity of speech processing circuitry to respond efficiently, dynamically, and adaptively to diverse auditory inputs in

multiple contexts and environments (Rankin et al., 2009; Samuel, 2011). Therefore, as exemplified by the primary neurodegenerative 'nexopathies' that cause dementia (Warren, Rohrer, Schott, et al., 2013), major dementias strike central auditory and language processing networks (Hardy et al., 2016). This presents itself as a potential early diagnostic marker (a neural computational 'stress test'). Further, it is plausible that impairments of degraded verbal and nonverbal processing present signature profiles of impairment according to the patterns of the network damage for each dementia syndrome. As described above and exemplified in a predicted model (see **Figure 1.4**) for each disease, the nature of auditory dysfunction (as reflected in the symptoms patients describe) varies between different forms of dementia. Through this, we can extrapolate the hypotheses for each chapter as below.

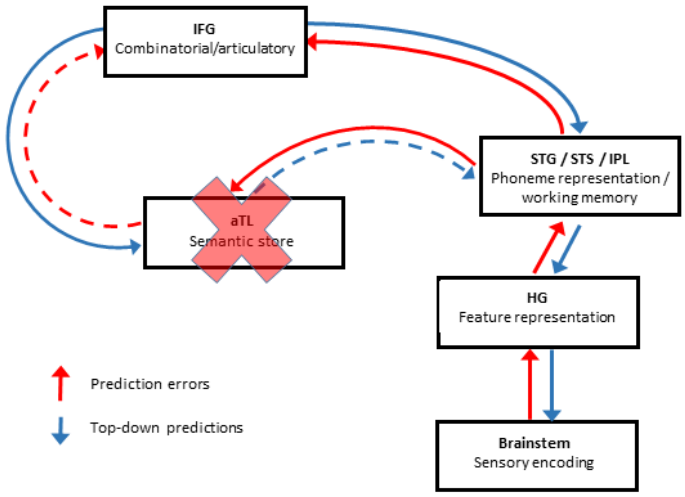
However, the processing of degraded speech, emotional prosody, and other paralinguistic auditory cues associated with speech in dementia remain poorly understood. We presently lack the following:

1. A comprehensive framework for interpreting and anticipating deficits, across different kinds of speech information and different dementia diseases.
2. Suitable tests to reliably measure the processing of degraded speech signals in AD and PPA
3. Data on the 'phenotypes' of degraded speech processing in these diseases
4. Data on how well performance on tests of degraded speech perception maps onto real-world listening, communicating, and social functions

nfvPPA



svPPA



lvPPA/AD

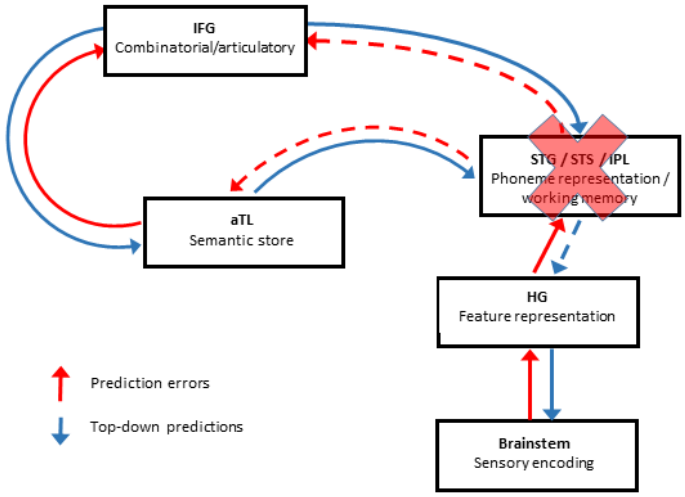


Figure 1.4. . A simplified hypothesised model of predictive coding of degraded speech processing in Alzheimer's disease and primary progressive aphasia.
Referenced to the healthy brain presented in **Figure 1.2**. Each syndrome is associated with a specific pattern of regional brain atrophy and/or dysfunction that is critical to the degraded speech processing network, implying that different dementias may be associated with specific profiles of degraded speech processing. AD and lvPPA groups are put together in this schematic as differences between these syndromes are likely to reflect disease stage and relative degree of involvement of left vs bi-parietal cortices (see **Section 1.5**). Boxes indicate processors that instantiate core speech decoding functions (see **Figure 1.2**), and arrows indicate their connections in the predictive coding framework, with the putative direction of information flow. In the case of nfvPPA, the emboldened descending arrow from IFG to STG signifies aberrantly increased precision of inflexible top-down priors (after (Cope et al., 2017)), to date the most secure evidence for a predictive coding mechanism in the PPA spectrum; the status of the IPL locus in this syndrome is more tentative. Implicitly in the model is the hypothesis that neurodegenerative pathologies will tend to disrupt stored neural templates and "prune" projections from heavily involved, higher-order association cortical areas due to neuronal dropout (promoting inflexible top-down predictions), but also degrade the fidelity of signal traffic through sensory cortices (reducing sensory precision and promoting over-precise prediction errors) (Kocagoncu et al., 2020). The relative prominence of these mechanisms will depend on the macro-network and local neural circuit anatomy of particular neurodegenerative pathologies. Proposed major loci of disruption caused by each disease are indicated with crosses; dashed arrows arising from these damaged modules indicate disrupted information flow. AD, Alzheimer's disease; aTL, anterior temporal lobe; HG, Heschl's gyrus; IFG, inferior frontal gyrus; IPL, inferior parietal lobule; lvPPA, logopenic variant primary progressive aphasia; nfvPPA, non-fluent primary progressive aphasia; STG, superior temporal gyrus; STS, superior temporal sulcus; svPPA, semantic variant primary progressive aphasia.

Therefore, using different degraded speech manipulations that capture relevant kinds of challenging listening conditions that are experienced frequently, this thesis aims to:

1. Design new measures to probe degraded speech perception in major AD and PPA
2. Assess degraded speech perception in these diseases relative to healthy older listeners
3. Investigate how both verbal and nonverbal signals associated with speech are affected in degraded speech paradigms
4. Stratify the auditory "profiles" or phenotypes seen in the diseases through performances on degraded speech paradigms

5. Compare and adjust degraded speech measures with peripheral hearing function to fully interpret results

Here I report the results of four linked experiments:

EXPERIMENT 1: PHONEMIC RESTORATION

How is phonemic restoration, a naturalistic/automatic auditory mechanism that 'repairs' interrupted speech signals, affected in dementia?

Phonemic restoration paradigms simulate the everyday scenario of tracking speech signals in the presence of fluctuating or intermittent background noise. Using a single-word signal detection paradigm, phonemic restoration mechanisms will be assessed in patients with AD and svPPA. To ascertain the top-down influences on degraded speech perception, bottom-up perceptual factors are controlled, and word conditions (real words versus pseudowords) will be manipulated to ascertain the differences in top-down compensatory influences between the two diseases. For this experiment, I predict that in comparison to healthy controls, patients with AD (benefitting from retained semantic 'repair' mechanisms) will show phonemic restoration of real words but reduced restoration of pseudowords. In contrast, patients with svPPA will show reduced phonemic restoration of both word classes.

EXPERIMENT 2: NOISE-VOCODED VERBAL MESSAGES

What is the speech intelligibility threshold for patients with dementia to understand degraded verbal messages?

Noise-vocoding paradigms simulate the daily life scenario of interpreting speech signals of suboptimal quality (e.g., a poor telephone or internet connection). Further, it allows a threshold for degraded speech comprehension to be

quantified. Thus, using noise-vocoding manipulation and psychometric modelling, a speech intelligibility threshold (set at 50%) is ascertained in AD and PPA. This will then be correlated with other demographic and disease characteristics, real-world hearing symptoms, and structural neuroanatomical associations. For this experiment, I predict that in comparison to healthy controls, patients with AD and PPA, particularly nfvPPA and lvPPA, will have an elevated threshold for comprehending vocoded speech. This elevated threshold will correspond with difficulties in daily life hearing symptoms and with using brain imaging analysis, anatomically correlate with regional grey matter atrophy in fronto-temporo-parietal network regions.

EXPERIMENT 3: NOISE-VOCODED EMOTIONAL PROSODY

How is the processing of emotional prosodic cues affected under degraded listening conditions?

Using noise-vocoding manipulation, the impact of degrading the identification of three canonical prosodic emotions (anger, surprise, sadness) at three levels of vocoding channels is assessed in AD and PPA. Confusion matrices and information transfer analyses will be conducted to understand different emotional prosodic cues and the effect of degradation on identification. Further, as the identification of prosody is crucial in social interactions and interpersonal relationships, results will be correlated with measures of social cognition. For this experiment, I predict that in comparison to healthy controls, AD and PPA will be impaired at identifying emotional prosody and that there will be an additional cost once the prosodic cues are noise-vocoded. I also predict that the accurate identification of emotional prosody will be correlated with measures of socio-emotional functioning in daily life.

EXPERIMENT 4: SINEWAVE ACCENTS

How are paralinguistic cues affected in comparison to verbal cues under degraded listening conditions?

Accent recognition is used here as a model of paralinguistic information extraction under non-ideal listening conditions. Building on the research conducted by Hardy and colleagues (2018), sinewave manipulation is used to investigate how the identification of three different accents (Standard Southern British, Standard American, and Standard Russian) is affected in AD and PPA. Confusion matrices and information transfer analyses are conducted to understand how the identification of each accent is affected by sinewave manipulation. For this chapter, I predict that in comparison to healthy controls, patients with AD and PPA will be impaired at identifying accents under sinewave manipulation, and svPPA will show a higher cost on the perception of paralinguistic cues in speech than the verbal content upon sinewave degradation.

2 GENERAL METHODS

2.1 PARTICIPANTS

A total of 119 participants were recruited for all the research presented here: 37 healthy control participants without any significant neurological or psychiatric diseases, and 82 patients (typical AD: 29, nfvPPA: 18, svPPA: 19, lvPPA: 16). For specified participant demographics for each experiment, please see the relevant participant sections in each chapter and corresponding tables. To see the participants in each experiment, see **Appendix Table 8.1**.

Patient participants were recruited from the tertiary specialist cognitive disorders clinic at the National Hospital for Neurology and Neurosurgery (NHNN), referrals from external clinicians, and through local Rare Dementia Support groups. All patients were assessed by a neurologist working at the Dementia Research Centre (DRC) to confirm they met the consensus diagnostic criteria (Dubois et al., 2014; Gorno-Tempini et al., 2011), had clinically mild-to-moderate disease, and determined their suitability for inclusion in research. Healthy control participants were recruited via the DRC's participant database.

2.1.1 Ethical Approval

All participants gave informed consent consistent with Declaration of Helsinki guidelines. Ethical approval for all experiments described in this thesis was granted by the University College London and NHNN Research Ethics Committee.

2.2 COVID-19 PANDEMIC

Due to the COVID-19 pandemic and associated social distancing and lockdown measures in the UK, in-person research that would have otherwise been conducted for my PhD was prevented. Many of our patients are at increased risk for COVID-19 and even when lockdown measures were not fully in place, many did not feel comfortable and/or safe to travel to Queen Square for research. Therefore, our research group strove to translate as much of our research procedure and content as possible to be administered remotely. In the next sections, the methods associated with in-person visits and remote visits are detailed.

2.3 IN-PERSON TESTING

2.3.1 Peripheral Audiometry

All patients were given a standard clinical audiometry protocol, used to assess participants' detection of pure tones at frequencies 250, 500, 1000, 2000, 4000, and 8000Hz (Audiology, 2018). A dual-channel GSI Audiostar Pro™ audiometer (Mark) and calibrated GSI Audiostar Pro headphones with noise-reducing earcups were used. Steady tones were presented in a quiet room separately into each ear: starting from 1000, then 2000, 4000, 8000, 500, 250, and again 1000Hz. The decibel hearing level was set typically at 50 dB, with decreases of 10dB if they could hear the tone, and increases of 5dB once they could not. For each participant, a composite hearing score was created by calculating the mean threshold across all frequencies in the best ear. Scores were reported for all participant groups for each chapter's method section (unless otherwise specified).

2.3.2 Clinical Assessment

When patients were seen for an in-person research visit, a clinical assessment was conducted by a neurologist from the DRC. They were seen alongside their named study partner, typically their primary caregiver, who acted as the research informant to provide reliable collateral information. The main purpose of the clinical assessment was to substantiate the syndromic diagnosis by collecting a detailed inventory of their symptoms, following a pre-defined structured format. The Mini Mental State Examination (MMSE; (Folstein et al., 1975)) was also conducted as a widely used index of disease severity, although it is less accurate in PPA due to linguistic demands in the test. During this assessment, the age of disease onset (estimated based on the patient's informant of first symptom onset) and medication usage were also recorded.

2.3.3 General Neuropsychology and Neurolinguistics Assessments

All patients were given the standard general DRC neuropsychological battery to provide crucial general neuropsychological information. These assessments aim to support the syndromic diagnosis, as well as provide covariates for analysis of experimental measures if needed. They comprised of standardised tests on general intellectual level and domain-specific tests. **Table 2.1** and **Table 2.2** lists the tests delivered in-person.

General neuropsychological tests were administered by myself or another trained psychologist to all recruited healthy control participants and patients. Additionally, general neurolinguistics assessments were conducted also by myself or another trained psychologist for all recruited healthy control participants and PPA patients. The general neurolinguistics assessments consist of more detailed language assessments to cover specific language-related domains that are relevant to the target dementia syndromes (e.g, PPAs).

Table 2.1. List of neuropsychological tests performed with research participants

NEUROPSYCHOLOGICAL BATTERY	
Episodic Memory	
Recognition Memory Test (RMT) for Words	Warrington (1984)
RMT for Faces*	Warrington (1984)
Paired Associate Learning	
Camden Paired Associate Learning	Warrington (1996)
Working Memory	
WMS-R Digit Span Reverse	Wechsler (1987)
Short-Term Verbal Memory	
WMS-R Digit Span Forward	Wechsler (1987)
Executive Function	
WASI Block Design	Wechsler (1997)
WASI Matrices	Wechsler (1997)
Letter Fluency (60s, "F")	In house test
Trials Making Test	Tombaugh (2004)
D-KEFS Colour-Word Inference Test	Delis, Kaplan, and Kramer (2001)
WAIS-R Digit Symbol	Wechsler (1997)
Language	
WASI Vocabulary	Wechsler (1997)
WASI Similarities	Wechsler (1997)
National Adult Reading Test	Nelson (1982)
Schonell Graded Word Reading Test	Schonell (1942)
British Picture Vocabulary Scale (BPVS)	Dunn and Whetton (1982)
Graded Naming Test	McKenna and Warrington (1980)
Category fluency (60s, "Animals")	In house test
Arithmetic	
Graded Difficulty Arithmetic	Jackson and Warrington (1986)
Visuospatial	
Visual Object and Space Perception	Warrington and James (1981)
Usual/Unusual Views	Warrington and Taylor (1973)

The column lists all the neuropsychological tests that were delivered in-person, with references to papers in which they were first described. **Bold** indicates that the tests were delivered remotely as well. * For the RMT (Faces), the shortened version was administered instead when conducted remotely. D-KEFS, Delis Kaplan Executive System; WAIS-R, Wechsler Adult Intelligence Scale-Revised; WASI, Wechsler Abbreviated Scale of Intelligence; WMS-R, Wechsler Memory Scale-Revised.

Table 2.2. List of neurolinguistic tests performed with research participants

NEUROLINGUISTIC BATTERY	
Auditory Input Processing	
PALPA-3	Kay et al. (1992)
Word retrieval	
Boston Naming Test (BNT)	Kaplan et al. (1983)
Language Comprehension	
Synonyms Test (Concrete and Abstract)	Warrington et al (1998)
PALPA-55	Kay et al (1992)
Modified Camel and Cactus	Bozeat et al (2000) Moore et al (2020)
Modified Kissing and Dancing	Bak and Hodges (2003)
Reading	
Grandfather passage	Darley, Aronson, and Brown (1975)
Non-words	In-house test
Regular Words	In-house test
Irregular Words	In-house test
Spelling	
Graded Difficulty Spelling Test	Baxter and Warrington (1994)
Speech Repetition	
Monosyllabic Single Word	Mccarthy and Warrington (1984)
Bisyllabic Single Word	Mccarthy and Warrington (1984)
Trisyllabic Single Word	Mccarthy and Warrington (1984)
Graded Difficulty Sentence Repetition Test	Mccarthy and Warrington (1984)
Spontaneous Speech	
Holiday	In-house test
Cookie Jar Theft	Goodglass and Kaplan (1972)
Sentence Construction	
Written	In-house test
Spoken	In-house test
Visuospatial Working Memory	
Spatial Span Forwards and Backwards	Corsi (1972)

The column lists all the neurolinguistic tests that were delivered in-person, with references to papers in which they were first described. **Bold** indicates that the tests were delivered remotely as well. PALPA, Psycholinguistic Assessment of Language Processing in Aphasia.

2.3.4 Questionnaires

A selection of questionnaires was administered in support of the experimental work conducted in this thesis, including the Modified Amsterdam Inventory of Auditory Disability and Handicap (mAIAD) (Bamiou et al., 2015; Kramer et al., 1995), the Modified Interpersonal Reactivity Index (mIRI) (Davis, 1983) and the Revised Self-Monitoring Scale (RSMS) (Lennox & Wolfe, 1984).

Patient participants had the questionnaires completed by their study informant to minimise the potential confound of disease-associated non-auditory cognitive change. For the mAIAD, healthy older control participants completed the questionnaires themselves.

The mAIAD is used to characterise and attempt to quantify auditory symptoms, disability, and handicap in dementia. It has a total of 28 questions (maximum score = 112), answered on a four-point scale ranging from one (“almost never is able to carry out that listening task”) to four (“almost always is able to carry out that listening task”). The questions are formatted such that a lower score signifies an increasing hearing disability. Scores obtained in the mAIAD were used in Experiment 2 (**Chapter 4**).

The mIRI, frequently used in dementia populations, is based on the Interpersonal Reactivity Index (Davis, 1983). It includes two seven-item subscales: the first subscale measures cognitive empathy in the form of perspective-taking, and the second subscale assesses emotional empathy in the form of empathic concern. The questions are formatted in a series of statements and responders are asked how well each statement describes the participant on a Likert response scale. Scores obtained in the mIRI were used in Experiment 3 (**Chapter 5**).

The RSMS (Lennox & Wolfe, 1984), also frequently used in dementia populations, is a 13-item questionnaire based on the Self-Monitoring Scale (Snyder, 1974). It is made up of two subscales: the first subscale measures participants’ sensitivity to expressive behaviour, and the second subscale measures the tendency to monitor self-presentation. The questions are formatted in a series of statements and responders are asked how well each statement describes the participant on

a Likert response scale. Scores obtained in the RSMS were used in Experiment 3 (**Chapter 5**).

2.4 REMOTE TESTING

Details written out in the next section are also published in a joint first-author paper from myself and another colleague (http://dx.doi.org/10.1136/bmjopen-2022-064576).

2.4.1 Participant Recruitment

Similar to our in-person protocol, potential patient participants were identified via the tertiary specialist cognitive disorders clinic at the NHNN, direct referrals from external clinicians, or through our local Rare Dementia Support groups. Healthy control participants were recruited via the DRC's participant database.

Considering the technological requirements to conduct remote research, an initial telephone screen and trial Zoom session were conducted with each participant to ascertain whether they could do remote research (e.g., if they had the correct equipment, internet connection, etc). Out of the 87 participants that were contacted for research between February and August 2021, only six declined due to not being comfortable with the technology required.

2.4.2 Remote "Audiometry": Bamford-Kowal-Bench (BKB)

Remote participants were unable to complete pure tone audiometry remotely, and thus a different hearing measure was designed using Bamford-Kowal-Bench (BKB), previously validated sentences used in hearing-impaired children (Bench et al., 1979). Each participant listened to a set of 10 sentences, in a fixed order, delivered one at a time via screen and sound share on Zoom. The stimuli itself was presented through Labvanced® (Finger et al., 2017). The participants were

asked to select the last word in the sentence they had just heard, with three possible options presented visually (see **Figure 2.1**). The reason for only requiring the selection of the last word was due to concerns about the effect of working memory in our patients with dementia. A perfect score in the final three trials was required for the participant to proceed with the remote testing session.

Most participants performed at ceiling across all 10 items, and no participant made an error on any of the final three items, meaning that none were rejected based on their BKB performance (see legend of **Table 4.1** in regards to BKB scores for Experiment 2 and Experiment 3).

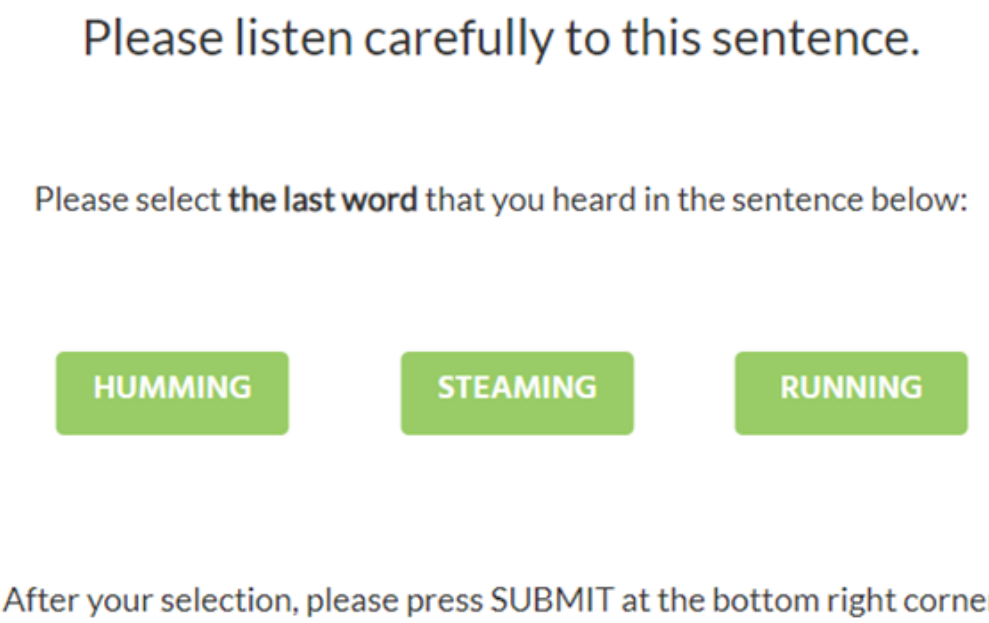


Figure 2.1. Example of Bamford-Kowal-Bench Task Remote Display.
The display was created through Labvanced® for the Bamford-Kowal-Bench (BKB) hearing screening measure. In this example, the sentence spoken was: "The car engine is running". For each sentence, two foils were created: (1) that made sense in the sentence and (2) loosely rhymes with the target word (e.g., "humming").

2.4.3 Remote Neuropsychology and Neurolinguistics Assessments

The selection of neuropsychology and neurolinguistics tests used in remote testing was done based on the in-person batteries. The tests chosen were feasible to be delivered remotely, while also preserving the original overall structure of the in-person batteries and sampling to encompass as many cognitive domains as possible (see **Table 2.1** and **Table 2.2** for the selection of tests). If the task required visual stimulus presentation, a high-quality copy of the stimuli was made and imported into Microsoft Powerpoint to present to participants via screen share.

A key factor in choosing only a selection of the assessments was the role of Zoom fatigue (Bailenson, 2021). The remote neuropsychology and neurolinguistics batteries were shortened to be each delivered in 60-minute sessions (versus the 120 minutes for neuropsychology battery, and 90 minutes for neurolinguistics battery when conducted in-person).

The differences in the standardised assessments between the in-person and remote cohorts were analysed and published in the joint first-author paper (Requena-Komuro et al., 2022). In our paper, we found that there was little evidence for an effect of assessment environment on general neuropsychological and neurolinguistics test performance in participant groups presented (which included AD and PPA patients). This is important as both Experiment 2 (**Chapter 4**) and 3 (**Chapter 5**) include a combination of both in-person and remote participants and therefore justifies combining both for group analyses.

2.4.4 Questionnaires

As a full clinical assessment could not be conducted remotely, part of the symptoms inventory was sent in the form of an online questionnaire to the

informant of the patient participant. Finally, the auditory symptoms and social cognition questionnaires were sent to the informants as well. The healthy controls did not receive them remotely.

Informants were also asked to provide an estimate of the age of disease onset (i.e., the first noticeable symptom), becoming the disease duration measure in this thesis. The telephone version of the MMSE, T-MMSE (Kennedy et al., 2014; Newkirk et al., 2004), was used instead for remote participants, and was administered by myself or another trained psychologist.

2.5 Generation and Validation of Degraded Speech Perception Battery

The clean and clear speech that is traditionally presented in experimental conditions for us to better understand speech perception is not the more accurate representation of day-to-day hearing. Rather, the speech we hear daily tends to be degraded in some form. Thus, the 'Degraded Speech Perception Battery' was created to involve degraded speech tasks that could be more specified, targeted, and probe speech processes within neurodegenerative diseases. The battery was initially trialled for overall effectiveness (i.e. consistency in answers, clear instructions, etc.) on 10 healthy young controls, and then further refined, ready to be administered to patients with dementia and their age-matched healthy controls.

The phonemic restoration test (Experiment 1, **Chapter 3**), the noise-vocoded verbal test (Experiment 2, **Chapter 4**), and the noise-vocoding emotional prosody test (Experiment 3, **Chapter 5**) are all included within the general 'Degraded Speech Perception Battery'. For more information on each of the specified methodologies in each of the tests, please see the methods section for each corresponding experiment chapter.

2.5.1 Remote Adaptation

Experiment 2 and Experiment 3 within the degraded speech perception battery were adapted for remote administration. The paradigm mirrored how it was presented in-person on Matlab, but due to the need to share audio and the screen through a video-conferencing call, the experiments were first implemented on Labvanced® (Finger et al., 2017) for easier administration.

2.6 Statistical Methods

Statistical analyses of behavioural data were performed using JASP® v15, STATA® v14, and R® v4. Brain imaging analysis was carried out using the SPM12 toolbox (http://www.fil.ion.ucl.ac.uk/spm/) in Matlab 2014b.

For continuous demographic and neuropsychological data, participant groups were compared using ANOVA or Kruskal Wallis tests, dependent on the normality of the data. Group categorical data were compared using Fisher's exact tests.

For the experimental and control tests in each chapter, they were analysed using either ANOVA or Kruskal Wallis (again, depending on the normality of the data). Where the omnibus test was significant, post hoc analyses were conducted (pairwise t-tests for ANOVA; Dunn test for Kruskal Wallis). Considering the small cohort sizes in each experiment, no multiple comparisons were conducted, to avoid inflating type II error. An alpha of 0.05 was adopted as a threshold for statistical significance on all tests. To see further details on statistical methods used for analyses in each experiment, please see the methods section for each chapter.

In Experiments 3 and 4, information transfer analysis was used to quantify confusion that occurred within the forced-choice paradigms. Information transfer

analysis was first introduced by Miller and Nicely (1955) and is defined as the ratio of transmitted information to input entropy. The true probabilities that comprise the input entropy are typically known *a priori*, while the probabilities that comprise the transmitted information need to be estimated based on the contents of a participant's error matrix. The information transfer score is therefore obtained by applying a maximum-likelihood estimate of the transmitted information to error matrices. If the participant had received the full "transfer" of information for a given stimulus, the information transfer score would be one (e.g., no errors appear in the error matrix); if the participant's response was independent of the stimuli (i.e., random guessing, where the participant received no information at all), then the information transfer score would be zero.

3 EXPERIMENT 1: PHONEMIC RESTORATION

3.1 SUMMARY

In daily life, spoken messages are often interrupted by extraneous sounds. Therefore, our brains automatically and efficiently 'repair' interrupted speech signals through phonemic restoration, a fundamental physiological process where speech sounds that are obscured by noise are 'filled-in' perceptually to reconstitute the underlying intended signal. As phonemic restoration is a dynamic and integrative process, it is potentially affected by neurodegenerative pathologies. As this has yet to be studied experimentally, the phonemic restoration mechanism within typical Alzheimer's disease (AD) and semantic variant primary progressive aphasia (svPPA) is investigated here.

Here, the phonemic restoration mechanism is studied through participants listening to isolated noise segments, spoken real words, and pseudowords. In both real words and pseudowords, noise bursts either overlaid or replaced the 'target' consonant and a tendency to "hear" the target consonant as present despite actual absence signifies that phonemic restoration has occurred.

All participant groups perceived and distinguished between the isolated noise segments well and showed retained phonemic restoration of real words. In the pseudowords condition, healthy controls showed no phonemic restoration, patients with AD showed a 'rejection' of phonemic restoration, and patients with svPPA showed phonemic restoration that was comparable to real words.

These findings provide the first evidence that phonemic restoration of real words is preserved or even enhanced in neurodegenerative diseases, with distinct syndromic profiles between AD and svPPA that may reflect the differences of

bottom-up phonological representation and top-down lexical disambiguation mechanisms in these diseases. This work has theoretical implications for predictive coding models of language and neurodegenerative disease, in understanding cognitive 'repair' processes in dementia, and implications for developing novel biomarkers and interventions in dementia based on these cognitive processes.

The work presented here in **Chapter 3** (Experiment 1) has been published in *Brain Communications* (https://doi.org/10.1093/braincomms/fcac118).

3.2 INTRODUCTION

The speech we hear in daily life is often interrupted by external sounds (e.g., a door closing during a conversation), yet we generally perceive spoken messages as continuous and coherent. Our brains 'repair' interrupted messages through phonemic restoration (see **Chapter 1.3.2.2**): a fundamental physiological process where speech sounds that are obscured by noise are filled in perceptually to reconstitute the underlying intended signal.

Phonemes are the smallest units of spoken language and are constituted by specific combinations of acoustic spectrotemporal features that define them as a special class of auditory objects (Griffiths & Warren, 2004). Therefore, phonemic perception is a touchstone for many of the fundamental mechanisms involved in auditory object processing.

In the original experiment to address the phonemic restoration mechanism, as introduced in **Chapter 1**, Richard Warren (1970) observed that when a key phoneme was artificially excised from a spoken sentence, listeners found it difficult to accurately identify the location of the missing phoneme when it was

“filled-in” with a coughing sound. On the other hand, participants could easily locate the missing phoneme when no replacement sound (i.e., silent gap) was presented instead. This key result has also been replicated with a variety of other replacement noises (Warren & Obusek, 1971; Warren & Sherman, 1974).

The original Warren paradigm was then further refined by Arthur Samuel in a series of experiments (A. G. Samuel, 1981; Arthur G. Samuel, 1981; Samuel, 1991; Samuel & Ressler, 1986). Instead of full sentences, single words were utilised, containing a white noise segment that either replaced or was added to a target phoneme. This allowed for quantification of phonemic restoration using the framework of signal detection theory. This paradigm also allowed further exploration of factors including phonemic class and position, word frequency, duration, and semantic predictability (real words versus pseudowords). Taken together, the findings from these experiments demonstrated that phonemic restoration depends on an adequate acoustic schema (e.g., a ‘speech-like’ noise) to provide the “filled-in” phenomena according to prior expectations established by lexical context. In neural terms, the component processes of phonemic restoration are mediated by ‘bottom-up’ perceptual mechanisms (spectrotemporal featural synthesis and template matching) that parse out the incoming auditory signals, and ‘top-down’ semantic mechanisms that predictively decode ambiguous signals based on stored knowledge of words (Başkent et al., 2010; Clarke et al., 2014; Guediche et al., 2016; Jaekel et al., 2018; Jiang et al., 2021; Sunami et al., 2013). These mechanisms are computationally demanding and depend on synchronised activity across large-scale neural networks, encompassing posterior superior temporal and inferior frontal cortices in the dominant hemisphere (Guediche et al., 2016; Jiang et al., 2021; Shahin et al., 2009; Sunami et al., 2013).

Presently, phonemic restoration has been little studied in clinical contexts. It appears to be unaffected by mild degrees of hearing loss (Başkent et al., 2010), and may further even be amplified in healthy older listeners due to increased reliance on top-down lexical mechanisms for processing speech signals (Bologna et al., 2018; Jaekel et al., 2018; Saija et al., 2014). An increased tendency for phonemic restoration has also been found in developmental dyslexia (Del Tufo & Myers, 2014), perhaps reflecting less stable acoustic phonological representations. Despite not having been studied in neurodegenerative dementias, on both physiological and neuroanatomical grounds, phonemic restoration is likely not only altered, but distinctive clinical and neuroanatomical profiles of different dementias can also predict the differing consequences for the phonemic restoration mechanism (Warren, Rohrer, Schott, et al., 2013). AD is associated with deficits of auditory scene analysis affecting sound segregation and streaming, spatial hearing, dichotic digit identification, and impaired understanding of sinewave degraded speech (Bouma & Gootjes, 2011; Gates et al., 2008; H. L. Golden et al., 2015; Goll et al., 2012; Hardy, Marshall, et al., 2018; Idrizbegovic et al., 2013; Utoomprurkporn et al., 2020) (see **Chapter 1.5.6** for more details). Bottom-up processes of perceptual analysis supporting phonemic restoration are therefore likely to be affected in AD. In contrast, svPPA has previously been shown to have a deficit in understanding sinewave degraded speech for semantically unpredictable messages (Hardy, Marshall, et al., 2018) (see **Chapter 1.5.6** for more details). This suggests that the top-down semantic disambiguation mechanisms in phonemic restoration may be affected in svPPA. These potentially distinct alterations of phonemic restoration in AD and svPPA (i.e., deficits in bottom-up processing for AD, top-down semantic processing for svPPA) might therefore be parsed and probed by varying the familiarity of the

spoken word stimulus (e.g., real word or pseudowords), thus modulating the degree to which lexical recognition mechanisms are engaged.

3.3 KEY PREDICTIONS

Here, I investigated phonemic restoration in patients with canonical syndromes of AD and svPPA, in relation to healthy controls. Using single real word and pseudoword stimuli (Del Tufo & Myers, 2014; A. G. Samuel, 1981), my hypotheses are:

H_1: In comparison to healthy controls, patients with AD will show increased phonemic restoration of real words but reduced restoration of pseudowords, due to impaired early perceptual analysis of phonemes and increased reliance on top-down processes of lexical recognition.

H_2: In contrast, patients with svPPA would show reduced phonemic restoration of both word classes, due to impaired top-down semantic influences on lexical processing and increased reliance on early perceptual mechanisms.

3.4 MATERIALS AND METHODS

3.4.1 Participants

In this experiment, four patients with svPPA, five with typical AD, and 23 healthy control participants were recruited (see **Appendix: Table 8.1** for participant breakdown and further details in **Chapter 2.1)**. No participant had abnormal peripheral hearing other than age-related hearing loss (see **Chapter 2.3.1** for details of audiometry procedure) or significant cerebrovascular burden on MRI. All participants had a comprehensive general neuropsychological assessment (**Table 3.1**).

3.4.2 Experimental Stimuli

Forty tri-syllabic words with a ‘target’ consonant /d/, /t/, /p/, /f/, or /s/ (which were never the initial or final phoneme of the word) were chosen and separated into two lists each comprising 20 words. The words were matched for phoneme, familiarity, concreteness, imageability, and written frequency using the MRC Psycholinguistic database (*MRC Psycholinguistic Database*) (see **Table 3.2**). Consonants were targeted for the noise manipulation, as they have been shown to produce stronger phonemic restoration effects in normal listeners due to their acoustic similarity to noise (Arthur G. Samuel, 1981). Forty matched, phonetically plausible pseudowords were created by changing specific phonemes in each of the real words (e.g., the real word ‘history’ became ‘bistoty’; see **Table 3.2**), but keeping the ‘target’ consonant unchanged in each case.

Recordings were made of each word by a male with a Standard Southern British English accent on a JoeMeek JM47a Meekrophone on a 2013 Macbook Air. Audio Software utilised were the Scarlett 2i2 First General Audio Interface, and Reaper digital audio workstation (DAW).

Recordings were edited using Praat software (Boersma & Weenink, 2023) to generate stimuli in which a white noise was inserted at the target consonant, either replacing or adding to the consonant. The white noise segments were created in Praat, setting the formula to randomGauss (0,0.25). In each case, the segment containing white noise was of equivalent duration and mean power to the original target consonant. Spectrographs of representative stimuli are shown in **Figure 3.1**.

Table 3.1. General demographic, clinical, and neuropsychological characteristics of participant groups.

	Healthy controls	svPPA	AD	Omnibus significance test
Demographic and clinical				
Sex (F:M)	10:12	0:4	2:3	Fisher's exact = 0.314
Age (years)	66.45 (6.34)	63.00 (8.33)	69.80 (7.95)	F(2,28)=1.12; p=0.342
Handedness (L:R)	2:18[b]	0:4	1:4	Fisher's exact = 0.845
Education (years)	15.65 (2.74)[b]	15.00 (2.00)	16.00 (4.00)	F(2,26)=0.14; p=0.874
Symptom duration (years)	N/A	5.25 (2.22)	5.20 (2.17)	t(7)=0.03; p=0.974
Peripheral hearing score (best ear; dB)	17.75 (8.43)[e]	20.25 (8.54)	25.20 (10.03)	F(2,18)=1.26; p=0.308
General intellect				
MMSE (/30)	29.67 (0.65)[e]	**25.00 (5.60)**	**25.20 (3.83)**	F(2,18)=6.11; p=0.010
Episodic memory				
RMT Words (/50)	47.95 (3.70)[c]	**35.50 (6.61)**	**34.40 (8.32)**	F(2,25)=19.92; p<0.001
RMT Faces (/50)	42.47 (3.91)[c]	**32.00 (4.55)**	**30.60 (4.88)**	F(2,25)=22.50, p<0.001
Working memory				
Digit span forwards (max)	6.79 (1.03)[c]	7.50 (0.58)	6.60 (0.84)	F(2,25)=0.94; p=0.405
Digit span backward (max)	5.63 (1.34)[c]	5.50 (1.91)	4.60 (0.55)	F(2,25)=1.19; p=0.322
Executive functions				
Stroop suppression (s)	55.68 (11.12)[c]	**87.33 (13.61)[a]**	**135.00 (41.75)[a,*]**	F(2,23)=31.79; p<0.001
Letter fluency (total)	17.78 (6.83)[d]	10.00 (8.19)[a]	12.60 (3.85)	F(2,23)=2.59; p=0.097
Category fluency (total)	24.94 (7.03)[d]	**5.75 (6.85)[a]**	**13.40 (6.23)**	F(2,24)=15.52; p<0.001
Language skills				
GNT (/30)	26.05 (2.37)	**3.75 (7.5)[†]**	**16.80 (8.34)**	F(2,25)=40.01; p<0.001
BPVS (/150)	147.63 (2.22)[c]	**82.50 (65.76)[†]**	146.40 (2.07)	F(2,25)=13.80; p<0.001
Posterior cortical functions				
Arithmetic (/24)	16.05 (4.82)[c]	15.50 (4.20)	**7.25 (4.57)[a,*]**	F(2,24)=5.82; p=0.009
VOSP (/20)	19.05 (1.43)[c]	17.67 (1.53)[a]	**15.60 (2.61)**	F(2,24)=8.44; p=0.002

Mean (standard deviation) values are given for variables; counts are given for categorical variables (maximum scores are indicated in parentheses where appropriate). **Bold,** significantly worse performance than healthy control group; [e]significantly worse than svPPA; [g]significantly worse than AD. AD, patient group with Alzheimer's disease; BPVS, British Picture Vocabulary Scale; dB, decibel; GNT, Graded Naming Test; MMSE, Mini-Mental State Examination; N/A, not applicable; RMT, Recognition Memory Test; svPPA, patient group with semantic variant primary progressive aphasia; VOSP, visual object space perception. [a]missing two participants, [b] missing 10 participants, [c] missing three participants, [d] missing one participant, [f] missing four participants.

Table 3.2. Words and pseudowords used in phonemic restoration experiment

Real words			Pseudowords			Manner	Placement
	A|A	R|A		A|A	R|A		
A/PP/EARANCE	21,4,5	1,0,3	I/PP/EAGANCE	12,3,1	4,0,0	Stop	Early
ASSI/S/TANCE	22,4,5	1,0,2	ABBI/S/TINCE	12,3,0	0,0,1	Fricative	Mid
ATMO/S/PHERE	21,4,4	9,1,4	ALMO/S/BERE	11,4,1	3,1,1	Fricative	Mid
ATTI/T/UDE	22,4,5	3,1,3	AFFI/T/UGE	13,4,0	2,0,0	Stop	Late
CA/P/ITAL	21,4,5	22,4,5	HA/P/IFAL	12,4,0	7,2,2	Stop	Mid
CEN/T/URY	20,4,5	7,2,1	CIN/T/URAB	18,4,1	5,1,1	Stop	Mid
CHARAC/T/ER	22,4,5	19,4,5	RARAC/T/ED	19,4,1,	8,2,0	Stop	Late
COM/P/ANY	21,4,4	19,4,5	DOM/P/ANED	18,4,0	17,4,1	Stop	Mid
CON/D/ITION	22,4,4	1,0,1	BON/D/ILON	17,4,0	8,2,0	Stop	Mid
CON/F/IDENCE	22,4,5	9,1,3	PON/F/IDENG	18,4,1	4,0,0	Fricative	Mid
CON/F/USION	20,4,5	22,4,5	FON/F/URON	18,4,1	18,4,1	Fricative	Mid
CONS/T/RUCTION	21,4,4	16,3,5	DONS/T/RUCFEN	16,4,0	7,3,1	Stop	Mid
DE/C/ISION	22,4,5	19,4,5	BE/C/IDON	17,4,1	11,4,1	Fricative	Mid
DEPAR/T/MENT	21,4,5	12,0,1	GEPAR/T/FENT	11,3,0	5,1,0	Stop	Mid
DESCRI/P/TION	22,4,5	18,2,5	DEFRI/P/BON	13,3,1	13,3,1	Stop	Mid
EM/P/LOYMENT	18,3,2	3,1,3	ED/P/LOFMENT	7,4,0	1,0,0	Stop	Mid
ENTER/P/RISE	22,4,5	7,0,2	ENFER/P/RASE	13,4,0	1,0,0	Stop	Late
EQUI/P/MENT	22,4,4	2,0,1	EBI/P/LENT	11,3,0	0,0,1	Stop	Mid
EX/P/RESSION	16,3,4	1,0,2	UX/P/REDDON	12,3,1	3,0,0	Stop	Mid
HIS/T/ORY	22,4,5	20,4,5	BIS/T/OTY	13,4,0	5,1,0	Stop	Mid
HOSPI/T/AL	20,4,5	21,4,5	HISPI/T/AD	19,4,1	20,4,0	Stop	Late
IMPOR/T/ANCE	20,4,5	21,3,5	AMPOR/T/ANE	19,3,1	6,2,2	Stop	Late
INDU/S/TRY	20,4,4	1,0,0	ILDU/S/TAY	17,4,2	1,0,0	Fricative	Mid
IN/S/TITUTE	20,4,5	18,1,4	IB/S/TITITE	11,3,1	8,1,0	Fricative	Early
MINI/S/TER	22,4,4	20,4,5	MUNI/S/GER	18,4,0	14,3,0	Fricative	Mid
NEW/S/PAPER	19,4,5	3,1,2	HEW/S/PADER	14,4,1	3,0,1	Fricative	Early
O/FF/ICER	21,4,5	5,0,2	U/FF/IYER	5,3,1	0,0,0	Fricative	Early
O/P/INION	21,3,5	12,4,4	U/P/IDION	20,4,0	2,0,1	Stop	Early
ORCHES/T/RA	22,4,5	21,3,5	ORFES/T/RID	15,4,1	7,2,2	Stop	Late
PERCE/P/TION	22,4,5	22,4,5	LERCE/P/RON	15,3,1	5,1,2	Fricative	Mid
PER/S/ONNEL	22,4,5	1,1,3	POR/S/OBBEL	14,4,1	5,0,0	Stop	Mid
POE/T/RY	22,4,5	16,2,5	HOE/T/ID	18,4,0	5,2,0	Stop	Late
PRINCI/P/LE	22,4,5	21,4,5	GRINCI/P/IT	18,4,1	9,3,0	Stop	Late
PRO/D/UCTION	21,4,4	19,4,5	PLO/D/UCFON	10,3,1	6,2,1	Stop	Mid
PRO/F/ESSOR	22,4,5	6,0,4	TRO/F/ETTOR	14,3,2	2,0,0	Fricative	Mid
PRO/P/ERTY	22,4,5	17,2,4	GRO/P/ERFY	18,4,1	16,3,1	Stop	Mid
PRO/T/ECTION	20,4,5	4,0,0	FRO/T/ECTAN	12,3,2	0,0,1	Stop	Mid
RA/D/IO	22,4,5	10,2,5	BA/D/IA	18,4,1	9,1,0	Stop	Mid
RESIS/T/ANCE	17,4,2	10,4,0	BESIS/T/ANG	22,4,5	22,4,5	Stop	Late
TEN/D/ENCY	22,4,5	2,0,2	REN/D/ENFY	17,3,1	5,0,0	Stop	Mid

For ease of reading, items are represented orthographically in the table. Pseudowords in each case were generated by modifying specific phonemes for the 'matching' real word. The 'target' phoneme that underwent noise modification in each word stimulus is framed by / /. For each stimulus, the 'A|A' column denotes the number of participants in each group (in order: control, svPPA, AD) who correctly identified the 'Added' version as 'Added'; 'R|A' gives the number of participants in each group who incorrectly identified the 'Replaced' version of each stimulus as 'Added' (i.e., phonemic restoration occurred). The column headed 'Manner' indicates the manner of articulation of the target phoneme, and 'Placement' refers to the part of the word in which the target phoneme was located: 'early' if it occurred in the first syllable; 'mid' if in the second; and 'late' if in the third. A two-tailed t-test showed that real and pseudowords did not differ significantly in target phoneme mean duration (real words mean = 105 (standard deviation = 21) ms; pseudowords mean = 105 (standard deviation = 21) ms; p = 0.99).

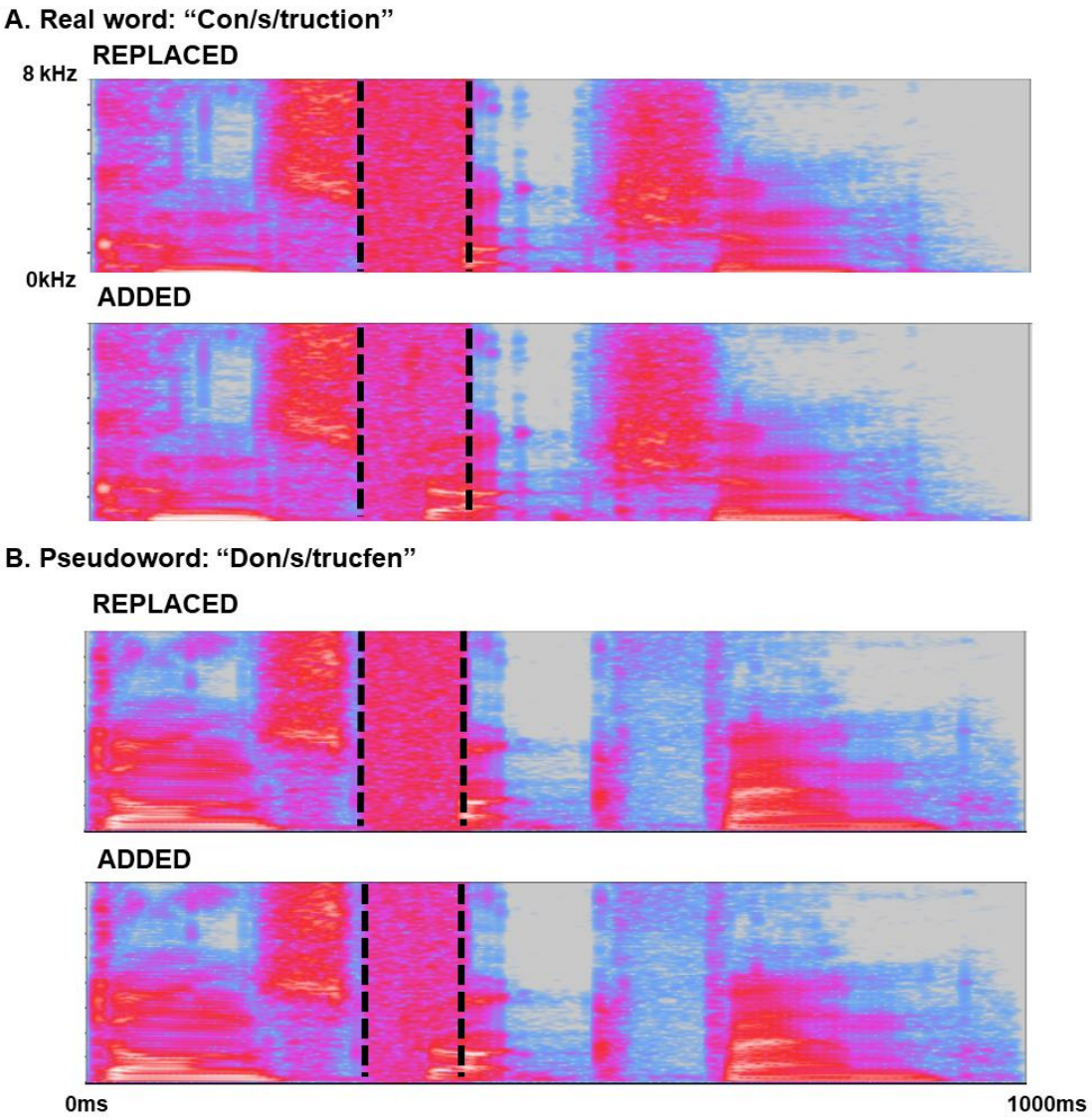


Figure 3.1. Representative time-frequency spectrograms of stimuli for the different experimental conditions based on word carriers.
The y-axis of each spectrogram codes frequency (kilohertz); the x-axis codes time (milliseconds). In all example spectrograms, vertical dotted lines show the boundaries of the target spoken phoneme/consonant (indicated in the word heading of each panel); the spoken word segment containing the target phoneme has been manipulated in each case with white noise. **(A)** Example stimuli based on real word carriers; **(B)** stimuli based on pseudoword carriers. In each panel, an example of a 'Replaced' stimulus (i.e. white noise replacing the spoken consonant) is shown above and an example of an 'Added' stimulus (i.e. white noise superimposed over the spoken consonant) is shown below. Spectrograms were generated in Audacity (v3.0.0) (https://audacityteam.org).

This resulted in half the recordings with white noise added to the consonant (e.g., real word, A/*PP*/EARANCE; pseudoword, I/*PP*/EAGANCE), while in the other half, white noise replaced the consonant completely (e.g. A/__/EARANCE or I/__/EAGANCE). This manipulation yielded a total of four-word stimulus conditions (two carrier conditions: Real words/Pseudowords) x (two noise conditions: Replaced/Added), each comprising 40 trials.

Separately, perceptual control stimuli were created to assess participants' ability to discriminate 'Replaced' versus 'Added' stimuli acoustically, without lexical influence. Therefore, the control stimuli comprised 40 isolated noise segments, each taken from the word stimuli created previously. Twenty of the control stimuli consisted of the white noise segments superimposed on the previous target consonant (e.g. '_S_'; equivalent to 'Added' noise segments in the spoken words) and 20 without an associated speech sound (e.g., '__'; equivalent to 'Replaced' noise segments in spoken words; i.e., white noise solely).

3.4.3 Procedure

All testing sessions took place in a quiet room and I administered the stimuli through MATLAB R2019b on a Windows laptop via headphones (Audio-Technica ATH-M50x) set at a comfortable listening volume (at least 70dB). During the experimental sessions, there was no time limit and I gave no feedback on performance.

3.4.3.1 Real word and pseudoword conditions

Following Del Tufo & Myers' (2014) procedure, both 'Added' and 'Replaced' stimuli were split into four blocks of 40 trials, each containing 10 trials from each stimulus condition (i.e., 10 'Added' real words, 10 'Added' pseudowords, 10 'Replaced' real words, 10 'Replaced' pseudowords). Trials were randomised

within each block, and the ‘Added’ and ‘Replaced’ versions of the same word and matched pseudowords never occurred within the same block.

I informed participants that they would hear a series of words, either ‘real’ or ‘made-up’, containing a noise and asked them to determine whether the word continues through the noise (‘Added’) or was interrupted by the noise (‘Replaced’). To ensure understanding, participants were first familiarised with some practice stimuli. I provided pictorial cue cards (**Figure 3.2**) as aids during the experimental session and the participants could choose whether to respond verbally or by pointing at the cue card to indicate their choice.

3.4.3.2 Perceptual control task on isolated noise segments

After the words, the perceptual control stimuli (isolated noise segments) were presented in a randomised order as a single block of 40 trials. Participants were told that they would hear a series of noises and their task on each trial was to decide whether it was ‘only noise’ or ‘noise-plus-letter’ (see **Figure 3.2**).

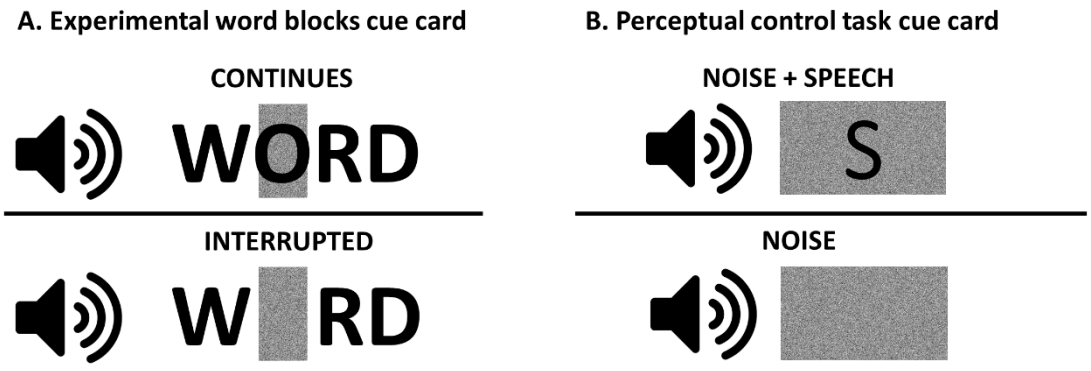


Figure 3.2. Cue cards for Experiment 1.
The cue cards presented here were each, **(A)** and **(B)**, printed on a full-size A4 size paper that was laminated (anti-glare) to help participants respond during the experiment.

3.4.4 Analysis of Data

Data were analysed using STATA v14 and JASP v15. Details on data statistical analyses methods for categorical and continuous demographic and neuropsychological data are detailed in **Chapter 2.6**.

While previous studies have used *d'*, upon inspection of our data (**Table 3.3** and **Appendix: Table 8.2**), it is apparent that individual participants across the healthy control and patient groups never mislabelled an 'Added' stimulus as 'Replaced' in the noise segment in the real word conditions, giving a value of zero for this response and rendering the use of *d'* untenable. Thus, a non-parametric *A*' as a measure of the sensitivity of discrimination between 'Added' and 'Replaced' stimuli (Stanislaw & Todorov, 1999) was used instead, and the criterion location *c* was used as the measure of bias in participants' responses.

For each participant, *A'* and *c* were calculated for each experimental condition separately using an Excel Workbook (Gaetano et al., 2017). Values of *A'* can range from zero to one – one indicates perfect discrimination (i.e. no phonemic restoration), 0.5 indicates that 'Added' and 'Replaced' versions of the presented words were indistinguishable (either meaning that all words labelled as 'Added' (i.e. complete phonemic restoration) or all words labelled as 'Replaced' (no phonemic restoration) or some combination of 'Replaced'/'Added' confusions), and <0.5 indicates a tendency to select the response opposite to what would be defined as an accurate 'hit' (e.g. a tendency to report 'Added' words as 'Replaced' and vice versa).

Table 3.3. Individual raw scores for patient groups across experimental conditions.

	Control	svPPA					AD					
	Mean (SD)	1	2	3	4	**Mean (SD)**	1	2	3	4	5	**Mean (SD)**
Isolated noise segments (/40)												
A|A	**19.3 (1.1)**	20	20	20	20	**20.0 (0.0)**	20	20	19	20	20	**19.8 (0.4)**
R|A	**1.1 (0.3)**	3	1	1	1	**1.5 (1.0)**	6	2	1	1	2	**2.4 (2.1)**
R|R	**18.9 (0.3)**	17	19	19	19	**18.5 (1.0)**	14	18	19	19	18	**17.6 (2.1)**
A|R	**0.7 (1.1)**	0	0	0	0	**0.0 (1.0)**	0	0	1	0	0	**0.2 (0.4)**
A'	**0.98 (0.02)**	0.96	0.99	0.99	0.99	**0.98(0.01)**	0.93	0.98	0.97	0.99	0.98	**0.97 (0.02)**
Real words (/80)												
A|A	**38.4 (2.2)**	40	38	40	39	**39.3 (1.0)**	37	39	39	40	33	**37.6 (2.8)**
R|A	**21.5 (4.3)**	27	13	19	19	**19.5 (5.7)**	30	33	21	28	29	**28.2 (4.4)**
R|R	**18.5 (4.3)**	13	27	21	21	**20.5 (5.7)**	10	7	19	12	11	**11.8 (4.4)**
A|R	**1.6 (2.2)**	0	2	0	1	**0.8 (1.0)**	3	1	1	0	7	**2.4 (2.8)**
A'	**0.84 (0.05)**	0.83	0.90	0.88	0.87	**0.87 (0.03)**	0.72	0.75	0.85	0.83	0.62	**0.75 (0.09)**
Pseudowords (/80)												
A|A	**26.8 (8.0)**	37	36	36	37	**36.5 (0.6)**	3	10	14	0	3	**6.0 (5.8)**
R|A	**11.6 (5.9)**	21	12	8	15	**14.0 (5.5)**	1	7	7	1	6	**4.4 (3.1)**
R|R	**28.4 (5.9)**	19	28	32	25	**26.0 (5.5)**	39	33	33	39	34	**35.6 (3.1)**
A|R	**13.2 (8.0)**	3	4	4	3	**3.5 (0.6)**	37	30	26	40	37	**34.0 (5.8)**
A'	**0.77 (0.11)**	0.82	0.88	0.91	0.87	**0.87 (0.04)**	0.68	0.60	0.68	0.24	0.36	**0.51 (0.20)**

This table shows individual patient participants' responses in each of the main experimental conditions (to see individual control participant responses, please see **Appendix Table 8.2**). Note that 20 trials of each stimulus type (Added / Replaced) were presented for each isolated noise segment condition and 40 trials for each word condition; the maximum score in each cell is therefore 20 for segments and 40 for real words/pseudowords. Stimulus conditions were delivered in randomised order during the experimental session. A|A denotes that the participant correctly identified an 'Added' stimulus as 'Added'; R|A denotes that the participant incorrectly identified a 'Replaced' stimulus as 'Added' (i.e., phonemic restoration occurred); R|R denotes that the participant correctly identified a 'Replaced' stimulus as 'Replaced'; A|R denotes that the participant incorrectly identified an 'Added' stimulus as 'Replaced'. Blue shading represents individual patients whose performance fell above the 95th percentile for the healthy control group; Red shading represents individual patients whose performance fell below the 5th percentile for the healthy control group. AD, Alzheimer's disease; SD, standard deviation; svPPA, semantic variant primary progressive.

The *A'* values themselves are not sufficient enough to interpret the results in full, rather the direction of *c*, the criterion location, is crucial. Negative values of *c* indicate a bias towards responding 'Added' over 'Replaced' (i.e. phonemic restoration occurred in the stimuli), and positive values indicate a bias towards responding 'Replaced' over 'Added'. Values near zero indicate no particular bias towards one response category over the other.

Given the disparate group sizes, non-parametric Kruskal-Wallis tests were used to assess whether there was an effect of each separate diagnostic group on *A'* and/or *c* in each experimental condition. In addition, difference scores between *A'* and *c* were generated and analysed between the pseudoword and real word conditions. Finally, within each diagnostic group, *A'* and *c* scores were compared between carrier conditions directly using Friedman's tests. Where the omnibus test was significant, post hoc analyses were conducted using two-tailed Wilcoxon rank-sum tests to compare groups directly and understand the direction of the effect.

To characterize the consistency and variability of individual patient performance profiles relative to healthy controls, the 5^{th} and 95^{th} percentiles for the healthy control group were calculated and patients in each dementia group who performed below the 5^{th} percentile or above the 95^{th} percentile were identified.

3.5 RESULTS

Sensitivity (*A'*) and bias (*c*) values, as well as statistical results for all experimental conditions in each participant group, are presented in **Table 3.4** and **Figure 3.3**. **Table 3.3** and Appendix **Table 8.2** show individual raw scores in each experimental test and condition.

Table 3.4. Summary of participant group performance on phonemic restoration conditions

	Healthy controls	**svPPA**	**AD**	**Omnibus significance test**
A'				
Isolated noise segments	0.98 (0.02)	0.98 (0.01)	0.97 (0.02)	H(2)=1.79, p=0.408
Real words	0.84 (0.05)	0.87 (0.03)*	**0.75 (0.09)**	**H(2)=7.48, p=0.023**
Pseudowords	0.77 (0.11)	**0.87 (0.04)***	**0.51 (0.20)**	**H(2)=13.38, p=0.001**
Difference between real words and pseudowords	0.07 (0.10)	-0.00 (0.02)*	**0.24 (0.20)**	**H(2)=9.24, p=0.010**
c				
Isolated noise segments	-0.16 (0.24)	-0.37 (0.15)	**-0.42 (0.31)**	**H(2)=6.09, p=0.048**
Real words	-1.04 (0.36)	-1.06 (0.38)	-1.16 (0.31)	H(2)=0.79, p=0.675
Pseudowords	0.06 (0.51)	**-0.48 (0.23)***	**1.33 (0.65)**	**H(2)=14.11, p<0.001**
Difference between real words and pseudowords	-1.10 (0.63)	-0.58 (0.35)*	**-2.48 (0.81)**	**H(2)=12.09, p=0.002**

Mean (standard deviation) phonemic restoration measures of sensitivity (***A'***) and bias (***c***) are shown for each participant group and word / sound condition. ***A'*** values typically lie between 0.5 (indicating the participant was unable to discriminate between 'Replaced' and 'Added' stimuli) and 1 (indicating perfect discrimination); values below 0.5 indicate response confusion (see text). For the measure of bias or criterion location (***c***), negative values indicate a bias toward responding 'Added' (i.e., phonemic restoration) while positive values indicate a bias toward responding 'Replaced'. **Bold,** significantly different from the healthy control group; *significantly different from AD group. AD, patient group with Alzheimer's disease; svPPA, patient group with semantic variant PPA.

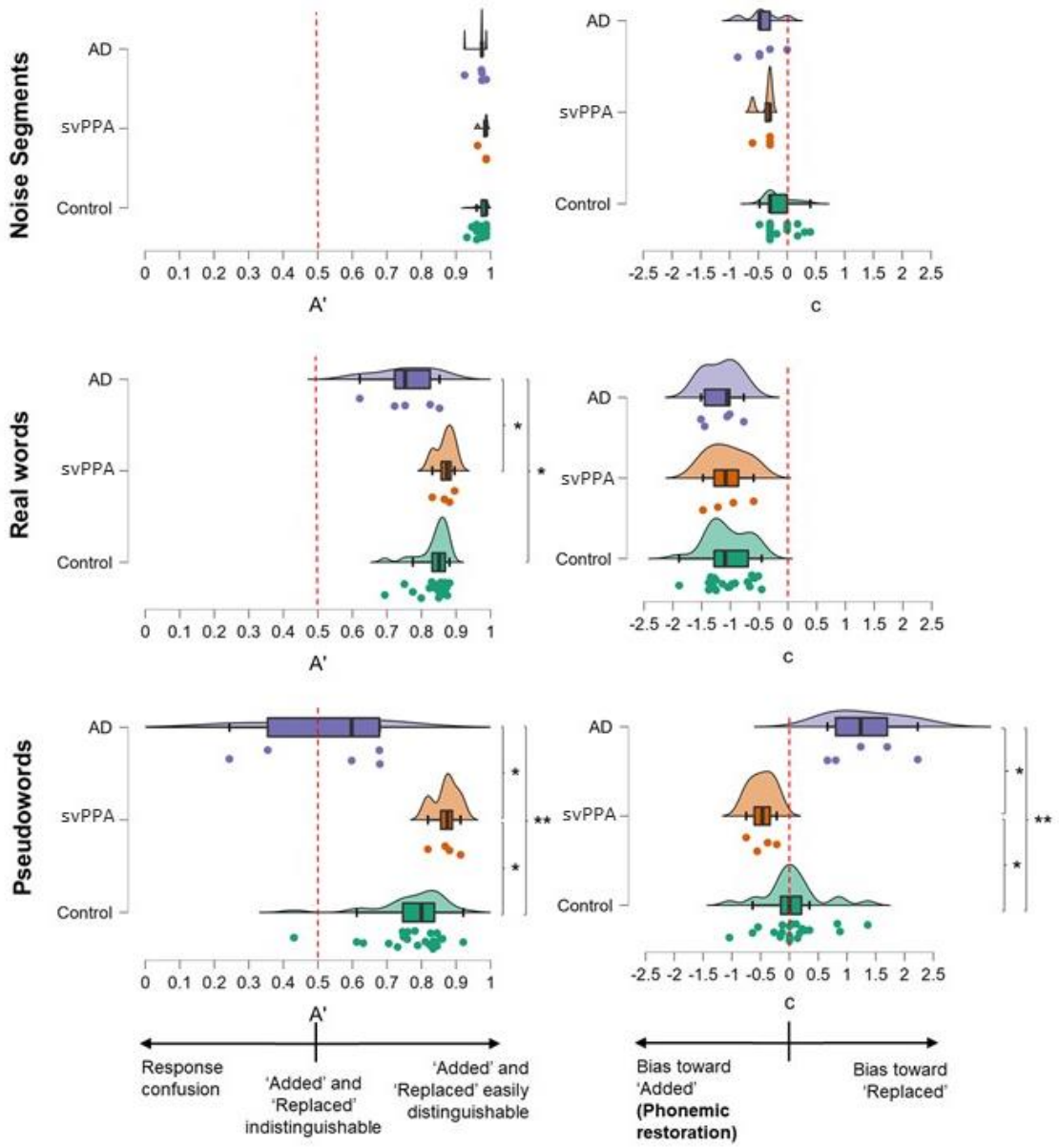


Figure 3.3. Summary of participant group profiles for response sensitivity and bias across experimental conditions.
Raincloud plots (Allen et al., 2019) of individual data for response sensitivity (left panels) and bias (right panels) for all experimental conditions and participant groups. Boxes represent the interquartile range, and whiskers indicate the overall range of values in each group; the vertical line in each box represents the median. The dots code values for individual participants. Sensitivity (***A'***) values typically lie between 0.5 (indicating the participant was unable to discriminate between 'Replaced' and 'Added' stimuli) and 1 (indicating perfect discrimination); values below 0.5 indicate response confusion (see text in the Methods section). For the measure of bias or criterion location (***c***), negative values indicate a bias toward responding 'Added' (i.e., phonemic restoration) while positive values indicate a bias toward responding 'Replaced'. AD, participant group with Alzheimer's disease; Control, healthy control participant group; svPPA, participant group with semantic variant primary progressive aphasia. *significant at $p<0.05$; **significant at $p<0.01$.

3.5.1 General characteristics of participant groups

Participant groups did not differ significantly in sex, age, handedness, years of education, or peripheral hearing function (all p>0.05, **Table 3.1**). Patient groups did not differ in the mean symptom duration (p=0.974). General neuropsychological profiles were in keeping with the syndromic diagnosis for each patient group (**Table 3.1**).

3.5.2 Word conditions

3.5.2.1 Real words

The effect of diagnosis on *A'* was significant (**Table 3.4**). The AD group had significantly lower median *A'* than both healthy controls (z=-2.22, p=0.027) and the svPPA group (z=-2.21, p=0.028). There was no significant difference between the healthy controls and the svPPA group (z=-1.42, p=0.155). All groups showed a clear bias (*c*) towards reporting words as 'Added' rather than 'Replaced', but there was no significant effect of diagnosis on *c*.

3.5.2.2 Pseudowords

There was a significant effect of diagnosis on *A'* (**Table 3.3**). The AD group, performing at chance, had significantly lower median *A'* than both healthy controls (z=-3.00, p=0.003) and the svPPA group (z=-2.45, p=0.014). The svPPA group had significantly higher *A'* than the healthy control group (z=2.20, p=0.03).

There was a significant effect of diagnosis on *c*. Healthy controls showed essentially no bias (*c*) in reporting pseudowords. Compared with healthy controls, the AD group showed a significantly greater bias toward reporting pseudowords as 'Replaced' over 'Added' (z=-3.00, p=0.003), while the svPPA group showed a significantly greater bias towards reporting pseudowords as 'Added' over 'Replaced' (z=2.42, p=0.016).

3.5.3 Comparisons of differences between word conditions across groups

There was a significant overall effect of diagnosis on comparing the *A'* value difference between real word and pseudoword conditions (**Table 3.3**). This was driven by a greater difference of *A'* value between word conditions in the AD group than in the healthy control (z=2.43, p=0.015) or svPPA groups (z=2.45, p=0.014). The performance of the svPPA and healthy control groups did not differ significantly (z=1.64, p=0.102).

The value of *c* also differed significantly between real word and pseudoword conditions according to diagnosis (**Table 3.3**). This was driven by a greater difference between word conditions in the AD group than in the healthy control (z=3.06, p=0.002) or svPPA (z=2.45, p=0.014) groups. Response bias in the svPPA and healthy control groups did not differ significantly (z=-1.64, p=0.102).

3.5.4 Perceptual control task on isolated noise segments

A' did not differ significantly across diagnoses and *A'* was uniformly high across participant groups, indicating that the 'Added' and 'Replaced' isolated noise segments were easily discriminable acoustically (**Figure 3.3** and **Table 3.4**). The effect of diagnosis on *c* was significant, with the AD group showing a greater bias towards reporting that there was an added letter in the noise than without (z=2.10, p=0.036).

3.5.5 Comparisons between stimuli conditions within groups

All groups showed significantly lower *A'* both for real words (controls: $\chi^2(1)=22.00$, $p<0.001$; svPPA: $\chi^2(1)=4.00$, $p=0.046$; AD: $\chi^2(1)=5.00$, $p=0.025$) and pseudowords (controls: $\chi^2(1)=22.00$, $p<0.001$; svPPA: $\chi^2(1)=4.00$, $p=0.046$; AD: $\chi^2(1)=5.00$, $p=0.025$), in comparison to the isolated noise segment. *A'* was significantly lower for pseudowords than real words in the healthy control group

($\chi^2(1)=4.55$, p=0.033) and AD group ($\chi^2(1)=5.00$, p=0.025), but not the svPPA group ($\chi^2(1)=0.00$, p=1.00).

All groups also showed significantly stronger bias towards reporting 'Added' (i.e. negative *c*) for real words compared with isolated noise segments (controls: $\chi^2(1)=22.00$, p<0.001; svPPA: $\chi^2(1)=4.00$, p=0.025; AD: $\chi^2(1)=5.00$, p=0.025). The AD group showed a significantly stronger bias towards reporting 'Replaced' (i.e. positive *c*) for pseudowords than isolated noise segments ($\chi^2(1)=5.00$, p=0.025). Response bias did not differ between the pseudoword and isolated noise segment conditions in healthy controls ($\chi^2(1)=1.64$, p=0.201) or the semantic dementia group ($\chi^2(1)=4.00$, p=0.046). All groups showed a significantly stronger bias for 'Added' for real words than pseudowords (controls: $\chi^2(1)=22.00$, p<0.001; svPPA: $\chi^2(1)=4.00$, p=0.046; AD: $\chi^2(1)=5.00$, p=0.025).

3.5.6 Individual patient performance profiles

For the isolated noise segment condition, one patient in the AD group (20% of the group) had an *A'* value below the healthy control 5th percentile. For the real word condition, two patients with svPPA (50% of the group) had *A'* values above the healthy control 95th percentile; whilst one patient with AD (20% of the group) had an *A'* value below the control 5th percentile. For the pseudoword condition, three patients with svPPA (75% of the group) had *A'* values above the healthy control 95th percentile, whilst two patients with AD (40% of the group) had *A'* values below the control 5th percentile.

3.6 DISCUSSION

In two distinct types of dementias, both groups showed evidence of phonemic restoration for real words. This was particularly more marked in patients with AD than in healthy controls or patients with svPPA. The results seen in the AD group

can be considered surprising, especially since past studies have suggested that the underlying mechanism for phonemic restoration is similar to auditory scene analysis (Başkent et al., 2010). Therefore, with auditory scene analysis being impaired in patients with AD (Goll et al., 2012), it could be theorised that the AD group would also have impaired phonemic restoration. However, as the results present otherwise, their retained phonemic restoration mechanism could potentially be a result of the lexical top-down mechanism associated with phonemic restoration being partly compensatory in patients with AD. Another possibility, similar to the results from participants with developmental dyslexia, is that the impaired phonological processing of speech sounds may reflect overly plastic and consequently unstable speech sound representations, leading to deficits in separating noise from the intended acoustic signal (Del Tufo & Myers, 2014).

In the healthy control group, not only did they show retained phonemic restoration, but also greater phonemic restoration for real words (c=-1.04) than for pseudowords (c=0.06). This profile of retained phonemic restoration modulated by top-down lexical context effects is in line both with prevailing models of auditory word processing (Samuel, 1997) and with previous work in older listeners using alternative phonemic restoration paradigms (Bologna et al., 2018; Jaekel et al., 2018; Saija et al., 2014).

The group profiles differed more substantially for phonemic restoration of the pseudowords. Patients with AD showed a marked tendency towards perceiving noise segments as replacing phonemes, and thus, could be interpreted as a ‘rejection’ of phonemic restoration. In contrast, patients with svPPA performed comparably on discrimination of noise conditions in both pseudowords and real words. Even in comparison to healthy controls, svPPA performed more

accurately for discriminating the noise conditions within pseudowords, whereas healthy controls not only showed less accurate discrimination between noise conditions for the pseudowords than for real words but also had no clear bias towards phonemic restoration.

A key factor to consider for these results is the processing of the 'bottom-up' auditory information for each group, and therefore the perceptual control task needs to be taken into account. Both the patient groups and healthy controls were highly accurate in discriminating whether or not isolated noise segments contained speech sounds. Therefore, the less accurate performance across groups in the word conditions is unlikely to be reflected in spectrotemporal feature discriminability (as the features were similar in the isolated noise segment and word conditions) or the proximity of additional spectrotemporal information surrounding the 'target' consonant (since performance differed between the real word and pseudoword conditions).

Taken together, the performance profiles in these dementia syndromes illuminate the underlying brain mechanisms of phonemic restoration. The findings in AD and svPPA are consistent with a phonemic restoration model in which phonological representations (likely situated in the posterior superior temporal cortex) interact with a modulatory, top-down mechanism of semantic prediction and disambiguation (likely mediated by more anterior cortical regions) (Shahin et al., 2009; Sunami et al., 2013). As in the healthy brain, the interaction of phonological and semantic mechanisms primes the 'repair' of real words over pseudowords (A. G. Samuel, 1981). In AD, the phonemic representations are damaged as part of a more general impairment of auditory object parsing (H. L. Golden et al., 2015; Goll et al., 2012), whereas the top-down semantic mechanism mediating lexical recognition is less impaired. Therefore, with an overriding effect of lexical

prediction, there is a strong ‘repair’ of real words and a rejection of pseudowords. Contrastingly, in svPPA, the profile of phonological and semantic effects is the opposite of AD. Thus, the semantic disadvantage of pseudowords relative to real words is no longer present considering the damaged lexical predictive mechanism in svPPA does not override the intact bottom-up phonological processing.

The phonemic restoration mechanism could be potentially understood through the matching of incoming speech signals to a stored lexical ‘template’ or through Gestalt continuity (the sensory expectation that phonological patterns corresponding to words tend to be spectrotemporally continuous) (Shahin et al., 2009). However, the strong word category effect in the AD group here suggests that top-down lexical template matching plays a dominant role in phonemic restoration. The comparable performance in the real word and pseudoword conditions in the svPPA group corroborates this interpretation as well.

Whereas perceptual completion of real words might be based on general lexical familiarity, the ‘rejection’ of pseudowords (seen in the AD group in comparison to healthy controls) depends on a more fine-grained semantic computation and can be interpreted as an ongoing ‘search’ of the stored semantic lexicon. If this process is deficient (like in svPPA), the pseudoword processing becomes relatively more dependent on the still intact bottom-up perceptual processing. The profile observed in the svPPA group here suggests that lexical predictive decoding is normally mediated by the anterior temporal lobe. While neuroanatomical models of phonemic restoration have not foregrounded this brain region (Sunami et al., 2013), it has been implicated in the predictive decoding of word identity (Cope et al., 2020).

This experiment is a preliminary investigation, as data collection was stopped prematurely due to the COVID-19 pandemic. Therefore, a major limitation seen here in this experiment is the small patient cohort size. It is likely that this resulted in limited power to detect disease effects, and more varied performance among individual patients (as well as healthy control participants; see **Figure 3.3** and **Appendix Table 8.2**). The individual variability seen could also be the result of other factors (see **Chapter 1.3.3**) that need to be further investigated and stratified. For example, a key factor amongst patient groups may be disease stage or severity. While it does not seem to appear to have affected phonemic restoration in this study, the traditional measures (e.g., symptom duration) utilised here are problematic in these syndromes, especially when applied across to compare diseases. Thus, any generalisation to group-level signatures should be considered cautiously.

What is important to consider is that the nature and factors influencing the phonemic restoration mechanism have not yet been fully defined. For example, stimulus properties such as manner of articulation and placement of target phonemes might affect phonemic restoration (A. G. Samuel, 1981), as well as interact with peripheral hearing function, attention, and other cognitive processes that are potentially altered with healthy ageing (see **Chapter 1.3.3.6**).

A related issue was the word lists, in particular elements of retained vocabulary in patients with svPPA, as it is unlikely that comprehension of the real word list was affected *uniformly* among the patients with svPPA studied here. Varying levels of lexical-semantic decoding may have influenced the individual patient profiles. On the other hand, the use of personalised stimulus lists would greatly complicate the interpretation of disease group profiles of phonemic restoration,

as the resulting stimuli would also vary widely in acoustic characteristics, making comparisons non-standardised.

More fundamentally, thresholds for the perception of speech-in-noise and the executive processes that guide perceptual decisions in interpreting degraded speech signals are likely to vary in dementia syndromes and between different neurodegenerative disorders as well (Jiang et al., 2021; Johnson et al., 2021). Therefore, a complete picture of phonemic restoration in these diseases will entail a better understanding of these processes. I hope that this experiment which is now published can be used to motivate and inform future work and fully characterise the processes of phonemic restoration in neurodegenerative disease.

Regardless of major caveats, it is encouraging that a neural mechanism for 'repairing' degraded speech can be preserved or relatively enhanced in certain dementias, as well as present itself as a mechanism in stratifying different pathologies. The striking polarity of phonemic restoration effects between real words versus pseudowords in the AD group reflects a compensatory mechanism that tends to maintain the intelligibility of speech, despite impaired auditory scene processing (Bouma & Gootjes, 2011; Gates et al., 2008; H. L. Golden et al., 2015; Goll et al., 2012; Idrizbegovic et al., 2013; Utoomprurkporn et al., 2020). Future work should test this hypothesis and extend the present findings to larger and more diverse patient cohorts, addressing the limits and influences on phonemic restoration in neurodegenerative disease and establishing its neural basis using functional neuroimaging. Within the field of neurolinguistics, combining neuroimaging, and studying different disease profiles, the phonemic restoration mechanism can potentially explore and parse out brain regions sensitive to certain speech-related auditory processes, such as articulation in phonemes

(which has been increasingly implicated in the STG (Bhaya-Grossman & Chang, 2021; Lakretz et al., 2021)). Dynamic neuroanatomical techniques such as MEG could also dissect the time courses of the component neural mechanisms that underpin phonemic restoration and reveal how these mechanisms contribute to a final percept and behavioural decision.

4 EXPERIMENT 2: NOISE-VOCODED VERBAL MESSAGES

4.1 SUMMARY

In **Chapter 1**, I summarised the complexity of degraded speech, noting that not only are verbal messages often "degraded" by competing sounds (as seen with phonemic restoration: **Chapter 3**) and/or vocal idiosyncrasies, but our speech can also be degraded through the carrier of the verbal message, like a telephone, or increasingly used, video-conferencing lines. Therefore, the comprehension of *all* forms of degraded speech demands intense computations across distributed neural networks, and is therefore likely to present different challenges to patients with different neurodegenerative pathologies. However, this issue has not been addressed systematically.

In this experiment, I studied the processing of degraded speech signals in a cohort of patients representing major variants of primary progressive aphasia (PPA) and patients with typical Alzheimer's disease (AD), and healthy age-matched controls. As a model paradigm for the degraded 'noisy' speech signals of daily life, I used noise-vocoded speech stimuli in this experiment. Noise-vocoding artificially divides the speech signal into a variable number of frequency channels constituted from amplitude-modulated white noise. This allows channels to be manipulated by increasing intelligibility with more channels or decreasing intelligibility with fewer channels. I investigated the impact of noise-vocoding on the recognition of spoken three-digit numbers and used psychometric modelling to ascertain the threshold of noise-vocoding channels required for 50% intelligibility by each participant. Further, I assessed the associations of noise-vocoded speech intelligibility threshold with general

demographic, clinical, neuropsychological characteristics, relevant daily hearing measures, and regional grey matter of patients' brain MR images.

Compared with healthy controls, all patient groups showed normal comprehension of clear speech, but a significantly elevated intelligibility threshold for noise-vocoded speech (i.e., needing more spectral detail to perceive 50% of the stimuli), particularly in lvPPA and nfvPPA, and significantly higher in AD than in svPPA. The intelligibility threshold did not correlate with measures of peripheral hearing or clear speech perception but correlated with overall disease severity. Neuroanatomically, after correcting for multiple voxel-wise comparisons in pre-defined regions of interest, impaired noise-vocoded speech comprehension across dementia syndromes was significantly associated with atrophy of left planum temporale, angular gyrus, and anterior cingulate gyrus. Taken together, the findings suggest that the comprehension of noise-vocoded speech captures a central process relevant to real-world hearing and communication in major dementia syndromes, with novel diagnostic and therapeutic implications.

The work presented here in **Chapter 4** (Experiment 2) has been published as a pre-print (https://doi.org/10.1101/2022.12.05.22283108).

4.2 INTRODUCTION

Successful communication in day-to-day situations typically requires the ability to understand spoken messages under non-ideal listening conditions. Because speech signals are critical for communication, perception of degraded speech is likely to be a functionally relevant index of auditory scene analysis and as an index of top-down compensatory mechanisms for ambiguous input conducted in daily life. This process, normally automatic and relatively effortless, is impaired in AD and PPA (Gates et al., 2011; Johnson et al., 2021; Lin et al., 2011).

As hearing impairment has recently been identified as a major risk factor for dementia and a driver of cognitive decline (Griffiths et al., 2020; Lin et al., 2011; Livingston et al., 2017), it is likely that measures of central hearing, such as the processing of degraded speech signals, is equally pertinent to peripheral hearing measures (Gates et al., 2008; Gates et al., 2011; Johnson et al., 2021) (see **Chapter 1.1.1**). Large cohort studies have found impaired comprehension of degraded messages as a harbinger of dementia (Gates et al., 2011; Pronk et al., 2019; Stevenson et al., 2022), building on the deficits in auditory and speech perception seen with a diagnosis of AD and PPA presented in **Chapter 1.5.6**. However, the specific neural mechanisms responsible, and the differentiating symptomatic effects in AD and PPA on central hearing (degraded speech signals), have not yet been fully clarified.

There are several grounds on which the processing of degraded speech may be especially vulnerable to neurodegenerative pathologies (see **Chapter 1**). Neuroanatomically, the processing of degraded speech signals engages distributed neural networks in peri-Sylvian and posterior temporo-parietal cortices. These brain networks are targeted preferentially in PPA, particularly in nfvPPA and lvPPA syndromes (Chris J. D. Hardy et al., 2017; C. J. D. Hardy, J. L. Agustus, et al., 2017; Hardy, Marshall, et al., 2018; Jiang et al., 2021). Considering that comprehension of degraded speech signals depends on precise yet dynamic integration of information across neural circuitry (Gates et al., 2010; Holmes et al., 2020; Jiang et al., 2021; Johnson et al., 2021; Lin et al., 2011), neurodegenerative pathologies are likely to affect these computations early and profoundly.

One widely used technique for altering speech signals experimentally is noise-vocoding (see **Chapter 1.3.2.7**), where a speech signal is divided digitally into

discrete frequency bands (‘channels’), each filled with white noise and modulated by the amplitude envelope of the original signal (Shannon et al., 1995). This procedure degrades the spectral content of the speech signal while preserving its overall longer-range temporal structure. Using noise-vocoding, intelligibility can be manipulated in a controlled manner: fewer channels is equivalent to less spectral detail and lead to less intelligibility of the speech, versus more channels is equivalent to more detail and increased intelligibility. As an exemplar of acoustic degradation based on the reduction of spectral information, noise-vocoding is applicable to a variety of daily listening scenarios requiring decoding of ‘noisy’ speech signals (e.g., a poor telephone or video-conferencing line). In contrast to speech-in-noise perception, comprehension of noise-vocoded speech depends intrinsically on auditory object (phonemic) decoding rather than selective attention (e.g., intrinsic degradation rather than extrinsic degradation (Mattys et al., 2012)). Further, noise-vocoding offers the substantial advantage of generating a quantifiable threshold for the intelligibility of the degraded speech signal, based on the number of vocoding channels. This potentially allows for a more sensitive, graded, and robust determination of deficit, enabling comparisons between diseases, tracking disease evolution, and assessing the impact of therapeutic interventions.

Noise-vocoding has been applied previously in a joint behavioural and MEG study on nfvPPA, to assess the brain mechanisms that mediate comprehension of degraded speech in the context of relatively focal cerebral atrophy (Cope et al., 2017). Patients in this study relied more on cross-modal cues to disambiguate vocoded speech signals and had inflexible predictive decoding mechanisms, instantiated in the left inferior frontal cortex. Noise-vocoding has not yet been exploited as a tool to study degraded speech perception in other

neurodegenerative syndromes. More generally, the cognitive and neuroanatomical mechanisms that mediate the processing of noise-vocoded speech and their clinical resonance in this disease spectrum remain poorly defined.

In this experiment, I addressed the comprehension of acoustically noise-vocoded spoken messages in patients with typical AD and PPA, referenced to healthy older listeners. I assessed how the understanding of noise-vocoded speech relates to other demographic and disease characteristics, as well as predicting real-world hearing functions in daily life situations that demand the processing of acoustically altered speech signals. I also looked at structural neuroanatomical associations of noise-vocoded speech intelligibility using voxel-based morphometry on patients' brain scans.

4.3 KEY PREDICTIONS

Based on available evidence from other noise-vocoding and degraded speech stimuli paradigms in AD and PPA (Cope et al., 2017; Chris J. D. Hardy et al., 2017; C. J. D. Hardy, J. L. Agustus, et al., 2017; Hardy, Bond, et al., 2018; Jiang et al., 2022), my hypotheses are:

H_1: Patients with AD and PPA will have an elevated threshold for comprehending vocoded speech compared with healthy controls, particularly nfvPPA and lvPPA.

H_2: Elevated intelligibility threshold in the patient groups will correspond with difficulties in daily life listening

H_3: Elevated intelligibility will anatomically correlate with regional grey matter atrophy in the left posterior superior temporal, inferior parietal, and inferior frontal cortices

4.4 Materials and Methods

4.4.1 Participants

In this experiment, 19 patients with typical AD, nine patients with lvPPA, 10 patients with nfvPPA, 12 patients with svPPA, and 25 healthy older control participants were recruited (see **Chapter 2.1** and **Appendix: Table 8.1** for participant breakdown per Chapter).

Due to the Covid-19 pandemic, 29 participants (four healthy controls, nine patients with AD, six with lvPPA, six with nfvPPA, and four with svPPA) were assessed remotely using Labvanced® (Finger et al., 2017) through a video link (see the details and descriptions in **Chapter 2.4** and **2.5.1**).

No participant had abnormal peripheral hearing other than age-related hearing loss (see **Chapter 2**). To assess daily-life hearing function in patients, the Modified Amsterdam Inventory for Auditory Disability and Handicap (mAIAD) (see **Chapter 2.3.4**) was adapted for completion by the primary caregiver or another close informant for each patient.

4.4.2 Experimental stimuli

To minimise top-down linguistic/semantic cues in speech (as likely to unequally affect svPPA than other disease groups), three-digit numbers were chosen as the target stimuli. Lists of 50 different three-digit numbers (of the form: ‘five hundred and eighty-seven’, not: ‘five-eight-seven’) were recorded by two adult female speakers in a Standard Southern British English accent with neutral prosody.

They were recorded on Audacity (v2.2.3) using a condenser microphone with a pop shield in a sound-proof booth.

Speech recordings were noise-vocoded using Matlab® (vR2019b) (https://uk.mathworks.com/) to generate acoustically altered stimuli with a prescribed level of degraded intelligibility (see **Figure 4.1** for spectrograms). The script used was written by Chris Darwin (Darwin).

The logspace function in Matlab was used to calculate log spacing between 50 and 8000Hz, which corresponded to logarithmically-spaced frequency bands that served as the basis for the noise-vocoding algorithm. The average (RMS) stimulus intensity was constant for all stimuli, as fixed in Matlab. All stimuli were windowed with 20ms onset-offset temporal ramps to prevent click artefacts. The algorithm was run iteratively to generate speech stimuli between one and 24 frequency bands ('channels'), sampling at each integer number of channels.

Considering that the vocoding intelligibility threshold for younger normal listeners is typically around three 'channels' (Shannon et al., 1995), the threshold is yet confirmed in patients and therefore this range was designed to accurately capture even markedly abnormal psychometric functions in the patient cohort.

Within each channel (ranging from one to 24), four three-digit number stimuli were presented. The final stimulus list comprised of 100 different spoken three-digit numbers: four non-vocoded (clear speech) and 96 noise-vocoded three-digit numbers.

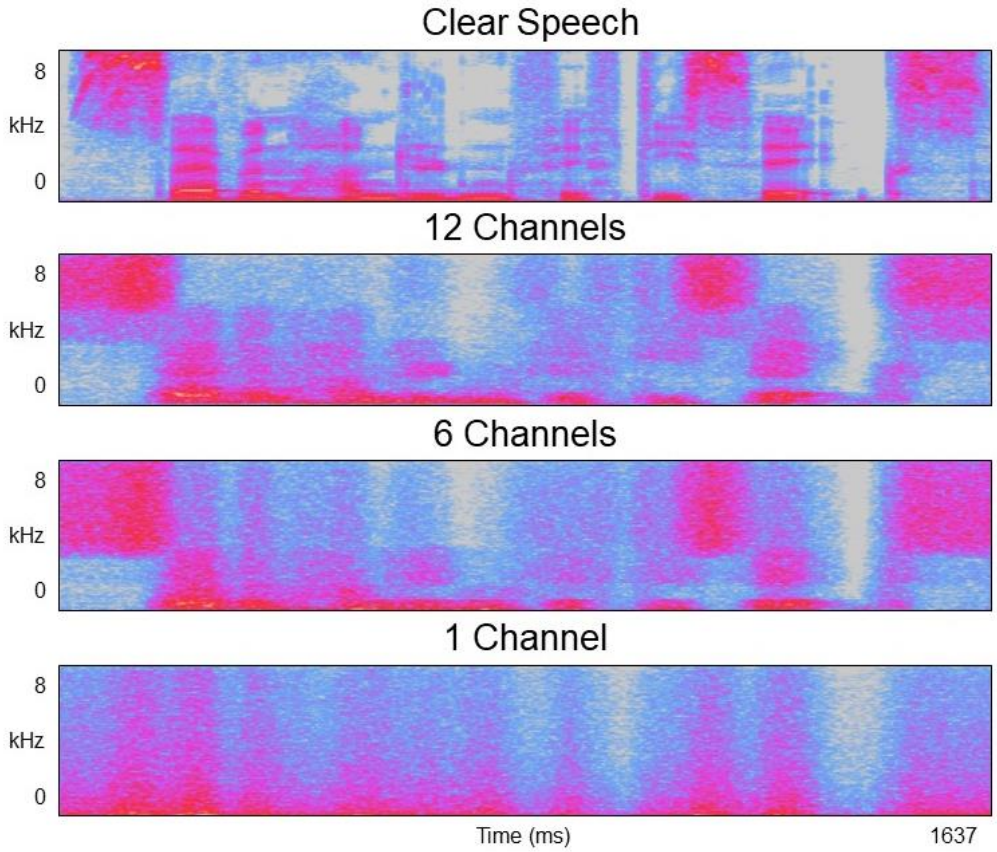


Figure 4.1. Spectrograms of clear and noise-vocoded speech resampled with differing numbers of discrete frequency bands (12, 6, and 1 channel).
The verbal message in each spectrogram is the spoken number 'seven hundred and fifty-six'. The spectrogram provides a visual indication of how the energy in different frequency bands of the speech signal (plotted on the y-axis) changes over time (plotted on the x-axis). The color scale used allows for blue to correspond to low amplitudes, and red corresponding to high amplitudes. Reducing the number of channels (frequency bands) reduces the amount of spectrotemporal fine structure in the speech signal.

4.4.3 Procedure

I administered the stimuli either in-person in a quiet room via Audio-Technica ATH-M50x headphones at a comfortable fixed listening level (at least 70 dB), or remotely via Labvanced and shared through a video link (see **Chapter 2.4** and **Table 4.1**). To familiarise participants with the experimental procedure, I first asked them to repeat five three-digit numbers spoken by myself (not included in the experimental session). Then, before presenting the experimental stimuli, I advised participants that the numbers they would hear will increase in difficulty,

Table 4.1. General demographic, clinical, and neuropsychological characteristics of all participant groups

Characteristic	Controls	AD	lvPPA	nfvPPA	svPPA
Demographic and clinical					
No. M:F	14:11	15:4	8:1	8:2	7:5
Age, years	68.3 (6.6)	70.1 (8.4)	70.4 (6.5)	72.7 (3.7)	63.1 (8.4)
Handedness (R/L/A)	21/1/1[b]	18/1/0	8/1/0	10/0/0	11/1/0
Education (y)	16.1 (2.7)	15.4 (3.8)	15.0 (3.1)	15.1 (2.6)	15.6 (2.1)
Symptom duration (y)	NA	5.9 (3.0)	6.6 (5.4)	3.3 (1.2)	5.4 (2.5)
Best ear average*	17.1 (8.7)[l]	27.7 (10.9)[i]	19.0 (10.7)[e]	29.3 (3.3)[e]	23.8 (8.1)[c]
Tested in-person/remote	21/4	10/9	3/6	4/6	8/4
Taking donepezil and/or memantine (%)	N/A	81.25%[c]	83.33%[c]	16.67%[d]	0.00%[a]
General intellect					
MMSE (/30)	29.8 (0.6)[f]	**20.4 (7.8)**	**22.7 (7.5)[a]**	26.5 (0.7)[b]	**22.9 (5.1)**
T-MMSE (/27)	26.1 (1.8)	**17.8 (4.5)**	**21.5 (4.5)[a]**	24.3 (2.3)[1]	24.0 (1.4)[1]
Episodic memory					
RMT Faces (Short) (/25)	23.8 (2.5)[m]	**16.1 (3.3)[i]**	21.4 (3.7)[d]	22.8 (3.5)[1d]	19.2 (3.7)[g]
RMT Faces (Long) (/50)	41.7 (3.7)[k]	**29.6 (5.6)[i]**	**29.0 (6.9)[e]**	35.5 (5.0)	**30.9 (3.4)[d]**
Working memory					
Digit span forward (max)	6.6 (1.0)[h]	5.8 (1.4)	**4.6 (1.4)**	5.6 (1.3)	6.7 (1.0)
Digit span reverse (max)	5.2 (1.2)[h]	**3.2 (1.4)**	**3.8 (1.3)**	**3.8 (1.9)**	4.9 (1.6)
Executive function					
WASI Matrices (/32)	26.8 (2.7)[h]	**11.8 (8.8)[b]**	23.1 (5.4)[1]	**19.1 (9.4)**	24.1 (6.7)[1]
Letter fluency (total)	15.9 (5.4)[i]	10.9 (5.9)[b]	8.9 (4.0)[a]	9.0 (9.2)[c]	**7.4 (6.4)**
Category fluency (total)	24.1 (6.3)[i]	**11.4 (6.7)[b]**	**11.1 (6.4)[a]**	**15.4 (11.6)[c]**	**6.7 (5.7)**
Auditory input processing					
PALPA-3 (/36)	34.6 (1.7)[j]	NA	31.2 (5.8)[c]	31.4 (5.5)	33.7 (2.3)
Speech repetition					
Polysyllabic words (/45)	44.0 (1.6)[h]	NA	**41.3 (3.8)[b]**	**38.9 (8.0)[a]**	**40.7 (5.4)**
Short sentences (/10)	9.5 (0.9)[j]	NA	**5.3 (1.6)[b]**	**6.6 (2.8)**	**7.5 (2.1)**
Other language skills					
GNT (/30)	25.8 (2.5)[h]	**13.0 (7.2)**	**10.9 (6.9)[a]**	**18.4 (8.8)**	**1.4 (4.3)**
BPVS (/150)	147.9 (2.1)[h]	**135.3 (23.4)[2]**	146.0 (3.3)[2a]	**127.5 (46.3)[2]**	**73.5 (50.8)**
PALPA-55 (/24)	23.5 (1.2)[j]	NA	**19.0 (4.5)[b]**	**19.8 (5.1)**	**18.8 (6.1)**
Other skills					
GDA calculation (/24)	14.8 (5.2)[h]	**5.1 (4.8)[c]**	**6.1 (4.4)[b]**	**7.0 (5.5)[a]**	10.2 (6.4)[a]
VOSP Object Decision(/20)	18.9 (1.5)[h]	**14.1 (3.6)**	**17.00 (1.4)[b]**	16.4 (4.6)[1a]	**15.5 (3.8)[a]**

Mean (standard deviation) values and raw scores are presented (maximum value possible in parentheses) unless otherwise indicated; significant differences from healthy controls ($p<0.05$) are in **bold**; [1]significantly different to AD ($p<0.05$); [2]significantly different to svPPA ($p<0.05$). See **Chapter 2** for details concerning the 'best ear average' measure. For remote participants, the Bamford-Kowal-Bench (BKB) sentence measure was used to assess peripheral hearing – most participants (97%) performed at ceiling, and none were rejected based on their BKB performance. Participants assessed in-person did the MMSE and RMT Faces (Long), while those assessed remotely did the T-MMSE and RMT Faces (Short). A, ambidextrous; AD, patient group with typical Alzheimer's disease; BPVS, British Picture Vocabulary Scale; Controls, healthy older control group; Digit span forward/reverse, maximum digit span recorded; F, female; GDA, Graded Difficulty Arithmetic; GNT, Graded Naming Test; L, left; lvPPA, patient group with logopenic variant primary progressive aphasia; M, male; MMSE, Mini-Mental State Examination; NA, not available/applicable; nfvPPA, patient group with nonfluent/agrammatic variant primary progressive aphasia; PALPA, Psycholinguistic Assessments of Language Processing in Aphasia; R, right; RMT, Recognition Memory Test; svPPA, patient group with semantic variant primary progressive aphasia; T-MMSE, tele-MMSE; VOSP, Visual Object and Space Perception; WASI, Wechsler Abbreviated Scale of Intelligence. [a] missing data for 1 participant; [b] missing data for 2 participants; [c] missing data for 3 participants; [d] missing data for 4 participants; [e] missing data for 6 participants; [f] missing data for 7 participants; [g] missing data for 8 participants; [h] missing data for 9 participants; [i] missing data for 10 participants; [j] missing data for 12 participants; [k] missing data for 13 participants; [l] missing data for 15 participants; [m] missing data for 21 participants.

but to guess the number if ever uncertain. Stimuli were presented in order of progressively decreasing channel number (intelligibility): starting at clear speech, then 24 channels to one vocoding channel. In each experimental trial, the task was to repeat the number (or as many of the three digits that the participant could identify). Participants were allowed to write down the number they heard rather than speaking if preferred. Even if imperfectly articulated, the intended target digit was accepted and no time limit was imposed. I recorded the responses for offline analysis and provided no feedback about performance to the participants.

4.4.4 Analysis of Data

Data was analysed in Matlab® (vR2019b) and R® (v4). See details on statistical analyses conducted in **Chapter 2.6**.

Performance profiles in seven healthy control participants who participated in the experiment both in person and remotely were very similar (V=11, p=0.688, see **Figure 4.2**), as well as a non-significant difference between the in-person and remote patient cohort (W=295.5, p=0.749), therefore justifying combining participants tested in person and remotely in the main analysis.

Identification of noise-vocoded spoken numbers was scored according to the number of digits correct for each three-digit number (e.g., if the target number was '587' and the participant responded '585', they would score two points on that trial).

Figure 4.2. Comparison of healthy older controls' performance on comprehension of noise-vocoded speech for in-person versus remote testing sessions
Seven healthy older control participants performed the noise-vocoded spoken number identification task both in-person at the research centre and remotely in their home environments, approximately 20 months later. Boxes represent the interquartile range, and whiskers indicate the overall range for each group; the horizontal line in each box represents the median. Threshold differences between the two sessions were non-significant (In person mean = 3.105; Remote mean = 3.003; p=0.69).

As three digits were presented on every trial, this system yielded a total of 12 (four trials x three digits) data points for each vocoding channel number per participant.

As the perceptual effect of noise-vocoding scales is exponential (i.e., the increase in intelligibility for normal listeners is much greater between two to four channels, than 20 and 24 channels), I applied a logarithmic (base 2) transformation to the data. The resulting data was then modelled using a Weibull sigmoid, a widely used function for fitting logarithmically scaled data (Schütt et al., 2016). I created psychometric curves for individual participants and a mean curve was created for each diagnostic group using the Matlab® psignifit package (Schütt et al., 2016).

For each function, the following parameters are given: (1) noise-vocoded speech intelligibility threshold (the number of vocoding channels at which 50% identification of noise-vocoded numbers was achieved, taking into account the lambda and gamma values at the upper and lower performance asymptotes, respectively), (2) the slope of the function at the threshold point, and (3) lambda (the lapse rate, or the number of incorrect responses at maximum performance asymptote). The gamma (the lower asymptote or guess rate) output was not considered in this analysis due to the responses being continuous and not restricted in choices.

I used Spearman's correlations to assess the relationship of noise-vocoded speech intelligibility threshold to forward digit span, a metric of each participant's overall ability to repeat (short-term memory and articulation) spoken numbers, over the whole patient cohort. I also conducted correlations to assess the relationship of intelligibility threshold with general demographic (age, sex, education), clinical (symptom duration, combined MMSE score), executive

performance (WASI Matrix), and auditory perceptual (PALPA-3, pure-tone audiometry) measures (where available) over the combined patient cohort.

Additionally, I assessed the relationship between noise-vocoded speech intelligibility threshold and daily life hearing functions. This was achieved by taking the combined responses of questions relating to hearing for communication in quiet and in background noise in the mAIAD (see **Table 4.2**) across the whole patient cohort. A few questionnaire items were excluded due to difficulties in interpreting in the target patient groups (reasons given in the notes of **Table 4.2**).

Table 4.2. Real-world hearing questionnaires and correlations with noise-vocoded speech intelligibility threshold in the patient groups

Hearing domain	Questions included	Reason for exclusion (if relevant)
Quiet	**Q2.** Can you carry on a conversation with someone in a quiet room?	
	Q8. Can you carry on a telephone conversation in a quiet room?	
	Q14. Can you understand the presenter of the news on TV?	
	Q19. Can you understand the presenter of the news on the radio?	
Noise	**Q1.** Can you understand a shop assistant in a crowded shop?	
	Q7. Can you carry on a conversation with someone in a crowded meeting?	Diminished relevance for retired people
	Q13. Can you easily carry on a conversation with somebody in a car or bus?	Difficult interpretation since both scenarios are quite different acoustically
	Q18. Can you follow a conversation between a few people during dinner?	
	Q24. Can you carry on a conversation with someone in a busy street?	

The table shows questions used from the Modified Amsterdam Inventory for Auditory Disability and Handicap (mAIAD) completed by each patient's primary caregiver (or another close informant).

Finally, I derived receiver operating characteristic (ROC) curves to assess the overall diagnostic utility of noise-vocoded speech comprehension in distinguishing each patient group from healthy controls. The binary classifier used was the speech intelligibility threshold (set at 50%) obtained from the psychometric function. I then calculated the area under the ROC curve (AUV) for each syndromic group using parametric estimates in the pROC R package (Hajian-Tilaki et al., 1997; Robin et al., 2011).

4.4.4.1 Brain image analysis

A volumetric T1 MR brain image was acquired on a Siemens Prisma 3T MRI scanner using a 32-channel phased array head-coil, following a T1-weighted sagittal 3D magnetisation prepared rapid gradient echo (MPRAGE) sequence (echo time / repetition time / inversion time respectively 2.9 / 2200 / 900ms, dimensions 256 x 256 x 208, voxel volume of 1.1 x 1.1 x 1.1mm). Before pre-processing, each scan was examined for quality control.

MR images from healthy control participants were not incorporated in the voxel-based morphometry (VBM) analyses. This was to avoid identifying spurious anatomical associations in brain areas with disease-related grey matter atrophy since it is likely that in the absence of neurodegeneration, factors other than regional brain volume changes can drive variance in experimental performance.

The brain images were first pre-processed and normalised to MNI space using SPM12 (www.fil.ion.ucl.ac.uk/spm/software/spm12/), and the Diffeomorphic Anatomical Registration using exponentiated lie algebra (DARTEL) toolbox with default parameters. Grey matter images were smoothed using a 6mm full width at half-maximum (FWHM) Gaussian kernel. To control for individual differences in total (pre-morbid) brain size, total intracranial volume was calculated by

summing grey matter, white matter, and cerebrospinal fluid volumes in each participant after segmentation of separate tissue types (Malone et al., 2015). An explicit brain mask was created using an automatic mask-creation strategy previously developed (Ridgway et al., 2009). A study-specific brain template brain image was created by warping all bias-corrected native space brain images to the final DARTEL template and calculating the average of the warped brain images.

I assessed grey matter association of noise-vocoded speech intelligibility threshold over the combined patient cohort. Voxel-wise grey matter intensity was modelled as a function of performance threshold in a multiple regression design, incorporating age, total intracranial volume, and diagnostic group memberships as covariates. Statistical parametric maps were assessed at peak-level significance threshold $p<0.05$, after family-wise error (FWE) correction for multiple voxel-wise comparisons within five pre-defined regions of interest, based on prior neuroanatomical hypotheses. I selected these regions based on functional neuroanatomical substrates in the healthy brain and defined them using the Harvard-Oxford Brian Atlas (http://fsl.fmrib.ox.ax.uk/fsl/fslwiki/Atlases).

These regions comprised left planum temporale (Griffiths & Warren, 2002; Warren et al., 2006), left angular gyrus (Davis & Johnsrude, 2003; Hartwigsen et al., 2015; Obleser et al., 2007), left anterior superior temporal gyrus (Hervais-Adelman et al., 2012; Obleser et al., 2007; Scott et al., 2006), left inferior frontal gyrus (Cope et al., 2017; Hervais-Adelman et al., 2012; Obleser et al., 2007), and cingulate gyrus (Gennari et al., 2018; Obleser et al., 2007). Anatomical volumes were derived from Oxford-Harvard cortical maps (Desikan et al., 2006), which are shown in **Figure 4.3**.

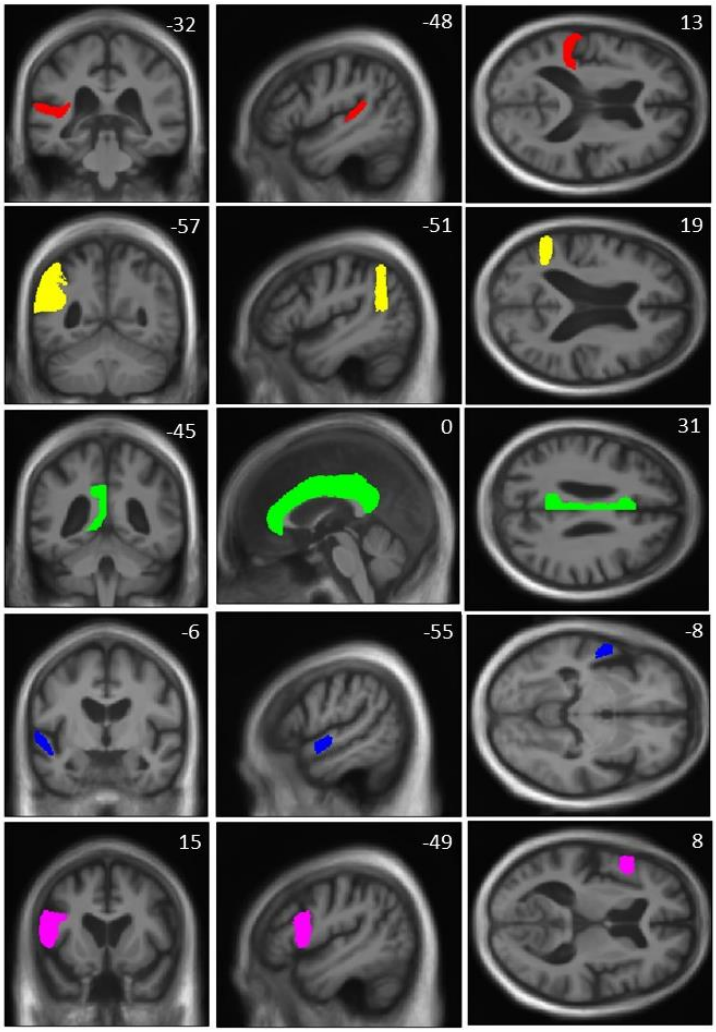


Figure 4.3. Representative sections of neuroanatomical regions in the left cerebral hemisphere that were used for multiple voxel-wise comparisons correction in region-of-interest analyses
See **Section 4.4.4.1**. Regions are rendered on coronal (left), sagittal (middle), and axial (right) sections of the mean normalised brain template for the patient cohort. MNI coordinates of the plane of each section are shown in the top right-hand corner. The neuroanatomical regions comprise the left planum temporale (red), left angular gyrus (yellow), left cingulate gyrus (green), left anterior superior temporal gyrus (blue), and left inferior frontal gyrus (purple).

4.5 RESULTS

4.5.1 General participant group characteristics

Participant groups did not differ significantly in age, sex distribution, handedness, or years of formal education (all p>0.05, **Table 4.1**). Patient groups did not differ in mean symptom duration (p=0.09) but did differ in the combined T-MMSE and

MMSE score ($X^2(3)=11.3$, p=0.01), with the AD group performing worse than the nfvPPA (z=-3.22, p=0.001) and svPPA (z=-2.10, p=0.04) groups. General neuropsychological profiles were in keeping with syndromic diagnosis for each patient group (**Table 4.1**).

Pure tone audiometry (in the participant subcohort assessed in-person) revealed no substantial peripheral hearing deficits nor any significant differences between participant groups (p>0.05). Basic speech discrimination (assessed using the PALPA-3) did not differ significantly from the healthy control group for any of the PPA syndromic groups (p>0.05).

4.5.2 Experimental behavioural data

Psychometric parameters for the participant groups are presented in **Table 4.3** and group mean psychometric functions are presented in **Figure 4.4**. Results from the full dataset are accordingly reported in-text below.

There was a significant main effect of diagnostic group on noise-vocoded speech intelligibility threshold (see **Table 4.3** and **Figure 4.5**). In post-hoc pairwise group comparisons versus healthy controls, the intelligibility threshold was significantly elevated in the lvPPA (z=3.87, p<0.001), nfvPPA (z=3.92, p<0.001), AD (z=5.01, p<0.001), and svPPA (z=2.20, p=0.03) groups.

Comparing patient groups, the intelligibility threshold was significantly elevated in the AD group (z=2.04, p=0.04) compared with the svPPA group. There was no significant effect of diagnostic group on the slope of the psychometric function (see **Table 4.3**).

A significant main effect of diagnostic group on the lapse rate (lambda) was found (see **Table 4.3**). In post-hoc pairwise group comparisons versus healthy controls,

Table 4.3. Mean psychometric function parameters for comprehension of noise-vocoded speech in each participant group

Parameter	Controls	AD	lvPPA	nfvPPA	svPPA	Omnibus significant test
Threshold	3.14 (0.27)	**4.33 (1.25)**	**5.04 (2.20)**	**4.68 (2.22)**	**3.55 (0.53)***	**$X^2(4)=34.35$, p<0.001**
Slope	1.08 (0.89)	0.79 (0.27)	0.77 (0.36)	0.81 (0.28)	0.95 (0.46)	$X^2(4)=5.42$, p=0.247
Lambda	0.00 (0.01)	**0.02 (0.02)**	**0.03 (0.03)**	**0.02 (0.02)**	**0.02 (0.03)**	**$X^2(4)=16.75$, p=0.002**

Parameters are based on mean psychometric functions for each participant group (see text and **Figure 4.4**); mean (standard deviation) values are shown. Threshold indicates 50% intelligibility of noise-vocoded spoken numbers, adjusted to take account of lambda and gamma values; slope indicates the slope of the psychometric function at this threshold point; lambda (lapse rate) indicates the number of incorrect responses at the maximum performance level. Significant differences (p<0.05) between patient groups and the healthy older control group are shown in **bold**; *significantly lower than the svPPA group. AD, patient group with typical Alzheimer's disease; Controls, healthy older control group; lvPPA, patient group with logopenic variant primary progressive aphasia; nfvPPA, patient group with nonfluent/agrammatic variant primary progressive aphasia; svPPA, patient group with semantic variant primary progressive aphasia.

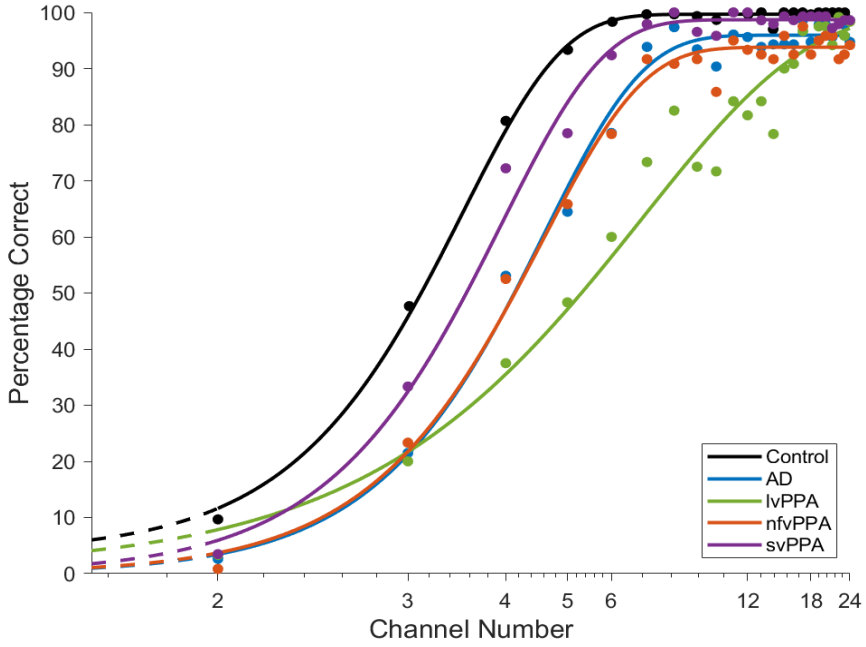


Figure 4.4. Average psychometric curves for comprehension of noise-vocoded speech in each participant group.
The y-axis here shows the percentage of digits identified correctly (from a total of 12 digits) at each noise-vocoding level; the x-axis shows the number of vocoding channels, plotted on a log scale. Mean psychometric functions were created for each diagnostic group (colour coded at lower right; see also text and **Table 4.3**); curves have been fitted through values (coloured dots) representing the mean score correct across individual participants in that group at each noise-vocoding level. AD, patient group with typical Alzheimer's disease; Control, healthy older control group; lvPPA, patient group with logopenic variant primary progressive aphasia; nfvPPA, patient group with nonfluent variant primary progressive aphasia; svPPA, patient group with semantic variant primary progressive aphasia.

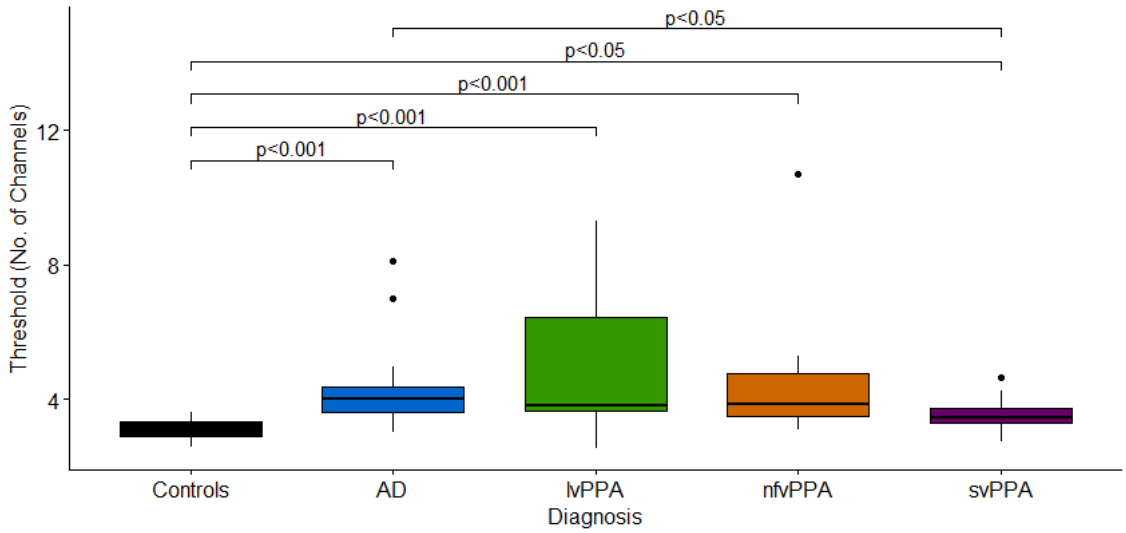


Figure 4.5. Plots of individual participant thresholds for comprehension of noise-vocoded speech within each diagnostic group.
Speech intelligibility threshold values are based on individual psychometric curves for the identification of noise-vocoded spoken numbers (see text for details). In this context, the threshold corresponds to the number of vocoding channels in the speech stimulus at which 50% intelligibility of spoken numbers was achieved, adjusted to take account of lambda value (the upper-performance asymptote; see **Table 4.3**). The line within each box indicates the median, with the boxes indicating the interquartile interval. AD, patient group with typical Alzheimer's disease; Control, healthy older control group; lvPPA, patient group with logopenic variant primary progressive aphasia; nfvPPA, patient group with nonfluent variant primary progressive aphasia; svPPA, patient group with semantic variant primary progressive aphasia.

there was a significantly higher lapse rate (more errors made at maximum performance) in all patient groups: this elevation was in the lvPPA (z=2.95, p=0.003), AD (z=2.61, p=0.009), nfvPPA (z=3.27, p=0.001), and svPPA (z=2.32, p=0.02) groups. There were no significant differences between patient groups for lapse rate.

Individual variability in psychometric parameters within participant groups was substantial (**Figure 4.5** and **Table 4.3**). Most pertinently, variation in noise-vocoded speech intelligibility threshold was wider in the AD group than in healthy controls and is most marked in the lvPPA and nfvPPA groups.

Over the combined patient cohort, noise-vocoded speech intelligibility threshold was not significantly correlated with peripheral hearing function (r=-0.04, p=0.856), phonological discrimination in clear speech (PALPA-3 score; r=-0.25, p=0.185), age (r=0.24, p=0.100) or symptom duration (r=-0.10, p=0.510). Intelligibility threshold in the patient cohort was significantly correlated with WASI Matrices score (r=-0.49, p<0.001), MMSE score (r=-0.53, p<0.001), and forward digit span (r=-0.66, p<0.001). Lapse rate was also significantly correlated with forward digit span across the combined patient cohort (r=-0.34, p=0.017).

When assessing the relation of noise-vocoded speech intelligibility threshold to relevant aspects of daily hearing function indexed on the mAIAD across the whole patient cohort, both correlations were negative with borderline significance for the questions revolving around hearing when communicating in noise (r=-0.3, p=0.067, see **Figure 4.6**).

Analysis of ROC curves revealed that noise-vocoded speech intelligibility discriminated all patient groups well from healthy controls (see **Figure 4.7**). Based on AUC values (where a value of 1 would indicate an ideal classifier and values >0.8 a clinically robust discriminatory (Carter et al., 2016; Ohman et al., 2000)), discrimination was ‘excellent’ for the AD (AUC 0.95) and nfvPPA (AUC 0.91) groups, ‘good’ for the lvPPA group (AUC 0.88), and ‘fair’ for the svPPA group (AUC 0.77).

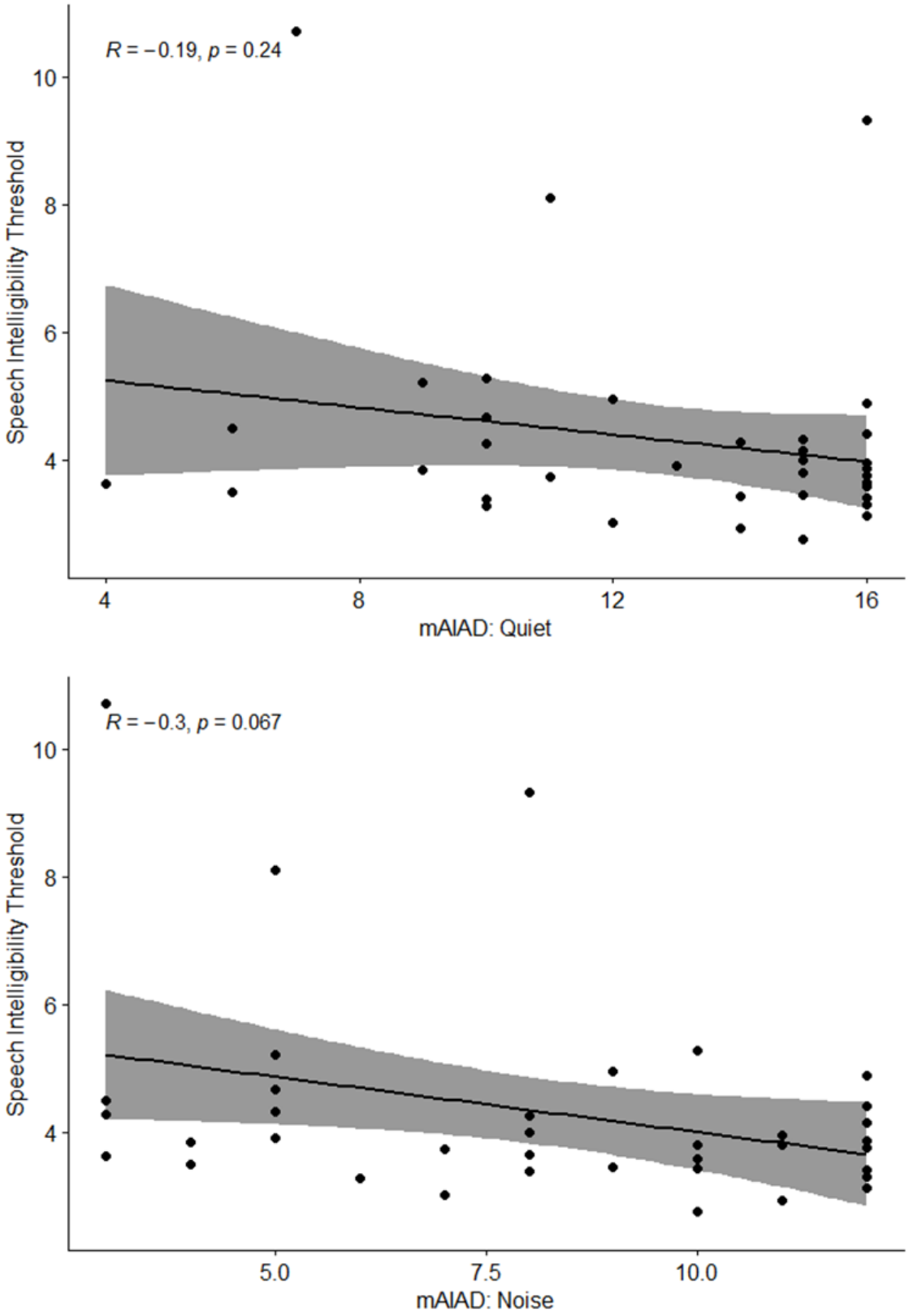


Figure 4.6. Scatter plot showing correlations between speech intelligibility threshold and measures of hearing in quiet and noise from the modified Amsterdam Inventory for Auditory Disability and Handicap across the patient cohort

See **Table 4.2**. Spearman's R and p-value are shown at the top left-hand corner of each plot. Dots represent each participant's performance.

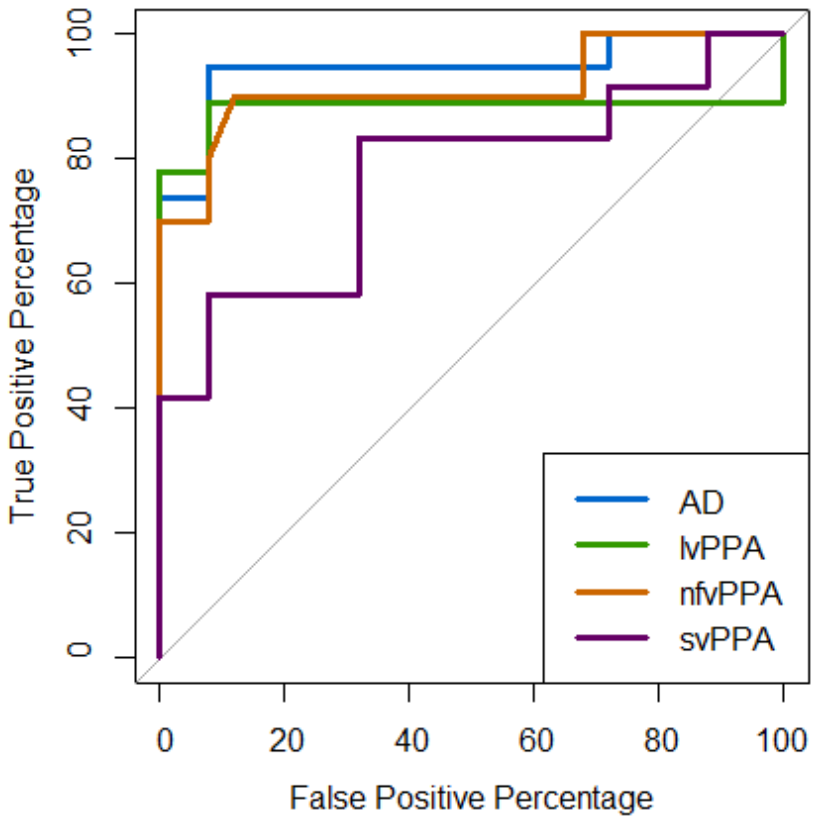


Figure 4.7. ROC curves for comprehension of noise-vocoded speech in patient groups versus healthy older controls
Receiver operating characteristic (ROC) curves for each syndromic group versus the healthy older control group are shown. The binary classifier used was the speech intelligibility threshold obtained in the psychometric functions (see **Table 4.3** and **Figure 4.4**). An area under the curve (AUC) of 1 would correspond to an ideal classifier. AUC values obtained were as follows: Alzheimer's disease, AUC = 0.95; nonfluent/agrammatic variant PPA, AUC = 0.91; logopenic variant primary progressive aphasia (PPA), AUC = 0.88; semantic variant PPA, AUC = 0.77.

4.5.2.1 Parallel Experimental Behavioural Data Results Without Outliers

The exclusion of two upper-bound outliers (>97.5 quantile) in parallel analyses left the results qualitatively unaltered. There was a significant main effect of diagnostic group on noise-vocoded speech intelligibility threshold ($H(4)=32.89$, $p<0.001$). In post-hoc pairwise group comparisons versus healthy controls, the intelligibility threshold was significantly elevated in the lvPPA ($z=3.48$, $p<0.001$), nfvPPA ($z=3.55$, $p<0.001$), AD ($z=5.15$, $p<0.001$) and svPPA ($z=2.26$, $p=0.02$) groups. Comparing patient groups, the intelligibility threshold was significantly elevated in the AD group ($z=2.10$, $p=0.04$) compared with the svPPA group.

There was no significant effect of diagnostic group on the slope of the psychometric function (p=0.320). There was, however, a significant main effect of diagnostic group on the lapse rate, lambda (H(4)=19.77, p=0.001). In post-hoc pairwise group comparisons versus healthy controls, there was a significantly higher lapse rate (more errors made at maximum performance) in all patient groups: this elevation was in the lvPPA (z=3.26, p=0.001), AD (z=2.60, p=0.009), nfvPPA (z=3.56, p<0.001), and svPPA (z=2.31, p=0.02) groups. No significant difference was seen between patient groups on lapse rate.

4.5.3 Neuroanatomical data

Statistical parametric maps of grey matter regions associated with speech intelligibility threshold are shown in **Figure 4.8** and local maxima are summarised in **Table 4.4**.

Across the combined patient cohort, the intelligibility threshold was significantly negatively associated with regional grey matter volume (i.e., associated with grey matter atrophy) in left planum temporale, left angular gyrus, and left anterior cingulate gyrus (all $p_{FWE}<0.05$) after correction for multiple voxel-wise comparisons within the relevant pre-specified neuroanatomical region of interest).

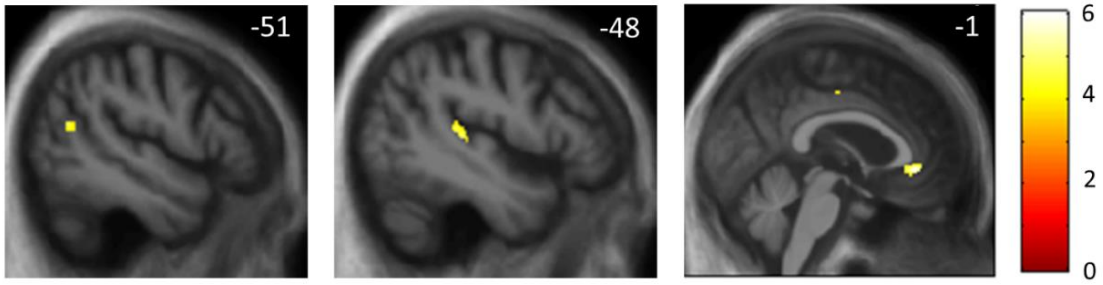


Figure 4.8. Statistical parametric maps of regional grey matter atrophy associated with elevated noise-vocoded speech intelligibility threshold in the combined patient cohort.
Maps are rendered on sagittal sections of the group mean T1-weighted MR image in MNI space, masked using the pre-specified neuroanatomical region of interests (as used in the small volume corrections) and thresholded at $p < 0.001$ uncorrected for multiple voxel-wise comparisons over the whole brain for display purposes (areas shown were significant at $p < 0.05_{FWE}$ for multiple comparisons within regions of interest). The colour bar (right) codes voxel-wise t-values. All sections are through the left cerebral hemisphere; the plane of each section is indicated using the corresponding MNI coordinate (mm).

Table 4.4. Neuroanatomical associations of noise-vocoded speech intelligibility threshold in the patient cohort

Region	Cluster size (voxels)	Peak (mm)			T score	P_{FWE}
		x	y	z		
Left planum temporale	131	-48	-31	6	4.65	0.019
Left angular gyrus	36	-51	-61	16	4.51	0.037
Left cingulate gyrus	142	-1	38	-5	5.68	0.012

The table shows significant negative associations between regional grey matter volume and intelligibility threshold for noise-vocoded speech, based on the voxel-based morphometric analysis of brain MR images for the combined patient cohort. Coordinates of peaks (local maxima) are in MNI standard space. Local maxima shown were significant ($p < 0.05$) after family-wise error (FWE) correction for multiple voxel-wise comparisons within the pre-specified anatomical regions of interest (see text and **Figure 4.3**).

4.6 DISCUSSION

In this experiment, I have shown that perception of acoustically degraded (noise-vocoded) speech is impaired in patients with AD and PPA syndromes relative to healthy older listeners, and further stratifies the different diseases, with the highest average speech intelligibility threshold in lvPPA and nfvPPA groups. While acoustically degrading speech did add different amounts of 'error' in detection (i.e., the 50% speech intelligibility threshold) across the patient groups, it did not alter the sensitivity of the detector (i.e., the slope was not significantly different across groups). The intelligibility threshold for noise-vocoded speech did not correlate with measures of pure tone detection or phoneme discrimination in clear speech, suggesting that the deficit shown here cannot simply be explained by an impairment in peripheral hearing or clear speech perception.

Neuroanatomically, impaired noise-vocoded speech comprehension across dementia syndromes was underpinned by atrophy of the left planum temporale, angular gyrus, and anterior cingulate gyrus. This cortical network is critical for processing speech signals under a range of noisy, real-world listening conditions (Davis & Johnsrude, 2003; Chris J. D. Hardy et al., 2017; C. J. D. Hardy, J. L. Agustus, et al., 2017; Jiang et al., 2021; Wild et al., 2012). Planum temporale is likely to play a fundamental role in the deconvolution of complex sound patterns and engagement of neural representations corresponding to phonemes and other auditory objects (Griffiths & Warren, 2002; Warren et al., 2006; Warren et al., 2005). Angular gyrus mediates the disambiguation of speech signals in challenging listening environments, working memory for speech signals, and transcoding of auditory inputs for motor responses (including orienting and repetition) (Golestani et al., 2013; Hartwigsen et al., 2015; Obleser & Kotz, 2010; Shahin et al., 2009; Warren et al., 2005). Both planum temporale and angular

gyrus are targeted in AD, lvPPA, and nfvPPA (Bejanin et al., 2017; Giannini et al., 2017; Lombardi et al., 2021; Ruksenaite et al., 2021), and have been implicated in the pathogenesis of impaired speech perception in these diseases (Chris J. D. Hardy et al., 2017; C. J. D. Hardy, Y. T. Hwang, et al., 2017; Hardy, Marshall, et al., 2018; Johnson et al., 2020).

The anterior cingulate cortex, with a more general role in cognitive control and allocation of attentional resources to salient stimuli (Abutalebi et al., 2011; Shenhav et al., 2013; Wild et al., 2012), works in conjunction with the planum temporale and angular gyrus in decoding spoken messages under challenging listening conditions (Gennari et al., 2018; Obleser et al., 2007). Reduced activation of the anterior cingulate cortex during tracking of information in degraded speech signals has previously been demonstrated in nfvPPA and svPPA (C. J. D. Hardy, J. L. Agustus, et al., 2017).

Noise-vocoding fundamentally reduces the availability of acoustic cues that define phonemes as auditory objects. Therefore, the impaired recognition of these degraded auditory objects could be a result of deficient encoding of acoustic features, damaged object-level representations (i.e., ‘apperceptive’ auditory deficits (Johanna C. Goll, Sebastian J. Crutch, & Jason D. Warren, 2010)), and/or impaired top-down, predictive disambiguation based on stored knowledge about speech signal characteristics. In AD and lvPPA, a core deficit of object-level representations has been demonstrated neuropsychologically and electrophysiologically using other procedures that alter acoustic detail in phonemes and nonverbal sounds (Goll et al., 2011; C. J. D. Hardy, J. L. Agustus, et al., 2017; Jiang et al., 2022; Stalpaert et al., 2021). Therefore, it is plausible that an analogous apperceptive deficit may have impacted the recognition of the noise-vocoded auditory objects in the AD and lvPPA groups here.

In nfvPPA, a previous study of noise-vocoded speech perception has foregrounded the role of inflexible top-down predictive decoding mechanisms instantiated in the frontal cortex (Cope et al., 2017). However, this is a clinically, neuroanatomically, and neuropathologically diverse syndrome and involvement of the posterior superior temporal cortex seen in patients could elicit deficits in early auditory pattern analysis, constituting a 'double hit' to phoneme recognition (Johanna C. Goll, Sebastian J. Crutch, & Jason D. Warren, 2010; Goll et al., 2011; Grube et al., 2016; Chris J. D. Hardy et al., 2017). Further, the paradigm used by Cope and colleagues (2017) was different from that used in this experiment: the vocoding channels were limited to four, eight, and 16 channels, resulting in a choice of four alternative forced-choice responses.

In svPPA, considering the relatively intact bottom-up perception of degraded speech stimuli (Hardy, Marshall, et al., 2018; Jiang et al., 2022), the elevated noise-vocoded intelligibility threshold is surprising. However, it is likely to reflect an extent of reduced activation of semantic mechanisms engaged in the predictive disambiguation of degraded speech signals, parsed out through the graded noise-vocoding methodology used here. Comprehension of other kinds of acoustically degraded speech signals by patients with svPPA has previously been shown to be sensitive to semantic predictability and engagement in the anterior cingulate cortex (C. J. D. Hardy, J. L. Agustus, et al., 2017; Hardy, Marshall, et al., 2018; Jiang et al., 2022).

Increasing intelligibility threshold was correlated with digit span over the combined patient cohort. This suggests that verbal working memory limitations may be integrally related to impaired processing of degraded speech, consistent with previous work highlighting the role of working memory in speech perception, particularly in older adults (Meister et al., 2013; Millman & Mattys, 2017).

As working memory demands did not vary across trials and the number of vocoding channels, the principal driver of the intelligibility threshold is likely to have been the level of acoustic alteration in the speech signal.

On the other hand, all patient groups showed an increased lapse rate (i.e., errors unrelated to the stimulus level (Schütt et al., 2016)) for minimally noise-vocoded speech signals. This was significantly correlated with digit span, echoing previous work demonstrating that active listening can be abnormal in lvPPA and nfvPPA even for clear speech and other sounds in quiet (Hardy et al., 2019; Johnson et al., 2020), and interactions with top-down mechanisms engaged in the predictive processing of speech (Cope et al., 2017).

Frontal processes are likely to play a broader role in the disambiguation of degraded speech signals as well, including the allocation of attentional and executive resources (Peelle, 2018), suggested by the observed correlation here between noise-vocoded speech intelligibility threshold and WASI Matrices score. Taken together, the present findings corroborate the profiles of deficits previously documented in AD and PPA syndromes for comprehension of sinewave speech and phonemic restoration in noise-interrupted speech (Hardy, Marshall, et al., 2018; Jiang et al., 2022).

From a clinical perspective, this study shows how acoustically altered speech perception can potentially serve as proxies for daily life hearing and communication. The correlation between communication in noisy environments and speech intelligibility threshold showed a slight trend toward significance. This should be extended with a larger cohort, inclusive of different dementia diseases. In any subjective reporting measures though (even when the mAIAD has been useful in assessing symptoms in stroke and adults with auditory

processing disorders (Bamiou et al., 2015; Bamiou et al., 2012)), individual factors that could influence reporting of symptoms (e.g., personality traits (Wöstmann et al., 2021)) need to be considered. Using these measures with caregivers of those living with dementia (i.e., third-party responses) has also not been validated, and therefore needs to be further investigated.

Crucially, building on other degraded speech manipulations (see **Chapter 1.3.2**), noise-vocoding is a good model for a diverse range of real-world scenarios in which decoding noisy or ambiguous speech signals is required. Past literature has shown that complex auditory tasks were better predictors of real-world auditory function than pure-tone audiometry (Füllgrabe et al., 2015; Holmes & Griffiths, 2019; Johnson, 2021; Utoomprurkporn et al., 2021). An audiological measure that could capture important dimensions of daily life communication function would have obvious value as a target for developing and evaluating novel hearing-based interventions. Therefore, clinical and environmental strategies need to be considered to help patients navigate an auditory difficulty that could build on other communication difficulties that they may already have.

The findings from this experiment suggest that markers of noise-vocoded speech comprehension may have potential utility, diagnostically and as a biomarker. The ROC analyses (**Figure 4.6**) suggest that it would constitute an 'excellent' clinical test (corresponding to an AUC > 0.9) for discriminating patients with AD and nfvPPA from healthy older individuals (Carter et al., 2016). Additionally, the noise-vocoded intelligibility threshold was correlated with overall disease severity (MMSE score) in the patient cohort. Therefore, these findings build on a growing body of work suggesting that markers of 'central' hearing may sensitively signal the functional integrity of cortical regions that are vulnerable to AD and other

neurodegenerative pathologies (Jiang et al., 2021; Johnson et al., 2021; Stevenson et al., 2022).

More work needs to be conducted to refine the paradigm used here. For example, the step-wise linear progression design was not optimally efficient. Rather, an adaptive staircase procedure would reduce testing time and allow individual thresholds to be captured without administering uninformative trials at higher channel numbers. However, as our natural listening environment tends to change gradually in time, the linear progression could be a more naturalistic approach. It is also possible that with the paradigm structure presented here, perceptual learning (see **Chapter 1.3.3.3**) could have acted as a nuisance factor.

It would be relevant to assess to what extent patients' comprehension of noise-vocoded speech can be modulated and modified pharmacologically, as seen in past research showing acetylcholinesterase inhibitors increasing sinewave speech perception in AD (C. J. D. Hardy, Y. T. Hwang, et al., 2017)), and/or by perceptual learning (Hardy, Marshall, et al., 2018). To establish how noise-vocoded speech perception and its modulatory factors related to neural circuit integrity in AD and PPA, functional neuroimaging (e.g., fMRI and MEG) will be required to capture dynamic network connectivity.

Lastly, considering that group sizes were relatively small, this paradigm should be extended to larger patient cohorts, which (given the comparative rarity of PPA) will likely entail multi-centre collaboration. Assessments in larger cohorts will allow for further characterisations of the sources of individual variation in degraded speech perception, but also within diagnostic groups (such as in the lvPPA group (see **Figure 4.5**), which could potentially be explained by differences in patterns of damage in the left temporo-parietal junction (Ramanan et al., 2022)).

If expanded to multi-centre collaborations, this paradigm could be extended internationally and cross-linguistically, but initial experiments will need to be conducted as it is likely that the acoustic properties of different languages could be differently and/or disproportionately affected by artificial manipulations such as noise-vocoding.

A natural direction for future research would be to assess how the noise-vocoding paradigm used here performs longitudinally. This would allow for assessments on how markers of degraded speech perception relate to disease course, and to determine how central auditory perception can signal underlying neurodegenerative pathologies. In particular, for AD, auditory measures based on degraded speech comprehension would be well suited to digital applications and large-scale screening of (and potentially international) populations at risk of incident dementia (Johnson et al., 2021; Stevenson et al., 2022).

5 EXPERIMENT 3: NOISE-VOCODED EMOTIONAL PROSODY

5.1 SUMMARY

Accurate comprehension of speech and successful communication is never solely dependent on the verbal content, but also on accurate interpretation of a speaker's nonverbal cues to elicit their intent. Particularly, understanding emotional vocal signals (their prosody) is essential. Emotional prosody plays a key role in social interactions and is crucial in maintaining interpersonal relationships. However, as our typically non-ideal listening conditions affect the perception of the verbal content in communication (as seen in **Chapter 4**), so can those conditions present a challenge for accurately processing nonverbal content as well. Considering these processes are equally difficult computational challenges, it is likely that degraded nonverbal auditory content may also be vulnerable to neurodegenerative pathologies. Therefore, this chapter aims to investigate degraded emotional prosody perception in Alzheimer's disease (AD) and primary progressive aphasias (PPA).

I assessed the impact of noise-vocoding on the identification of three canonical prosodic emotions (anger, surprise, sadness), at three levels of noise-vocoding (six, 12, and 18 channels), versus clear speech, in AD and PPA patients. Compared with healthy controls, all patient groups were impaired in the identification of clear emotional prosody. This impairment in all groups was sustained when the emotional prosody stimuli were vocoded, with differences seen between 18 and six channels, and 12 and six channels. Cost analyses were conducted to understand the additional effect of noise-vocoding (from clear

speech) on the identification of emotional prosody, and patients with lvPPA showed a significantly higher cost than healthy controls.

Crucially, accurate identification of degraded emotional prosody in patients was significantly correlated with more measures of social cognition than clear emotional prosody. The findings presented here open a window on a dimension of real-world emotional communication that has often been overlooked in dementia, with particular relevance to social cognition, and begin to suggest a novel candidate paradigm for investigating and quantifying this systematically.

5.2 INTRODUCTION

Despite verbal speech being the main component of human communication for hearing individuals, there is much more to successful communication than solely the verbal message. To understand the contextual form of the word and gain full insights into the speaker's communicative intent, the nonverbal auditory information needs to be taken into consideration as well (Liebenthal et al., 2016; Morningstar et al., 2022; Wilson & Sperber, 2002).

Prosody (see **Chapter 1.4**) is a complex nonverbal feature associated with speech that considers individual speech sounds, pitch, intonation, stress, duration, and intensity. It conveys multidimensional information and diverse functions, including distinguishing word meanings in tone languages, disambiguating sentence structure, highlighting or emphasising elements in a sentence, and signalling emotion (Zatorre & Baum, 2012). Modulations in vocal pitch (fundamental frequency (f0)), syllable length, duration, intensity, and voice quality perceived by listeners to convey emotional states are collectively known as 'emotional prosody' (Everhardt et al., 2020; Horley et al., 2010). Different emotions tend to create different prosody "profiles" in speech: joy is typically

characterised by faster speech rate, high intensity, and increases in f0 mean and variability, resulting in sounds both more melodic and energetic; sadness is typically characterised by slow speech, low intensity, and decreases in f0 mean and variability (Banse & Scherer, 1996; Schirmer & Kotz, 2006).

Past research has shown impairments in nonverbal auditory perception in AD and PPA (see **Chapter 1.5.7**). Specifically in emotional prosody, impaired comprehension has been documented in AD (Amlerova et al., 2022; Arroyo-Anlló et al., 2019; Horley et al., 2010; Taler et al., 2008; Testa et al., 2001) and in nfvPPA and lvPPA (Jonathan D. Rohrer, Disa Sauter, et al., 2010; Shany-Ur & Rankin, 2011). Research in svPPA has been less conclusive, with some studies findings no deficits in emotional prosody (Rankin et al., 2009), while others finding an impairment (Macoir et al., 2019). This is particularly crucial as an impaired ability to identify prosody could have implications for social interactions and interpersonal relationships (Everhardt et al., 2020). Less research has been conducted on perceiving emotional prosody once degraded – some research has been conducted on children and individuals with cochlear implants (Chatterjee et al., 2019; Everhardt et al., 2020; Tinnemore et al., 2018).

As in **Chapter 4**, this experiment will also be using noise-vocoding, a technique in which the speech signal is divided digitally into discrete frequency bands (‘channels’), each filled with white noise and modulated by the amplitude envelope of the original signal (Shannon et al., 1995). As the fronto-temporo-parietal brain network engages in emotional prosody perception (Kotz et al., 2013; Kotz et al., 2003; Jonathan D. Rohrer, Disa Sauter, et al., 2010; Wildgruber et al., 2006), it is likely that by degrading emotional prosodic cues, the same network will be strained and will require additional neural resources for accurate identification. This processing is likely to not only be vulnerable in

neurodegenerative pathologies but to also be affected early and significantly. Therefore, this experiment is designed to look at the perception of emotional prosody in all PPA syndromes and AD, both in 'clear' and 'degraded' forms.

5.3 KEY PREDICTIONS

Based on available evidence from past emotional prosody research (Amlerova et al., 2022; Arroyo-Anlló et al., 2019; Horley et al., 2010; Macoir et al., 2019; Omar et al., 2011; Rankin et al., 2009; Jonathan D. Rohrer, Disa Sauter, et al., 2010; Shany-Ur & Rankin, 2011; Taler et al., 2008; Testa et al., 2001), my hypotheses are:

H_1: AD and PPA patients, particularly nfvPPA, will perform worse than healthy controls at identifying clear emotional prosody

H_2: AD and PPA patients will have an additional cost on the identification of emotions from degrading emotional prosodic cues than healthy controls

H_3: Accurate identification of emotional prosody is correlated with measures of daily life socio-emotional functioning

5.4 MATERIALS AND METHODS

5.4.1 Participants

For this experiment, 18 patients with typical amnestic AD, 10 patients with lvPPA, 11 patients with nfvPPA, and 11 patients with svPPA, and 23 healthy older control participants with no history of neurological or psychiatric disorders were recruited (see **Appendix: Table 8.1** for participant breakdown and **Chapter 2.1** for details).

Due to the Covid-19 pandemic, 29 participants (four healthy controls, nine patients with AD, six with lvPPA, six with nfvPPA, and four with svPPA) were

assessed remotely via Labvanced and shared through a video link (see **Chapter 2.4** for details).

No participant had abnormal peripheral hearing, other than age-related hearing loss. To assess patients' social cognition, the questionnaires Modified Interpersonal Reactivity Index (mIRI) and the Revised Self-Monitoring Scale (RSMS) were completed by the primary caregiver or another close informant for each patient (see **Chapter 2.3.4** for details).

5.4.2 Experimental Stimuli

Forty-five three-digit numbers (of the form: 'three-hundred-and-seventy-three', not: 'three-seven-three') were spoken by two adult male and two adult female speakers (all with Standard Southern British English accents). These were taken from a previously normed set of vocal emotional stimuli (Sauter, 2006). The numbers selected portrayed one of three emotions: anger, surprise, and sadness. The three emotions were chosen based on the consistency of results in Rohrer and colleagues (2010) and additional piloting on 10 young healthy controls.

Speech recordings were noise-vocoded using Matlab® (vR2019b) (https://uk.mathworks.com/) to generate acoustically altered stimuli at either six, 12, or 18 channels (details on the script used for noise-vocoding can be seen in **Chapter 4.4.2**). Three levels of noise-vocoding was used instead of thresholding as seen in **Chapter 4** due to considerations of how much longer the test would be and concerns of participant fatigue. Therefore, three channel levels were chosen and designated to signify hard (six channels), medium (12 channels), and easy (18 channels) perception. This was informed by previous work showing that

vocoded speech at 10 channels is readily intelligible by healthy listeners, whereas four channels only become intelligible after hours of training (Davis et al., 2005).

At each noise-vocoding level, 15 three-digit spoken number stimuli were presented, with five numbers in each of the three emotions. Separately, 21 three-digit word stimuli were kept in clear speech, six (two for each emotion) were to be used unscored for familiarising the participant with the stimuli before the experimental test, and 15 (five per emotion) were used as a clear speech control condition. Thus, a total of 60 stimuli (20 for each emotion) were presented during the experimental test session.

5.4.3 Procedure

The stimuli were administered either in-person in a quiet room via Audio-Technica ATH-M50x headphones at a comfortable fixed listening level (at least 70 dB), or remotely via Labvanced and shared through a video link (see **Chapter 2.5.1** and **Table 5.1**).

To be familiarised with the experimental procedure, participants first practiced with six clear three-digit number stimuli. I asked them to determine which emotion was being portrayed in the voice, using a provided cue card as a guide (**Figure 5.1**). The familiarisation section was completed once the participants understood the task at hand.

I then advised the participants that for the experimental task, the three-digit numbers will sound distorted and will vary in how difficult they are to understand, and to guess the emotion if uncertain. Unlike in **Chapter 4**, the 45 trials were presented in randomised order to each participant. The 45 noise-vocoded trials were followed by 15 randomised trials of emotions presented in clear speech (the clear speech stimuli were presented last to avoid priming effects on vocoded

Table 5.1. General demographic, clinical, and neuropsychological characteristics of all participant groups

Characteristic	Controls	AD	lvPPA	nfvPPA	svPPA
Demographic and clinical					
No. M:F	13:11	13:5	8:2	8:3	7:4
Age, years	68.7 (6.4)	70.0 (8.9)	69.4 (6.9)	68.7 (6.1)	62.5 (8.6)
Handedness (R/L/A)	21/1/1[a]	17/1	9/1	11/0	10/1
Education (y)	16.1 (2.7)[a]	15.4 (3.9)[a]	15.1 (2.9)[a]	14.6 (2.7)[a]	15.7 (2.1)[a]
Symptom duration (y)	NA	6.0 (3.1)	6.7 (5.1)	3.4 (1.6)	5.5 (2.6)
Best ear average*	17.1 (8.7)[k]	27.7 (10.9)[h]	18.2 (10.7)[e]	29.0 (4.0)[g]	23.3 (8.6)
Tested in-person/remote	19/4	9/9	4/6	5/6	7/4
General intellect					
MMSE (/30)	29.8 (0.6)[j]	**21.8 (7.5)[j]**	**22.7 (7.5)[f]**	26.5 (0.7)[h]	**23.6 (5.1)[c]**
T-MMSE (/27)	26.1 (1.8)[l]	**17.4 (4.5)[h]**	**21.5 (4.5)[e]**	22.8 (4.0)[d]*	24.0 (1.4)[f]*
Episodic memory					
RMT Faces (Short) (/25)	23.8 (2.5)[m]	**16.1 (3.3)[h]**	21.4 (3.7)[d]	22.8 (3.5)[d]*	19.2 (3.7)[f]
RMT Faces (Long) (/50)	41.7 (3.7)[j]	**29.6 (5.6)[h]**	**28.5 (5.7)[e]**	36.8 (7.3)[f]	**31.1 (3.5)[c]**
Working memory					
Digit span forward (max)	6.6 (1.0)[g]	5.8 (1.4)	**4.5 (1.4)*[+]**	**4.9 (1.8)[+]**	6.6 (1.0)
Digit span reverse (max)	5.2 (1.2)[g]	**3.3 (1.4)[+]**	**3.4 (1.7)**	**3.6 (1.7)[+]**	5.0 (1.6)
Executive function					
WASI Matrices (/32)	26.8 (2.7)[g]	**12.4 (8.6)[b]**	**22.6 (5.4)***	**22.1 (6.9)***	25.5 (5.0)*
Letter fluency (total)	15.9 (5.4)[h]	10.9 (5.9)[a]	**8.2 (4.2)[a]**	9.1 (8.4)[b]	**7.7 (6.6)**
Category fluency (total)	24.1 (6.3)[h]	**11.4 (6.7)[a]**	**10.6 (6.2)[a]**	16.8 (9.0)[b]	**7.3 (5.6)**
Auditory input processing					
PALPA-3 (/36)	34.6 (1.7)[i]	NA	31.8 (5.6)[a]	32.9 (2.6)	33.6 (2.4)
Other language skills					
GNT (/30)	25.8 (2.5)[g]	**13.2 (7.4)[+]**	**9.7 (7.4)[a]**	**17.7 (10.8)[+]**	**1.6 (4.5)**
BPVS (/150)	148.0 (2.1)[g]	**139.0 (15.6)[+]**	141.0 (13.2)[a+]	**135.0 (19.2)[+]**	**79.8 (48.1)**
PALPA-55 (/24)	23.5 (1.20)[i]	NA	**17.0 (47.0)[b]**	**19.3 (3.6)**	**19.9 (5.1)**
Modified Camel and Cactus (/32)	30.6 (1.1)[g]	NA	**26.4 (7.6)[b]**	**28.5 (3.4)[+]**	**21.5 (7.9)[a]**
Other skills					
GDA calculation (/24)	14.8 (5.2)[g]	**5.1 (4.8)[b+]**	**5.4 (4.6)[b+]**	**6.1 (5.4)**	11.2 (5.7)[a]
VOSP Object Decision(/20)	18.9 (1.5)[g]	**14.6 (2.9)**	**16.2 (2.5)[b]***	17.9 (1.9)[a]	**15.7 (4.0)[a]**

Mean (standard deviation) values and raw scores are presented (maximum value possible in parentheses) unless otherwise indicated; significant differences from healthy controls ($p<0.05$) are in bold; *significantly different to AD ($p<0.05$); [+]significantly different to svPPA ($p<0.05$). *See **Chapter 2** for details concerning the 'best ear average'

measure. Participants assessed in-person did the MMSE and RMT Faces (Long), while those assessed remotely did the T-MMSE and RMT Faces (Short). A, ambidextrous; AD, patient group with typical Alzheimer's disease; BPVS, British Picture Vocabulary Scale; Controls, healthy older control group; Digit span forward/reverse maximum digit span recorded; F, female; GDA, Graded Difficulty Arithmetic; GNT, Graded Naming Test; L, left; lvPPA, patient group with logopenic variant primary progressive aphasia; M, male; MMSE, Mini-Mental State Examination; NA, not available/applicable; nfvPPA , patient group with nonfluent/agrammtaic variant primary progressive aphasia; PALPA, Psycholinguistic Assessments of Language Processing in Aphasia; R, right; RMT, Recognition Memory Test; svPPA, patient group with semantic variant primary progressive aphasia; T-MMSE, tele-MMSE; VOSP, Visual Object and Space Perception; WASI, Wechsler Abbreviated Scale of Intelligence. [a] missing data for 1 participant; [b] missing data for 2 participants; [c] missing data for 4 participants; [d] missing data for 5 participants; [e] missing data for 6 participants; [f] missing data for 7 participants; [g] missing data for 8 participants; [h] missing data for 9 participants; [i] missing data for 11 participants; [j] missing data for 12 participants; [k] missing data for 14 participants; [l] missing data for 15 participants; [m] missing data for 19 participants.

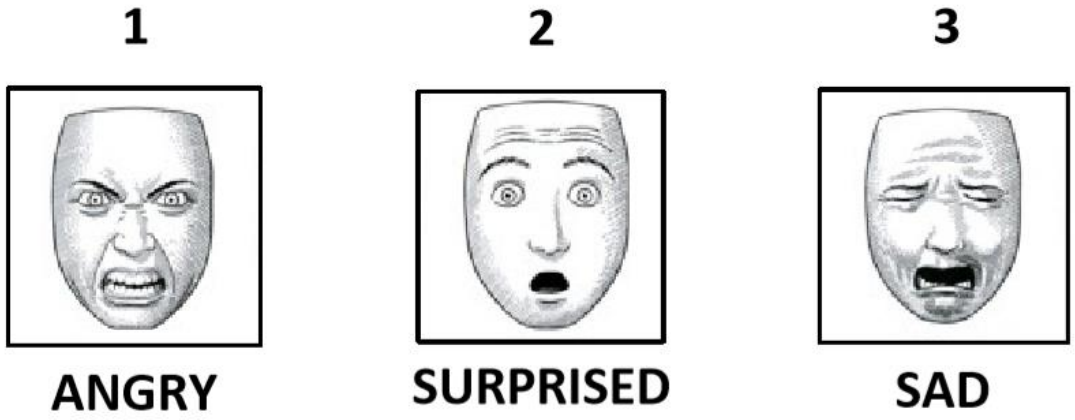


Figure 5.1. Sex-neutral cue card used in Experiment 3.
The cue card displays the three emotions on masks (i.e., sex-neutral faces), each labelled with their corresponding emotion.

emotion recognition). On each trial, the participant's task was to indicate (either verbally or by pointing at the cue card) which emotion was being portrayed. Responses were recorded for offline analysis. During the experiment, no feedback about the performance was given and no time limits were imposed.

5.4.4 Analysis of Data

Data was analysed in R® (v4). See details on the statistical analyses conducted in **Chapter 2.6**. Data for participants tested in-person and remotely were combined due to similarity in the perception of noise-vocoded stimuli (see **Chapter 4.4.4** and **Figure 4.2**). Data for the identification of clear and noise-vocoded numbers are shown in **Chapter 4**.

All participant responses were recorded to allow for error matrices to be created. In calculating the accuracy of identification of the emotions, a binary score was given: one if correct, zero if incorrect. Pearson's correlation was used to assess the relationship between accurate identification of degraded emotional prosody and general demographic (age, sex, education), clinical (symptom duration, combined MMSE score), executive performance (WASI Matrix), auditory perceptual (PALPA-3, pure-tone audiometry), and working memory (digit span) factors.

For each participant, difference scores were generated to understand the cost of each noise-vocoding level on the accurate identification of the stimuli. Prior, due to bias effects as a result of chance performance, participants who scored at or below chance on clear accent identification were excluded from cost analyses. To ascertain whether participants scored at or below chance (i.e., a score that could have been achieved by random guessing, the cumulative probability function was adopted. For the clear emotional prosody performance, 15 trials with

probability 0.33' suggested that a hit rate (k) of 9 or above was unlikely to be achieved by chance (p=0.029 [k=8, p=0.084]).

The difference between clear and noise-vocoded emotion identification performance was calculated by subtracting the noise-vocoded (at each channel level) emotion identification score from the clear emotion identification score. Similar cost scores were generated for the differences between the clear and noise-vocoded numbers (at six, 12, and 18 channels) perception. To compare the results of identifying the emotions in comparison to the perception of the spoken numbers, z-scores were created from the means of the healthy control groups' cost performance.

Error matrices were created to understand the distribution of answers across the three choices of emotions between each group. Information transfer analysis was also conducted to understand the confusion of error matrices (see **Chapter 2.6** for details on information transfer analysis).

In addition, the relationships between accurate identification of clear and degraded emotional prosody and measures of socio-emotional awareness and functioning in daily life were assessed. This was through using responses on the mIRI and RSMS in the combined patient cohort. The mIRI includes two subscales, one measuring cognitive empathy in the form of perspective taking, the other assessing emotional empathy in the form of empathic concern. For this experiment, the more pertinent scale was the emotional empathy subscale. The RSMS also includes two subscales, one indexing participants' sensitivity to expressive behaviour, and the other measuring the tendency to monitor self-presentation. In this experiment, the more pertinent scale was the participants' sensitivity to expressive behaviour.

5.5 RESULTS

5.5.1 General participant group characteristics

Participant groups did not differ significantly in age, sex distribution, handedness, or years of formal education (all p>0.05, **Table 5.1**). Patient groups did not differ in mean symptom duration (p=0.086) but did differ in the combined MMSE score ($X^2(3)=8.10$, p=0.044), with the AD group performing worse than the nfvPPA (z=-2.53, p=0.011) and svPPA (z=-2.16, p=0.031). General neuropsychology profiles were in keeping with syndromic diagnosis for each patient group (**Table 5.1**).

Pure tone audiometry (in the participant subcohort assessed in-person) revealed no substantial peripheral hearing deficits nor any significant differences between participant groups (p>0.05). Basic speech discrimination (assessed using the PALPA-3) did not differ significantly from the healthy control group for any of the PPA syndromic groups (p>0.05).

5.5.2 Experimental Behavioural Data

5.5.2.1 Raw Data

5.5.2.1.1 Clear Emotional Prosody

There was a significant difference across participant groups in accurately identifying the three clear prosodic emotions (**Table 5.2** and **Figure 5.2**). In post-hoc pairwise group comparisons versus healthy controls, the worst performance was seen in the nfvPPA group (z=-3.57, p<0.001), AD (z=-3.93, p<0.001), svPPA (z=-3.16, p<0.001), and lvPPA (z=-2.93, p=0.002) groups.

Table 5.2. Average correct raw responses and information transfer scores for comprehension of emotional prosody at clear and noise-vocoding channels, in each participant group

	Control	AD	lvPPA	nfvPPA	svPPA	Omnibus significant test
Correct Raw Responses						
6 Channels (Total Correct /15)	11.1 (1.9)	**8.9 (2.6)**	**8.5 (3.1)**	9.5 (3.1)	9.7 (3.4)	H(4) = 9.51, p=0.05
12 Channels (Total Correct /15)	12.5 (1.5)	**9.9 (3.0)**	**9.8 (1.6)**	10.4 (3.5)	**10.1 (3.2)**	H(4) = 16.24, p<0.001
18 Channels (Total Correct /15)	12.8 (1.4)	**10.4 (2.6)**	**10.6 (2.5)**	**10.5 (2.8)**	11.4 (2.2)	H(4) = 14.49, p=0.01
Clear (Total Correct /15)	14.7 (0.5)	**12.7 (2.3)**	**13.6 (1.0)**	**12.6 (2.3)**	**13.0 (2.7)**	H(4) = 23.62, p<0.001
Information Transfer						
6 Channels	0.57 (0.14)	0.40 (0.16)	0.47 (0.26)	0.44 (0.22)	0.49 (0.20)	F(4)=2.34, p=0.063
12 Channels	0.70 (0.17)	**0.46 (0.20)**	**0.47 (0.16)**	0.57 (0.22)	**0.54 (0.19)**	H(4)=16.92, p=0.002
18 Channels	0.76 (0.15)	**0.48 (0.28)**	**0.54 (0.22)**	**0.53 (0.21)**	**0.55 (0.23)**	H(4)=17.05, p=0.002
Clear	0.95 (0.08)	**0.71 (0.24)**	**0.80 (0.14)**	**0.70 (0.24)**	**0.77 (0.17)**	H(4)=23.51, p<0.001

Mean (standard deviation) values are shown. Raw scores are presented (maximum value possible in parentheses). Significantly different from healthy controls ($p<0.05$) are in **bold.** There was also a significant difference across the raw scores of vocoding channels ($F(2)=7.72$, $p<0.001$) across participants, particularly between 12 and 6 channels ($t=2.39$, $p=0.018$) and 18 and 6 channels ($t=3.63$, $p<0.001$). For the information transfer score, a score of 1 represents perfect "transfer" of information from the stimuli to the participants (i.e., no confusion). A score of 0 represents no "transfer" of the stimuli to the participants' response (i.e., a response independent of the stimuli). AD, patient group with Alzheimer's disease; lvPPA, patient group with logopenic variant primary progressive aphasia; nfvPPA, patient group with nonfluent variant primary progressive aphasia; svPPA, patient group with semantic variant primary progressive aphasia.

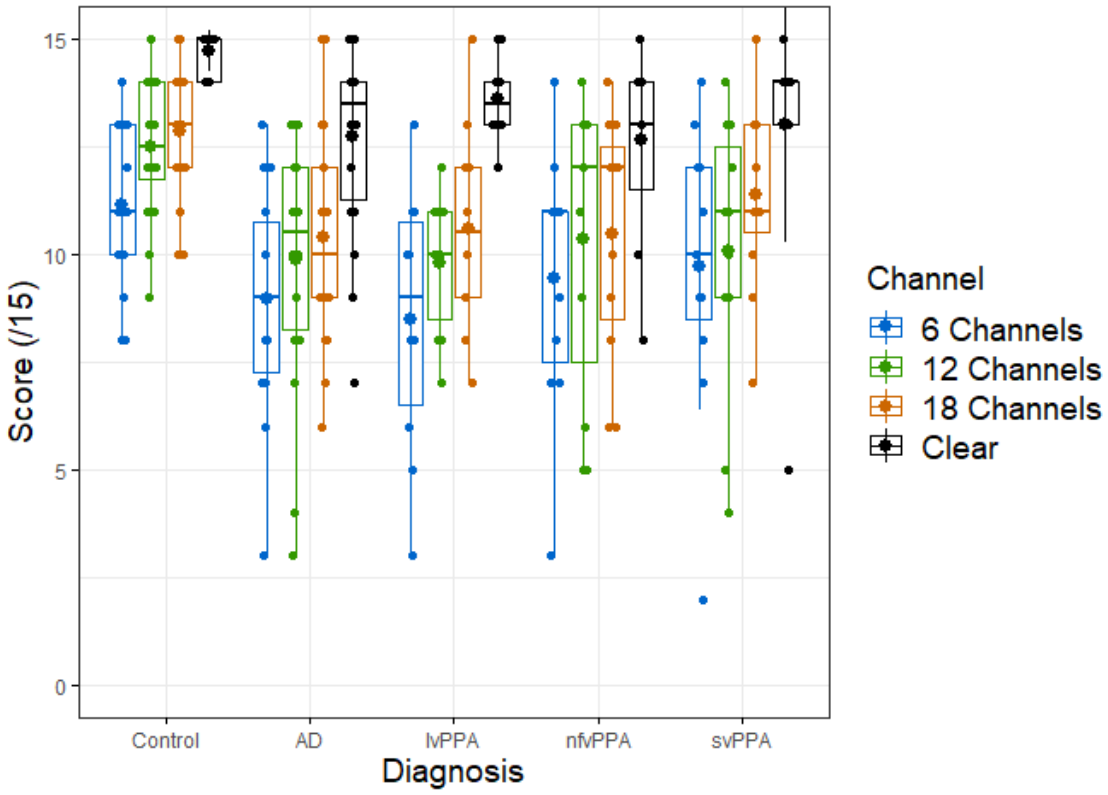


Figure 5.2. Box plots of individual participants' raw scores on correct identification of the emotion within each diagnostic group split between the clear stimuli and noise-vocoded stimuli at six, 12, and 18 channels.
Clear stimuli are in black. The line within each box indicates the median, with the boxes indicating the interquartile interval. The dots represent each participant's performance. AD, patient group with Alzheimer's disease, Control, healthy older control groups; lvPPA, patient group with logopenic variant primary progressive aphasia; nfvPPA, patient group with nonfluent variant primary progressive aphasia; svPPA, patient group with semantic variant primary progressive aphasia.

5.5.2.1.2 Noise-Vocoded Emotional Prosody

Within the noise-vocoded stimuli (combined across the three-channel levels), there was a significant difference in the accuracy of identifying the emotions across participant groups ($F(4)=10.13$, $p<0.001$), with all patient groups performing significantly worse than healthy controls (AD: $t=-5.30$, $p<0.001$; lvPPA: $t=-4.59$, $p<0.001$; nfvPPA: $t=-3.88$, $p<0.001$; svPPA: $t=-3.31$, $p=0.001$). There was no interaction effect between diagnostic group and channel level.

Total performance on the identification of noise-vocoded emotions in the patient cohort was not significantly correlated with the combined MMSE (r(39)=0.27, p=0.084), age (r(48)=0.07, p=0.616), disease duration (r(46)=-0.07, p=0.625), or pure-tone audiometry (r(21)=-0.14, p=0.521). Total performance on the identification of noise-vocoded emotions in the patient cohort was significantly correlated with WASI Matrices (r(46)=0.32, p=0.028), years in education (r(44)=0.34, p=0.022), digit span forward (r(48)=0.34, p=0.017), digit span reverse (r(48)=0.41, p=0.003), and PALPA-3 (r(29)=0.40, p=0.024).

5.5.2.2 Error Matrices

To examine the identification of individual emotions across the channels (see **Figure 5.3** for emotion recognition 'trajectory'), error matrices are created for clear speech and each vocoding channel (see **Table 5.3** and **Figure 5.4**).

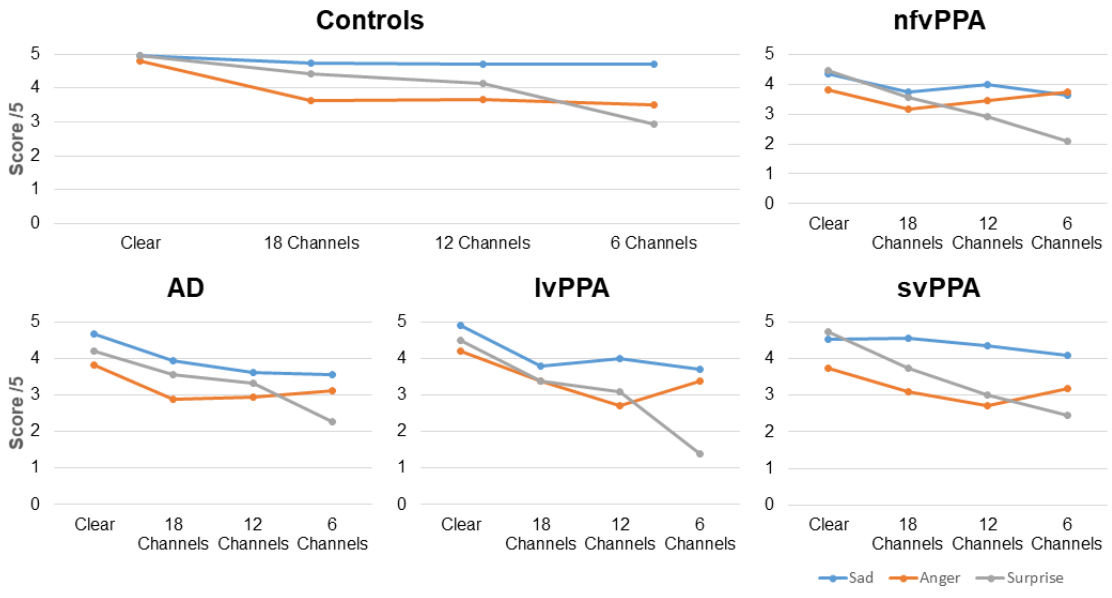


Figure 5.3. Plots showing the accuracy trajectory of each emotion through degradation, with means of each participant group.
AD, patient group with Alzheimer's disease, Control, healthy older control groups; lvPPA, patient group with logopenic variant primary progressive aphasia; nfvPPA, patient group with nonfluent variant primary progressive aphasia; svPPA, patient group with semantic variant primary progressive aphasia.

Table 5.3. Percentage of answer selection for each intended emotion in clear, 18, 12, and six channels, averaged over each participant group.

		Controls			AD			lvPPA			nfvPPA			svPPA		
	Answered/ Expected	**Anger**	**Surprise**	**Sadness**	**Anger**	**Surprise**	**Sadness**	**Anger**	**Surprise**	**Sadness**	**Anger**	**Surprise**	**Sadness**	**Anger**	**Surprise**	**Sadness**
Clear	**Anger**	**95**	5	0	**76.67**	18.89	4.44	**84**	14	0	**76.36**	18.8	5.46	**72.73**	25.46	1.82
	Surprise	0	**99.17**	0.83	7.78	**84.44**	7.78	2	**94**	4	1.82	**89.09**	9.09	1.82	**92.73**	3.64
	Sadness	0	0	**100**	0	6.67	**93.34**	0	6	**94**	9.09	10.01	**87.27**	3.64	1.82	**94.55**
18 Channels	**Anger**	**74.17**	25.83	0	**55.56**	40	4.44	**66**	26	6	**63.64**	30.91	5.46	**58.18**	38.18	3.64
	Surprise	10.83	**88.33**	0.83	18.89	**72.22**	8.89	18	**64**	10	16.36	**70.91**	12.73	18.18	**72.73**	9.09
	Sadness	0.83	0.83	**93.33**	8.89	12.22	**78.89**	10	18	**72**	3.64	21.82	**74.55**	0	9.09	**90.91**
12 Channels	**Anger**	**73.33**	25.83	0.83	**60**	32.22	7.78	**52**	38	2	**69.09**	29.09	1.82	**56.36**	38.18	3.64
	Surprise	19.17	**80.83**	0	24.44	**65.56**	10	22	**62**	12	32.73	**58.18**	9.09	27.27	**61.82**	9.09
	Sadness	4.17	1.67	**94.17**	13.33	12.22	**74.44**	6	12	**78**	5.46	7.27	**87.27**	5.46	7.27	**87.27**
6 Channels	**Anger**	**69.17**	29.17	1.67	**61.11**	32.22	6.67	**68**	24	6	**73**	18.18	9.09	**65.46**	29.09	5.46
	Surprise	38.33	**58.33**	3.33	44.44	**41.11**	14.44	44	**28**	16	43.64	**43.64**	14.55	36.36	**49.09**	14.55
	Sadness	6.66	0	**93.33**	14.44	10	**75.56**	16	4	**74**	14.55	10.91	**74.55**	10.91	5.46	**81.82**

Correct answers are on the diagonal in bold. Columns represent participants' answered responses, with rows representing the correct answer (i.e., if all participants responded to all the stimuli, the rows across their sections should equal 100%). AD, patients with Alzheimer's disease; lvPPA, logopenic variant primary progressive aphasia; nfvPPA, nonfluent variant primary progressive aphasia; svPPA, semantic variant primary progressive aphasia.

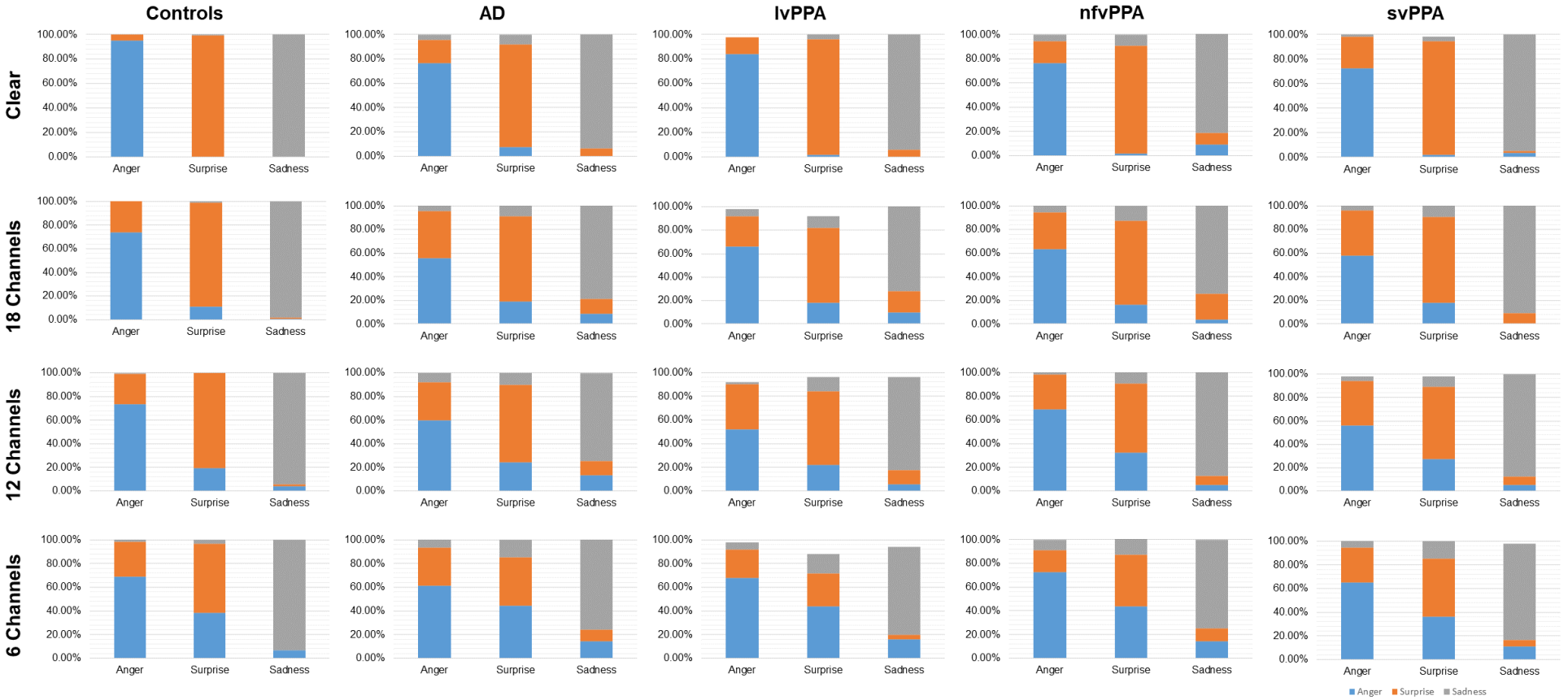


Figure 5.4. Error matrices for Experiment 3.

The colours indicate each groups' answered responses, while the columns in each stacked graph represents the correct answer. AD, patients with Alzheimer's disease; lvPPA, logopenic variant primary progressive aphasia; nfvPPA, nonfluent variant primary progressive aphasia; svPPA, semantic variant primary progressive aphasia.

For clear speech, diagnostic groups show similar patterns of identifying sadness most accurately, followed by surprise, and then anger. At 18 channels, AD, nfvPPA, and svPPA groups showed similar patterns to healthy controls on identifying sadness most accurately, followed by surprise, and then anger (with lvPPA more accurately identifying anger over surprise). Patient participants were also more likely to select 'surprise' for sadness stimuli. Error matrices for 12 channels, show AD, lvPPA, and svPPA groups displaying similar patterns to healthy controls on identifying sadness most accurately, followed by surprise, and then anger (with nfvPPA more accurately identifying anger over surprise). Healthy controls and AD participants were more likely to select 'anger' for sadness stimuli, whilst PPAs were more likely to select 'surprise' for sadness stimuli. Finally, error matrices at six channels show similar patterns across groups at identifying sadness most accurately, followed by anger, and then surprise.

5.5.2.3 Information Transfer

Using information transfer analysis, the information transfer score in diagnostic groups at clear, 18 channels, 12 channels, and six channels were analysed (**Table 5.2** and **Figure 5.5**). The differences between diagnostic groups were significant at clear, 18, and 12 channels ($p<0.01$).

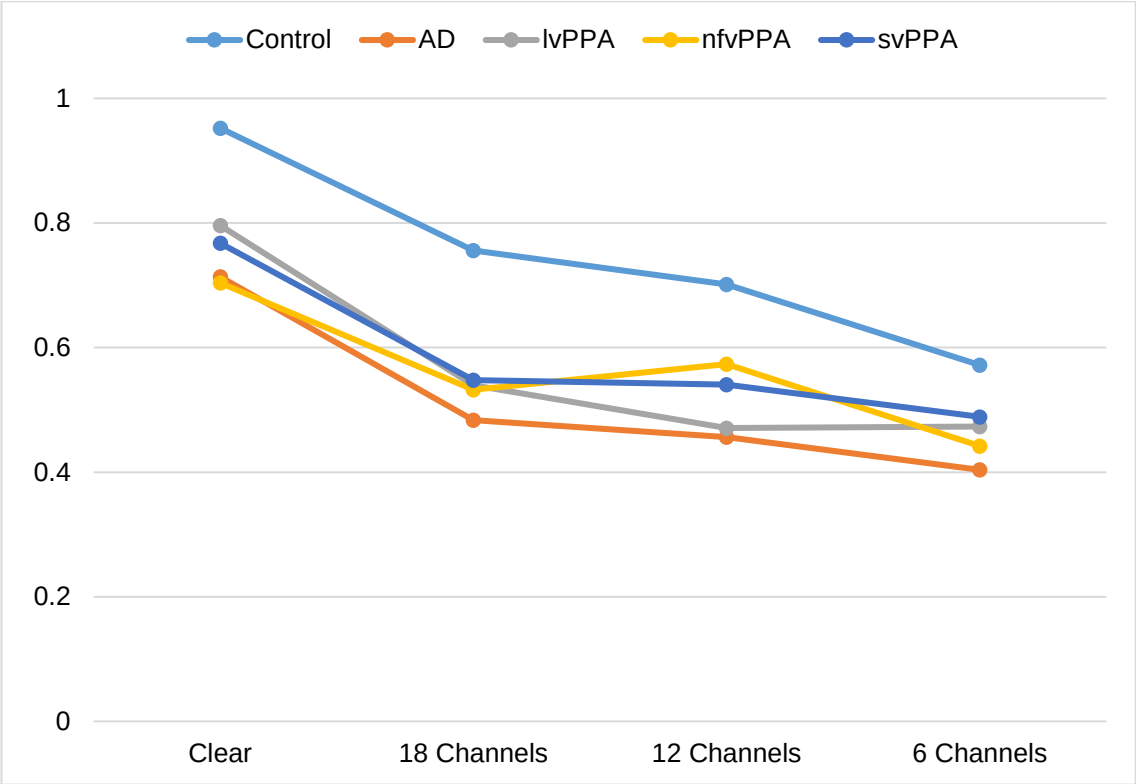


Figure 5.5. Line graph displaying average information transfer scores in each diagnostic group across stimuli conditions.
Mean scores are plotted on the line graph (to see standard deviations, refer to **Table 5.2**). A score of 1 represents perfect "transfer" of the stimuli to the participants' response (i.e., no confusion). A score of 0 represents no "transfer" of the stimuli to the participants' response (i.e., independent response to the stimuli). AD, patient group with typical Alzheimer's disease; Controls, healthy older control group; lvPPA, patient group with logopenic variant primary progressive aphasia; nfvPPA, patient group with nonfluent/agrammatic variant primary progressive aphasia; svPPA, patient group with semantic variant primary progressive aphasia.

5.5.2.4 Cost Analyses

To eliminate potential bias effects due to chance performance, participants who scored at or below chance in the control test were excluded from cost analyses. This excluded two patients with AD, one with nfvPPA, and one with svPPA. Another AD patient was excluded due to not detecting any of the angry stimuli. Considering the ceiling scores within the spoken numbers paradigm at six, 12, and 18 channels, z-score comparisons were not made between the identification of emotional prosody versus spoken numbers.

5.5.2.4.1 Cost on Emotional Prosody Performance

Comparing the cost of each noise-vocoding channel versus clear speech, there was a significant diagnostic effect (F(4)=3.01, p=0.019; see **Table 5.4** and **Figure 5.6**), particularly with lvPPA having a higher cost than controls (t=2.98, p=0.003), nfvPPA (t=2.66, p=0.008), and svPPA (t=2.01, p=0.046). There was a significant effect on the cost for the level of vocoding (F(2)=10.11, p<0.001), particularly that there was a significant difference between clear to 18 channels and clear to six channels (t=4.37, p<0.001), and between clear to 12 channels and clear to six channels (t=3.02, p=0.003).

Table 5.4. Average cost scores for comprehension of emotional prosody at each channel from clear speech, across each participant group

	Control	**AD**	**lvPPA**	**nfvPPA**	**svPPA**	**Omnibus significant test**
Emotional Prosody						
Clear – 18 Channels	1.88 (1.45)	2.56 (2.39)	3 (2.21)	2.2 (2.57)	2 (1.41)	H(4)=3.97, p=0.41
Clear - 12 Channels	2.21 (1.59)	2.94 (2.35)	3.8 (1.55)	2.2 (2.66)	3.2 (2.53)	H(4)=8.55, p=0.07
Clear – 6 Channels	3.58 (1.86)	4.25 (2.49)	5.1 (2.85)	3 (2.26)	3.3 (2.16)	H(4)=3.22, p=0.53=2
Numbers						
Clear – 18 Channels	0.01 (0.05)	0 (0.29)	0.2 (0.90)	0.18 (0.24)	0 (0)	H(4)=4.18, p=0.38
Clear - 12 Channels	0.01 (0.05)	0 (0.24)	0.3 (0.95)	0.11 (0.24)	0 (0)	H(4)=3.19, p=0.53
Clear – 6 Channels	0.04 (0.10)	**0.45 (0.52)**	**0.95 (0.89)**$^{+}$	**0.57 (0.61)**	0.2 (0.47)	H(4)=20.85, p<0.001

Mean (standard deviation) values are shown. Raw scores of the cost are presented. Significantly different from healthy controls (p<0.05) are in **bold.** $^{+}$significantly different to svPPA (p<0.05). Three nfvPPA participants that completed the verbal vocoding task (**Chapter 4**) did not complete the task in this experiment. AD, patient group with Alzheimer's disease; lvPPA, patient group with logopenic variant primary progressive aphasia; nfvPPA, patient group with nonfluent variant primary progressive aphasia; svPPA, patient group with semantic variant primary progressive aphasia.

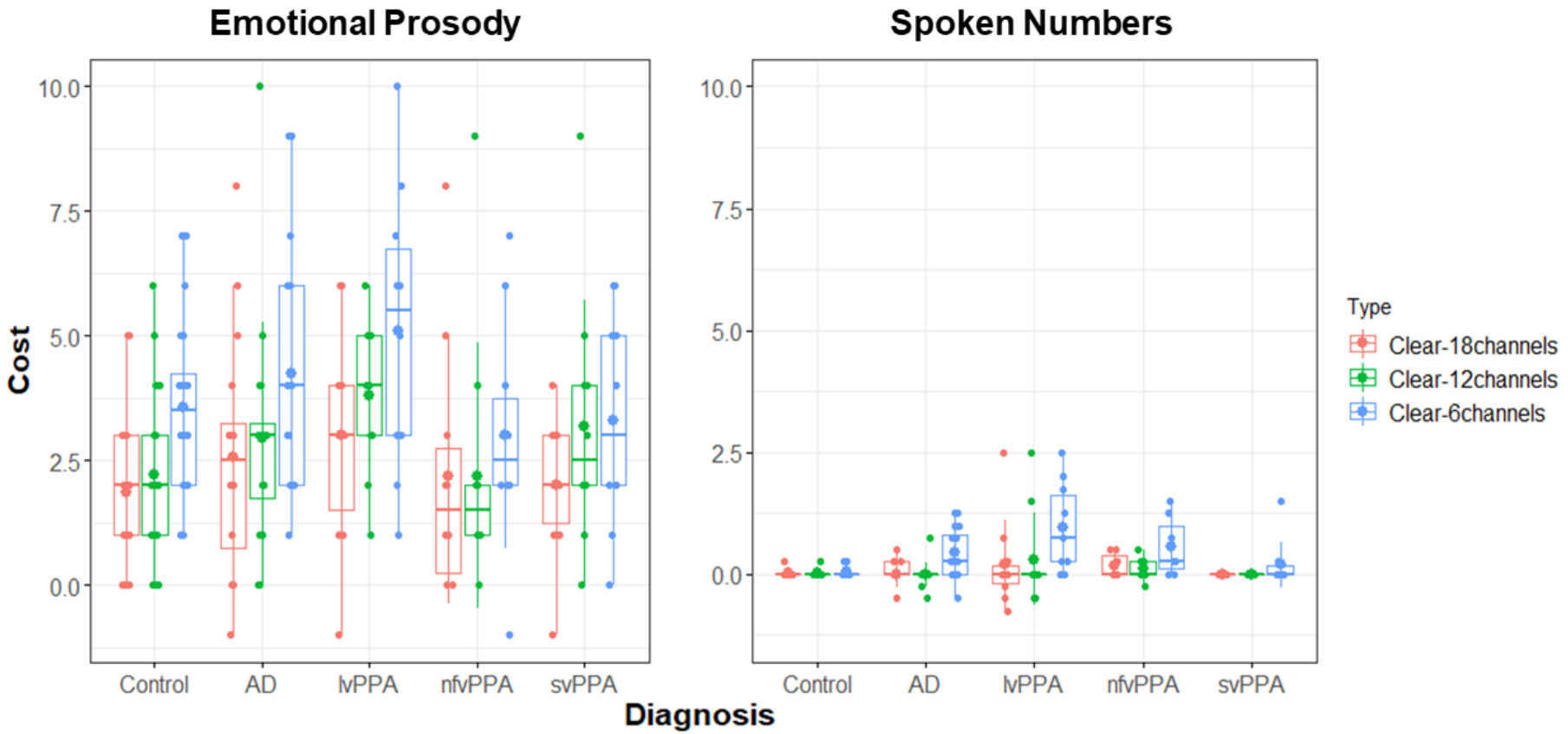


Figure 5.6. Boxplots of participant groups' cost scores from clear speech to noise-vocoded six, 12, and 18 channels, in both the emotional prosody and spoken numbers task.

The line within each box indicates the median, with the boxes indicating the interquartile interval. The dots represent each participant's performance. AD, patient group with Alzheimer's disease, Control, healthy older control groups; lvPPA, patient group with logopenic variant primary progressive aphasia; nfvPPA, patient group with nonfluent variant primary progressive aphasia; svPPA, patient group with semantic variant primary progressive aphasia.

5.5.2.4.2 Cost on Spoken Numbers Perception

Comparing the cost of each noise-vocoding channel versus clear speech, there was a significant diagnostic effect ($F(4)=7.03$, $p<0.001$; see **Table 5.4** and **Figure 5.6**), particularly with lvPPA and nfvPPA groups having a higher cost than healthy controls (lvPPA: $t=4.68$, $p<0.001$; nfvPPA: $t=2.35$, $p=0.020$). Further, lvPPA has a significantly higher cost than AD ($t=3.14$, $p=0.002$) and svPPA ($t=3.55$, $p=0.001$) groups. There was a significant effect on the cost for the level of vocoding ($F(2,186)=10.92$, $p<0.001$), particularly that there was a significant difference between clear to 18 channels and clear to six channels ($t=3.81$, $p<0.001$), and between clear to 12 and clear to six channels ($t=3.72$, $p<0.001$).

5.5.2.5 Correlations with Measures of Social Cognition

To assess the relation of emotional prosody, both in clear and noise-vocoded, to relevant aspects of social cognition (**Table 5.5**), the total scores of accurate identification of clear and noise-vocoded emotional prosody were correlated with scores on the mIRI and RSMS across the combined patient cohort. Accurate identification of clear emotional prosody was significantly correlated with the mIRI subscale: emotional empathy ($r(26)=0.45$, $p=0.015$) (**Figure 5.7**). Accurate identification of noise-vocoded emotional prosody was significantly correlated with the full mIRI ($r(26)=0.50$, $p=0.007$), mIRI subscale: emotional empathy ($r(26)=0.57$, $p=0.002$), the full RSMS ($r(25)=0.41$, $p=0.035$), and the RSMS subscale: sensitivity to expressive behaviour ($r(25)=0.44$, $p=0.023$) (**Figure 5.8**).

To assess any noise-vocoding effects on social cognition, correlations were also conducted on the clear and noise-vocoded spoken numbers. None of the correlations with mIRI and RSMS for the spoken numbers were significant (all $p>0.05$).

Table 5.5. Correlations of emotional prosody and spoken numbers with social cognition questionnaires in the combined patient cohort

		Clear Emotional Prosody	Vocoded Emotional Prosody	Clear Spoken Numbers	Vocoded Spoken Numbers
		r	r	r	r
mIRI	**Total**	0.330	**0.502**	0.077	-0.114
	Subscale: Empathy	**0.454**	**0.570**	0.186	-0.166
	Subscale: Perspective	0.125	0.335	-0.058	-0.031
RSMS	**Total**	0.152	**0.408**	0.007	-0.017
	Subscale: Modify self-presentation	0.129	0.120	0.035	-0.291
	Subscale: Sensitivity to Socio-emotional expressiveness	0.173	**0.435**	-0.013	0.005

The table shows results from the Modified Interpersonal Reactivity Index (mIRI) and Revise Self-Monitoring Scale (RSMS), as well as each of their two subscales, completed by each patient's primary caregiver (or another close informant). Correlations (Pearson's r value) are conducted on both the total clear emotional prosody identification score and the total noise-vocoded emotional prosody identification score. Significant correlations ($p<0.05$) are indicated in **bold**.

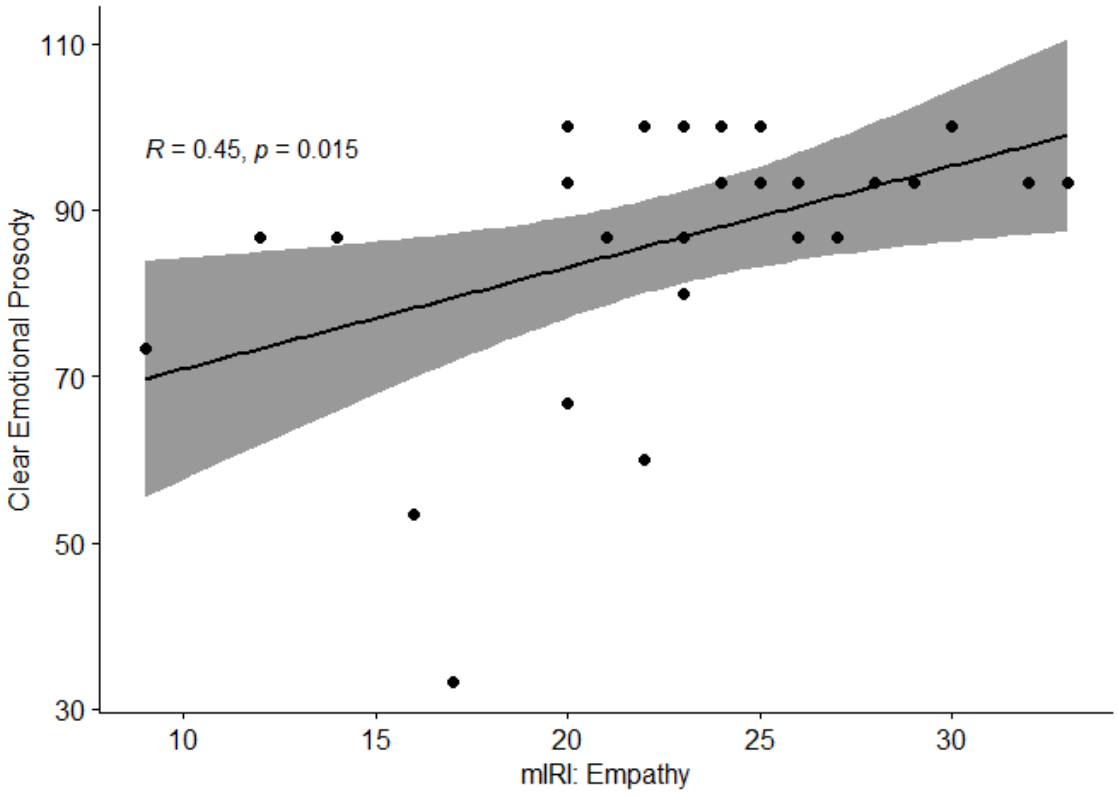


Figure 5.7. Correlation of clear emotional prosody with the modified interpersonal reactivity index, empathy subscale, across the patient cohort.
See **Table 5.5**. Pearson's R and p-value are shown at the top left-hand corner of the plot. Dots represent each participant's performance.

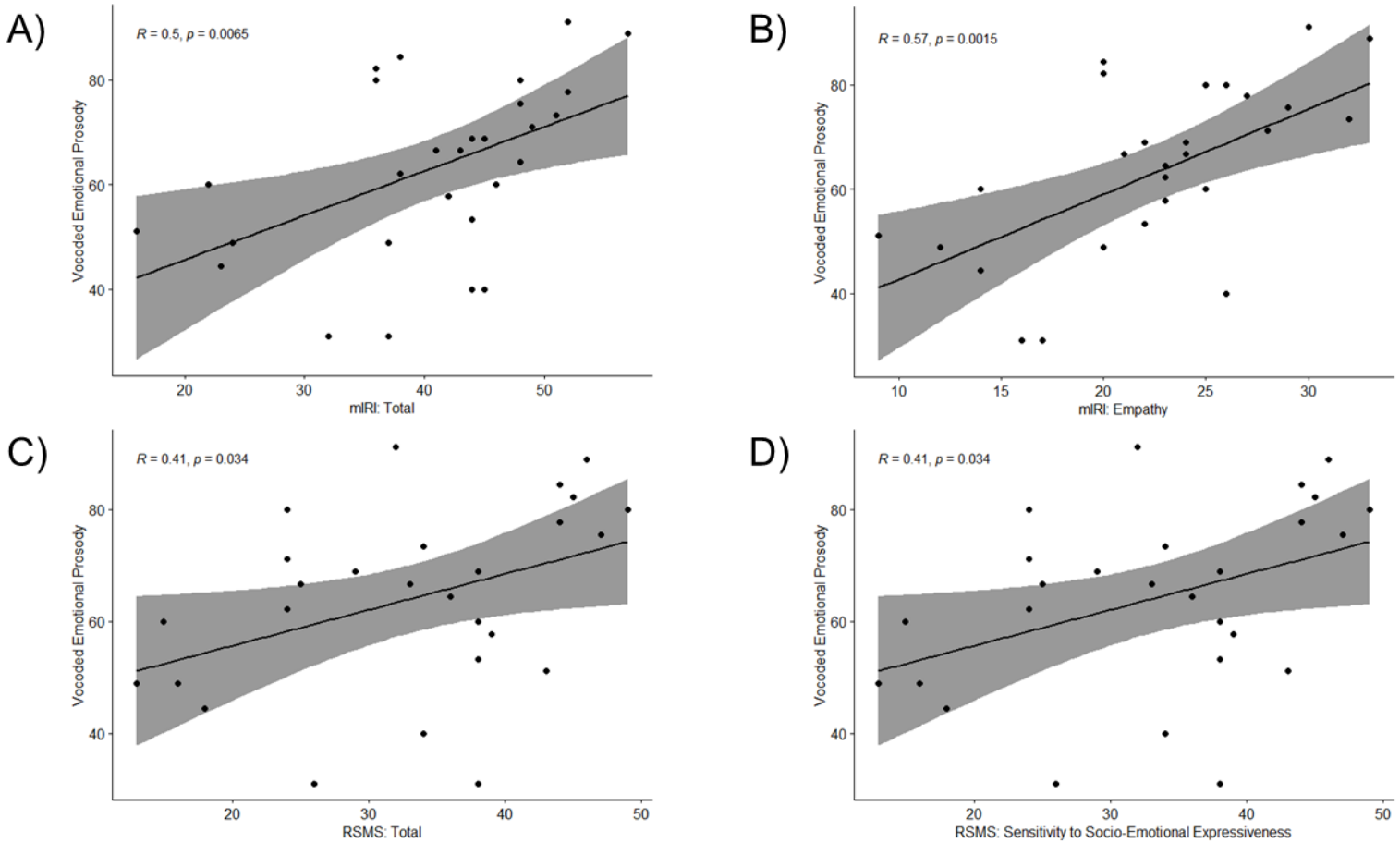


Figure 5.8. Correlation of vocoded emotional prosody with measures of social cognition across the patient cohort.
See Table 5.5. Pearson's R and p-value shown at the top left-hand corner of each plot. Dots represent each participant's performance.

5.6 DISCUSSION

Fitting with past research, all patient groups performed significantly worse than healthy controls at identifying emotional prosody in clear speech. This was particularly marked in patients with nfvPPA, as they also had the lowest information transfer (i.e., most confusion) of all groups for clear speech.

When emotional prosody stimuli were degraded through noise-vocoding, all patient groups performed significantly worse than healthy controls, particularly in AD and lvPPA. Even at each channel (six, 12, and 18 channels), AD and lvPPA groups performed significantly worse than healthy controls, in comparison to nfvPPA and svPPA having more variability depending on the channel. This could be attributed to AD and lvPPA both having the biggest decrease in the "transfer" of stimuli from clear speech to 18 channels, reflecting the difficulty with perceiving noise-vocoded auditory stimuli in general, as seen in **Chapter 4**. Further, lvPPA had a significantly higher cost from degrading emotional prosodic cues than healthy controls, nfvPPA, and svPPA. Again, this might plausibly reflect a core deficit in lvPPA and AD of apperceptive processing (e.g., the representation and decoding of auditory objects). There are likely to be stored neural 'templates' corresponding to the perceptual characteristics of the prosodic signatures of particular emotions, and under acoustic degradation, the neural template matching is stressed, similar to how AD and lvPPA show an impaired 'template activation' for phonemes (C. J. D. Hardy, J. L. Agustus, et al., 2017; Jiang et al., 2022).

Separately, differences were seen between six and 12 channels, and between six and 18 channels. The non-significant result and 'flattening' between 12 and

18 channels is unsurprising as it fits with how the perception of the noise-vocoding scale is exponential. This is also reflected in the cost difference scores on the type of channels and in comparing between the noise-vocoded channels as well. This is likely reflecting a high dependence of prosody perception on the accurate encoding of acoustic spectral features (Warren et al., 2006), regardless of channel number (noise-vocoding level). It also could be reflecting that prosodic information is lost in relatively “mild” vocoding.

The error matrices speculatively illustrate the acoustic confusability of the three different emotions (Banse & Scherer, 1996), particularly once degraded across different channels. In noise-vocoding manipulations, the “roughness” (as an example of voice quality/ timbre) of emotional prosody is likely to be impacted first, and therefore, identifying anger in comparison to the other two emotions is likely to be most affected initially. However, as degradation continues, more prosodic cues, such as voice pitch will be increasingly impacted, and therefore emotions such as surprise (which rely heavily on pitch to ascertain) would then be increasingly impacted at lower channels. On the other hand, cues such as intensity for anger would be retained, making it easier to perceive than surprise. In general, sadness is typically less affected by noise-vocoding due to its retained longer duration prosodic intonational cues, linked to slow f0 shifts.

The correlation found with executive functioning and working memory fits with past research implicating the involvement of fronto-parietal circuitry (Breitenstein et al., 2001; Darki & Klingberg, 2014; Sauseng et al., 2005). The correlation with PALPA-3 is surprising, especially considering that it was not correlated with the verbal vocoding analyses (see **Chapter 4**). In this experiment, the significant correlation could be showcasing that when spectral cues are eroded, prosody recognition is typically unable to heavily rely on the ‘regular’ cues associated with

emotional prosody (Everhardt et al., 2020), and is more reliant on other cues such as enunciation and temporal dynamics (aspects that are more heavily aligned with PALPA-3). Finally, the significant correlation with education could also be reflecting greater experiences with ‘puzzles’ and odd ‘tests’ that this paradigm is reflecting.

Most crucially, a significant correlation was found between accurate emotional identification and social cognition questionnaires (mIRI and RSMS). In clear emotional prosody, there was a significant correlation for the mIRI subscale: emotional empathy. For noise-vocoded emotional prosody, there was also a significant correlation in the mIRI subscale: emotional empathy, but additionally, the full mIRI, the full RSMS, and the RSMS subscale: sensitivity to socio-emotional expressiveness. In people living with dementia, noise-vocoding could present as a ‘stress test’ for the vocal emotion decoder that simulates more closely with the understanding of nonverbal emotional vocal signals in the noisy world at large in patients. This could be interpreted as due to the difficulty hearing emotional cues within our daily hearing environments (i.e. an intrinsic auditory deficit), patients are therefore less likely to engage with emotional cues being spoken and utilised towards them. Alternatively, the findings could be interpreted to show that if patients have an impairment in social cognition, they also have a decreased top-down processing that can affect their perception of degraded emotional prosodic cues. These two interpretations are of course not mutually exclusive – both mechanisms might operate jointly and could be mutually reinforcing.

Increasingly, especially within the FTD syndromes, it is important to be able to have an easily administered measure in clinic to characterise social cognition deficits. Potentially, with the significant correlations found with the perception of

degraded emotional prosody, this could be used as a clinical marker. However, prior to that, the mechanisms and relationships involved need to be further explored, and additionally, this paradigm should be studied in patients with bvFTD.

There are several limitations in this experiment and more work needs to be conducted to not only refine the paradigm used here but to also further understand the underlying mechanisms of degraded emotional prosody perception in healthy ageing and dementia, as well the implications for daily life function for patients. Firstly, given the results seen in **Chapter 4**, the selection of channels could be refined here to better represent levels of easy, medium, and hard difficulty per each diagnostic group (i.e., perceiving stimuli at six channels is more difficult for those with lvPPA than svPPA). Further, given the different testing paradigms in this experiment compared to **Chapter 4**, it did not provide an ideal direct comparison with the verbal (spoken numbers) stimuli presented there (i.e., the three-forced choice used here versus free direct verbal/written responses in **Chapter 4**). It would also be interesting to see whether comprehension of noise-vocoded emotional prosody can be modulated and improved pharmacologically, by dopaminergic or cholinergic stimulation as seen in the schizophrenia (Breitenstein et al., 2001; Kee et al., 1998), and/or perceptual learning (Hardy, Marshall, et al., 2018).

Further, to establish the brain basis and neural mechanisms for processing degraded prosodic and other emotional signals in dementia, further imaging analyses need to be conducted, both structural (e.g., MRI) and more pertinently, functional (e.g., fMRI and MEG). The marked deficit in nfvPPA could be reflecting the role of the fronto-temporo-parietal brain network found in past imaging studies on emotional prosody perception (Kotz et al., 2013; Jonathan D. Rohrer, Disa Sauter, et al., 2010; Wildgruber et al., 2006). Separately, recent research has

further found that the right IFG (Kotz et al., 2013) is particularly enhanced for vocal expressions of surprise (in comparison to other emotions), likely due to surprise being more acoustically and situationally ambiguous and requiring more top-down, contextual processes. Thus, the results seen in nfvPPA, particularly the differences in errors made on anger and surprise at six and 12 channels (when other cues are not as widely available), speak to the particular role of top-down processing in detecting surprise.

The group sizes were also relatively small in this paradigm, and considering the rarity of PPA, the collection of substantially larger datasets would require multi-centre collaboration. Particularly in the case of emotional prosody, this could potentially be extended to investigate which emotional nonverbal vocal signals are more easily transferred transculturally and cross-linguistically. Research should also be conducted in tonal language speakers with dementia as the processing of emotional vocalisations (e.g., prosodic cues needed) in these different populations may well differ in informative and clinically relevant ways. Likely, nonverbal vocalisations such as laughter or crying would be particularly suitable to study in cross-cultural and cross-linguistic populations. Finally, this work should be expanded to be conducted in bvFTD (as mentioned above) and the right temporal variant FTD (rtvFTD). Past research began to explore the impact of the disease on emotional prosody in rtvFTD (Perry et al., 2001), but more research should be conducted with the group, using imaging, and also comparing different types of prosody (e.g., linguistic and emotional) among rtvFTD and svPPA. This is key as there is a profound loss of emotional awareness documented in rtvFTD and could have a particular difficulty interpreting vocal emotional signals under non-ideal, real-world listening conditions.

Overall, the work presented in this chapter generally supported my three hypotheses (see **Section 5.3**). Firstly, patients with AD and PPA do perform worse than healthy controls at identifying clear emotional prosody. Secondly, AD and lvPPA patients showed an additional cost in identifying emotions once degraded, but in general not as large as the cost on verbal message comprehension. Thirdly, accurate identification of emotional prosody was significantly correlated with relevant measures of socio-emotional functioning in daily life.

6 EXPERIMENT 4: SINEWAVE ACCENT IDENTIFICATION

6.1 SUMMARY

Accents is a good model to study the role of paralinguistic cues, as it reveals a multitude of nonverbal information, such as a speaker's geographical and sociocultural background (Fletcher et al., 2013; Hailstone et al., 2012). However, how paralinguistic cues (relevant to accent processing) are affected in non-ideal auditory environments is yet to be understood. As a model for the comprehension of degraded paralinguistic signals, accent processing may be particularly pertinent to people with dementia, who (as shown by other work presented in this thesis) often struggle to process verbal and nonverbal messages under non-ideal listening conditions. Therefore, in this chapter, I investigate how accent identification is affected under sinewave manipulations (a radical loss of perceptual detail) in Alzheimer's disease (AD) and primary progressive aphasia (PPA). This expands on the work conducted by Hardy and colleagues (2018), where AD and PPA participants showed deficits in perceiving sinewave-degraded spoken numbers and place names (verbal content) in comparison to healthy age-matched control participants.

In this experiment, participants were tasked to determine the accent (Standard Southern British English, Standard American, Standard Russian) of a sinewave-manipulated spoken sentence. The results showed an accent identification deficit in patients with AD and logopenic variant PPA (lvPPA) against healthy controls in clear speech, and an accent identification deficit in all patient groups in sinewave degraded speech. Patients with lvPPA and semantic variant PPA (svPPA) had a significantly higher cost to accent identification under sinewave degradation in comparison to healthy controls. As patients with svPPA only

showed a deficit for accent identification in comparison to healthy controls when the stimuli were sinewave degraded, this suggests additional and substantial stress to the top-down semantic processing needed in decoding paralinguistic acoustic information. Further, the svPPA group, in comparison to all other patient groups, showed a trend for having a higher cost of sinewave degradation on accent identification than perception of spoken numbers. This speaks to the role of semantics in processing paralinguistic cues and suggests a higher demand to identify accents than highly familiar phonemic objects (e.g., spoken numbers), thus being particularly vulnerable to the primary semantic deficit in svPPA.

6.2 INTRODUCTION

In both **Chapter 3** and **Chapter 4**, I studied the identification of degraded speech in patients with AD and PPA using phonemic restoration and noise-vocoding, respectively. Hardy and colleagues (2018) studied the identification of degraded speech in these patient groups as well but used a different form of degraded speech: sinewave speech (see **Chapter 1.3.2.8**). Sinewave speech reduces speech signals by tracking the major formants of the speech signal and replacing them with sinewaves, making the speech signals sound like “whistled” tones (Remez et al., 1981). Typically, this radical perceptual alteration makes it difficult to understand as speech, but through perceptual learning, sinewave speech becomes intelligible to trained listeners (Hardy, Marshall, et al., 2018; Remez et al., 1981).

In Hardy and colleagues’ (2018) study, participants listened to both three-digit numbers and geographical place names and were told to repeat and/or write down the sinewave spoken word as accurately as possible. The study showed that patients with AD and PPA had deficits in perceiving the sinewave speech in

comparison to healthy age-matched control participants (Hardy, Marshall, et al., 2018). This was most marked in nfvPPA and lvPPA (paralleling the results seen in **Chapter 4**), with svPPA having a significant advantage at perceiving sinewave speech with highly predictable content (numbers) rather than less predictable content (geographical place names).

However, in daily communication, the information received from speech is never solely and strictly linguistic: rather speech signals can provide an abundance of other information as well (e.g., the emotional state of the speaker, see **Chapter 5**). This additional information is conveyed by paralinguistic acoustic perceptual cues, such as pitch, volume, tempo, and others.

Accents provide a model for how pattern of nonverbal auditory features are processed by the brain to derive semantic information about the speaker (e.g., geographical and/or socio-cultural characteristics). In addition, however, accents constitute paralinguistic' 'auditory objects' in their own right (e.g., 'English spoken with an American accent', 'English spoken with a French accent') (see **Chapter 1.3.2.1**) (Fletcher et al., 2013; Hailstone et al., 2012). Accents modify acoustic properties (e.g., pitch contour, rhythm, and stress patterns) of spoken phonemes and phrases in different ways, and interact with individual vocal characteristics and prosody (Hailstone et al., 2012). Therefore, identifation of accents presents itself as a suitable methodology on understanding the processing of nonverbal auditory semantic information of a speaker within neurodegenerative diseases.

In a past study, AD and nfvPPA patients had difficulty identifying both international and regional British accents compared to healthy controls (Hailstone et al., 2012). In a separate study comparing one nfvPPA and another svPPA patient on accent processing, both patients had impaired (albeit milder for the

svPPA patient) accent identification (Fletcher et al., 2013). Similar research has been conducted with voice identification, as the processing associated also occurs hierarchically in the same manner as accent identification: perceptual representations or templates derive from modality-specific recognition units, evoking semantic knowledge about the voice/accent. On voice identification, AD and svPPA patients were both impaired, more significantly in svPPA due to semantic retrieval (Hailstone et al., 2011).

Further research needs to be conducted on how paralinguistic semantic information about speakers can still be extracted even when the acoustic input is degraded, and how it is affected in neurodegenerative diseases. Therefore, in this experiment, I investigated how accent identification is affected by degraded auditory stimuli, using sinewave manipulations.

6.3 KEY PREDICTIONS

Expanding on Hardy and colleagues' (2018) work, accents that are sinewave degraded are used to investigate the role of perceiving paralinguistic cues under non-ideal listening conditions. My hypotheses are:

H_1: In line with impairments seen previously in AD and nfvPPA in accent identification, these groups and additionally lvPPA will have impaired clear accent identification in comparison to healthy controls.

H_2: Once sinewave degraded, impairments in accent identification will be seen in all patient groups, in comparison to healthy controls.

H_3: As accents are patterns of paralinguistic features that convey nonverbal semantic information about speakers, once degraded, svPPA patients will

disproportionately have difficulty with distinguishing accents than for spoken numbers stimuli (verbal) perception, as seen by Hardy and colleagues (2018).

6.4 METHODS AND MATERIALS

6.4.1 Participants

In this experiment, 29 patients and 17 healthy older controls were recruited, following general recruitment requirements detailed in **Chapter 2.1** (see **Appendix Table 8.1** for participant breakdown in each chapter). However, one control performed at chance in the control task (i.e., clear accent identification) and was therefore not included in the analyses. Five patients (two AD and three svPPA patients) struggled to complete the experimental paradigm and were therefore removed from the analyses as well. This resulted in the performance of eight patients with AD, six patients with lvPPA, four patients with nfvPPA, six patients with svPPA, and 16 healthy older controls, being used in this chapter (see **Table 6.1**).

6.4.2 Experimental Stimuli

Twelve English sentences ranging from four to nine words in length (e.g., 'Cover your mouth when you cough') were recorded in a soundproof booth by the same professional vocal coach in a technique known as the 'matched guise', where the speaker produces speech in different styles – in this case, three different accents: Standard Southern English, Standard American, or Russian (see **Table 6.2**). The sentences were used purely as a vehicle for conveying accent and comprehension of the verbal message was not assessed. Each of the recorded wavefiles was sinewave-transformed using a previously described procedure (Hardy, Marshall, et al., 2018).

Table 6.1. General demographic, clinical, and neuropsychological data for all participant groups

Characteristic	Controls	AD	lvPPA	nfvPPA	svPPA
Demographic and clinical					
No. M:F	8:8	3:5	5:1	0:4	4:2
Age, years	67.4 (5.30)	72.5 (8.83)	67.5 (5.65)	70.8 (9.29)	63.7 (4.23)
Handedness (R/L/A)	15:0:1	8:1:0	6:0:0	3:1:0	7:0:0
Education (y)	16.2 (2.71)[b]	14.1 (1.90)	15.2 (2.48)	14.2 (4.35)	14 (3.61)
Symptom duration (y)	NA	6.62 (3.29)	3.17 (1.33)*+	3.5 (1.91)+	6.33 (1.86)
Best ear average[1]	26.3 (5.32)	29.7 (7.94)[c]	32.5 (10.8)	23.9 (3.47)[a]	25.3 (6.00)
General intellect					
MMSE (/30)	29.7 (0.48)	**19.8 (5.52)**	**20.2 (7.22)**	**25.2 (3.77)**	**26 (5.93)**
Episodic memory					
RMT Faces (Short) (/25)	NA	17.6 (3.02)	24 (NA)	NA	NA
RMT Faces (Long) (/50)	44.2 (4.64)	NA	**33.2 (10.5)**	38.2 (3.86)[a]	**34.6 (3.91)[a]**
RMT Words (Short) (/25)	NA	15.2 (2.92)	23 (NA)	NA	NA
RMT Words (Long) (/50)	48.6 (1.89)	NA	**33.4 (12.2)**	**37.8 (9.07)[a]**	**36.6 (10.4)[a]**
Working memory					
Digit span forward (max)	7.31 (0.793)	**6 (0.76)**	**3.83 (2.48)***	**5 (1.83)***	7 (1.55)
Digit span reverse (max)	4.88 (1.20)	3.75 (0.71)*	**2.67 (1.75)***	**2.25 (0.5)***	5.83 (2.04)
Executive function					
WASI Matrices (/32)	26.8 (2.93)[a]	**12.5 (8.09)***	**15.8 (9.95)**	**13.2 (8.85)***	26.2 (3.19)
Letter fluency (total)	18.2 (.11)	**9.75 (4.13)**	**4.67 (5.92)***	**1.75 (1.71)*+**	**10 (3.10)**
Category fluency (total)	25.1 (6.03)	**5.62 (3.16)**	**9.17 (10.8)**	**9.5 (2.65)**	**7.83 (4.07)**
Auditory input processing					
PALPA-3 (/36)	35.4 (1.09)[b]	NA	**29.8 (6.94)*[a]**	35 (1.41)	35.5 (0.84)
Other language skills					
GNT (/30)	27.2 (1.97)	**11.6 (8.04)**	**11.4 (11.5)**	**14 (6.48)**	**0.17 (0.41)**
BPVS (/150)	149 (0.96)	**139 (7.63)***	**124 (40.6)**	**135 (17.9)**	**87 (32.7)**
PALPA-55 (/24)	23.9 (0.36)[a]	NA	**17.8 (2.36)*[b]**	19.8 (5.06)	22.8 (2.17)[a]
Other skills					
GDA calculation (/24)	14.7 (4.98)	**1.5 (1.31)***	**2.33 (2.88)***	**1.75 (2.06)***	14.2 (7.73)
VOSP Object Decision(/20)	18.9 (1.06)	**15.4 (2.56)**	**17 (2.28)**	**17.2 (0.5)**	17.8 (1.83)

Mean (standard deviation) values and raw scores are presented (maximum value possible in parentheses) unless otherwise indicated; significant differences from healthy controls ($p<0.05$) are in **bold;** *significantly different to svPPA ($p<0.05$); +significantly different to AD ($p<0.05$). [1]see Chapter 2 for details concerning the 'best ear average' measure. All AD patients and one lvPPA patient did the short version of the RMT Faces

and Words. A, ambidextrous; AD, patient group with Alzheimer's disease; BPVS, British Picture Vocabulary Scale; Controls; healthy older control group; Digit span forward/reverse, maximum digit span recorded; F, female; GDA, Graded Difficulty Arithmetic; GNT, Graded Naming Test; L, left; lvPPA, patient group with logopenic variant primary progressive aphasia; M, male; MMSE, Mini-Mental State Examination; NA, not available/applicable; nfvPPA, patient group with nonfluent/agrammatic variant primary progressive aphasia; PALPA, Psycholinguistic Assessments of Language Processing in Aphasia; R, right; RMT, Recognition memory Test; svPPA, patient group with semantic variant primary progressive aphasia; VOSP, Visual Object and Space Perception; WASI, Wechsler Abbreviated Scale of Intelligence. [a] missing data for 1 participant; [b] missing data for 2 participants; [c] missing data for 3 participants; [d] missing data for 4 participants.

Table 6.2. Spoken sentences used in Experiment 4

Standard Southern British English	**Standard American**	**Standard Russian**
Familiarisation (clear sentences)		
Pass the bread and butter	Don't spend too much	Computers make typing reports much easier
My friend was just fired from his job	Please don't give the dog that bone	Let the grill get really hot before you barbecue
Sinewave sentences		
I just had lunch	Clean the kitchen first	The office is busy
The lawn needs mowing	Cover your mouth when you cough	The tennis match was rained off
Don't go outside if it's too cold	He plays chamber music on Wednesday nights	My doctor says I'll be better soon
Don't run for a couple of days	He put his savings into the bank	We're going trick or treating on Halloween
Buy champagne for the New Year's Eve party	Iron that blouse before you go to school	I always pay my credit card bills on time
Don't take a long lunch	It's very windy outside today	They bought a new house
I have ten cousins	Let the dog out	That's a beautiful dress
Give the ducks some brown bread	Put these lights on the tree	It's going to be sunny tomorrow
Iron all of your shirts	They like to go fishing	That dog likes to run
You have to stop smoking before next year	Some baby pigeons were resting on the windowsill	I have a sore throat and a bad cough
He plays in a jazz band every Monday night	Walk the horses and clean their stalls before dinner	Visit your mother at least once a week please
I have to call the bank about my statement	Remember to keep all your receipts when going shopping	Professional musicians must practice 30 hours each week
Clear sentences		
Visit your mother at least once a week please	He plays in a jazz band every Monday night	Buy yourself a new jacket while they're on sale
My doctor says I'll be better soon	I always pay my credit card bills promptly	Put these lights on the tree
The office is busy	I just had lunch	Clean the kitchen first
The tennis match was rained off	I need to buy a suit	It's very windy outside today

The sentences were taken from the Boothroyd et al. (1985) sentences of topic-related conversation, used for assessing speech perception in adults with cochlear implants. Sentence length was counterbalanced across conditions.

6.4.3 Procedure

During the experiment, 36 sinewave-transformed sentences (12 representing each accent) and 12 clear sentences (four representing each accent) were presented. Stimuli were delivered from a laptop running Matlab (vR2014a) binaurally via headphones (Audio-Technica®) at a constant, comfortable listening level (at least 70 dB).

The sinewave stimuli were all delivered before the clear stimuli, to minimise priming effects. In each block (sinewave and clear speech), the order of accents was randomised. Participants were asked on each trial to decide the accent in which the sentence was spoken and responses were stored for offline analysis.

All participants were first familiarised with the experiment and given six practice trials of clear sentences (not presented in the subsequent experiment; see **Table 6.2**) to ensure they understood and could comply with the procedure. Throughout the experiment, no feedback about performance was given, and no time limits on responses were imposed.

All participants in this experiment had previously participated in Hardy and colleagues' (2018) experimental procedure, where participants were tasked on each clear and sinewave stimuli trial to repeat and/or write down the numbers as accurately as possible. Therefore, prior to the experiment presented here, the participants were already acclimated to sinewave-degraded speech.

6.4.4 Analysis of Data

Data was analysed using R© v4. To see details on statistical analyses conducted, see **Chapter 2.6**. Data for the identification of clear and sinewave-transformed numbers was previously published by Hardy and colleagues (2018) but has been reanalysed here for consistency with the experimental and control data.

The difference between clear and sinewave-transformed accent identification performance was calculated by converting scores on both tasks into percentages and subtracting the sinewave-transformed accent score from the clear accent score. Similar cost scores were generated from the difference between the clear and sinewave-transformed number task performance previously published (Hardy, Marshall, et al., 2018). To compare the results of identifying the accents in comparison to the perception of the spoken numbers, I created z-scores from the means and standard deviations of the healthy control groups' cost performance ($z = (x-\mu)/\sigma$, where x is the raw score, μ is the control groups' mean, and σ is the control groups' standard deviation).

Prior to the cost calculations, due to concerns over floor biasing effects as a result of chance performance, participants who scored at or below chance on clear accent identification were excluded from cost analyses. To ascertain whether participants scored at or below chance (i.e., a score that could have been achieved by random guessing), the cumulative probability function was adopted (https://stattrek.com/online-calculator/binomial). For the clear accent performance, 12 trials with probability 0.33' suggested that a hit rate (k) of eight or above was unlikely to be achieved by chance ($p=0.019$ [$k=7$, $p=0.066$]).

I created error matrices to understand the distribution of answers across the three choices of accents between each group. Information transfer analysis was also conducted (see **Chapter 2.6** for details) to attribute a score of confusion from stimuli choices.

Finally, I used Pearson's correlations to assess the relationship between accurate identification of degraded accents and general demographic (age, education), clinical (symptom duration, combined MMSE score), executive performance

(WASI Matrix), auditory perceptual (pure-tone audiometry), and working memory (digit span) factors.

6.5 RESULTS

6.5.1 General Participant Group Characteristics

General demographic, clinical, and neuropsychological data is given in **Table 6.1.** Participant groups did not differ in sex, age, handedness, education, or peripheral hearing function (all $p>0.05$). Patient groups did not differ significantly in MMSE score ($H(3)=5.76$, $p=0.12$), but did significantly differ for symptom duration ($F(3,20)=3.56$, $p=0.033$), with lvPPA having a significantly shorter duration than AD ($t=-2.69$, $p=0.014$) and svPPA ($t=-2.31$, $p=0.032$), and nfvPPA having a significantly shorter duration than AD ($t=-2.15$, $p=0.044$).

6.5.2 Experimental Behavioural Raw Data Analyses

Experimental data on accent and spoken number identification for all participant groups is presented in **Table 6.3**.

6.5.2.1 Accent Performance

There was a significant effect of diagnosis on clear accent identification performance (see **Table 6.3** and **Figure 6.1**), with the AD ($z=-3.93$, $p<0.001$) and lvPPA ($z=-1.70$, $p=0.045$) groups performing significantly worse than the healthy control group. The AD group also performed significantly worse than the svPPA group ($z=-2.69$, $p=0.004$). There was a significant effect of diagnosis on sinewave accent identification performance (see **Table 6.3** and **Figure 6.1**), with AD ($t=-5.17$, $p<0.001$), lvPPA ($t=-3.79$, $p<0.001$), nfvPPA ($t=-2.46$, $p=0.019$), and svPPA ($t=-2.94$, $p=0.006$) groups performing significantly worse than healthy controls.

Table 6.3. Performance of participant groups on experimental tasks

	Controls	AD	lvPPA	nfvPPA	svPPA	Omnibus Significance Test
Raw Data Performance						
Clear accent (/12)	10.7 (0.95)	**7.75 (1.49)**	**9.83 (0.75)**	9.5 (1.73)	9.83 (2.99)$^{+}$	H(4)=16.58, p=0.002
Sinewave accent (/36)	24.8 (4.11)	**13.8 (3.69)**	**16.3 (6.15)**	**18 (4.97)**	**17.8 (6.49)**	F(4,35)=8.42, p<0.001
Clear spoken numbers (/30)	29.9 (0.25)	29.8 (0.46)	**23.8 (7.0)*$^{+}$**	**22.2 (6.85)*$^{+}$**	30 (0)	H(4)=21.93, p<0.001
Sinewave spoken numbers (/120)	112 (6.27)	**97.8 (9.94)***	**66.8 (42.4)***	**76.8 (29.7)***	112 (3.14)	H(4)=22.16, p<0.001
Information Transfer Analyses						
Clear accent	0.80 (0.12)	**0.49 (0.12)***	**0.67 (0.08)***	0.66 (0.15)*	0.88 (0.14)	H(4)=22.09, p<0.001
Sinewave accent	0.41 (0.17)	**0.22 (0.09)**	**0.16 (0.13)**	**0.15 (0.09)**	**0.25 (0.15)**	H(4)=15.90, p<0.001
Cost Analyses						
Cost of sinewave on accent identification (%)	20.3 (15.3)	36.1 (7.17)	**38.0 (18.9)**	NA	**36.7 (14.8)**	F(3,27)=3.10, p=0.04
Z-scores for cost of sinewave accents	NA	1.03 (0.47)	1.16 (1.24)	NA	1.07 (0.97)	F(2,12)=0.02, p=0.98
Cost of sinewave on spoken number perception (%)	6.15 (5.31)	**18.1 (10.3)***	**23.8 (18.5)***	NA	5.67 (2.16)	H(3)=12.11, p=0.01
Z-scores for cost of sinewave spoken numbers	NA	2.25 (1.94)*	3.31 (3.49)*	NA	-0.09 (0.41)	H(2)=7.62, p=0.02

Participant group performance data for the key experimental tasks of interest assessing accent identification and spoken number perception, under clear and sinewave manipulation. Mean (standard deviation) values presented. Raw scores for experimental tasks are presented (maximum value in parentheses). Percentage scores are presented for cost analyses on participants who scored above chance on clear accent identification task. Z-scores created from cost analyses are also presented (see **Section 6.4.4**). Significant differences from healthy controls ($p<0.05$) are in **bold;** *significantly different to svPPA ($p<0.05$); $^{+}$significantly different to AD ($p<0.05$). AD, patient group with Alzheimer's disease; Controls; healthy older control group; lvPPA, patient group with logopenic variant primary progressive aphasia; NA, not available; nfvPPA, patient group with nonfluent/agrammatic variant primary progressive aphasia; svPPA, patient group with semantic variant primary progressive aphasia

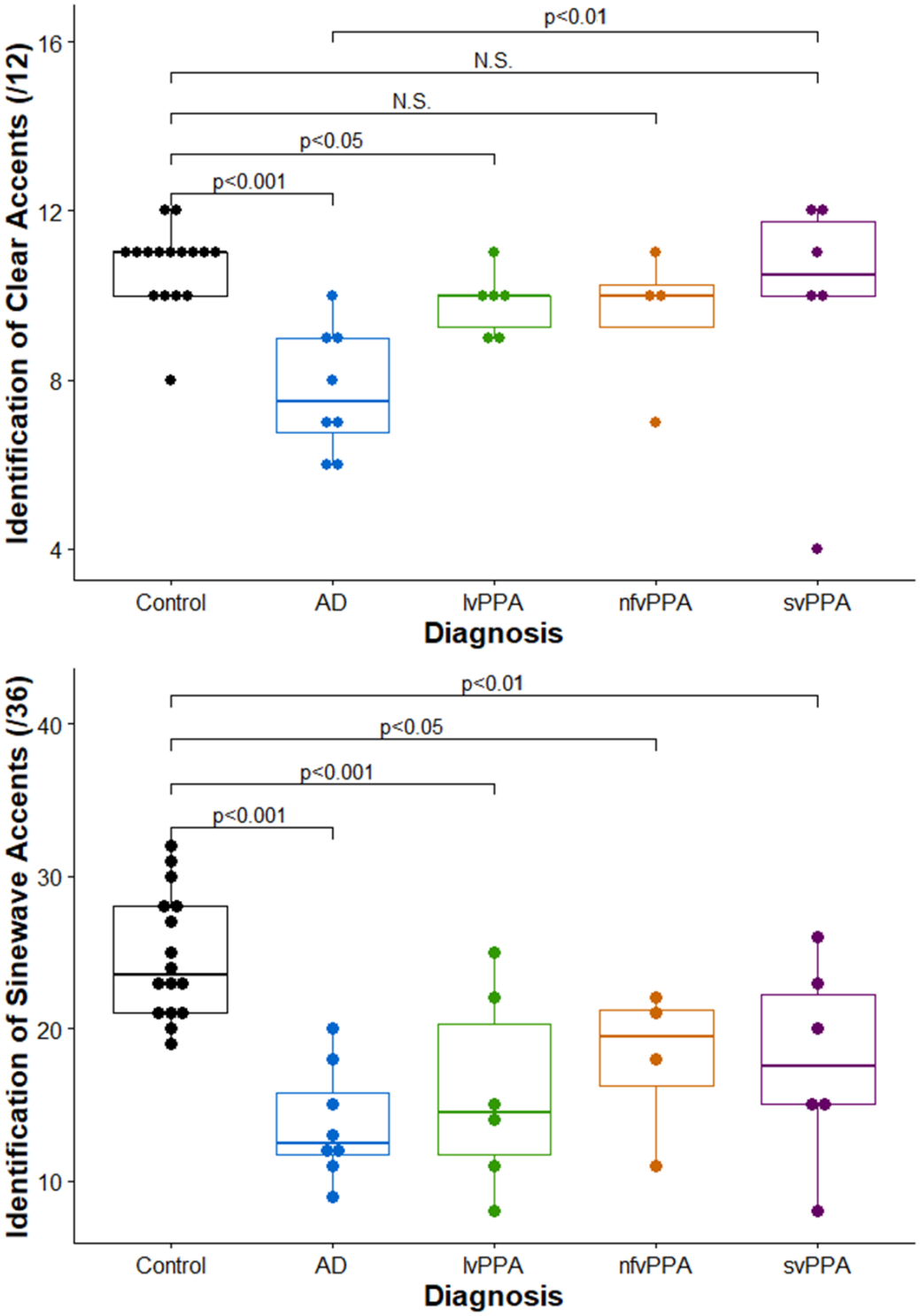


Figure 6.1. Boxplots showing performance on experimental (sinewave) and control (clear) accent tasks, for each participant group

Boxes represent the interquartile range, and whiskers indicate the overall range of values in each group; the horizontal line in each box represents the median. The dots code values for individual participants. AD, participant group with Alzheimer's disease; Control, healthy control participant group; lvPPA, participant group with logopenic variant primary progressive aphasia; nfvPPA, nonfluent/agrammatic variant primary progressive aphasia; N.S., not significant. svPPA, participant group with semantic variant primary progressive aphasia.

Total performance on the identification of sinewave accents in the patient cohort was not significantly correlated with age (r(22)=-0.20, p=0.348), education (r(22)=-0.04, p=0.852), disease duration (r(22)=-0.22, p=0.294), audiometry (r(18)=0.17, p=0.484), digit span forward (r(22)=0.31, p=0.146), and digit span backward (r(22)=0.26, p=0.21). It was correlated with MMSE (r(22)=0.58, p=0.003) and WASI Matrices (r(22)=0.42, p=0.039).

6.5.2.2 Spoken Number Performance

There was a significant effect of diagnosis on clear spoken number identification performance (see **Table 6.3** and **Figure 6.2**), with lvPPA and nfvPPA groups performing significantly worse than healthy controls groups (lvPPA: z=-3.73, p<0.001; nfvPPA: z=-3.01, p=0.001), AD (lvPPA: z=-2.68, p=0.004; nfvPPA: z=2.20, p=0.014), and svPPA (lvPPA: z=-3.28, p=0.001; nfvPPA: z=-2.78, p=0.003). The clear spoken number identification is significantly correlated with peripheral hearing (r=-0.46, p=0.006).

There was a significant effect of diagnosis on sinewave spoken number recognition performance (see **Table 6.3** and **Figure 6.2**), with the AD, lvPPA, and nfvPPA groups performing significantly worse than healthy controls groups (AD: z=-3.09, p=0.001; lvPPA: z=-3.39, p<0.001; nfvPPA: z=-3.01, p=0.001) and svPPA (AD: z=-2.27, p=0.011; lvPPA: z=-2.62, p=0.004; nfvPPA: z=-2.43, p=0.008).

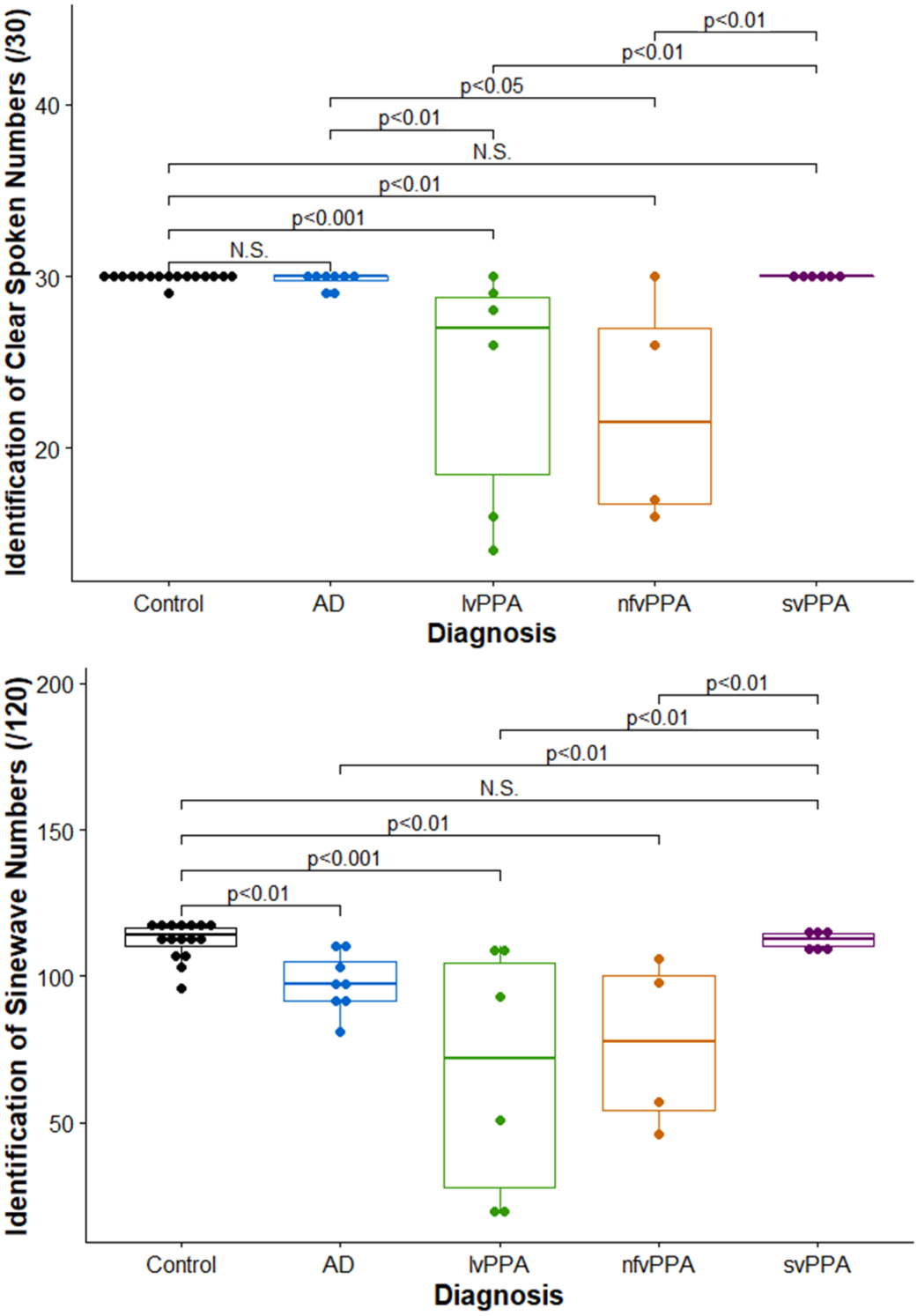


Figure 6.2. Boxplots showing performance on experimental (sinewave) and control (clear) spoken numbers tasks, for each participant group.
Boxes represent the interquartile range, and whiskers indicate the overall range of values in each group; the horizontal line in each box represents the median. The dots code values for individual participants. AD, participant group with Alzheimer's disease; Control, healthy control participant group; lvPPA, participant group with logopenic variant primary progressive aphasia; nfvPPA, nonfluent/agrammatic variant primary progressive aphasia; N.S., not significant. svPPA, participant group with semantic variant primary progressive aphasia.

6.5.2.3 Cost Analyses

To eliminate potential bias effects due to chance performance, participants who scored at or below chance were excluded from cost analyses. This excluded four patients with AD, one with nfvPPA, and one with svPPA. Since there were only three nfvPPA patients left, they were removed from the group analyses (see **Table 6.4**).

6.5.2.3.1 Cost on Accent Identification

There was a significant effect on diagnosis on the cost of sinewave degradation on the identification of different accents (see **Table 6.3** and **Figure 6.3**) with lvPPA (t=2.41, p=0.023) and svPPA (t=2.09, p=0.046) groups having a significantly higher cost of sinewave degradation than healthy controls.

6.5.2.3.2 Cost on Spoken Number Perception

There was a significant effect of diagnosis on the difference between clear versus sinewave spoken numbers recognition performance (see **Table 6.3** and **Figure 6.3**), with the AD and lvPPA groups having a significantly higher cost than healthy controls groups (AD: z=2.45, p=0.007; lvPPA: z=2.77, p=0.003) and svPPA (AD: z=1.93, p=0.027; lvPPA: z=2.07, p=0.020).

6.5.2.3.3 Cost of Accents VS Spoken Numbers

Using z-scores of the cost analyses (**Table 6.3**), there was a borderline significant difference between the cost of sinewave degradation on identifying the accents versus the spoken numbers in the svPPA group (V=15, p=0.063) (see **Figure 6.4**), where all five patients with svPPA had a higher cost in accent identification than the perception of spoken numbers.

Table 6.4. Individual patient performance profiles in Experiment 4

	Cost of Accents		Cost of Spoken Numbers	
	PERCENTAGE	Z-SCORES	PERCENTAGE	Z-SCORES
HEALTHY CONTROLS				
1	25		10.83	
2	33		3.33	
3	6		2.5	
4	19		1.67	
5	3		0.83	
6	44		10.83	
7	-17		1.67	
8	17		5.83	
9	14		2.5	
10	28		3.33	
11	33		14.17	
12	6		3.33	
13	28		20	
14	33		5	
15	22		7.5	
16	31		5	
ALZHEIMER'S DISEASE				
1	44	1.58	32.5	4.96*
2	28	0.49*	8.33	0.41
3	33	0.85	14.17	1.51*
4	39	1.22	17.5	2.14*
LOGOPENIC VARIANT PRIMARY PROGRESSIVE APHASIA				
1	53	2.13	30	4.49*
2	53	2.13	50.83	8.41*
3	50	1.94	37.5	5.90*
4	14	-0.42	9.17	0.57*
5	44	1.58*	5	-0.22
6	14	-0.42	10	0.73*
NONFLUENT VARIANT PRIMARY PROGRESSIVE APHASIA				
1	25	0.31	9.17	0.57*
2	22	0.13*	-1.66	-1.47
3	42	1.40	15	1.67*
SEMANTIC VARIANT PRIMARY PROGRESSIVE APHASIA				
1	42	1.40*	8.33	0.41
2	28	0.49*	4.17	-0.37
3	36	1.03*	7.5	0.25
4	19	-0.06*	3.33	-0.53
5	58	2.49*	5	-0.22

All cost scores (percentages) are presented for each participant. The higher the number, the larger the cost of the sinewave degradation. Z-scores were created from the average and standard deviation of the healthy control group. *signifies the whether accents or spoken numbers had the higher cost (i.e., the higher cost has a * designated beside it).

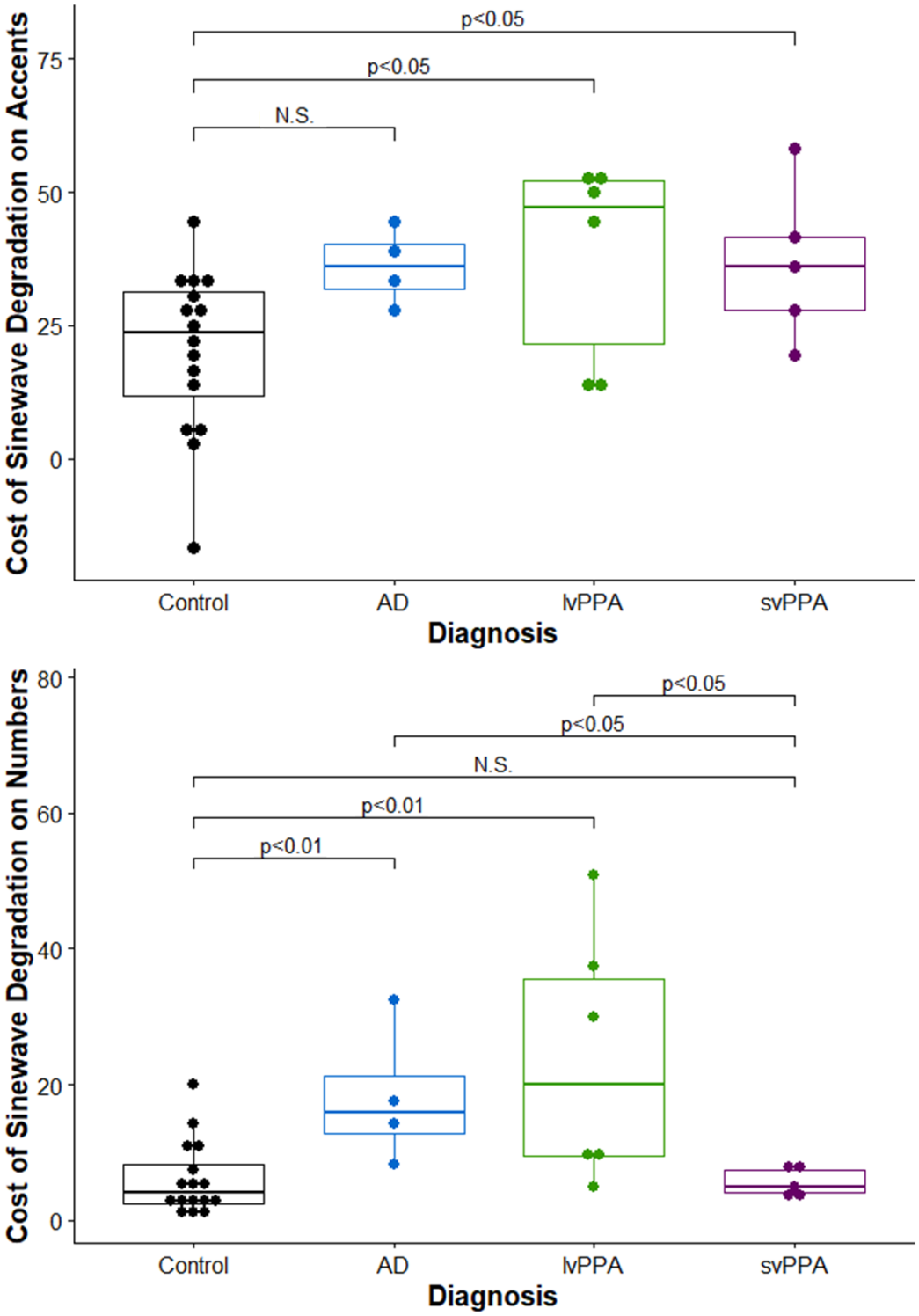


Figure 6.3. Boxplots showing cost of sinewave degradation on accent identification and spoken number perception, for each participant group

A positive cost indicates a higher amount of cost because of the sinewave manipulation. Boxes represent the interquartile range, and whiskers indicate the overall range of values in each group; the horizontal line in each box represents the median. The dots code values for individual participants. AD, participant group with Alzheimer's disease; Control, healthy control participant group; lvPPA, participant group with logopenic variant primary progressive aphasia; N.S., not significant. svPPA, participant group with semantic variant primary progressive aphasia.

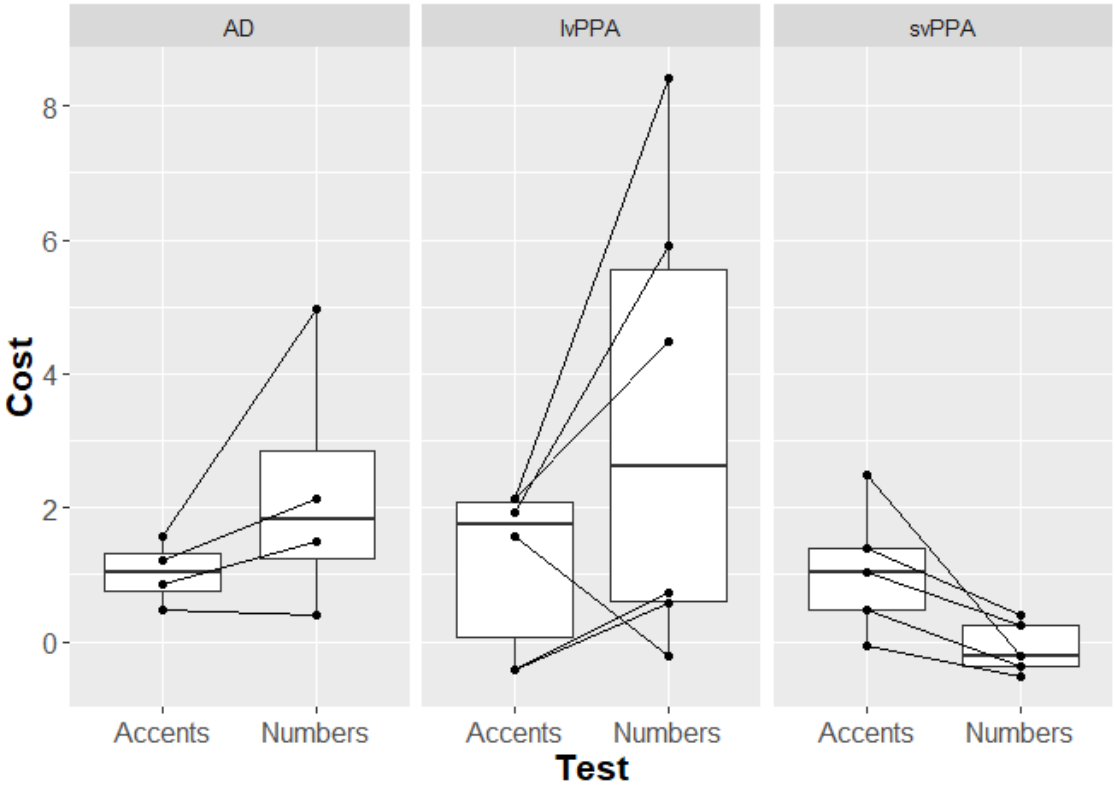


Figure 6.4. Paired boxplots showing the z-score cost of sinewave degradation on accent identification and spoken number perception, for AD, lvPPA, and svPPA. A positive cost indicates a higher amount of cost as a result of the sinewave manipulation (i.e., the z-scores on sinewave-manipulated stimuli are worse than the z-scores in clear stimuli). Boxes represent the interquartile range, and whiskers indicate the overall range of values in each group; the horizontal line in each box represents the median. The dots code values for individual participants, with the line linking each participant's performance in the different tests. AD, participant group with Alzheimer's disease; Control, healthy control participant group; lvPPA, participant group with logopenic variant primary progressive aphasia; svPPA, participant group with semantic variant primary progressive aphasia.

6.5.2.3.4 Individual Patient Performance Profiles on Cost Analyses

Individual patient performances are seen in **Table 6.4**. While five patients with svPPA had a higher cost in accent identification than their perception of spoken numbers, three out of four AD patients, five out of six lvPPA patients, and two out of three nfvPPA patients had a higher cost on the perception of spoken numbers than accent identification.

6.5.2.4 Error Matrices and Information Transfer Analysis

Error matrices are presented in **Table 6.5** and **Figure 6.5**. Overall, the lvPPA and nfvPPA groups identified Russian accents, and the AD group identified British accents most consistently, in clear and sinewave-degraded conditions. Patients with svPPA best identified British accents in clear speech, and Russian once sinewave degraded. Healthy controls identified Russian most consistently in clear speech and British once sinewave degraded.

Using information transfer analysis (see **Table 6.3**), the AD and lvPPA patient groups showed significantly higher confusion (i.e. lower information transfer score) for clear accent identification in comparison to healthy controls and svPPA patients. The nfvPPA patient group also showed a significantly lower information transfer score for clear accent identification than svPPA patients. Under sinewave degradation, all patient groups had a significantly higher confusion for accent identification in comparison to healthy controls.

Table 6.5. Percentage of answer types for each intended accent in clear and sinewave speech averaged over the participants in each group

		Controls			AD			lvPPA			nfvPPA			svPPA		
	Answered/ Expected	English	American	Russian	English	American	Russian	English	American	Russian	English	American	Russian	English	American	Russian
Clear	English	**81.25**	14.06	4.69	**84.38**	12.5	3.13	**62.5**	20.83	16.67	**68.75**	18.75	12.5	**95.83**	4.17	0
	American	9.38	**85.94**	4.69	43.75	**53.13**	3.13	20.83	**70.83**	8.33	25	**68.75**	6.25	33.33	**66.67**	0
	Russian	0	0	**100**	9.38	28.13	**56.25**	4.17	0	**95.83**	0	0	**100**	0	16.67	**83.33**
Sinewave	English	**76.04**	17.19	4.69	**45.83**	29.17	20.83	**43.06**	29.17	27.78	**50**	33.33	16.67	**58.33**	22.22	19.44
	American	31.25	**56.77**	11.46	42.71	**32.29**	21.88	43.06	**36.11**	20.83	39.58	**39.58**	20.83	40.28	**37.5**	22.22
	Russian	17.19	8.33	**73.96**	29.17	27.08	**42.71**	22.22	25	**52.78**	22.92	16.67	**60.42**	13.89	26.39	**59.72**

Correct answers are on the diagonal in bold. Columns represent participants' answered responses, with rows representing the correct answer (i.e., if all participants responded to all the stimuli, the rows across their sections should equal 100%). AD, patients with Alzheimer's disease; lvPPA, logopenic variant primary progressive aphasia; nfvPPA, nonfluent variant primary progressive aphasia; svPPA, semantic variant primary progressive aphasia.

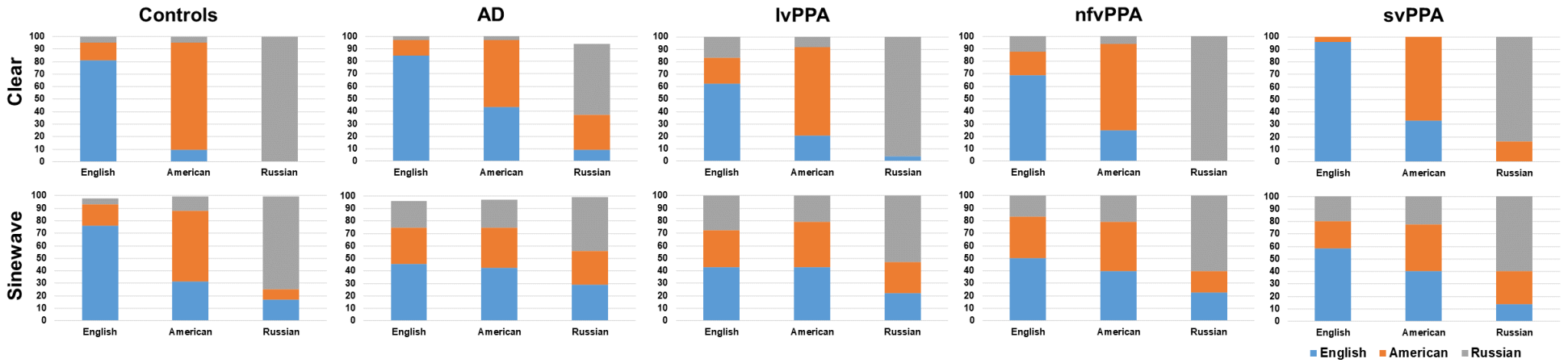


Figure 6.5. Error matrices for Experiment 4

The colours indicate each participant's choice response in percentage, while the columns in each stacked graph represent the correct answer. AD, patients with Alzheimer's disease; lvPPA, logopenic variant primary progressive aphasia; nfvPPA, nonfluent variant primary progressive aphasia; svPPA, semantic variant primary progressive aphasia.

6.6 DISCUSSION

Following the findings from Hailstone et al. (2012), I've shown here that there is an accent identification in clear speech deficit in AD (and now including the language variant of AD, lvPPA) patients against healthy controls. For nfvPPA, it is likely that due to small numbers and insufficient power, the clear accent raw score did not come out to be significantly less than healthy controls. However, using information transfer analysis, the nfvPPA patient group did show a significantly lower information transfer score (i.e., higher confusion) for clear accent identification in comparison to svPPA patients. AD and lvPPA patient groups also had significantly lower information transfer scores than svPPA patients and healthy controls.

Once the accents were sinewave degraded, correct identification of accents was impaired in all patient groups in comparison to healthy controls. This was particularly evident in lvPPA and svPPA groups as they had a significantly higher cost of sinewave degradation of accent identification than healthy controls. To fully decode sinewave degraded accents (as a form of paralinguistic acoustic information), it depends on an interplay of bottom-up perceptual and top-down semantic processing interacting (i.e., an 'apperceptive' template matching) that enables identification of accents (Hailstone et al., 2012).

The difficulties seen in lvPPA and AD are likely due to bottom-up perceptual dysfunction and impaired template matching, previously also seen in the other experiments in this thesis. In svPPA, even though top-down semantic mechanisms are sufficient enough to recognise clear accents, the deficit appears under sinewave transformation, where the manipulations had significantly stressed the semantic mechanisms enough to result in impairments in the

identification of the accents. In parallel to the verbal paradigm conducted previously (Hardy, Marshall, et al., 2018) and in visual object decision (Hovius et al., 2003), patients with svPPA have a deficit in the very fundamental process where sensory object information is used to engage and update stored templates for object concepts (Ruksenaite et al., 2021).

Following the spoken numbers results (Hardy, Marshall, et al., 2018), AD, lvPPA, and nfvPPA groups performed significantly worse than healthy controls and svPPA patients at the sinewave spoken numbers recognition. The lvPPA and nfvPPA group did perform significantly worse than the healthy controls at clear spoken number perception and as this was significantly correlated with pure-tone audiometry, this deficit is likely led by peripheral hearing impairment. Similar to the cost on accent identification, the lvPPA group also had a significantly larger cost from the sinewave degradation of the spoken numbers than healthy controls. The AD group separately showed a significantly higher cost for the verbal spoken numbers paradigm, but not for accent identification. Particularly in comparing directly the cost of sinewave manipulation on accent identification and spoken number paradigm, all patient groups, except for svPPA, showed a trend of having a higher cost of sinewave degradation on verbal content than the paralinguistic cues.

This is likely to reflect disease-specific mechanisms affecting either perception of paralinguistic or verbal content under degraded listening conditions. Accents are likely more demanding to identify the highly familiar spoken numbers, and therefore are more vulnerable to the primary semantic deficit in svPPA. On the other hand, the identification of spoken numbers depends more on fine-grained decoding of spectrotemporal details and therefore is more vulnerable to the deficit in spectrotemporal feature analysis in nfvPPA, lvPPA, and AD.

The error matrices for the sinewave degraded speech provide speculative insight into the difficulties with template matching to particular accents. Standard American accents were the hardest to identify in comparison to the other two accents across participant groups. This may be due to despite being a foreign accent to all of our participants, in comparison to Russian, it is more ambiguously similar to Standard Southern British accents and our participants are likely to be frequently exposed to American accents due to mass media. In other words, the 'templates' of an American accent are likely to be more blurry in comparison to a Russian foreign accent, resulting in larger difficulties with identification and frequent confusion with the Standard Southern British accents. However, this is difficult to fully implicate due to similarities in acoustic-phonetic parameters as a result of the same speaker and an "acted" foreign accents in English.

The correlation found with executive function could be reflecting difficulties with cognitive flexibility, particularly in listening and identifying a 'deviant' accent due to the sinewave degradation (Adank & Janse, 2010), decrease in speed on adaptation to accents (Banks et al., 2015), as well as allocation of executive resources with more effortful listening (Peelle, 2018). Additionally, with the correlation to overall disease severity, indexed by the MMSE, the processing of paralinguistic cues can add to measures of central hearing being a suitable signal for functional integrity in neurodegenerative pathologies (Johnson et al., 2021).

A limitation of this experiment is the small sample size, particularly in the nfvPPA group. This was mostly due to the paradigm being too difficult for many of the participants: many did not pass the training and therefore the experiment could not be conducted. However, this could, be interpreted as the enhanced difficulty with paralinguistic cues processing under a degraded context for patients (in comparison to degraded verbal content), and belies the real-world

communication difficulties that they may have, particularly for social purposes (as seen in **Chapter 5**). Another limitation is that the paradigm did not allow for the best direct comparison to the verbal (spoken numbers) stimuli results of Hardy and colleagues (2018), as the response method in this experiment for each participant was a three-way forced choice, in comparison to a free direct verbal/written response in the spoken numbers paradigm.

Future directions of this research should be conducted in larger cohorts where possible, expanding to other neurodegenerative diseases (particularly rtvFTD patients), as well as creating a paradigm that will allow for a more direct comparison of processing of sinewave degraded verbal and paralinguistic cues in speech among neurodegenerative diseases. However, creating paradigms to allow for direct comparisons is tricky as even the 'lexicons' and the perceptual characteristics of spoken words and accents are quite dissimilar. Therefore, what may be a better paradigm is to incorporate a verbal comprehension of accents, as this would allow for an assessment of the effect of noisy listening conditions on both linguistic and paralinguistic information conveyed by accents. Further, future studies could assess the role of perceptual learning in processing paralinguistic cues associated with speech (Hardy, Marshall, et al., 2018).

Similarly to **Chapter 5**, future research should also be conducted with structural and functional neuroimaging in conjunction with behavioural analysis, to understand the underlying neural mechanisms involved. As accents are paralinguistic phenomena and not strictly nonverbal (i.e., linguistic information may be needed to fully ascertain the accent type), the neural mechanisms may be more similar to linguistic prosody, incorporating both right and left-hemisphere neural correlates.

Overall, the work presented in this chapter generally supported my three hypotheses. Firstly, AD and lvPPA patients did show significant impairments in clear accent identification. Secondly, once sinewave degraded, all patient groups had a significantly impaired identification performance than healthy controls. Thirdly, svPPA patients had a significantly higher additional cost from sinewave degrading accent identification than healthy controls. They also had borderline significantly higher cost on sinewave accent identification than spoken numbers perception. Crucially, this was flipped to other patient groups' profiles (in terms of higher costs on paralinguistic versus verbal content), suggesting a stronger top-down semantic involvement is needed for patterns of paralinguistic information.

7 GENERAL DISCUSSION

7.1 SUMMARY OF EXPERIMENTAL FINDINGS

This thesis set out to explore degraded verbal and nonverbal dimensions of speech perception in AD and PPA, and stratify the differences in auditory symptoms seen in the diseases through new central hearing paradigms. This was conducted by designing new measures to probe cortical mechanisms of degraded speech perception, comparing each patient group against healthy older listeners and other patient groups, as well as assessing performance in comparison to measures of peripheral hearing function.

7.1.1 Chapter 3: Phonemic Restoration

Experiment 1 aimed to address how phonemic restoration, a naturalistic and automatic auditory mechanism that 'repairs' interrupted speech signals, is affected in dementia. The results showed that the healthy controls, and both AD and svPPA patients showed a retained phonemic restoration of real words, but differed in performance on pseudowords. In healthy controls, there was no clear bias for phonemic restoration of pseudowords. In contrast, patients with AD showed a marked tendency to perceive noise segments as replacing phonemes in pseudowords, interpreted as a 'suppression of phonemic restoration'. On the other hand, patients with svPPA showed comparable performance between pseudowords and real words.

These different "auditory profiles" in the two diseases together illuminate the underlying brain mechanisms of phonemic restoration: phonological representation (situated in the posterior STS/STG) is likely to interact with a modulatory, top-down mechanism of semantic prediction and disambiguation

(Shahin et al., 2009; Sunami et al., 2013). The findings in the chapter speak to a partially intact capacity for neural plasticity, as a neural mechanism for 'repairing' degraded speech is found to be retained in certain dementias.

7.1.2 Chapter 4: Noise-Vocoded Verbal Messages

Experiment 2 aimed to determine a degraded speech intelligibility threshold in different dementia syndromes, correlating them with demographic and disease characteristics, real-world hearing functions, and structural neuroanatomical associations. The results showed that AD and PPA syndromes have an elevated threshold for intelligibility of noise-vocoded speech signals, in comparison to healthy controls, particularly in lvPPA and nfvPPA. This elevated threshold did not correlate with measures of pure tone detection or phoneme discrimination in clear speech, suggesting that the deficits shown could not be explained by an impairment in peripheral hearing or clear speech perception. Neuroanatomically, the elevated threshold was correlated with atrophy of the left planum temporale, angular gyrus, and anterior cingulate gyrus (a cortical network that is critical for processing degraded speech signals (Davis & Johnsrude, 2003; Chris J. D. Hardy et al., 2017; C. J. D. Hardy, J. L. Agustus, et al., 2017; Jiang et al., 2021; Wild et al., 2012)).

The impairment seen in the perception of vocoded verbal messages in AD and PPA speaks to different pathophysiology in each disease that affects degraded speech processing (see **Figure 7.1**). Clinically, the experiment suggests a potential future application of the vocoded speech intelligibility threshold, along with other sensitive measures of detecting dementia, as a quantifiable biomarker or 'stress test' for early diagnosis of AD and PPA, considering the promising results seen with the ROC curve analyses, and the real-world relevance of this

measure to the processing of spoken messages under challenging listening conditions in daily life.

7.1.3 Chapter 5: Noise-Vocoded Emotional Prosody

Experiment 3 aimed to address how nonverbal cues associated with speech, like emotional prosody, are affected, particularly and novelly, in a non-ideal listening environment. The results showed that AD and PPA patients were impaired at identifying emotional prosody in speech, in comparison to healthy controls. The impairment was sustained when the prosodic cues were degraded, with a significantly higher cost for lvPPA patients than healthy controls. This is likely to reflect a deficit seen in AD and lvPPA with apperceptive processing (e.g., neural template matching; also seen as a potential mechanism in **Chapter 3** and **Chapter 4**).

Error matrices also showcase certain differential vocoding effects on the audibility of prosodic cues, such as sad prosodic cues being less affected by noise-vocoding, compared to other prosodic cues that rely more heavily on other acoustic characteristics (e.g., dynamic intensity variations). This impairment significantly correlated with daily life measures of social cognition, providing a potential window on real-world emotional communication under non-ideal listening conditions.

7.1.4 Chapter 6: Sinewave Accents

Experiment 4 aimed to investigate the identification of accents, as a model of how patterns of paralinguistic features convey nonverbal semantic information about speakers, and how their processing is affected in non-ideal listening environments. As in Experiment 3, this experiment is one of the first paradigms to address how paralinguistic information is processed under degraded listening

conditions. The results showed that there was a significant impairment in clear accent identification in AD and lvPPA compared with healthy controls, and an identification deficit was also seen in all patient groups when the accents were sinewave degraded. Patients with lvPPA and svPPA had a significantly higher cost in the identification of sinewave-transformed versus clear accents, in comparison to healthy controls. In comparing the effects of sinewave manipulation on spoken numbers versus accent identification, svPPA patients had a higher cost on identifying accents than spoken numbers, versus the other patient groups having a higher cost on the perception of spoken numbers.

These findings may speak to disease-specific mechanisms affecting accent identification under degraded listening conditions. As nonverbal auditory objects, identifying accents is likely to be more of a demanding task on the semantic system than solely recognising highly familiar phonemic objects (e.g., spoken numbers), and therefore, more vulnerable to the primary semantic deficit in svPPA. On the other hand, spoken number identification may depend on more fine-grained decoding of spectrotemporal detail than accent identification, and therefore, more vulnerable in nfvPPA and lvPPA, where the primary deficit lies with spectrotemporal feature analysis.

7.2 Degraded Speech Perception in AD and PPA

The perception, and ultimately, the understanding of degraded speech relies upon flexible and dynamic neural interactions across distributed brain networks. These physiological and anatomical substrates are intrinsically vulnerable to the disruptive effects of neurodegenerative diseases.

Experiments 1 and 2 provide insights into the perception of verbal messages in our naturalistic non-ideal listening environments. My thesis has shown that there

is a retained mechanism to repair speech in certain dementias, and has illuminated some of the mechanisms involved in degraded speech perception – particularly phoneme perception and semantic disambiguation. My findings also suggest differentiable profiles of degraded speech perception in AD and PPA syndromes. Using noise-vocoding, my experiment highlighted additional deficits that can be seen in dementia, generally relevant to the comprehension of everyday speech signals conveyed by acoustically degraded carriers (such as a noisy telephone line).

The results in the svPPA patient group showcased in Experiment 2 (i.e., a need for a significantly higher threshold in comparison to healthy controls) are clarified with Experiment 4, which illustrates the role and need for top-down semantic disambiguation to fully understand degraded speech (Cai et al., 2017; Hardy, Marshall, et al., 2018). This applies both for comprehension of the verbal message, but also the associated and important paralinguistic cues. In comparison to other dementia groups, the svPPA group showed a higher cost for paralinguistic information degradation than verbal content. Such a deficit might contribute to the more general difficulty experienced by svPPA patients in socio-emotional cognition and communication in daily life (Marshall et al., 2019; Rankin et al., 2009). This interpretation is corroborated by the finding that measures of degraded emotional prosody processing were correlated with measures of social cognition across the whole patient cohort.

All four experiments presented in this thesis have provided data on certain 'phenotypes' of degraded speech processing in different diseases (see a suggested framework in **Section 7.2.1**). In lvPPA and AD, impaired extraction of both verbal and nonverbal information from degraded speech signals is likely to reflect the computational challenge posed by degrading auditory apperceptive

mechanisms that are primarily targeted by AD pathology (Golden et al., 2017; C. J. D. Hardy, J. L. Agustus, et al., 2017; Ruksenaite et al., 2021). In nfvPPA, in line with past research (Cope et al., 2017; Grube et al., 2016; Hardy, Marshall, et al., 2018), we can see difficulties with degraded speech perception that likely reflect impairments both in bottom-up perceptual difficulties and top-down mechanisms. The results of Experiments 3 and 4 may also speak to an impairment of rhythm perception in nfvPPA (Ruksenaite et al., 2021), as recognition of emotional prosody is probably more reliant on dynamic structure than recognition of accents. In svPPA, the findings highlight how acoustic factors may interact with or expose a primary semantic deficit: in this syndrome, the impact of spoken message degradation is modulated by the top-down predictability of the message.

Parts of the next few sections, which review the predictive coding model as well as therapeutic approaches, have been published in a review in *Brain Sciences* (https://doi.org/10.3390/brainsci11030394).

7.2.1 Predictive Coding Model in AD and PPA

As suggested in Chapter 1, predictive coding offers a potential framework to consider degraded speech processing in the healthy brain (**Figure 1.2**). It could therefore also be used for interpreting and anticipating deficits, across different kinds of speech information, as well as in different dementia diseases.

Taken together with previous evidence, the experiments presented in this thesis can be applied to the healthy brain predictive coding model, to formulate explicit pathophysiological hypotheses in AD and PPA. This framework could then serve as a model for interpreting abnormalities of degraded speech processing in a wider range of brain disorders. This model is outlined in **Figure 7.1**.

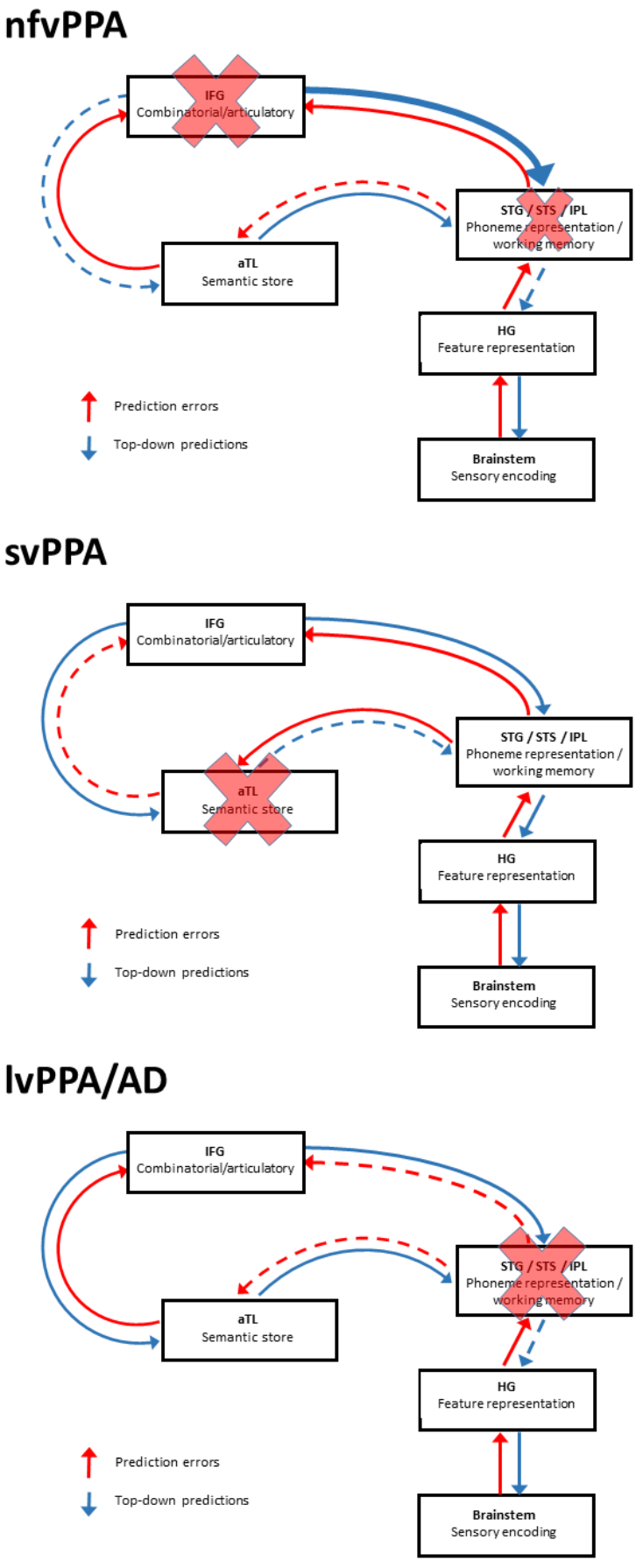


Figure 7.1. A simplified model of predictive coding of degraded speech processing in Alzheimer's disease and primary progressive aphasia.

Referenced to the healthy brain presented in **Figure 1.2** and presented again from **Figure 1.4.** Each syndrome is associated with a specific pattern of regional brain atrophy and/or dysfunction that is critical to the degraded speech processing network, implying that different dementias may be associated with specific profiles of degraded speech processing (see **Section 7.2.1** for details). AD and lvPPA groups are put together in this schematic as differences between these syndromes are likely to reflect disease stage and relative degree of involvement of left vs bi-parietal cortices (see **Section 7.2.1**). Boxes indicate processors that instantiate core speech decoding functions (see **Figure 1.2**), and arrows indicate their connections in the predictive coding framework, with the putative direction of information flow. In the case of nfvPPA, the emboldened descending arrow from IFG to STG signifies aberrantly increased precision of inflexible top-down priors (after (Cope et al., 2017)), to date the most secure evidence for a predictive coding mechanism in the PPA spectrum; the status of the IPL locus in this syndrome is more tentative. Implicitly in the model is the hypothesis that neurodegenerative pathologies will tend to disrupt stored neural templates and "prune" projections from heavily involved, higher-order association cortical areas due to neuronal dropout (promoting inflexible top-down predictions), but also degrade the fidelity of signal traffic through sensory cortices (reducing sensory precision and promoting over-precise prediction errors) (Kocagoncu et al., 2020). The relative prominence of these mechanisms will depend on the macro-network and local neural circuit anatomy of particular neurodegenerative pathologies. Proposed major loci of disruption caused by each disease are indicated with crosses; dashed arrows arising from these damaged modules indicate disrupted information flow. AD, Alzheimer's disease; aTL, anterior temporal lobe; HG, Heschl's gyrus; IFG, inferior frontal gyrus; IPL, inferior parietal lobule; lvPPA, logopenic variant primary progressive aphasia; nfvPPA, non-fluent primary progressive aphasia; STG, superior temporal gyrus; STS, superior temporal sulcus; svPPA, semantic variant primary progressive aphasia.

According to this model, different dementia syndromes produce distinct profiles of abnormal predictive coding of degraded speech signals. Firstly, nfvPPA is associated with a "double-hit" to the degraded speech processing network. The most clearly established consequence is overly precise, top-down predictions due to neuronal dysfunction and loss in inferior frontal cortex (Cope et al., 2017). The top-down mechanism may be compounded by decreased signal fidelity (precision) due to abnormal auditory cortical representations, as seen in Chapter 4 (Experiment 2) and other past studies (Chris J. D. Hardy et al., 2017; C. J. D. Hardy, J. L. Agustus, et al., 2017; Hardy, Marshall, et al., 2018). The clinic-anatomical heterogeneity of nfvPPA is an important consideration here

(see **Chapter 1.5.3**), implying that the mechanism may not be uniform between patients.

In svPPA, the primary focus of atrophy in the anterior temporal lobe principally affects the top-down integration of contextual and stored semantic information. This reduces the neural capacity to modify semantic predictions on less predictable verbal signals (i.e., priors are inaccurate), in line with experimental observations in Chapter 3 (Experiment 1), Chapter 6 (Experiment 4), and past research (Hardy et al., 2018). Particularly, the primary deficit in svPPA is semantic representation, regardless of modality.

Atrophy in both lvPPA and AD predominantly involves the temporo-parietal cortex. In lvPPA, phonemic decoding and earlier stages in the representation of acoustic features in the auditory cortex and brainstem are impaired due to altered top-down influences from the temporal parietal junction on the auditory cortex and brainstem: this could be via altered precision weighting of prediction errors conveyed by the auditory efferent pathways, or inaccurate priors. This formulation is supported in Experiment 2 (**Chapter 4**), Experiment 3 (**Chapter 5**), Experiment 4 (**Chapter 6**), and other past research (C. J. D. Hardy, J. L. Agustus, et al., 2017; Hardy, Marshall, et al., 2018; Johnson et al., 2020).

Most cases of lvPPA have AD pathology and lvPPA and typical AD lie on a continuum pathophysiologically and neuroanatomically (see **Chapter 1.5** and Ramanan et al. (2022)). Therefore, it is likely that lvPPA and AD affect the predictive decoding of degraded speech via similar mechanisms, however, at a comparable disease stage, lvPPA has more severe and more focal involvement of left temporo-parietal cortical mechanisms, whereas, in AD, temporo-parietal involvement is more symmetrical between the left and right hemispheres.

In AD, there is a more generic disorder of auditory scene analysis (H. L. Golden et al., 2015; Hannah L. Golden, Jennifer M. Nicholas, et al., 2015; Goll et al., 2012), resulting in difficulty disambiguating auditory sound objects as well as identifying and understanding sound objects under degraded listening conditions. This likely reflects impaired bottom-up perception and altered/overriding effects of top-down predictions, which is supported in all the experiments presented in this thesis and past research (Hailstone et al., 2012; Hardy, Marshall, et al., 2018).

The predictive coding framework presented here could direct future research within this field. However, the experiments in this thesis were not explicitly designed under a predictive coding model, but are generally compatible. Notably, this framework is missing crucial elements, such as the specific roles of nonlinguistic information encoded in the acoustic (paralinguistic) characteristics of speakers (as explored in **Chapter 6**), and the social importance of the role of speech (as explored in **Chapter 5**). Aspects such as hemispheric differences and the role of social cognition networks when engaging in conversation should be elaborated and explored. Finally, as with any other scientific paradigm, predictive coding demands a critical evaluation of falsifiable hypotheses (see Heilbron and Chait (2018) for a review of the issues concerning the auditory system). Future research will need to address how the processing of degraded speech is generically underpinned by predictive coding, but also how predictive coding of degraded speech is represented in the brain.

7.3 Limitations Of This Work

Due to Covid-19, there were significant limitations on data collection. Data collection for Experiment 1 (**Chapter 3**) was halted prematurely, and therefore the cohort numbers were small within the group. Experiment 2 (**Chapter 4**) and

Experiment 3 (**Chapter 5**) utilised remote testing. While our research suggests negligible differences with respect to in-person testing (Requena-Komuro et al., 2022), more research needs to be conducted to clarify differences that could result due to different testing modalities (e.g., inconsistent hardware use, inconsistent audio equipment use, inability to control remote testing environment, inability to check for peripheral hearing function). Nonetheless, it is encouraging that the future of research seems to be moving to incorporate more of a 'hybrid' approach.

Furthermore, especially in the context of PPA, the cohort numbers are naturally quite small due to the rarity of the disease. As more is being understood about the diseases, the variants within PPA could be stratified to accommodate for variability seen within patient cohorts. Larger group sizes should involve multi-centre collaborations, an aspect that is currently being increasingly encouraged internationally (Russell, 2022).

Even within 'typical' amnestic AD, there is considerable individual variation (Snowden et al., 2007). This could easily have affected my results, as signified by the variability seen among the AD patients participating in this thesis. Therefore, there is a need to establish pathophysiological markers of AD pathology that transcend the conventional markers previously used. With growing evidence of central hearing function (Stevenson et al., 2022), more research needs to be conducted to establish a viable biomarker to be used in clinics. A limitation of this thesis is demonstrating real-world resonance, with limited correlations to measures of daily life listening, communication, and social functioning (e.g., in Experiments 1 and 4, correlations with measures of daily life listening were not conducted; in Experiment 2, the results were borderline significant). Future studies need to address real-world applications, and

longitudinal studies of healthy ageing and dementia cohorts will be essential to assess the development and impact of central auditory dysfunction over time.

There is a growing need for robust measures of disease severity in the field. Standard measures that have been used in this thesis, such as the MMSE, are heavily weighted toward typical AD. The recently developed Mini Linguistic State Examination (Patel et al., 2021) may help address some of the issues, but overall there is still a lack of adequate non-linguistic assessments. Ultimately, what is needed is better staging markers across diseases for accurate classification of severity and disease progression.

Outside of **Chapter 4** (Experiment 2), the other experiments in this thesis did not include imaging analyses. Particularly in speech paradigms, not only would structural imaging help with anatomical correlates, but functional imaging would allow for delineating the networks engaged in degraded speech processing. Temporally sensitive neurophysiological and functional neuroimaging techniques such as EEG and MEG will be required to define the dynamic neural mechanisms by which brain pathologies disrupt degraded speech perception. Proteinopathies are anticipated to have separable MEG signatures based on differential patterns of cortical laminar involvements (Shaw et al., 2021).

7.4 Future Directions

As Experiments 2 and 3 (and other past research) showed correlations with factors such as working memory, future work should explore the relationship between working memory and other factors that can influence perception in day-to-day communication (e.g., musical influences, languages, see **Chapter 1.3.3**). Particularly for languages, degraded speech perception work should be extended to languages that are dissimilar to English (e.g., tonal languages).

Further research, using other degraded speech manipulations and paradigms, can also be conducted within the patient groups tested in this thesis here. For example, with some evidential support (Cope et al., 2017; Hardy et al., 2019), patients with nfvPPA have disordered efferent regulation of auditory signal analysis, thus this could be explored using dichotic listening techniques. As for lvPPA, being associated with the 'blurring' of phonemic representational boundaries (Johnson et al., 2020), phonemic restoration would also likely be critically impaired.

7.5 Clinical Translation of Thesis Results

From a clinical perspective, the work presented in this thesis could have a number of applications. Firstly, measures of degraded speech perception could potentially be rapid readouts on therapeutic effects on neural circuit function. This will be useful as disease-modifying therapies for dementia become feasible (van Dyck et al., 2023). There is an urgent need to harness dynamic and fundamental neurophysiological processes (e.g., phonemic restoration) that could be targeted for intervention and provide good markers for clinical trials.

The processing of degraded speech (as a sensitive index of neural circuit integrity) could also potentially facilitate the early diagnosis of neurodegenerative diseases. As dementia is often difficult to diagnose in its early stages, paradigms utilising central auditory perception (as represented here in this thesis) might be adapted to constitute dynamic, neurophysiological "stress tests" to detect pathologies (Stevenson et al., 2022).

The findings presented in this thesis also have implications for being used as 'real-world' models, indexing daily life hearing functions beyond sound detection, including socio-emotional communicative abilities, and semantic cues for

interpreting vocal messages. As mentioned in **Section 7.3**, a limitation of this thesis is demonstrating real-world resonance, and while Experiment 2 and Experiment 3 had begun to explore this avenue, more research needs to relate degraded speech measures with real-world hearing function, and compare the results with other established measures typically used in clinics (e.g., peripheral hearing assessments for hearing function).

In a clinical context, the results shown here suggest a difficulty with communication in AD and PPA patients, particularly in non-ideal listening conditions. Although, the degradations used were rather severe and artificial, listening under challenging conditions is arguably more reflective of everyday listening than clear speech in silence. It is important not to overlook nonverbal strategies to compensate for reduced capacity to process degraded verbal and nonverbal messages, such as minimisation of environmental noise, training speakers to face the patient to maximise visual support, aid sound discrimination, and using gestures to support semantic context (Conway & Chenery, 2016; Liddle et al., 2012; Pralus et al., 2021; Ritter & Vongpaisal, 2018).

Future research could look to compare and contrast other therapeutic approaches with the ones currently used and inform and build on behavioural pharmacological therapies, such as those that harness neural plasticity (see **Chapter 1.3.3.3** on perceptual learning). Past research has found that perceptual learning of degraded speech is retained in dementia (Hardy, Marshall, et al., 2018). This offers exciting prospects for designing training interventions to harness neural plasticity in these conditions. So far, most work in this line has been directed at improving the understanding of challenging speech (in particular, speech-in-noise) in older adults with peripheral hearing loss (Bieber & Gordon-Salant, 2021). On the other hand, there is some evidence that training on

degraded environmental sounds (Shafiro et al., 2012), auditory interaural level difference, and fundamental frequency discrimination (Gao et al., 2020) may generalise to improve the perception of degraded speech. Enhanced perceptual learning through the facilitation of regional neuronal plasticity also provides a rationale for the transcranial stimulation of key cortical language areas, such as the inferior frontal gyrus (Sehm et al., 2013). Potentially, a technique such as transcranial temporal interference stimulation could selectively target deep brain circuitry and feedforward or feedback connections (Rampersad et al., 2019) to probe specific pathophysiological mechanisms of degraded speech processing.

Other therapeutic approaches to improve degraded auditory perception have focused on using music. Melodic Intonation Therapy (MIT) is a prominent rehabilitation program originally developed for individuals with nonfluent aphasia (Albert et al., 1973), but a recent meta-analysis has suggested that there are limited positive effects in specific domains (e.g., repetition) (Popescu et al., 2022). However, focusing on certain aspects of music has been shown to improve degraded speech perception. Training of musical working memory has shown potential crossover benefits for speech-in-noise recognition (Escobar et al., 2020; Zhang et al., 2020). A study using harmonic training increased the temporal processing of pitch pattern sequence test and consonant-vowel in noise in hearing-impaired children (Moossavi et al., 2021), and another music training program (inclusive of group music therapy and take-home music applications) resulted in children with hearing loss having an increase in perception of speech-in-noise, linguistic prosody, music timbre, and spectral resolution, albeit no improvement in emotional prosody and pitch perception (Lo et al., 2020). Music-based interventions should have specified focuses (e.g., pitch, timbre),

targeted depending on the clinical population, and on what auditory cues are most relevant (Paquette et al., 2018).

Particularly in the case of perception of emotional prosodic cues, pitch-based music interventions could be particularly beneficial. Individuals with developmental amusia, who have shown difficulties with certain emotional prosodic cues (Pralus et al., 2019), have shown consistent patterns of top-down processing and accuracy on discrimination of high-frequency syllables, high-probability tones, and tone errors similar to those of control listeners (Zhu et al., 2022). This suggests that amusics can learn syllable and tone statistical regularities in a language context, with the potential for rehabilitation programs aiming at improving sensitivity to pitch and thus, paralinguistic cues. In translating for usage in patients with dementia, it is possible that by using preserved musical capacities in patients, this could be harnessed to improve prosody recognition.

A combined auditory cognitive training programme, potentially incorporating musical skills, may be the most rational strategy (Bieber & Gordon-Salant, 2021; Zendel, West, Belleville, & Peretz, 2019). However, future studies need to aim to evaluate the association between speech perception and music training more explicitly (McKay, 2021).

Finally, future studies could look to incorporate pharmacological approaches to potentially complement behavioural interventions or transcranial stimulation. In healthy individuals, dopamine has been shown to enhance the perception of spectrally shifted noise-vocoded speech (Cardin et al., 2020). In patients with AD, acetylcholinesterase inhibition ameliorates the understanding of sinewave speech (C. J. D. Hardy, Y. T. Hwang, et al., 2017). Indeed, degraded speech

processing might prove to be a rapid and sensitive biomarker of therapeutic efficacy in brain disorders.

It is crucial to develop future studies and interventions that enhance degraded speech processing (and other ecologically relevant aspects of communication), not only to maximise daily life functioning in patients but also with a future view to using such techniques adjunctively with disease-modifying therapies as these become available. Ultimately, irrespective of the brain pathology, it will be essential to determine how far improvements on degraded speech processing tasks translate to improving communication in daily life.

8 APPENDIX

Table 8.1. Participant Involvement by Chapter

ID	GROUP	3	4	5	6
1	CONTROL		X	X	
2	CONTROL				X
3	CONTROL	X			
4	CONTROL	X			
5	CONTROL	X	X	X	X
6	CONTROL	X	X	X	X
7	CONTROL				X
8	CONTROL	X	X		
9	CONTROL	X	X	X	
10	CONTROL				X
11	CONTROL				X
12	CONTROL	X	X	X	X
13	CONTROL				X
14	CONTROL	X	X	X	X
15	CONTROL		X	X	X
16	CONTROL	X	X	X	
17	CONTROL	X	X	X	
18	CONTROL				X
19	CONTROL	X	X	X	X
20	CONTROL	X	X	X	
21	CONTROL	X	X	X	
22	CONTROL	X	X	X	
23	CONTROL	X	X	X	
24	CONTROL				X
25	CONTROL				X
26	CONTROL	X	X	X	
27	CONTROL	X	X	X	
28	CONTROL	X	X	X	
29	CONTROL		X	X	
30	CONTROL	X	X	X	
31	CONTROL	X	X	X	
32	CONTROL	X	X	X	
33	CONTROL	X	X	X	
34	CONTROL		X	X	X
35	CONTROL		X	X	
36	CONTROL				X
37	CONTROL				X
38	AD		X		
39	AD				X
40	AD				X
41	AD				X

42	AD	X	X	X	
43	AD		X	X	
44	AD	X	X	X	
45	AD		X	X	
46	AD				X
47	AD		X	X	
48	AD				X
49	AD		X	X	
50	AD		X	X	
51	AD		X	X	
52	AD				X
53	AD		X	X	
54	AD	X	X	X	
55	AD				X
56	AD		X	X	
57	AD		X	X	
58	AD				X
59	AD				X
60	AD				X
61	AD		X	X	
62	AD		X	X	
63	AD	X	X	X	
64	AD		X	X	
65	AD		X	X	
66	AD	X	X	X	
67	LVPPA				X
68	LVPPA		X	X	
69	LVPPA		X	X	
70	LVPPA		X	X	
71	LVPPA		X	X	
72	LVPPA		X	X	
73	LVPPA			X	
74	LVPPA				X
75	LVPPA				X
76	LVPPA		X	X	
77	LVPPA				X
78	LVPPA		X	X	
79	LVPPA				X
80	LVPPA		X	X	
81	LVPPA				X
82	LVPPA		X	X	
83	NFVPPA		X	X	
84	NFVPPA		X		
85	NFVPPA		X		
86	NFVPPA			X	
87	NFVPPA		X	X	

88	NFVPPA		X	X	
89	NFVPPA			X	
90	NFVPPA		X	X	
91	NFVPPA		X	X	
92	NFVPPA				X
93	NFVPPA				X
94	NFVPPA		X	X	
95	NFVPPA		X	X	
96	NFVPPA				X
97	NFVPPA				X
98	NFVPPA				X
99	NFVPPA			X	
100	NFVPPA		X	X	
101	SVPPA		X	X	X
102	SVPPA				X
103	SVPPA		X	X	
104	SVPPA	X	X	X	
105	SVPPA		X	X	
106	SVPPA	X	X	X	
107	SVPPA				X
108	SVPPA				X
109	SVPPA				X
110	SVPPA		X	X	
111	SVPPA		X	X	
112	SVPPA				X
113	SVPPA				X
114	SVPPA		X	X	
115	SVPPA	X	X	X	
116	SVPPA	X	X	X	
117	SVPPA				X
118	SVPPA		X	X	X
119	SVPPA		X		

Participants are ordered by group. An ‘X’ indicates that the participant was recruited into the specific experiment, corresponding with the Chapter.

Table 8.2. Individual raw scores for healthy controls across phonemic restoration experimental conditions

	Control																					
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**
Isolated noise segments (/40)																						
A|A	20	20	16	20	20	20	20	19	17	20	20	19	20	20	20	18	20	18	19	19	20	20
R|A	1	1	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1	2	1	1	1
R|R	19	19	19	19	18	19	19	19	19	19	19	19	19	19	19	19	19	19	18	19	19	19
A|R	0	0	4	0	0	0	0	1	3	0	0	1	0	0	0	2	0	2	1	1	0	0
A'	0.99	0.99	0.93	0.99	0.98	0.99	0.99	0.97	0.95	0.99	0.99	0.97	0.99	0.99	0.99	0.96	0.99	0.96	0.96	0.97	0.99	0.99
Real words (/80)																						
A|A	40	40	40	37	40	34	39	39	35	39	40	39	40	40	39	33	40	36	40	40	36	38
R|A	21	21	36	19	22	18	22	25	24	18	20	20	21	24	21	25	23	16	24	20	18	15
R|R	19	19	4	21	18	22	18	15	16	22	20	20	19	16	19	15	17	24	16	20	22	25
A|R	0	0	0	3	0	6	1	1	5	1	0	1	0	0	1	7	0	4	0	0	4	2
A'	0.87	0.87	0.78	0.84	0.86	0.80	0.85	0.82	0.75	0.87	0.88	0.86	0.87	0.85	0.85	0.69	0.86	0.85	0.85	0.88	0.83	0.88
Pseudowords (/80)																						
A|A	28	36	25	14	27	26	29	39	27	33	32	26	25	25	36	25	3	31	26	15	30	32
R|A	9	17	20	4	5	7	19	22	13	4	11	14	9	11	20	19	4	9	13	3	10	12
R|R	31	23	20	36	35	33	21	18	28	36	29	26	31	29	20	21	36	31	27	37	30	28
A|R	12	4	15	26	13	14	11	1	13	7	8	14	15	15	4	15	37	9	14	25	10	8
A'	0.82	0.84	0.61	0.75	0.86	0.86	0.71	0.85	0.76	0.92	0.85	0.73	0.79	0.76	0.81	0.63	0.43	0.85	0.75	0.78	0.83	0.83

The table shows individual healthy control participant response totals in each of the main phonemic restoration experimental conditions. Note that 20 trials of each stimulus type (Added / Replaced) were presented for each isolated noise segment condition and 40 trials for each word condition; the maximum score in each cell is therefore 20 for segments and 40 for real words / pseudowords. Stimulus conditions were delivered in randomised order during the experimental session. A|A denotes that the participant correctly identified an 'Added' stimulus as 'Added'; R|A denotes that the participant incorrectly identified a 'Replaced' stimulus as 'Added' (i.e., phonemic restoration); R|R denotes that the participant correctly identified a 'Replaced' stimulus as 'Replaced'; A|R denotes that the participant incorrectly identified an 'Added' stimulus as 'Replaced'.

8.1 DIVISION OF LABOUR

The work described in this thesis was conducted by JJ with assistance from other researchers based at the Dementia Research Centre, UCL. Contributors are detailed below:

8.1.1 Chapter 3: Phonemic Restoration

Experimental design: JJ, CJDH, JDW

Construction of tests: JJ, CJDH

Data collection: JJ, CJDH, EB, HS, JCSJ, MCRK

Data analysis: JJ, CJDH

8.1.2 Chapter 4: Noise-Vocoded Verbal Messages

Experimental design: JJ, CJDH, JDW

Construction of tests: JJ, CJDH

Data collection: JJ, CJDH, EB, HS, JCSJ, MCRK

Data analysis: JJ

8.1.3 Chapter 5: Noise-Vocoded Emotional Prosody

Experimental design: JJ, CJDH, JDW

Construction of tests: JJ, CJDH

Data collection: JJ, CJDH, EB, HS, JCSJ, MCRK

Data analysis: JJ

8.1.4 Chapter 6: Sinewave Accent Identification

Experimental design: CJDH, JDW

Construction of tests: CJDH, SJR

Data collection: CJDH, RLB, CRM, LLR

Data analysis: JJ

8.2 PUBLICATIONS

8.2.1 Publications arising as a direct result of the work conducted in this thesis

Jiang, J., Benhamou, E., Waters, S., Johnson, J. C. S., Volkmer, A., Weil, R. S., Marshall, C. R., Warren, J. D., & Hardy, C. J. D. (2021). Processing of degraded speech in brain disorders. *Brain Sciences, 11*(3), 394.

Requena-Komuro, M. C.*, **Jiang, J.***, Dobson, L., Benhamou, E., Russell, L., Bond, Rebecca, Brotherhood, E.V., Greaves, C., Barker, S., Rohrer, J.D., Crutch, S., Warren, J.D., & Hardy, C.J.D. (2022). Remote versus face-to-face neuropsychological testing for dementia research: a comparative study in people with Alzheimer's disease, frontotemporal dementia and healthy older individuals. *BMJ Open, 12*. doi:10.1136/bmjopen-2022-064576

*__Joint first author__

Jiang, J., Johnson, J. C. S., Requena-Komuro, M-C., Benhamou, E., Sivasathiaseelan, H., Sheppard, D., Volkmer, A., Crutch, S. J., Warren, J. D., & Hardy, C. J. D. (2022). Phonemic restoration in neurodegenerative disease. *Brain Communications, 4*(3). https://doi.org/10.1093/braincomms/fcac118

Jiang, J., Johnson, J. C.S., Requena-Komuro, M-C., Benhamou, E., Sivasathiaseelan, H., Chokesuwattanaskul, A., Nelson, A., Nortley, R., Weil, R. S., Volkmer, A., Marshall, C. R., Bamiou, D-E., Warren, J. D., Hardy, C. J. D. (2022). Comprehension of acoustically degraded speech in Alzheimer's disease and primary progressive aphasia. *medRxiv*. https://doi.org/10.1101/2022.12.05.22283108

8.2.2 Other substantial contributions

Johnson, J. C. S., **Jiang, J.**, Bond, R. L., Benhamou, E., Requena-Komuro, M.-C., Russell, L. L., Greaves, C., Nelson, A., Sivasathiaseelan, H., Marshall, C. R., Volkmer, A. P., Rohrer, J. D., Warren, J. D., & Hardy, C. J. D. (2020). Impaired phonemic discrimination in logopenic variant primary progressive aphasia. *Annals of Clinical and Translational Neurology, 7*(7), 1252-1257. doi:10.1002/acn3.51101.

Ruksenaite, J., Volkmer, A., **Jiang, J.**, Johnson, J. C. S., Marshall, C. R., Warren, J. D., & Hardy, C. J. D. (2021). Primary Progressive Aphasia: Toward a pathophysiological synthesis. *Current Neurology and Neuroscience Reports, 21*(7). https://doi.org/10.1007/s11910-021-01097-z.

Bolton, L. M., **Jiang, J.**, & Warren, J. D. (2022). Music as a person centred intervention for dementia. *BMJ*, 376. https://doi.org/10.1136/bmj.o518

9 REFERENCES

Abutalebi, J., Della Rosa, P. A., Green, D. W., Hernandez, M., Scifo, P., Keim, R., Cappa, S. F., & Costa, A. (2011). Bilingualism Tunes the Anterior Cingulate Cortex for Conflict Monitoring. *Cerebral Cortex, 22*(9), 2076-2086. https://doi.org/10.1093/cercor/bhr287

Adank, P., Davis, M. H., & Hagoort, P. (2012). Neural dissociation in processing noise and accent in spoken language comprehension. *Neuropsychologia, 50*(1), 77-84. https://doi.org/10.1016/j.neuropsychologia.2011.10.024

Adank, P., & Devlin, J. T. (2010, Jan 1). On-line plasticity in spoken sentence comprehension: Adapting to time-compressed speech. *NeuroImage, 49*(1), 1124-1132. https://doi.org/10.1016/j.neuroimage.2009.07.032

Adank, P., & Janse, E. (2010). Comprehension of a Novel Accent by Young and Older Listeners. *Psychology and Aging, 25*, 736-740. https://doi.org/doi: 10.1037/a0020054.

Adank, P., Nuttall, H. E., Banks, B., & Kennedy-Higgins, D. (2015). Neural bases of accented speech perception. *Frontiers in Human Neuroscience, 9*. https://doi.org/10.3389/fnhum.2015.00558

Agosta, F., Pievani, M., Geroldi, C., Copetti, M., Frisoni, G. B., & Filippi, M. (2012, Aug). Resting state fMRI in Alzheimer's disease: beyond the default mode network. *Neurobiol Aging, 33*(8), 1564-1578. https://doi.org/10.1016/j.neurobiolaging.2011.06.007

Agustus, J. L., Mahoney, C. J., Downey, L. E., Omar, R., Cohen, M., White, M. J., Scott, S. K., Mancini, L., & Warren, J. D. (2015). Functional MRI of music emotion processing in frontotemporal dementia. *Annals of the New York Academy of Sciences, 1337*(1), 8. https://doi.org/10.1111/nyas.12620

Ainley, V., Apps, M. A. J., Fotopoulou, A., & Tsakiris, M. (2016). 'Bodily precision': a predictive coding account of individual differences in interoceptive accuracy. *Philosophical Transactions of the Royal Society B: Biological Sciences, 371*(1708), 20160003. https://doi.org/10.1098/rstb.2016.0003

Akeroyd, M. A. (2008, Nov). 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. *Int J Audiol, 47 Suppl 2*, S53-71. https://doi.org/10.1080/14992020802301142

Alain, C., Du, Y., Bernstein, L. J., Barten, T., & Banai, K. (2018). Listening under difficult conditions: An activation likelihood estimation meta-analysis. *Human Brain Mapping, 39*(7), 2695-2709. https://doi.org/10.1002/hbm.24031

Alain, C., Zendel, B. R., Hutka, S., & Bidelman, G. M. (2014, Feb). Turning down the noise: the benefit of musical training on the aging auditory brain. *Hear Res, 308*, 162-173. https://doi.org/10.1016/j.heares.2013.06.008

Alba-Ferrara, L., Hausmann, M., Mitchell, R. L., & Weis, S. (2011). The Neural Correlates of Emotional Prosody Comprehension: Disentangling Simple from Complex Emotion. *PLoS One, 6*(12), e28701. https://doi.org/10.1371/journal.pone.0028701

Albert, M. L., Sparks, R. W., & Helm, N. A. (1973). Melodic Intonation Therapy for Aphasia. *Archives of Neurology, 29*(2), 130-131. https://doi.org/10.1001/archneur.1973.00490260074018

Allen, M., Poggiali, D., Whitaker, K., Marshall, T. R., & Kievit, R. A. (2019). Raincloud plots: a multi-platform tool for robust data visualization. *Wellcome Open Research, 4*, 63. https://doi.org/10.12688/wellcomeopenres.15191.1

Amlerova, J., Laczó, J., Nedelska, Z., Laczó, M., Vyhnálek, M., Zhang, B., Sheardova, K., Angelucci, F., Andel, R., & Hort, J. (2022, 2022/04/05). Emotional prosody recognition is impaired in Alzheimer's disease. *Alzheimer's Research & Therapy, 14*(1), 50. https://doi.org/10.1186/s13195-022-00989-7

Anderson, S., White-Schwoch, T., Parbery-Clark, A., & Kraus, N. (2013, Jun). A dynamic auditory-cognitive system supports speech-in-noise perception in older adults. *Hear Res, 300*, 18-32. https://doi.org/10.1016/j.heares.2013.03.006

Andrews, G., Howie, P. M., Dozsa, M., & Guitar, B. E. (1982). Stuttering: Speech Pattern Characteristics Under Fluency-Inducing Conditions. *Journal of Speech, Language, and Hearing Research, 25*(2), 208-216. https://doi.org/https://doi.org/10.1044/jshr.2502.208

Arlinger, S., Lunner, T., Lyxell, B., & Pichora-Fuller, M. K. (2009, Oct). The emergence of cognitive hearing science. *Scand J Psychol, 50*(5), 371-384. https://doi.org/10.1111/j.1467-9450.2009.00753.x

Arnal, L. H., Doelling, K. B., & Poeppel, D. (2015, Sep). Delta-Beta Coupled Oscillations Underlie Temporal Prediction Accuracy. *Cereb Cortex, 25*(9), 3077-3085. https://doi.org/10.1093/cercor/bhu103

Arnal, L. H., Wyart, V., & Giraud, A.-L. (2011). Transitions in neural oscillations reflect prediction errors generated in audiovisual speech. *Nature Neuroscience, 14*(6), 797-801. https://doi.org/10.1038/nn.2810

Arroyo-Anlló, E. M., Dauphin, S., Fargeau, M. N., Ingrand, P., & Gil, R. (2019, 2019/08/07). Music and emotion in Alzheimer's disease. *Alzheimer's Research & Therapy, 11*(1), 69. https://doi.org/10.1186/s13195-019-0523-y

Audiology, B. S. o. (2018). *Recommended Procedure: Pure-tone air-conduction and bone-conduction threshold audiometry with and without masking*.

Awad, M., Warren, J. E., Scott, S. K., Turkheimer, F. E., & Wise, R. J. S. (2007). A Common System for the Comprehension and Production of Narrative Speech. *Journal of Neuroscience, 27*(43), 11455-11464. https://doi.org/10.1523/jneurosci.5257-06.2007

Bailenson, J. N. (2021). Nonverbal overload: A theoretical argument for the causes of Zoom fatigue. *Technology, Mind, and Behavior, 2*(1). https://doi.org/10.1037/tmb0000030

Bamiou, D.-E., Iliadou, V. V., Zanchetta, S., & Spyridakou, C. (2015). What Can We Learn about Auditory Processing from Adult Hearing Questionnaires? *Journal of the American Academy of Audiology, 26*(10), 824-837. https://doi.org/10.3766/jaaa.15009

Bamiou, D.-E., Werring, D., Cox, K., Stevens, J., Musiek, F. E., Brown, M. M., & Luxon, L. M. (2012). Patient-Reported Auditory Functions After Stroke of the Central Auditory Pathway. *Stroke, 43*(5), 1285-1289. https://doi.org/doi:10.1161/STROKEAHA.111.644039

Banai, K., & Lavner, Y. (2014, Oct). The effects of training length on the perceptual learning of time-compressed speech and its generalization. *J Acoust Soc Am, 136*(4), 1908-1917. https://doi.org/10.1121/1.4895684

Banks, B., Gowen, E., Munro, K. J., & Adank, P. (2015, Apr). Cognitive predictors of perceptual adaptation to accented speech. *J Acoust Soc Am, 137*(4), 2015-2024. https://doi.org/10.1121/1.4916265

Banse, R., & Scherer, K. R. (1996). Acoustic Profiles in Vocal Emotion Expression. *Journal of Personality and Social Psychology, 70*(3), 614-636.

Barker, J., & Cooke, M. (1999). Is the sine-wave speech cocktail party worth attending? *Speech Communication, 27*(3-4), 159-174. https://doi.org/10.1016/s0167-6393(98)00081-8

Başkent, D., Eiler, C. L., & Edwards, B. (2010, 2010/02/01/). Phonemic restoration by hearing-impaired listeners with mild to moderate sensorineural hearing loss. *Hearing Research, 260*(1), 54-62. https://doi.org/https://doi.org/10.1016/j.heares.2009.11.007

Başkent, D., & Gaudrain, E. (2016, Mar). Musician advantage for speech-on-speech perception. *J Acoust Soc Am, 139*(3), El51-56. https://doi.org/10.1121/1.4942628

Bateman, R. J., Aisen, P. S., De Strooper, B., Fox, N. C., Lemere, C. A., Ringman, J. M., Salloway, S., Sperling, R. A., Windisch, M., & Xiong, C. (2011, 2011/01/06). Autosomal-dominant Alzheimer's disease: a review and proposal for the prevention of Alzheimer's disease. *Alzheimer's Research & Therapy, 3*(1), 1. https://doi.org/10.1186/alzrt59

Becker, R., Pefkou, M., Michel, C. M., & Hervais-Adelman, A. G. (2013). Left temporal alpha-band activity reflects single word intelligibility. *Frontiers in Systems Neuroscience, 7*. https://doi.org/10.3389/fnsys.2013.00121

Bejanin, A., Schonhaut, D. R., La Joie, R., Kramer, J. H., Baker, S. L., Sosa, N., Ayakta, N., Cantwell, A., Janabi, M., Lauriola, M., O'Neil, J. P., Gorno-Tempini, M. L., Miller, Z. A., Rosen, H. J., Miller, B. L., Jagust, W. J., & Rabinovici, G. D. (2017). Tau pathology and neurodegeneration contribute to cognitive impairment in Alzheimer's disease. *Brain, 140*(12), 3286-3300. https://doi.org/10.1093/brain/awx243

Bench, J., Kowal, Å., & Bamford, J. (1979, 1979/01/01). The Bkb (Bamford-Kowal-Bench) Sentence Lists for Partially-Hearing Children. *British Journal of Audiology, 13*(3), 108-112. https://doi.org/10.3109/03005367909078884

Bent, T., & Bradlow, A. R. (2003, Sep). The interlanguage speech intelligibility benefit. *J Acoust Soc Am, 114*(3), 1600-1610. https://doi.org/10.1121/1.1603234

Bhaya-Grossman, I., & Chang, E. F. (2021). Speech Computations of the Human Superior Temporal Gyrus. *Annual Review of Psychology, 73*, 79-102. https://doi.org/https://doi.org/10.1146/annurev-psych-022321-035256

Bidelman, G. M., & Howell, M. (2016, Jan 1). Functional changes in inter- and intra-hemispheric cortical processing underlying degraded speech perception. *NeuroImage, 124*(Pt A), 581-590. https://doi.org/10.1016/j.neuroimage.2015.09.020

Bidelman, G. M., & Krishnan, A. (2010, Oct 8). Effects of reverberation on brainstem representation of speech in musicians and non-musicians. *Brain Res, 1355*, 112-125. https://doi.org/10.1016/j.brainres.2010.07.100

Bieber, R. E., & Gordon-Salant, S. (2021, Mar 15). Improving older adults' understanding of challenging speech: Auditory training, rapid adaptation and perceptual learning. *Hear Res, 402*, 108054. https://doi.org/10.1016/j.heares.2020.108054

Binney, R. J., Embleton, K. V., Jefferies, E., Parker, G. J., & Ralph, M. A. (2010, Nov). The ventral and inferolateral aspects of the anterior temporal lobe are crucial in semantic memory: evidence from a novel direct comparison of distortion-corrected fMRI, rTMS, and semantic dementia. *Cereb Cortex, 20*(11), 2728-2738. https://doi.org/10.1093/cercor/bhq019

Blanco-Elorrieta, E., Ding, N., Pylkkänen, L., & Poeppel, D. (2020, Oct). Understanding Requires Tracking: Noise and Knowledge Interact in Bilingual Comprehension. *J Cogn Neurosci, 32*(10), 1975-1983. https://doi.org/10.1162/jocn_a_01610

Bocchetta, M., Iglesias Espinosa, M. D. M., Lashley, T., Warren, J. D., & Rohrer, J. D. (2020, Mar 27). In vivo staging of frontotemporal lobar degeneration TDP-43 type C pathology. *Alzheimers Res Ther, 12*(1), 34. https://doi.org/10.1186/s13195-020-00600-x

Boersma, P., & Weenink, D. (2023). *Praat: doing phonetics by computer.* In (Version 6.3.09) http://www.praat.org/

Bologna, W. J., VadenJr., K. I., Ahlstrom, J. B., & Dubno, J. R. (2018). Age effects on perceptual organization of speech: Contributions of glimpsing, phonemic restoration, and speech segregation. *The Journal of the Acoustical Society of America, 144*(1), 267-281. https://doi.org/10.1121/1.5044397

Boothroyd, A., Hanin, L., & Hnath, T. (1985). A sentence test of speech perception: reliability, set equivalence, and short term learning. *CUNY Academic Works*.

Botha, H., Duffy, J. R., Strand, E. A., Machulda, M. M., Whitwell, J. L., & Josephs, K. A. (2014). Nonverbal oral apraxia in primary progressive aphasia and apraxia of speech. *Neurology, 82*(19), 1729-1735. https://doi.org/10.1212/wnl.0000000000000412

Bouma, A., & Gootjes, L. (2011, Jul). Effects of attention on dichotic listening in elderly and patients with dementia of the Alzheimer type. *Brain Cogn, 76*(2), 286-293. https://doi.org/10.1016/j.bandc.2011.02.008

Bozeat, S., Lambon Ralph, M. A., Patterson, K., Garrard, P., & Hodges, J. R. (2000). Non-verbal semantic impairment in semantic dementia. *Neuropsychologia, 38*(9), 8. https://doi.org/10.1016/S0028-3932(00)00034-8

Braak, H., Alafuzoff, I., Arzberger, T., Kretzschmar, H., & Del Tredici, K. (2006, 2006/10/01). Staging of Alzheimer disease-associated neurofibrillary pathology using paraffin sections and immunocytochemistry. *Acta Neuropathologica, 112*(4), 389-404. https://doi.org/10.1007/s00401-006-0127-z

Bradlow, A. R., & Bent, T. (2008, Feb). Perceptual adaptation to non-native speech. *Cognition, 106*(2), 707-729. https://doi.org/10.1016/j.cognition.2007.04.005

Bregman, A. S. (1990). *Auditory scene analysis: The perceptual organization of sound*. MIT press.

Breitenstein, C., Lancker, D. V., Daum, I., & Waters, C. H. (2001). Impaired Perception of Vocal Emotions in Parkinson's Disease: Influence of Speech Time Processing and Executive Functioning. *Brain and Cognition, 45*, 277-314. https://doi.org/doi:10.1006/brcg.2000.1246

Bsharat-Maalouf, D., & Karawani, H. (2022). Learning and bilingualism in challenging listening conditions: How challenging can it be? *Cognition, 222*. https://doi.org/https://doi.org/10.1016/j.cognition.2022.105018

Buchanan, T. W., Lutz, K., Mirzazade, S., Specht, K., Shah, N. J., Zilles, K., & Jancke, L. (2000). Recognition of emotional prosody and verbal components of spoken language: an fMRI study. *Cognitive Brain Research, 9*(3), 11. https://doi.org/10.1016/S0926-6410(99)00060-9

Burda, A. N., Hageman, C. F., Brousard, K. T., & Miller, A. L. (2004, Jun). Dementia and identification of words and sentences produced by native and nonnative English speakers. *Percept Mot Skills, 98*(3 Pt 2), 1359-1362. https://doi.org/10.2466/pms.98.3c.1359-1362

Burnett, T., Freedland, M. B., & Larson, C. R. (1998). Voice F0 responses to manipulations in pitch feedback. *The Journal of the Acoustical Society of America, 103*(6). https://doi.org/https://doi.org/10.1121/1.423073

Busche, M. A., & Hyman, B. T. (2020, 2020/10/01). Synergy between amyloid-β and tau in Alzheimer's disease. *Nature Neuroscience, 23*(10), 1183-1193. https://doi.org/10.1038/s41593-020-0687-6

Cabeza, R., Albert, M., Belleville, S., Craik, F. I. M., Duarte, A., Grady, C. L., Lindenberger, U., Nyberg, L., Park, D. C., Reuter-Lorenz, P. A., Rugg, M. D., Steffener, J., & Rajah, M. N. (2018). Maintenance, reserve and compensation: the cognitive neuroscience of healthy ageing. *Nature Reviews Neuroscience, 19*(11), 701-710. https://doi.org/10.1038/s41583-018-0068-2

Cai, Z. G., Gilbert, R. A., Davis, M. H., Gaskell, M. G., Farrar, L., Adler, S., & Rodd, J. M. (2017, Nov). Accent modulates access to word meaning: Evidence for a speaker-model account of spoken word recognition. *Cogn Psychol, 98*, 73-101. https://doi.org/10.1016/j.cogpsych.2017.08.003

Cannon, S. A., & Chatterjee, M. (2022, Mar/Apr). Age-Related Changes in Voice Emotion Recognition by Postlingually Deafened Listeners With Cochlear Implants. *Ear Hear, 43*(2), 323-334. https://doi.org/10.1097/aud.0000000000001095

Carter, J. V., Pan, J., Rai, S. N., & Galandiuk, S. (2016, 2016/06/01/). ROC-ing along: Evaluation and interpretation of receiver operating characteristic curves. *Surgery, 159*(6), 1638-1645. https://doi.org/https://doi.org/10.1016/j.surg.2015.12.029

Chang, E. F., Niziolek, C. A., Knight, R. T., Nagarajan, S. S., & Houde, J. F. (2013). Human cortical sensorimotor network underlying feedback control of vocal pitch. *PNAS, 110*(7), 2653-2658. https://doi.org/https://doi.org/10.1073/pnas.1216827110

Chatterjee, M., Kulkarni, A. M., Siddiqui, R. M., Christensen, J. A., Hozan, M., Sis, J. L., & Damm, S. A. (2019). Acoustic of Emotional Prosody Produced by Prelingually Deaf Children with Cochlear Implants *Frontiers in Psychology*. https://doi.org/https://doi.org/10.3389/fpsyg.2019.02190

Cherry, E. C. (1953). Some Experiments on the Recognition of Speech, with One and with Two Ears. *The Journal of the Acoustical Society of America, 25*(5), 975-979. https://doi.org/10.1121/1.1907229

Cherry, E. C., Sayers, B. M., & Marland, P. M. (1955). Experiments on the Complete Suppression of Stammering. *Nature, 176*, 874-875.

Choi, J. Y., & Perrachione, T. K. (2019, 2019/09/01/). Noninvasive neurostimulation of left temporal lobe disrupts rapid talker adaptation in speech processing. *Brain and Language, 196*, 104655. https://doi.org/https://doi.org/10.1016/j.bandl.2019.104655

Chon, H., Kraft, S. J., Zhang, J., Loucks, T., & Ambrose, N. G. (2013). Individual Variability in Delayed Auditory Feedback Effects on Speech Fluency and Rate in Normally Fluent Adults. *Journal of Speech, Language, and Hearing Research, 56*(2), 489-504. https://doi.org/https://doi.org/10.1044/1092-4388(2012/11-0303)

Christensen, J., Sis, J., Kulkarni, A. M., & Chatterjee, M. (2019). Effects of Age and Hearing Loss on the Recognition of Emotions in Speech. *Ear & Hearing, 40*(5), 1069-1083. https://doi.org/DOI: 10.1097/AUD.0000000000000694

Clark, C. N., Nicholas, J. M., Agustus, J. L., Hardy, C. J. D., Russell, L. L., Brotherhood, E. V., Dick, K. M., Marshall, C. R., Mummery, C. J., Rohrer, J. D., & Warren, J. D. (2017). Auditory conflict and congruence in frontotemporal dementia. *Neuropsychologia, 104*, 12. https://doi.org/10.1016/j.neuropsychologia.2017.08.009

Clarke, C. M., & Garrett, M. F. (2004). Rapid adaptation to foreign-accented English. *The Journal of the Acoustical Society of America, 116*(6), 3647-3658. https://doi.org/10.1121/1.1815131

Clarke, J., Gaudrain, E., Chatterjee, M., & Başkent, D. (2014, 2014/09/01/). T'ain't the way you say it, it's what you say – Perceptual continuity of voice and top–down restoration of speech. *Hearing Research, 315*, 80-87. https://doi.org/https://doi.org/10.1016/j.heares.2014.07.002

Clopper, C. G., & Pisoni, D. B. (2004). Effects of talker variability on perceptual learning of dialects. *Lang Speech, 47*(Pt 3), 207-239. https://doi.org/10.1177/00238309040470030101

Conway, E. R., & Chenery, H. J. (2016). Evaluating the MESSAGE Communication Strategies in Dementia training for use with community-based aged care staff working with people with dementia: a controlled pretest-post-test study. *Journal of Clinical Nursing, 25*(7-8), 1145-1155. https://doi.org/10.1111/jocn.13134

Cooke, M., King, S., Garnier, M., & Aubanel, V. (2014). The listening talker: A review of human and algorithmic context-induced modifications of speech. *Computer Speech & Language, 28*(2), 543-571. https://doi.org/https://doi.org/10.1016/j.csl.2013.08.003

Cope, T. E., Baguley, D. M., & Griffiths, T. D. (2015). The functional anatomy of central auditory processing. *Practical Neurology, 15*(4), 302-308. https://doi.org/10.1136/practneurol-2014-001073

Cope, T. E., Shtyrov, Y., MacGregor, L. J., Holland, R., Pulvermüller, F., Rowe, J. B., & Patterson, K. (2020, 2020/05/01/). Anterior temporal lobe is necessary for efficient lateralised processing of spoken word identity. *Cortex, 126*, 107-118. https://doi.org/https://doi.org/10.1016/j.cortex.2019.12.025

Cope, T. E., Sohoglu, E., Sedley, W., Patterson, K., Jones, P. S., Wiggins, J., Dawson, C., Grube, M., Carlyon, R. P., Griffiths, T. D., Davis, M. H., & Rowe, J. B. (2017, Dec 18). Evidence for causal top-down frontal contributions to predictive processes in speech perception. *Nat Commun, 8*(1), 2154. https://doi.org/10.1038/s41467-017-01958-7

Cordella, C., Dickerson, B. C., Quimby, M., Yunusova, Y., & Green, J. R. (2017). Slowed articulation rate is a sensitive diagnostic marker for identifying non-fluent primary progressive aphasia. *Aphasiology, 31*(2), 241-260. https://doi.org/10.1080/02687038.2016.1191054

Cutler, A., Mehler, J., Norris, D., & Segui, J. (1986, 1986/08/01/). The syllable's differing role in the segmentation of French and English. *Journal of Memory and Language, 25*(4), 385-400. https://doi.org/https://doi.org/10.1016/0749-596X(86)90033-1

Cutler, A., & Otake, T. (1994, 1994/12/01/). Mora or Phoneme? Further Evidence for Language-Specific Listening. *Journal of Memory and Language, 33*(6), 824-844. https://doi.org/https://doi.org/10.1006/jmla.1994.1039

D'Ausilio, A., Pulvermüller, F., Salmas, P., Bufalari, I., Begliomini, C., & Fadiga, L. (2009, 2009/03/10/). The Motor Somatotopy of Speech Perception. *Current Biology, 19*(5), 381-385. https://doi.org/https://doi.org/10.1016/j.cub.2009.01.017

Darki, F., & Klingberg, T. (2014). The Role of Fronto-Parietal and Fronto-Striatal Networks in the Development of Working Memory: A Longitudinal Study. *Cerebral Cortex, 25*(6), 1587-1595. https://doi.org/10.1093/cercor/bht352

Darwin, C. *Shannon*.

Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of Personality and Social Psychology, 44*, 113-126.

Davis, M. H., & Johnsrude, I. S. (2003). Hierarchical Processing in Spoken Language Comprehension. *The Journal of Neuroscience, 23*(8), 3423-3431. https://doi.org/10.1523/jneurosci.23-08-03423.2003

Davis, M. H., & Johnsrude, I. S. (2007, Jul). Hearing speech sounds: top-down influences on the interface between audition and speech perception. *Hear Res, 229*(1-2), 132-147. https://doi.org/10.1016/j.heares.2007.01.014

Davis, M. H., Johnsrude, I. S., Hervais-Adelman, A., Taylor, K., & McGettigan, C. (2005, May). Lexical information drives perceptual learning of distorted speech: evidence from the comprehension of noise-vocoded sentences. *J Exp Psychol Gen, 134*(2), 222-241. https://doi.org/10.1037/0096-3445.134.2.222

Davis, M. H., & Sohoglu, E. (2019). Three Functions of Prediction Error for Bayesian Inference in Speech Perception. In M. Gazzaniga, R. Mangun, & D. Poeppel (Eds.), *The Cognitive Neurosciences* (6th ed.). MIT Press. https://doi.org/10.31234/osf.io/qc4u6

de Pablo-Fernández, E., González-Herrero, B., Cerdán Santacruz, D., Rossor, M. N., Schott, J. M., Lashley, T., Holton, J. L., Fox, N. C., Revesz, T., Warren, J. D., Jaunmuktane, Z., Rohrer, J. D., & Warner, T. T. (2021). A Clinicopathologic Study of Movement Disorders in Frontotemporal Lobar Degeneration. *Movement Disorders, 36*(3), 632-641. https://doi.org/https://doi.org/10.1002/mds.28356

Del Tufo, S. N., & Myers, E. B. (2014, 2014-June-04). Phonemic restoration in developmental dyslexia [Original Research]. *Frontiers in Neuroscience, 8*. https://doi.org/10.3389/fnins.2014.00134

Desikan, R. S., Ségonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., Buckner, R. L., Dale, A. M., Maguire, R. P., Hyman, B. T., Albert, M. S., & Killiany, R. J. (2006, 2006/07/01/). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *NeuroImage, 31*(3), 968-980. https://doi.org/https://doi.org/10.1016/j.neuroimage.2006.01.021

Di Liberto, G. M., Lalor, E. C., & Millman, R. E. (2018). Causal cortical dynamics of a predictive enhancement of speech intelligibility. *NeuroImage, 166*, 247-258. https://doi.org/10.1016/j.neuroimage.2017.10.066

Ding, N., & Simon, J. Z. (2012, Jan). Neural coding of continuous speech in auditory cortex during monaural and dichotic listening. *J Neurophysiol, 107*(1), 78-89. https://doi.org/10.1152/jn.00297.2011

Doherty, K. M., Rohrer, J. D., Lees, A. J., Holton, J. L., & Warren, J. (2013, Jun). Primary progressive aphasia with parkinsonism. *Mov Disord, 28*(6), 741-746. https://doi.org/10.1002/mds.25341

Donhauser, P. W., & Baillet, S. (2020, Jan 22). Two Distinct Neural Timescales for Predictive Speech Processing. *Neuron, 105*(2), 385-393.e389. https://doi.org/10.1016/j.neuron.2019.10.019

Dryden, A., Allen, H. A., Henshaw, H., & Heinrich, A. (2017). The Association Between Cognitive Performance and Speech-in-Noise Perception for Adult Listeners: A Systematic Literature Review and Meta-Analysis. *Trends in Hearing, 21*, 2331216517744675. https://doi.org/10.1177/2331216517744675

Dubois, B., Feldman, H. H., Jacova, C., Hampel, H., Molinuevo, J. L., Blennow, K., Dekosky, S. T., Gauthier, S., Selkoe, D., Bateman, R., Cappa, S., Crutch, S., Engelborghs, S., Frisoni, G. B., Fox, N. C., Galasko, D., Habert, M.-O., Jicha, G. A., Nordberg, A., Pasquier, F., Rabinovici, G., Robert, P., Rowe, C., Salloway, S., Sarazin, M., Epelbaum, S., De Souza, L. C., Vellas, B., Visser, P. J., Schneider, L., Stern, Y., Scheltens, P., & Cummings, J. L. (2014). Advancing research diagnostic criteria for Alzheimer's disease: the IWG-2 criteria. *The Lancet Neurology, 13*(6), 614-629. https://doi.org/10.1016/s1474-4422(14)70090-0

Dupoux, E., & Green, K. (1997, Jun). Perceptual adjustment to highly compressed speech: effects of talker and rate changes. *J Exp Psychol Hum Percept Perform, 23*(3), 914-927. https://doi.org/10.1037//0096-1523.23.3.914

Eisner, F., McGettigan, C., Faulkner, A., Rosen, S., & Scott, S. K. (2010, May 26). Inferior frontal gyrus activation predicts individual differences in perceptual learning of cochlear-implant simulations. *J Neurosci, 30*(21), 7179-7186. https://doi.org/10.1523/jneurosci.4040-09.2010

Eisner, F., & McQueen, J. M. (2005, Feb). The specificity of perceptual learning in speech processing. *Percept Psychophys, 67*(2), 224-238. https://doi.org/10.3758/bf03206487

Eisner, F., & McQueen, J. M. (2006). Perceptual learning in speech: Stability over time. *The Journal of the Acoustical Society of America, 119*(4), 1950-1953. https://doi.org/10.1121/1.2178721

Erb, J., Henry, M. J., Eisner, F., & Obleser, J. (2012). Auditory skills and brain morphology predict individual differences in adaptation to degraded speech. *Neuropsychologia, 50*(9), 2154-2164. https://doi.org/10.1016/j.neuropsychologia.2012.05.013

Erb, J., Henry, M. J., Eisner, F., & Obleser, J. (2013). The Brain Dynamics of Rapid Perceptual Adaptation to Adverse Listening Conditions. *Journal of Neuroscience, 33*(26), 10688-10697. https://doi.org/10.1523/jneurosci.4596-12.2013

Escobar, J., Mussoi, B. S., & Silberer, A. B. (2020, Mar/Apr). The Effect of Musical Training and Working Memory in Adverse Listening Situations. *Ear Hear, 41*(2), 278-288. https://doi.org/10.1097/aud.0000000000000754

Ethofer, T., Anders, S., Erb, M., Herbert, C., Wiethoff, S., Kissler, J., Grodd, W., & Wildgruber, D. (2006, 2006/04/01/). Cerebral pathways in processing of affective prosody: A dynamic causal modeling study. *NeuroImage, 30*(2), 580-587. https://doi.org/https://doi.org/10.1016/j.neuroimage.2005.09.059

Everhardt, M. K., Sarampalis, A., Coler, M., Baskent, D., & Lowie, W. (2020). Meta-Analysis on the Identification of Linguistic and Emotional Prosody in Cochlear Implant Users and Vocoder Simulations. *Ear and Hearing, 41*(5), 1092-1102. https://doi.org/10.1097/aud.0000000000000863

Fausto, B. A., Badana, A. N. S., Arnold, M. L., Lister, J. J., & Edwards, J. D. (2018, Apr 17). Comparison of Subjective and Objective Measures of Hearing, Auditory Processing, and Cognition Among Older Adults With and Without Mild Cognitive Impairment. *J Speech Lang Hear Res, 61*(4), 945-956. https://doi.org/10.1044/2017_jslhr-h-17-0263

Fellowes, J. M., Remez, R. E., & Rubin, P. E. (1997, 1997/01/01). Perceiving the sex and identity of a talker without natural vocal timbre. *Perception & Psychophysics, 59*(6), 839-849. https://doi.org/10.3758/BF03205502

Fernández, M., Gobartt, A. L., Balañá, M., & the, C. S. G. (2010, 2010/09/28). Behavioural symptoms in patients with Alzheimer's disease and their association with cognitive impairment. *BMC Neurology, 10*(1), 87. https://doi.org/10.1186/1471-2377-10-87

Finger, H., Goeke, C., Diekamp, D., Standvoß, K., & König, P. (2017). LabVanced: a unified JavaScript framework for online studies. International Conference on Computational Social Science (Cologne),

Fletcher, P. D., Downey, L. E., Agustus, J. L., Hailstone, J. C., Tyndall, M. H., Cifelli, A., Schott, J. M., Warrington, E. K., & Warren, J. D. (2013, Aug). Agnosia for accents in primary progressive aphasia. *Neuropsychologia, 51*(9), 1709-1715. https://doi.org/10.1016/j.neuropsychologia.2013.05.013

Floccia, C., Butler, J., Goslin, J., & Ellis, L. (2009). Regional and Foreign Accent Processing in English: Can Listeners Adapt? *Journal of Psycholinguistic Research, 38*(4), 379-412. https://doi.org/10.1007/s10936-008-9097-8

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975, Nov). "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res, 12*(3), 189-198. https://doi.org/10.1016/0022-3956(75)90026-6

Foulke, E., Amster, C. H., Nolan, C. Y., & Bixler, R. H. (1962). The Comprehension of Rapid Speech by the Blind. *Exceptional Children, 29*(3), 134-142.

Foulke, E., & Sticht, T. G. (1969, Jul). Review of research on the intelligibility and comprehension of accelerated speech. *Psychol Bull, 72*(1), 50-62. https://doi.org/10.1037/h0027575

Foundas, A. L., Mock, J. R., Corey, D. M., Golob, E. J., & Conture, E. G. (2013). The SpeechEasy device in stuttering and nonstuttering adults: fluency effects while speaking and reading. *Brain and Language 126*(2), 141-150. https://doi.org/https://doi.org/10.1016/j.bandl.2013.04.004

Francis, A. L., Ciocca, V., Ma, L., & Fenn, K. (2008). Perceptual learning of Cantonese lexical tones by tone and non-tone language speakers. *Journal of Phonetics, 36*(2), 268-294. https://doi.org/10.1016/j.wocn.2007.06.005

Frisina, D. R., & Frisina, R. D. (1997). Speech recognition in noise and presbycusis: relations to possible neural mechanisms. *Hearing Research, 106*(1-2), 95-104. https://doi.org/10.1016/s0378-5955(97)00006-3

Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society B: Biological Sciences, 360*(1456), 815-836. https://doi.org/10.1098/rstb.2005.1622

Frühholz, S., Ceravolo, L., & Grandjean, D. (2011). Specific Brain Networks during Explicit and Implicit Decoding of Emotional Prosody. *Cerebral Cortex, 22*(5), 1107-1117. https://doi.org/10.1093/cercor/bhr184

Fuller, C. D., Galvin, J. J., 3rd, Maat, B., Free, R. H., & Başkent, D. (2014). The musician effect: does it persist under degraded pitch conditions of cochlear implant simulations? *Front Neurosci, 8*, 179. https://doi.org/10.3389/fnins.2014.00179

Füllgrabe, C., Moore, B. C. J., & Stone, M. A. (2015, 2015-January-13). Age-group differences in speech identification despite matched audiometrically normal hearing: contributions from auditory temporal processing and cognition [Original Research]. *Frontiers in Aging Neuroscience, 6*. https://doi.org/10.3389/fnagi.2014.00347

Füllgrabe, C., & Rosen, S. (2016, 2016-August-30). On The (Un)importance of Working Memory in Speech-in-Noise Processing for Listeners with Normal Hearing Thresholds [Perspective]. *Frontiers in Psychology, 7*. https://doi.org/10.3389/fpsyg.2016.01268

Gaetano, J., Lancaster, S., & Tindle, R. (2017). Signal detection theory calculator 1.2. *Retrieved from Excel workbook. See https://www. researchgate. net/profile/Justin_Gaetano2*.

Galantucci, B., Fowler, C. A., & Turvey, M. T. (2006, 2006/06/01). The motor theory of speech perception reviewed. *Psychonomic Bulletin & Review, 13*(3), 361-377. https://doi.org/10.3758/BF03193857

Gao, X., Yan, T., Huang, T., Li, X., & Zhang, Y. X. (2020, Nov 9). Speech in noise perception improved by training fine auditory discrimination: far and applicable transfer of perceptual learning. *Sci Rep, 10*(1), 19320. https://doi.org/10.1038/s41598-020-76295-9

Garrard, P., & Carroll, E. (2006, May). Lost in semantic space: a multi-modal, non-verbal assessment of feature knowledge in semantic dementia. *Brain, 129*(Pt 5), 1152-1163. https://doi.org/10.1093/brain/awl069

Gates, G. A., Anderson, M. L., Feeney, M. P., McCurry, S. M., & Larson, E. B. (2008). Central Auditory Dysfunction in Older Persons With Memory Impairment or Alzheimer Dementia. *Archives of Otolaryngology–Head & Neck Surgery, 134*(7), 771. https://doi.org/10.1001/archotol.134.7.771

Gates, G. A., Anderson, M. L., McCurry, S. M., Feeney, M. P., & Larson, E. B. (2011, Apr). Central auditory dysfunction as a harbinger of Alzheimer dementia. *Arch Otolaryngol Head Neck Surg, 137*(4), 390-395. https://doi.org/10.1001/archoto.2011.28

Gates, G. A., Gibbons, L. E., McCurry, S. M., Crane, P. K., Feeney, M. P., & Larson, E. B. (2010, Dec). Executive dysfunction and presbycusis in older persons with and without memory loss and dementia. *Cogn Behav Neurol, 23*(4), 218-223. https://doi.org/10.1097/WNN.0b013e3181d748d7

Gates, G. A., & Mills, J. H. (2005, Sep 24-30). Presbycusis. *Lancet, 366*(9491), 1111-1120. https://doi.org/10.1016/s0140-6736(05)67423-5

Gennari, S. P., Millman, R. E., Hymers, M., & Mattys, S. L. (2018, 2018/09/01/). Anterior paracingulate and cingulate cortex mediates the effects of cognitive load on speech sound discrimination. *NeuroImage, 178*, 735-743. https://doi.org/https://doi.org/10.1016/j.neuroimage.2018.06.035

George, M. S., Parekh, P. I., & Rosinsky, N. (1996). Understanding Emotional Prosody Activates Right Hemisphere Regions. *JAMA Neurology, 53*(7), 5. https://doi.org/10.1001/archneur.1996.00550070103017

Giannini, L. A. A., Irwin, D. J., McMillan, C. T., Ash, S., Rascovsky, K., Wolk, D. A., Van Deerlin, V. M., Lee, E. B., Trojanowski, J. Q., & Grossman, M. (2017). Clinical marker for Alzheimer disease pathology in logopenic primary progressive aphasia. *Neurology, 88*(24), 2276-2284. https://doi.org/10.1212/wnl.0000000000004034

Golden, H. L., Agustus, J. L., Goll, J. C., Downey, L. E., Mummery, C. J., Schott, J. M., Crutch, S. J., & Warren, J. D. (2015). Functional neuroanatomy of auditory scene analysis in Alzheimer's disease. *Neuroimage Clin, 7*, 699-708. https://doi.org/10.1016/j.nicl.2015.02.019

Golden, H. L., Clark, C. N., Nicholas, J. M., Cohen, M. H., Slattery, C. F., Paterson, R. W., Foulkes, A. J. M., Schott, J. M., Mummery, C. J., Crutch, S. J., & Warren, J. D. (2017). Music Perception in Dementia. *J Alzheimers Dis, 55*(3), 933-949. https://doi.org/10.3233/jad-160359

Golden, H. L., Downey, L. E., Fletcher, P. D., Mahoney, C. J., Schott, J. M., Mummery, C. J., Crutch, S. J., & Warren, J. D. (2015). Identification of environmental sounds and melodies in syndromes of anterior temporal lobe degeneration. *Journal of the Neurological Sciences, 352*, 4. https://doi.org/10.1016/j.jns.2015.03.007

Golden, H. L., Nicholas, J. M., Yong, K. X. X., Downey, L. E., Schott, J. M., Mummery, C. J., Crutch, S. J., & Warren, J. D. (2015). Auditory spatial processing in Alzheimer's disease. *Brain, 138*(1), 189-202. https://doi.org/10.1093/brain/awu337

Goldstone, R. L. (1998). PERCEPTUAL LEARNING. *Annual Review of Psychology, 49*(1), 585-612. https://doi.org/10.1146/annurev.psych.49.1.585

Golestani, N., Hervais-Adelman, A., Obleser, J., & Scott, S. K. (2013, 2013/10/01/). Semantic versus perceptual interactions in neural processing of speech-in-noise. *NeuroImage, 79*, 52-61. https://doi.org/https://doi.org/10.1016/j.neuroimage.2013.04.049

Golestani, N., Rosen, S., & Scott, S. K. (2009). Native-language benefit for understanding speech-in-noise: The contribution of semantics. *Bilingualism: Language and Cognition, 12*(3), 385-392. https://doi.org/10.1017/S1366728909990150

Goll, J. C., Crutch, S. J., Loo, J. H., Rohrer, J. D., Frost, C., Bamiou, D. E., & Warren, J. D. (2010, Jan). Non-verbal sound processing in the primary progressive aphasias. *Brain, 133*(Pt 1), 272-285. https://doi.org/10.1093/brain/awp235

Goll, J. C., Crutch, S. J., Loo, J. H. Y., Rohrer, J. D., Frost, C., Bamious, D.-E., & Warren, J. D. (2010). Non-verbal sound processing in the primary progressive aphasias. *Brain, 133*(1), 13. https://doi.org/10.1093/brain/awp235

Goll, J. C., Crutch, S. J., & Warren, J. D. (2010). Central auditory disorders: toward a neuropsychology of auditory objects. *Current Opinion in Neurology, 23*(6), 10. https://doi.org/10.1097/WCO.0b013e32834027f6

Goll, J. C., Kim, L. G., Hailstone, J. C., Lehmann, M., Buckley, A., Crutch, S. J., & Warren, J. D. (2011, Jul). Auditory object cognition in dementia. *Neuropsychologia, 49*(9), 2755-2765. https://doi.org/10.1016/j.neuropsychologia.2011.06.004

Goll, J. C., Kim, L. G., Ridgway, G. R., Hailstone, J. C., Lehmann, M., Buckley, A. H., Crutch, S. J., & Warren, J. D. (2012, Jan). Impairments of auditory scene analysis in Alzheimer's disease. *Brain, 135*(Pt 1), 190-200. https://doi.org/10.1093/brain/awr260

Gordon-Salant, S., & Fitzgibbons, P. J. (1993, Dec). Temporal factors and speech recognition performance in young and elderly listeners. *J Speech Hear Res, 36*(6), 1276-1285. https://doi.org/10.1044/jshr.3606.1276

Gordon-Salant, S., Yeni-Komshian, G. H., Fitzgibbons, P. J., & Schurman, J. (2010, Oct). Short-term adaptation to accented English by younger and older adults. *J Acoust Soc Am, 128*(4), El200-204. https://doi.org/10.1121/1.3486199

Gorno-Tempini, M. L., Hillis, A. E., Weintraub, S., Kertesz, A., Mendez, M., Cappa, S. F., Ogar, J. M., Rohrer, J. D., Black, S., Boeve, B. F., Manes, F., Dronkers, N. F., Vandenberghe, R., Rascovsky, K., Patterson, K., Miller, B. L., Knopman, D. S., Hodges, J. R., Mesulam, M. M., & Grossman, M. (2011). Classification of primary progressive aphasia and its variants. *Neurology, 76*(11), 1006-1014. https://doi.org/10.1212/wnl.0b013e31821103e6

Graff-Radford, J., Duffy, J. R., Strand, E. A., & Josephs, K. A. (2012, Aug). Parkinsonian motor features distinguish the agrammatic from logopenic variant of primary progressive aphasia. *Parkinsonism Relat Disord, 18*(7), 890-892. https://doi.org/10.1016/j.parkreldis.2012.04.011

Griffiths, T. D., Lad, M., Kumar, S., Holmes, E., McMurray, B., Maguire, E. A., Billig, A. J., & Sedley, W. (2020, 2020/11/11/). How Can Hearing Loss Cause Dementia? *Neuron, 108*(3), 401-412. https://doi.org/https://doi.org/10.1016/j.neuron.2020.08.003

Griffiths, T. D., & Warren, J. D. (2002, 2002/07/01/). The planum temporale as a computational hub. *Trends in Neurosciences, 25*(7), 348-353. https://doi.org/https://doi.org/10.1016/S0166-2236(02)02191-4

Griffiths, T. D., & Warren, J. D. (2004). What is an auditory object? *Nature Reviews Neuroscience, 5*, 5.

Grube, M., Bruffaerts, R., Schaeverbeke, J., Neyens, V., De Weer, A. S., Seghers, A., Bergmans, B., Dries, E., Griffiths, T. D., & Vandenberghe, R. (2016, Jun). Core auditory processing deficits in primary progressive aphasia. *Brain, 139*(Pt 6), 1817-1829. https://doi.org/10.1093/brain/aww067

Guediche, S., Blumstein, S. E., Fiez, J. A., & Holt, L. L. (2014). Speech perception under adverse conditions: insights from behavioral, computational, and neuroscience research. *Frontiers System Neuroscience*. https://doi.org/https://doi.org/10.3389/fnsys.2013.00126

Guediche, S., Reilly, M., Santiago, C., Laurent, P., & Blumstein, S. E. (2016). An fMRI study investigating effects of conceptually related sentences on the perception of degraded speech. *Cortex, 79*, 57-74. https://doi.org/10.1016/j.cortex.2016.03.014

Gunawardena, D., Ash, S., McMillan, C., Avants, B., Gee, J., & Grossman, M. (2010, Aug 17). Why are patients with progressive nonfluent aphasia nonfluent? *Neurology, 75*(7), 588-594. https://doi.org/10.1212/WNL.0b013e3181ed9c7d

Hagoort, P. (2005). On Broca, brain, and binding: a new framework. *Trends in Cognitive Sciences, 9*(9), 416-423. https://doi.org/10.1016/j.tics.2005.07.004

Hailstone, J. C., Omar, R., & Warren, J. D. (2009, Jul). Relatively preserved knowledge of music in semantic dementia. *J Neurol Neurosurg Psychiatry, 80*(7), 808-809. https://doi.org/10.1136/jnnp.2008.153130

Hailstone, J. C., Ridgway, G. R., Bartlett, J. W., Goll, J. C., Buckley, A. H., Crutch, S. J., & Warren, J. D. (2011). Voice processing in dementia: a neuropsychological and neuroanatomical analysis. *Brain, 134*(9), 2535-2547. https://doi.org/10.1093/brain/awr205

Hailstone, J. C., Ridgway, G. R., Bartlett, J. W., Goll, J. C., Crutch, S. J., & Warren, J. D. (2012, Jul). Accent processing in dementia. *Neuropsychologia, 50*(9), 2233-2244. https://doi.org/10.1016/j.neuropsychologia.2012.05.027

Hajian-Tilaki, K. O., Hanley, J. A., Joseph, L., & Collet, J.-P. (1997). A Comparison of Parametric and Nonparametric Approaches to ROC Analysis of Quantitative Diagnostic Tests. *Medical Decision Making, 17*(1), 94-102. https://doi.org/10.1177/0272989x9701700111

Hardy, C. J., Marshall, C. R., Golden, H. L., Clark, C. N., Mummery, C. J., Griffiths, T. D., Bamiou, D. E., & Warren, J. D. (2016, Nov). Hearing and dementia. *J Neurol, 263*(11), 2339-2354. https://doi.org/10.1007/s00415-016-8208-y

Hardy, C. J. D., Agustus, J. L., Marshall, C. R., Clark, C. N., Russell, L. L., Bond, R. L., Brotherhood, E. V., Thomas, D. L., Crutch, S. J., Rohrer, J. D., & Warren, J. D. (2017). Behavioural and neuroanatomical correlates of auditory speech analysis in primary progressive aphasias. *Alzheimer's Research & Therapy, 9*(1). https://doi.org/10.1186/s13195-017-0278-2

Hardy, C. J. D., Agustus, J. L., Marshall, C. R., Clark, C. N., Russell, L. L., Brotherhood, E. V., Bond, R. L., Fiford, C. M., Ondobaka, S., Thomas, D. L., Crutch, S. J., Rohrer, J. D., & Warren, J. D. (2017, Aug). Functional neuroanatomy of speech signal decoding in primary progressive aphasias. *Neurobiol Aging, 56*, 190-201. https://doi.org/10.1016/j.neurobiolaging.2017.04.026

Hardy, C. J. D., Bond, R. L., Jaisin, K., Marshall, C. R., Russell, L. L., Dick, K., Crutch, S. J., Rohrer, J. D., & Warren, J. D. (2018). Sensitivity of Speech Output to Delayed Auditory Feedback in Primary Progressive Aphasias. *Front Neurol, 9*, 894. https://doi.org/10.3389/fneur.2018.00894

Hardy, C. J. D., Frost, C., Sivasathiaseelan, H., Johnson, J. C. S., Agustus, J. L., Bond, R. L., Benhamou, E., Russell, L. L., Marshall, C. R., Rohrer, J. D., Bamiou, D. E., & Warren, J. D. (2019, May 1). Findings of Impaired Hearing in Patients With Nonfluent/Agrammatic Variant Primary Progressive Aphasia. *JAMA Neurol, 76*(5), 607-611. https://doi.org/10.1001/jamaneurol.2018.4799

Hardy, C. J. D., Hwang, Y. T., Bond, R. L., Marshall, C. R., Ridha, B. H., Crutch, S. J., Rossor, M. N., & Warren, J. D. (2017, Nov). Donepezil enhances understanding of degraded speech in Alzheimer's disease. *Ann Clin Transl Neurol, 4*(11), 835-840. https://doi.org/10.1002/acn3.471

Hardy, C. J. D., Marshall, C. R., Bond, R. L., Russell, L. L., Dick, K., Ariti, C., Thomas, D. L., Ross, S. J., Agustus, J. L., Crutch, S. J., Rohrer, J. D., Bamiou, D. E., & Warren, J. D. (2018, Jul 25). Retained capacity for perceptual learning of degraded speech in primary progressive aphasia and Alzheimer's disease. *Alzheimers Res Ther, 10*(1), 70. https://doi.org/10.1186/s13195-018-0399-2

Hardy, C. J. D., Yong, K. X. X., Goll, J. C., Crutch, S. J., & Warren, J. D. (2020, Sep 1). Impairments of auditory scene analysis in posterior cortical atrophy. *Brain, 143*(9), 2689-2695. https://doi.org/10.1093/brain/awaa221

Hartwigsen, G., Golombek, T., & Obleser, J. (2015, Jul). Repetitive transcranial magnetic stimulation over left angular gyrus modulates the predictability gain in degraded speech comprehension. *Cortex, 68*, 100-110. https://doi.org/10.1016/j.cortex.2014.08.027

Hashimoto, Y., & Sakai, K. L. (2003). Brain activations during conscious self-monitoring of speech production with delayed auditory feedback: An fMRI study. *Human Brain Mapping, 20*(1), 22-28. https://doi.org/10.1002/hbm.10119

Heilbron, M., & Chait, M. (2018, Oct 1). Great Expectations: Is there Evidence for Predictive Coding in Auditory Cortex? *Neuroscience, 389*, 54-73. https://doi.org/10.1016/j.neuroscience.2017.07.061

Hennessy, S., Mack, W. J., & Habibi, A. (2022, Mar 15). Speech-in-noise perception in musicians and non-musicians: A multi-level meta-analysis. *Hear Res, 416*, 108442. https://doi.org/10.1016/j.heares.2022.108442

Henry, M. J., Herrmann, B., Kunke, D., & Obleser, J. (2017). Aging affects the balance of neural entrainment and top-down neural modulation in the listening brain. *Nature Communications, 8*(1). https://doi.org/10.1038/ncomms15801

Hervais-Adelman, A., Davis, M. H., Johnsrude, I. S., & Carlyon, R. P. (2008). Perceptual learning of noise vocoded words: Effects of feedback and lexicality. *Journal of Experimental Psychology: Human Perception and Performance, 34*(2), 460-474. https://doi.org/10.1037/0096-1523.34.2.460

Hervais-Adelman, A., Pefkou, M., & Golestani, N. (2014, May). Bilingual speech-in-noise: neural bases of semantic context use in the native language. *Brain Lang, 132*, 1-6. https://doi.org/10.1016/j.bandl.2014.01.009

Hervais-Adelman, A. G., Carlyon, R. P., Johnsrude, I. S., & Davis, M. H. (2012). Brain regions recruited for the effortful comprehension of noise-vocoded words. *Language and Cognitive Processes, 27*(7-8), 1145-1166. https://doi.org/10.1080/01690965.2012.662280

Hickok, G. (2012, Nov-Dec). The cortical organization of speech processing: feedback control and predictive coding the context of a dual-stream model. *J Commun Disord, 45*(6), 393-402. https://doi.org/10.1016/j.jcomdis.2012.06.004

Hickok, G., Costanzo, M., Capasso, R., & Miceli, G. (2011, 2011/12/01/). The role of Broca's area in speech perception: Evidence from aphasia revisited. *Brain and Language, 119*(3), 214-220. https://doi.org/https://doi.org/10.1016/j.bandl.2011.08.001

Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews Neuroscience, 8*(5), 393-402. https://doi.org/10.1038/nrn2113

Hirano, S., Kojima, H., Naito, Y., Honjo, I., Kamoto, Y., Okazawa, H., Ishizu, K., Yonekura, Y., Nagahama, Y., Fukuyama, H., & Konishi, J. (1997). Cortical processing mechanism for vocalization with auditory verbal feedback. *Neuroreport, 8*(9), 2379-2382.

Hirnstein, M., Westerhausen, R., Korsnes, M. S., & Hugdahl, K. (2013, Jul-Aug). Sex differences in language asymmetry are age-dependent and small: a large-scale, consonant-vowel dichotic listening study with behavioral and fMRI data. *Cortex, 49*(7), 1910-1921. https://doi.org/10.1016/j.cortex.2012.08.002

Hodges, J. R., Bozeat, S., Lambon Ralph, M. A., Patterson, K., & Spatt, J. (2000, Sep). The role of conceptual knowledge in object use evidence from semantic dementia. *Brain, 123 (Pt 9)*, 1913-1925. https://doi.org/10.1093/brain/123.9.1913

Holmes, E., & Griffiths, T. D. (2019, 2019/11/14). 'Normal' hearing thresholds and fundamental auditory grouping processes predict difficulties with speech-in-noise perception. *Scientific Reports, 9*(1), 16771. https://doi.org/10.1038/s41598-019-53353-5

Holmes, E., Zeidman, P., Friston, K. J., & Griffiths, T. D. (2020). Difficulties with Speech-in-Noise Perception Related to Fundamental Grouping Processes in Auditory Cortex. *Cerebral Cortex, 31*(3), 1582-1596. https://doi.org/10.1093/cercor/bhaa311

Horley, K., Reid, A., & Burnham, D. (2010, Oct). Emotional prosody perception and production in dementia of the Alzheimer's type. *J Speech Lang Hear Res, 53*(5), 1132-1146. https://doi.org/10.1044/1092-4388(2010/09-0030)

Houde, J. F., & Jordan, M. I. (1998). Sensorimotor Adaptation in Speech Production. *Science, 279*(5354), 1213-1216. https://doi.org/DOI: 10.1126/science.279.5354.1213

Hovius, M., Kellenbach, M. L., Graham, K. S., Hodges, J. R., & Patterson, K. (2003, Jan). What does the object decision task measure? Reflections on the basis of evidence from semantic dementia. *Neuropsychology, 17*(1), 100-107.

Hovsepyan, S., Olasagasti, I., & Giraud, A.-L. (2020). Combining predictive coding and neural oscillations enables online syllable recognition in natural speech. *Nature Communications, 11*(1). https://doi.org/10.1038/s41467-020-16956-5

Hsieh, S., Hornberger, M., Piguet, O., & Hodges, J. R. (2011). Neural basis of music knowledge: evidence from the dementias. *Brain, 134*(9), 2523-2534. https://doi.org/10.1093/brain/awr190

Hsieh, S., Hornberger, M., Piguet, O., & Hodges, J. R. (2012, Jul). Brain correlates of musical and facial emotion recognition: evidence from the dementias. *Neuropsychologia, 50*(8), 1814-1822. https://doi.org/10.1016/j.neuropsychologia.2012.04.006

Humes, L. E., Dubno, J. R., Gordon-Salant, S., Lister, J. J., Cacace, A. T., Cruickshanks, K. J., Gates, G. A., Wilson, R. H., & Wingfield, A. (2012). Central Presbycusis: A Review and Evaluation of the Evidence. *Journal of the American Academy of Audiology, 23*(08), 635-666. https://doi.org/10.3766/jaaa.23.8.5

Huyck, J. J., & Johnsrude, I. S. (2012, Mar). Rapid perceptual learning of noise-vocoded speech requires attention. *J Acoust Soc Am, 131*(3), El236-242. https://doi.org/10.1121/1.3685511

Hwang, J.-H., Li, C.-W., Wu, C.-W., Chen, J.-H., & Liu, T.-C. (2007). Aging Effects on the Activation of the Auditory Cortex during Binaural Speech Listening in White Noise: An fMRI Study. *Audiology and Neurotology, 12*(5), 285-294. https://doi.org/10.1159/000103209

Hyde, K. L., Lerch, J., Norton, A., Forgeard, M., Winner, E., Evans, A. C., & Schlaug, G. (2009). Musical Training Shapes Structural Brain Development. *Journal of Neuroscience, 29*(10), 3019-3025. https://doi.org/10.1523/jneurosci.5118-08.2009

Idrizbegovic, E., Hederstierna, C., Dahlquist, M., & Rosenhall, U. (2013, Jan). Short-term longitudinal study of central auditory function in Alzheimer's disease and mild cognitive impairment. *Dement Geriatr Cogn Dis Extra, 3*(1), 468-471. https://doi.org/10.1159/000355371

Ikeda, M., Tashiro, Y., Takai, E., Kurose, S., Fugami, N., Tsuda, K., Arisaka, Y., Kodaira, S., Fujita, Y., Makioka, K., Mizuno, Y., Shimada, H., Harigaya, Y., Takatama, M., Amari, M., Yamazaki, T., Yamaguchi, H., Higuchi, T., Okamoto, K., Tsushima, Y., & Ikeda, Y. (2014, 2014/12/01). CSF levels of Aβ1-38/Aβ1-40/Aβ1-42 and 11C PiB-PET studies in three clinical variants of primary progressive aphasia and Alzheimer's disease. *Amyloid, 21*(4), 238-245. https://doi.org/10.3109/13506129.2014.949231

Jack, C. R., Jr., Knopman, D. S., Jagust, W. J., Shaw, L. M., Aisen, P. S., Weiner, M. W., Petersen, R. C., & Trojanowski, J. Q. (2010, Jan). Hypothetical model of dynamic biomarkers of the Alzheimer's pathological cascade. *Lancet Neurol, 9*(1), 119-128. https://doi.org/10.1016/s1474-4422(09)70299-6

Jaekel, B. N., Newman, R. S., & Goupell, M. J. (2018). Age effects on perceptual restoration of degraded interrupted sentences. *The Journal of the Acoustical Society of America, 143*(1), 84-97. https://doi.org/10.1121/1.5016968

Jiang, J., Benhamou, E., Waters, S., Johnson, J. C. S., Volkmer, A., Weil, R. S., Marshall, C. R., Warren, J. D., & Hardy, C. J. D. (2021). Processing of Degraded Speech in Brain Disorders. *Brain Sciences, 11*(394), 28. https://doi.org/10.3390/brainsci11030394

Jiang, J., Johnson, J. C. S., Requena-Komuro, M.-C., Benhamou, E., Sivasathiaseelan, H., Sheppard, D. L., Volkmer, A., Crutch, S. J., Hardy, C. J. D., & Warren, J. D. (2022). Phonemic restoration in Alzheimer's disease and semantic dementia: a preliminary investigation. *Brain Communications, 4*(3). https://doi.org/10.1093/braincomms/fcac118

Jin, S.-H., & Liu, C. (2012). English sentence recognition in speech-shaped noise and multi-talker babble for English-, Chinese-, and Korean-native listeners. *The Journal of the Acoustical Society of America, 132*(5), EL391-EL397. https://doi.org/10.1121/1.4757730

Johnson, J. C. S. (2021). *Hearing impairment in dementia: defining deficits and assessing impact* UCL].

Johnson, J. C. S., Jiang, J., Bond, R. L., Benhamou, E., Requena-Komuro, M. C., Russell, L. L., Greaves, C., Nelson, A., Sivasathiaseelan, H., Marshall, C. R., Volkmer, A. P., Rohrer, J. D., Warren, J. D., & Hardy, C. J. D. (2020, Jul). Impaired phonemic discrimination in logopenic variant primary progressive aphasia. *Ann Clin Transl Neurol, 7*(7), 1252-1257. https://doi.org/10.1002/acn3.51101

Johnson, J. C. S., Marshall, C. R., Weil, R. S., Bamiou, D.-E., Hardy, C. J. D., & Warren, J. D. (2021). Hearing and dementia: from ears to brain. *Brain, 144*(2), 391-401. https://doi.org/10.1093/brain/awaa429

Johnsrude, I. S. (2002). The neuropsychological consequences of temporal lobe lesions. In J. E. Harrison & A. M. Owen (Eds.), *Cognitive deficits in brain disorders* (pp. 37-57). Martin Dunitz Ltd.

Jones, J. A., & Munhall, K. G. (2000, Sep). Perceptual calibration of F0 production: evidence from feedback perturbation. *J Acoust Soc Am, 108*(3 Pt 1), 1246-1251. https://doi.org/10.1121/1.1288414

Josephs, K. A., Duffy, J. R., Strand, E. A., Machulda, M. M., Senjem, M. L., Lowe, V. J., Jack, C. R., & Whitwell, J. L. (2013). Syndromes dominated by apraxia of speech show distinct characteristics from agrammatic PPA. *Neurology, 81*(4), 337-345. https://doi.org/10.1212/WNL.0b013e31829c5ed5

Kamath, V., Sutherland, E. R., & Chaney, G.-A. (2020). A Meta-Analysis of Neuropsychological Functioning in the Logopenic Variant of Primary Progressive Aphasia: Comparison with the Semantic and Non-Fluent Variants. *Journal of the International Neuropsychological Society, 26*(3), 322-330. https://doi.org/10.1017/S1355617719001115

Kashino, M. (2006). Phonemic restoration: The brain creates missing speech sounds. *Acoustical Science and Technology, 6*, 4. https://doi.org/https://doi.org/10.1250/ast.27.318

Kee, K. S., Kern, R. S., Marshall, B. D., & Green, M. F. (1998, 1998/05/25/). Risperidone versus haloperidol for perception of emotion in treatment-resistant schizophrenia: preliminary findings. *Schizophrenia Research, 31*(2), 159-165. https://doi.org/https://doi.org/10.1016/S0920-9964(98)00026-7

Kennedy, R. E., Williams, C. P., Sawyer, P., Allman, R. M., & Crowe, M. (2014). Comparison of In-Person and Telephone Administration of the Mini-Mental State Examination in the University of Alabama at Birmingham Study of Aging. *Journal of the American Geriatrics Society, 62*(10), 1928-1932. https://doi.org/https://doi.org/10.1111/jgs.13027

Khoshkhoo, S., Leonard, M. K., Mesgarani, N., & Chang, E. F. (2018). Neural correlates of sine-wave speech intelligibility in human frontal and temporal cortex. *Brain and Language, 187*, 83-91. https://doi.org/https://doi.org/10.1016/j.bandl.2018.01.007

Kim, S., Choi, I., Schwalje, A. T., Kim, K., & Lee, J. H. (2020). Auditory Working Memory Explains Variance in Speech Recognition in Older Listeners Under Adverse Listening Conditions. *Clin Interv Aging, 15*, 395-406. https://doi.org/10.2147/CIA.S241976

Knibb, J. A., & Hodges, J. R. (2005, Oct-Dec). Semantic dementia and primary progressive aphasia: a problem of categorization? *Alzheimer Dis Assoc Disord, 19 Suppl 1*, S7-14. https://doi.org/10.1097/01.wad.0000183085.22562.13

Kocagoncu, E., Klimovich-Gray, A., Hughes, L., & Rowe, J. B. (2020, 2020-06-11). Evidence and implications of abnormal predictive coding in dementia. *arXiv pre-print server*, 16. https://doi.org/None

arxiv:2006.06311

Koelsch, S., Schröger, E., & Tervaniemi, M. (1999, Apr 26). Superior pre-attentive auditory processing in musicians. *Neuroreport, 10*(6), 1309-1313. https://doi.org/10.1097/00001756-199904260-00029

Kort, N. S., Cuesta, P., Houde, J. F., & Nagarajan, S. S. (2016). Bihemispheric network dynamics coordinating vocal feedback control. *Human Brain Mapping, 37*(4), 1474-1485. https://doi.org/https://doi.org/10.1002/hbm.23114

Kort, N. S., Nagarajan, S. S., & Houde, J. F. (2014). A bilateral cortical network responds to pitch perturbations in speech feedback. *NeuroImage, 86*, 525-535. https://doi.org/https://doi.org/10.1016/j.neuroimage.2013.09.042

Kotz, S. A., Kalberlah, C., Bahlmann, J., Friederici, A. D., & Haynes, J. D. (2013, Aug). Predicting vocal emotion expressions from the human brain. *Hum Brain Mapp, 34*(8), 1971-1981. https://doi.org/10.1002/hbm.22041

Kotz, S. A., Meyer, M., Alter, K., Besson, M., Cramon, D. Y. v., & Friederici, A. D. (2003). On the lateralization of emotional prosody: An event-related functional MR investigation. *Brain and Language, 86*(3), 10. https://doi.org/10.1016/S0093-934X(02)00532-1

Kousaie, S., Baum, S., Phillips, N. A., Gracco, V., Titone, D., Chen, J. K., Chai, X. J., & Klein, D. (2019, Sep). Language learning experience and mastering the challenges of perceiving speech in noise. *Brain Lang, 196*, 104645. https://doi.org/10.1016/j.bandl.2019.104645

Kraljic, T., & Samuel, A. G. (2005). Perceptual learning for speech: Is there a return to normal? *Cognitive Psychology, 51*(2), 141-178. https://doi.org/10.1016/j.cogpsych.2005.05.001

Kramer, S. E., Kapteyn, T. S., Festen, J. M., & Tobi, H. (1995, 1995/01/01). Factors in Subjective Hearing Disability. *Audiology, 34*(6), 311-320. https://doi.org/10.3109/00206099509071921

Kraus, N., Skoe, E., Parbery-Clark, A., & Ashley, R. (2009, Jul). Experience-induced malleability in neural encoding of pitch, timbre, and timing. *Ann N Y Acad Sci, 1169*, 543-557. https://doi.org/10.1111/j.1749-6632.2009.04549.x

Kremen, S. A., Mendez, M. F., Tsai, P. H., & Teng, E. (2011, Feb). Extrapyramidal signs in the primary progressive aphasias. *Am J Alzheimers Dis Other Demen, 26*(1), 72-77. https://doi.org/10.1177/1533317510391239

Kyong, J. S., Scott, S. K., Rosen, S., Howe, T. B., Agnew, Z. K., & McGettigan, C. (2014). Exploring the Roles of Spectral Detail and Intonation Contour in Speech Intelligibility: An fMRI Study. *Journal of Cognitive Neuroscience, 26*(8), 1748-1763. https://doi.org/10.1162/jocn_a_00583

Lakretz, Y., Ossmy, O., Friedmann, N., Mukamel, R., & Fried, I. (2021). Single-cell activity in human STG during perception of phonemes is organized according to manner of articulation. *NeuroImage, 226*. https://doi.org/https://doi.org/10.1016/j.neuroimage.2020.117499

Lambon Ralph, M. A., & Patterson, K. (2008). Generalization and Differentiation in Semantic Memory. *Annals of the New York Academy of Sciences, 1124*(1), 61-76. https://doi.org/10.1196/annals.1440.006

Lee, B. S. (1950). Effects of Delayed Speech Feedback. *The Journal of the Acoustical Society of America, 22*(6). https://doi.org/https://doi.org/10.1121/1.1906696

Lehet, M. I., Fenn, K. M., & Nusbaum, H. C. (2020, 2020/10/01). Shaping perceptual learning of synthetic speech through feedback. *Psychonomic Bulletin & Review, 27*(5), 1043-1051. https://doi.org/10.3758/s13423-020-01743-6

Lennox, R. D., & Wolfe, R. N. (1984). Revision of the Self-Monitoring Scale. *Journal of Personality and Social Psychology, 46*, 1349-1364.

Leonard, M. K., Baud, M. O., Sjerps, M. J., & Chang, E. F. (2016). Perceptual restoration of masked speech in human cortex. *Nature Communications, 7*(1), 13619. https://doi.org/10.1038/ncomms13619

Lewis, J. L. (1970, Aug). Semantic processing of unattended messages using dichotic listening. *J Exp Psychol, 85*(2), 225-228. https://doi.org/10.1037/h0029518

Liang, B., & Du, Y. (2018). The Functional Neuroanatomy of Lexical Tone Perception: An Activation Likelihood Estimation Meta-Analysis. *Frontiers in Neuroscience*. https://doi.org/https://doi.org/10.3389/fnins.2018.00495

Liberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological Review, 74*(6), 431-461. https://doi.org/https://doi.org/10.1037/h0020279

Liberman, A. M., & Mattingly, I. G. (1985, 1985/10/01/). The motor theory of speech perception revised. *Cognition, 21*(1), 1-36. https://doi.org/https://doi.org/10.1016/0010-0277(85)90021-6

Liddle, J., Smith-Conway, E. R., Baker, R., Angwin, A. J., Gallois, C., Copland, D. A., Pachana, N. A., Humphreys, M. S., Byrne, G. J., & Chenery, H. J. (2012, Dec). Memory and communication support strategies in dementia: effect of a training program for informal caregivers. *Int Psychogeriatr, 24*(12), 1927-1942. https://doi.org/10.1017/s1041610212001366

Lidestam, B., Holgersson, J., & Moradi, S. (2014). Comparison of informational vs. energetic masking effects on speechreading performance. *Front Psychol, 5*, 639. https://doi.org/10.3389/fpsyg.2014.00639

Liebenthal, E., Silbersweig, D. A., & Stern, E. (2016). The Language, Tone and Prosody of Emotions: Neural Substrates and Dynamics of Spoken-Word Emotion Perception. *Frontiers in Neuroscience, 10*(506). https://doi.org/10.3389/fnins.2016.00506

Lin, F. R., Metter, E. J., O'Brien, R. J., Resnick, S. M., Zonderman, A. B., & Ferrucci, L. (2011). Hearing Loss and Incident Dementia. *Archives of Neurology, 68*(2), 214-220. https://doi.org/10.1001/archneurol.2010.362

Lincoln, M., Packman, A., & Onslow, M. (2006). Altered auditory feedback and the treatment of stuttering: a review. *J Fluency Disord, 31*(2), 71-89. https://doi.org/10.1016/j.jfludis.2006.04.001

Livingston, G., Sommerlad, A., Orgeta, V., Costafreda, S. G., Huntley, J., Ames, D., Ballard, C., Banerjee, S., Burns, A., Cohen-Mansfield, J., Cooper, C., Fox, N., Gitlin, L. N., Howard, R., Kales, H. C., Larson, E. B., Ritchie, K., Rockwood, K., Sampson, E. L., Samus, Q., Schneider, L. S., Selbæk, G., Teri, L., & Mukadam, N. (2017). Dementia prevention, intervention, and care. *The Lancet, 390*(10113), 2673-2734. https://doi.org/10.1016/s0140-6736(17)31363-6

Lo, C. Y., Looi, V., Thompson, W. F., & McMahon, C. M. (2020, Jun 22). Music Training for Children With Sensorineural Hearing Loss Improves Speech-in-Noise Perception. *J Speech Lang Hear Res, 63*(6), 1990-2015. https://doi.org/10.1044/2020_jslhr-19-00391

Lombardi, J., Mayer, B., Semler, E., Anderl-Straub, S., Uttner, I., Kassubek, J., Diehl-Schmid, J., Danek, A., Levin, J., Fassbender, K., Fliessbach, K., Schneider, A., Huppertz, H. J., Jahn, H., Volk, A., Kornhuber, J., Landwehrmeyer, B., Lauer, M., Prudlo, J., Wiltfang, J., Schroeter, M. L., Ludolph, A., & Otto, M. (2021). Quantifying progression in primary progressive aphasia with structural neuroimaging. *Alzheimer's & Dementia, 17*(10), 1595-1609. https://doi.org/10.1002/alz.12323

Loughrey, D. G., Kelly, M. E., Kelley, G. A., Brennan, S., & Lawlor, B. A. (2018). Association of Age-Related Hearing Loss With Cognitive Function, Cognitive Impairment, and Dementia: A Systematic Review and Meta-analysis. *JAMA Otolaryngology–Head & Neck Surgery, 144*(2), 115-126. https://doi.org/10.1001/jamaoto.2017.2513

Lucks Mendel, L., & Widner, H. (2016). Speech perception in noise for bilingual listeners with normal hearing. *Int J Audiol, 55*(2), 126-134. https://doi.org/10.3109/14992027.2015.1061710

Macoir, J., Hudon, C., Tremblay, M.-P., Laforce, R. J., & Wilson, M. A. (2019). The contribution of semantic memory to the recognition of basic emotions and emotional valence:

Evidence from the semantic variant of primary progressive aphasia. *Social Neuroscience, 14*(6), 11. https://doi.org/10.1080/17470919.2019.1577295

Magdalinou, N. K., Golden, H. L., Nicholas, J. M., Witoonpanich, P., Mummery, C. J., Morris, H. R., Djamshidian, A., Warner, T. T., Warrington, E. K., Lees, A. J., & Warren, J. D. (2018, Aug). Verbal adynamia in parkinsonian syndromes: behavioral correlates and neuroanatomical substrate. *Neurocase, 24*(4), 204-212. https://doi.org/10.1080/13554794.2018.1527368

Magne, C., Schön, D., & Besson, M. (2006, Feb). Musician children detect pitch violations in both music and language better than nonmusician children: behavioral and electrophysiological approaches. *J Cogn Neurosci, 18*(2), 199-211. https://doi.org/10.1162/089892906775783660

Malone, I. B., Leung, K. K., Clegg, S., Barnes, J., Whitwell, J. L., Ashburner, J., Fox, N. C., & Ridgway, G. R. (2015, Jan 1). Accurate automatic estimation of total intracranial volume: a nuisance variable with less nuisance. *NeuroImage, 104*, 366-372. https://doi.org/10.1016/j.neuroimage.2014.09.034

Mark, G. https://www.guymark.com/product-information/hearing-assessment/audiometers/clinical/gsi-audiostar-pro

Marshall, C. R., Hardy, C. J. D., Russell, L. L., Bond, R. L., Sivasathiaseelan, H., Greaves, C., Moore, K. M., Agustus, J. L., van Leeuwen, J. E. P., Wastling, S. J., Rohrer, J. D., Kilner, J. M., & Warren, J. D. (2019). The functional neuroanatomy of emotion processing in frontotemporal dementias. *Brain, 142*(9), 2873-2887. https://doi.org/10.1093/brain/awz204

Marshall, C. R., Hardy, C. J. D., Volkmer, A., Russell, L. L., Bond, R. L., Fletcher, P. D., Clark, C. N., Mummery, C. J., Schott, J. M., Rossor, M. N., Fox, N. C., Crutch, S. J., Rohrer, J. D., & Warren, J. D. (2018). Primary progressive aphasia: a clinical approach. *Journal of Neurology, 265*(6), 1474-1490. https://doi.org/10.1007/s00415-018-8762-6

Maruta, C., Makhmood, S., Downey, L. E., Golden, H. L., Fletcher, P. D., Witoonpanich, P., Rohrer, J. D., & Warren, J. D. (2014). Delayed auditory feedback simulates features of nonfluent primary progressive aphasia. *Journal of the Neurological Sciences, 347*(1-2), 345-348. https://doi.org/https://doi.org/10.1016/j.jns.2014.09.039

Mattys, S. L., Davis, M. H., Bradlow, A. R., & Scott, S. K. (2012). Speech recognition in adverse conditions: A review. *Language and Cognitive Processes, 27*(7-8), 953-978. https://doi.org/10.1080/01690965.2012.705006

McGettigan, C., Evans, S., Rosen, S., Agnew, Z. K., Shah, P., & Scott, S. K. (2012, Mar). An application of univariate and multivariate approaches in FMRI to quantifying the hemispheric lateralization of acoustic and linguistic processes. *J Cogn Neurosci, 24*(3), 636-652. https://doi.org/10.1162/jocn_a_00161

McGettigan, C., Rosen, S., & Scott, S. (2014). Lexico-semantic and acoustic-phonetic processes in the perception of noise-vocoded speech: implications for cochlear implantation. *Frontiers System Neuroscience*. https://doi.org/https://doi.org/10.3389/fnsys.2014.00018

McGettigan, C., & Scott, S. K. (2012, 2012/05/01/). Cortical asymmetries in speech perception: what's wrong, what's right and what's left? *Trends in Cognitive Sciences, 16*(5), 269-276. https://doi.org/https://doi.org/10.1016/j.tics.2012.04.006

McGettigan, C., & Tremblay, P. (2018). Links Between Perception and Production: Examining the roles of motor and premotor cortices in understanding speech. In. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780198786825.013.14

McGuire, P. K., Silbersweig, D. A., & Frith, C. D. (1996). Functional neuroanatomy of verbal self-monitoring. *Brain, 119*(3), 907-917. https://doi.org/https://doi.org/10.1093/brain/119.3.907

Meister, H., Schreitmüller, S., Grugel, L., Beutner, D., Walger, M., & Meister, I. (2013). Examining Speech Perception in Noise and Cognitive Functions in the Elderly. *American Journal of Audiology, 22*(2), 310-312. https://doi.org/doi:10.1044/1059-0889(2012/12-0067)

Meister, I. G., Wilson, S. M., Deblieck, C., Wu, A. D., & Iacoboni, M. (2007, 2007/10/09/). The Essential Role of Premotor Cortex in Speech Perception. *Current Biology, 17*(19), 1692-1696. https://doi.org/https://doi.org/10.1016/j.cub.2007.08.064

Mendez, M. F., Monserratt, L. H., Liang, L.-J., Chavez, D., Jimenez, E. E., Maurer, J. J., & Laffey, M. (2019). Neuropsychological Similarities and Differences Between Amnestic Alzheimer's Disease and its Non-Amnestic Variants. *Journal of Alzheimer's Disease, 69*, 849-855. https://doi.org/10.3233/JAD-190124

Merten, N., Fischer, M. E., Dillard, L. K., Klein, B. E. K., Tweed, T. S., & Cruickshanks, K. J. (2021, Jul 16). Benefit of Musical Training for Speech Perception and Cognition Later in Life. *J Speech Lang Hear Res, 64*(7), 2885-2896. https://doi.org/10.1044/2021_jslhr-20-00588

Mill, A., Allik, J., Realo, A., & Valk, R. (2009). Age-related differences in emotion recognition ability: A cross-sectional study. *Emotion, 9*(5), 619-630. https://doi.org/https://doi.org/10.1037/a0016562

Miller, G. A., & Nicely, P. E. (1955). An Analysis of Perceptual Confusions Among Some English Consonants. *The Journal of the Acoustical Society of America, 27*(2), 338-352. https://doi.org/10.1121/1.1907526

Millman, R. E., & Mattys, S. L. (2017). Auditory Verbal Working Memory as a Predictor of Speech Perception in Modulated Maskers in Listeners With Normal Hearing. *Journal of Speech, Language, and Hearing Research, 60*(5), 1236-1245. https://doi.org/doi:10.1044/2017_JSLHR-S-16-0105

Mole, J. A., Baker, I. W., Ottley Munoz, J. M., Danby, M., Warren, J. D., & Butler, C. R. (2019). Avian agnosia: A window into auditory semantics. *Neuropsychologica, 134*. https://doi.org/10.1016/j.neuropsychologica.2019.107219

Moossavi, A., Mehrkian, S., Gohari, N., Nazari, M. A., Bakhshi, E., & Alain, C. (2021, Oct). The effect of harmonic training on speech perception in noise in hearing-impaired children. *Int J Pediatr Otorhinolaryngol, 149*, 110845. https://doi.org/10.1016/j.ijporl.2021.110845

Moray, N. (1959). Attention in Dichotic Listening: Affective Cues and the Influence of Instructions. *Quarterly Journal of Experimental Psychology, 11*(1), 56-60. https://doi.org/10.1080/17470215908416289

Moreno, S., Marques, C., Santos, A., Santos, M., Castro, S. L., & Besson, M. (2008). Musical Training Influences Linguistic Abilities in 8-Year-Old Children: More Evidence for Brain Plasticity. *Cerebral Cortex, 19*(3), 712-723. https://doi.org/10.1093/cercor/bhn120

Morningstar, M., Mattson, W. I., & Nelson, E. E. (2022). Longitudinal change in neural response to vocal emotion in adolescence. *Social Cognitive and Affective Neuroscience, 17*(10), 890-903. https://doi.org/10.1093/scan/nsac021

Möttönen, R., Calvert, G. A., Jääskeläinen, I. P., Matthews, P. M., Thesen, T., Tuomainen, J., & Sams, M. (2006, Apr 1). Perceiving identical sounds as speech or non-speech modulates activity in the left posterior superior temporal sulcus. *NeuroImage, 30*(2), 563-569. https://doi.org/10.1016/j.neuroimage.2005.10.002

MRC Psycholinguistic Database. https://websites.psychology.uwa.edu.au/school/mrcdatabase/uwa_mrc.htm

Muhammed, L., Hardy, C. J. D., Russell, L. L., Marshall, C. R., Clark, C. N., Bond, R. L., Warrington, E. K., & Warren, J. D. (2018). Agnosia for bird calls. *Neuropsychologia, 113*, 6. https://doi.org/10.1016/j.neuropsychologica.2018.03.024

Musiek, F. E., & Baran, J. A. (2020). *The auditory system: Anatomy, physiology, and clinical correlates* (Second ed.). Plural Publishing.

Musiek, F. E., & Chermak, G. D. (2015). Chapter 18 - Psychophysical and behavioral peripheral and central auditory tests. In M. J. Aminoff, F. Boller, & D. F. Swaab (Eds.), *Handbook of Clinical Neurology* (Vol. 129, pp. 313-332). Elsevier. https://doi.org/https://doi.org/10.1016/B978-0-444-62630-1.00018-4

Neves, L., Correia, A. I., Castro, S. L., Martins, D., & Lima, C. F. (2022, 2022/09/01/). Does music training enhance auditory and linguistic processing? A systematic review and meta-analysis of behavioral and brain evidence. *Neuroscience & Biobehavioral Reviews, 140*, 104777. https://doi.org/https://doi.org/10.1016/j.neubiorev.2022.104777

Newkirk, L. A., Kim, J. M., Thompson, J. M., Tinklenberg, J. R., Yesavage, J. A., & Taylor, J. L. (2004). Validation of a 26-Point Telephone Version of the Mini-Mental State Examination. *Journal of Geriatric Psychiatry and Neurology, 17*(2), 81-87. https://doi.org/10.1177/0891988704264534

Norris, D. (2003). Perceptual learning in speech. *Cognitive Psychology, 47*(2), 204-238. https://doi.org/10.1016/s0010-0285(03)00006-9

Nygaard, L. C., & Pisoni, D. B. (1998). Talker-specific learning in speech perception. *Perception & Psychophysics, 60*(3), 355-376. https://doi.org/10.3758/bf03206860

Obleser, J., & Kotz, S. A. (2010, Mar). Expectancy constraints in degraded speech modulate the language comprehension network. *Cereb Cortex, 20*(3), 633-640. https://doi.org/10.1093/cercor/bhp128

Obleser, J., Wise, R. J. S., Alex Dresner, M., & Scott, S. K. (2007). Functional Integration across Brain Regions Improves Speech Perception under Adverse Listening Conditions. *The Journal of Neuroscience, 27*(9), 2283-2289. https://doi.org/10.1523/jneurosci.4663-06.2007

Ohman, E. M., Granger, C. B., Harrington, R. A., & Lee, K. L. (2000). Risk Stratification and Therapeutic Decision Making in Acute Coronary Syndromes. *JAMA, 284*(7), 876-878. https://doi.org/10.1001/jama.284.7.876

Okada, K., Rong, F., Venezia, J., Matchin, W., Hsieh, I. H., Saberi, K., Serences, J. T., & Hickok, G. (2010, Oct). Hierarchical organization of human auditory cortex: evidence from acoustic invariance in the response to intelligible speech. *Cereb Cortex, 20*(10), 2486-2495. https://doi.org/10.1093/cercor/bhp318

Omar, R., Hailstone, J. C., Warren, J. E., Crutch, S. J., & Warren, J. D. (2010, Apr). The cognitive organization of music knowledge: a clinical analysis. *Brain, 133*(Pt 4), 1200-1213. https://doi.org/10.1093/brain/awp345

Omar, R., Henley, S. M., Bartlett, J. W., Hailstone, J. C., Gordon, E., Sauter, D. A., Frost, C., Scott, S. K., & Warren, J. D. (2011). The structural neuroanatomy of music emotion recognition: evidence from frontotemporal lobar degeneration. *NeuroImage, 56*(3), 7. https://doi.org/10.1016/j.neuroimage.2011.03.002

Panza, F., Solfrizzi, V., & Logroscino, G. (2015, 2015/03/01). Age-related hearing impairment—a risk factor and frailty marker for dementia and AD. *Nature Reviews Neurology, 11*(3), 166-175. https://doi.org/10.1038/nrneurol.2015.12

Paquette, S., Ahmed, G. D., Goffi-Gomez, M. V., Hoshino, A. C. H., Peretz, I., & Lehmann, A. (2018, 2018/12/01/). Musical and vocal emotion perception for cochlear implants users. *Hearing Research, 370*, 272-282. https://doi.org/https://doi.org/10.1016/j.heares.2018.08.009

Parbery-Clark, A., Skoe, E., Lam, C., & Kraus, N. (2009, Dec). Musician enhancement for speech-in-noise. *Ear Hear, 30*(6), 653-661. https://doi.org/10.1097/AUD.0b013e3181b412e9

Patel, A. D. (2011). Why would musical training benefit the neural encoding of speech? The OPERA hypothesis. *Frontiers in Psychology*. https://doi.org/https://doi.org/10.3389/fpsyg.2011.00142

Patel, N., Peterson, K. A., Ingram, R. U., Storey, I., Cappa, S. F., Catricala, E., Halai, A., Patterson, K. E., Lambon Ralph, M. A., Rowe, J. B., & Garrard, P. (2021). A 'Mini Linguistic State Examination' to classify primary progressive aphasia. *Brain Communications, 4*(2). https://doi.org/10.1093/braincomms/fcab299

Paulmann, S. (2016). The Neurocognition of Prosody. In *Neurobiology of Language* (pp. 11). Academic Press.

Paulus, M., Hazan, V., & Adank, P. (2020). The relationship between talker acoustics, intelligibility, and effort in degraded listening conditions. *The Journal of the Acoustical Society of America, 147*(5). https://doi.org/https://doi.org/10.1121/10.0001212

Peelle, J. E. (2010). Hierarchical processing for speech in human auditory cortex and beyond. *Frontiers in Human Neuroscience*. https://doi.org/10.3389/fnhum.2010.00051

Peelle, J. E. (2018, Mar/Apr). Listening Effort: How the Cognitive Consequences of Acoustic Challenge Are Reflected in Brain and Behavior. *Ear Hear, 39*(2), 204-214. https://doi.org/10.1097/aud.0000000000000494

Peelle, J. E., McMillan, C., Moore, P., Grossman, M., & Wingfield, A. (2004, Dec). Dissociable patterns of brain activity during comprehension of rapid and syntactically complex speech: evidence from fMRI. *Brain Lang, 91*(3), 315-325. https://doi.org/10.1016/j.bandl.2004.05.007

Peelle, J. E., & Wingfield, A. (2005, Dec). Dissociations in perceptual learning revealed by adult age differences in adaptation to time-compressed speech. *J Exp Psychol Hum Percept Perform, 31*(6), 1315-1330. https://doi.org/10.1037/0096-1523.31.6.1315

Peretz, I., Vuvan, D., Lagrois, M.-É., & Armony, J. L. (2015). Neural overlap in processing music and speech. *Philosophical Transactions of the Royal Society B: Biological Sciences, 370*(1664), 20140090. https://doi.org/doi:10.1098/rstb.2014.0090

Perrachione, T. K., Lee, J., Ha, L. Y. Y., & Wong, P. C. M. (2011). Learning a novel phonological contrast depends on interactions between individual differences and training paradigm design. *The Journal of the Acoustical Society of America, 130*(1), 461-472. https://doi.org/10.1121/1.3593366

Perry, R. J., Rosen, H. R., Kramer, J. H., Beer, J. S., Levenson, R. L., & Miller, B. L. (2001). Hemispheric dominance for emotions, empathy and social behaviour: evidence from right and left handers with frontotemporal dementia. *Neurocase, 7*(2), 15. https://doi.org/10.1093/neucase/7.2.145

Petrov, A. A., Dosher, B. A., & Lu, Z. L. (2005, Oct). The dynamics of perceptual learning: an incremental reweighting model. *Psychol Rev, 112*(4), 715-743. https://doi.org/10.1037/0033-295x.112.4.715

Pichora-Fuller, M. K. (2008). Use of supportive context by younger and older adult listeners: Balancing bottom-up and top-down information processing. *International Journal of Audiology, 47*(sup2), S72-S82. https://doi.org/10.1080/14992020802307404

Pichora-Fuller, M. K., Schneider, B. A., & Daneman, M. (1995, Jan). How young and old adults listen to and remember speech in noise. *J Acoust Soc Am, 97*(1), 593-608. https://doi.org/10.1121/1.412282

Pickles, J. O. (2015). Chapter 1 - Auditory pathways: anatomy and physiology. In M. J. Aminoff, F. Boller, & D. F. Swaab (Eds.), *Handbook of Clinical Neurology* (Vol. 129, pp. 3-25). Elsevier. https://doi.org/https://doi.org/10.1016/B978-0-444-62630-1.00001-9

Pihans, H., Altenmuller, E., & Ackermann, H. (1997). The cortical processing of perceived emotion a DC-potential study on affective speech prosody. *Cognitive Neuroscience and Neuropsychology, 8*(3), 4.

Pobric, G., Jefferies, E., & Ralph, M. A. L. (2007). Anterior temporal lobes mediate semantic representation: Mimicking semantic dementia by using rTMS in normal participants. *Proceedings of the National Academy of Sciences, 104*(50), 20137-20141. https://doi.org/10.1073/pnas.0707383104

Poldrack, R. A., Temple, E., Protopapas, A., Nagarajan, S., Tallal, P., Merzenich, M., & Gabrieli, J. D. (2001, Jul 1). Relations between the neural bases of dynamic auditory processing and phonological processing: evidence from fMRI. *J Cogn Neurosci, 13*(5), 687-697. https://doi.org/10.1162/089892901750363235

Popescu, T., Stahl, B., Wiernik, B. M., Haiduk, F., Zemanek, M., Helm, H., Matzinger, T., Beisteiner, R., & Fitch, W. T. (2022). Melodic Intonation Therapy for aphasia: A multi-level meta-analysis of randomized controlled trials and individual participant data. *Annals*

of the New York Academy of Sciences, 1516(1), 76-84. https://doi.org/https://doi.org/10.1111/nyas.14848

Powell, D. S., Oh, E. S., Reed, N. S., Lin, F. R., & Deal, J. A. (2021). Hearing Loss and Cognition: What We Know and Where We Need to Go. *Front Aging Neurosci, 13*, 769405. https://doi.org/10.3389/fnagi.2021.769405

Pralus, A., Fornoni, L., Bouet, R., Gomot, M., Bhatara, A., Tillmann, B., & Caclin, A. (2019). Emotional prosody in congenital amusia: Impaired and spared processes. *Neuropsychologia, 134*. https://doi.org/https://doi.org/10.1016/j.neuropsychologia.2019.107234

Pralus, A., Hermann, R., Cholvy, F., Aguera, P. E., Moulin, A., Barone, P., Grimault, N., Truy, E., Tillmann, B., & Caclin, A. (2021, May 13). Rapid Assessment of Non-Verbal Auditory Perception in Normal-Hearing Participants and Cochlear Implant Users. *J Clin Med, 10*(10). https://doi.org/10.3390/jcm10102093

Prince, M., Bryce, R., Albanese, E., Wimo, A., Ribeiro, W., & Ferri, C. P. (2013, 2013/01/01/). The global prevalence of dementia: A systematic review and metaanalysis. *Alzheimer's & Dementia, 9*(1), 63-75.e62. https://doi.org/https://doi.org/10.1016/j.jalz.2012.11.007

Pronk, M., Lissenberg-Witte, B. I., van der Aa, H. P. A., Comijs, H. C., Smits, C., Lemke, U., Zekveld, A. A., & Kramer, S. E. (2019, Apr 26). Longitudinal Relationships Between Decline in Speech-in-Noise Recognition Ability and Cognitive Functioning: The Longitudinal Aging Study Amsterdam. *J Speech Lang Hear Res, 62*(4s), 1167-1187. https://doi.org/10.1044/2018_jslhr-h-ascc7-18-0120

Purcell, D. W., & Munhall, K. G. (2006). Adaptive control of vowel formant frequency: Evidence from real-time formant manipulation. *The Journal of the Acoustical Society of America, 120*(2). https://doi.org/https://doi.org/10.1121/1.2217714

Rabinovici, G. D., & Miller, B. L. (2010, 2010/05/01). Frontotemporal lobar degeneration. *CNS Drugs, 24*(5), 375-398. https://doi.org/10.2165/11533100-000000000-00000

Ramanan, S., Irish, M., Patterson, K., Rowe, J. B., Gorno-Tempini, M. L., & Lambon Ralph, M. A. (2022, Sep 14). Understanding the multidimensional cognitive deficits of logopenic variant primary progressive aphasia. *Brain, 145*(9), 2955-2966. https://doi.org/10.1093/brain/awac208

Rammell, C. S., Cheng, H., Pisoni, D. B., & Newman, S. D. (2019, Oct 1). L2 speech perception in noise: An fMRI study of advanced Spanish learners. *Brain Res, 1720*, 146316. https://doi.org/10.1016/j.brainres.2019.146316

Rampersad, S., Roig-Solvas, B., Yarossi, M., Kulkarni, P. P., Santarnecchi, E., Dorval, A. D., & Brooks, D. H. (2019, Nov 15). Prospects for transcranial temporal interference stimulation in humans: A computational study. *NeuroImage, 202*, 116124. https://doi.org/10.1016/j.neuroimage.2019.116124

Ranasinghe, K. G., Gill, J. S., Kothare, H., Beagle, A. J., Mizuiri, D., Honma, S. M., Gorno-Tempini, M. L., Miller, B. L., Vossel, K. A., Nagarajan, S. S., & Houde, J. F. (2017, Apr). Abnormal vocal behavior predicts executive and memory deficits in Alzheimer's disease. *Neurobiol Aging, 52*, 71-80. https://doi.org/10.1016/j.neurobiolaging.2016.12.020

Ranasinghe, K. G., Kothare, H., Kort, N., Hinkley, L. B., Beagle, A. J., Mizuiri, D., Honma, S. M., Lee, R., Miller, B. L., Gorno-Tempini, M. L., Vossel, K. A., Houde, J. F., & Nagarajan, S. S. (2019, Apr 5). Neural correlates of abnormal auditory feedback processing during speech production in Alzheimer's disease. *Sci Rep, 9*(1), 5686. https://doi.org/10.1038/s41598-019-41794-x

Rankin, K. P., Salazar, A., Gorno-Tempini, M. L., Sollberger, M., Wilson, S. M., Pavlic, D., & Miller, B. L. (2009). Detecting sarcasm from paralinguistic cues: anatomic and cognitive correlates in neurodegenerative disease. *NeuroImage, 47*(4), 10.

Remez, R., Rubin, P., Pisoni, D., & Carrell, T. (1981). Speech perception without traditional speech cues. *Science, 212*(4497), 947-949. https://doi.org/10.1126/science.7233191

Remez, R. E., Fellowes, J. M., & Rubin, P. E. (1997). Talker identification based on phonetic information. *Journal of Experimental Psychology: Human Perception and Performance, 23*(3), 651-666. https://doi.org/https://doi.org/10.1037/0096-1523.23.3.651

Requena-Komuro, M.-C., Jiang, J., Dobson, L., Benhamou, E., Russell, L., Bond, R. L., Brotherhood, E. V., Greaves, C., Barker, S., Rohrer, J. D., Crutch, S. J., Warren, J. D., & Hardy, C. J. (2022). Remote versus face-to-face neuropsychological testing for dementia research: a comparative study in patients with Alzheimer's disease, patients with frontotemporal dementia, and healthy older individuals. *BMJ Open*.

Ridgway, G. R., Omar, R., Ourselin, S., Hill, D. L. G., Warren, J. D., & Fox, N. C. (2009, 2009/01/01/). Issues with threshold masking in voxel-based morphometry of atrophied brains. *NeuroImage, 44*(1), 99-111. https://doi.org/https://doi.org/10.1016/j.neuroimage.2008.08.045

Ritter, C., & Vongpaisal, T. (2018, Jan-Dec). Multimodal and Spectral Degradation Effects on Speech and Emotion Recognition in Adult Listeners. *Trends Hear, 22*, 2331216518804966. https://doi.org/10.1177/2331216518804966

Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J.-C., & Muller, M. (2011). pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics, 12*, 77. https://doi.org/10.1186/1471-2105-12-77

Rodd, J. M., Davis, M. H., & Johnsrude, I. S. (2005, Aug). The neural mechanisms of speech comprehension: fMRI studies of semantic ambiguity. *Cereb Cortex, 15*(8), 1261-1269. https://doi.org/10.1093/cercor/bhi009

Rohrer, J. D., Crutch, S. J., Warrington, E. K., & Warren, J. D. (2010, Jan). Progranulin-associated primary progressive aphasia: a distinct phenotype? *Neuropsychologia, 48*(1), 288-297. https://doi.org/10.1016/j.neuropsychologia.2009.09.017

Rohrer, J. D., McNaught, E., Foster, J., Clegg, S. L., Barnes, J., Omar, R., Warrington, E. K., Rossor, M. N., Warren, J. D., & Fox, N. C. (2008, Oct 28). Tracking progression in frontotemporal lobar degeneration: serial MRI in semantic dementia. *Neurology, 71*(18), 1445-1451. https://doi.org/10.1212/01.wnl.0000327889.13734.cd

Rohrer, J. D., Ridgway, G. R., Crutch, S. J., Hailstone, J., Goll, J. C., Clarkson, M. J., Mead, S., Beck, J., Mummery, C., Ourselin, S., Warrington, E. K., Rossor, M. N., & Warren, J. D. (2010, 2010/01/01/). Progressive logopenic/phonological aphasia: Erosion of the language network. *NeuroImage, 49*(1), 984-993. https://doi.org/https://doi.org/10.1016/j.neuroimage.2009.08.002

Rohrer, J. D., Sauter, D., Scott, S., Rossor, M. N., & Warren, J. D. (2010). Receptive prosody in nonfluent primary progressive aphasias. *Cortex, 48*(3), 8. https://doi.org/10.1016/j.cortex.2010.09.004

Rohrer, J. D., & Schott, J. M. (2011, May). Primary progressive aphasia: defining genetic and pathological subtypes. *Curr Alzheimer Res, 8*(3), 266-272. https://doi.org/10.2174/156720511795563728

Rosen, H. J., Allison, S. C., Ogar, J. M., Amici, S., Rose, K., Dronkers, N., Miller, B. L., & Gorno-Tempini, M. L. (2006, Nov 28). Behavioral features in semantic dementia vs other forms of progressive aphasias. *Neurology, 67*(10), 1752-1756. https://doi.org/10.1212/01.wnl.0000247630.29222.34

Ross, B., Dobri, S., & Schumann, A. (2020, Feb). Speech-in-noise understanding in older age: The role of inhibitory cortical responses. *Eur J Neurosci, 51*(3), 891-908. https://doi.org/10.1111/ejn.14573

Ross, E. D. (1981). The Aprosodias: Functional-Anatomic Organization of the Affective Components of Language in the Right Hemisphere. *JAMA Neurology, 38*(9), 8. https://doi.org/10.1001/archneur.1981.00510090055006

Roth, T. N. (2015). Aging of the auditory system. In *Handbook of Clinical Neurology* (Vol. 129, pp. 357-373). Elsevier. https://doi.org/10.1016/b978-0-444-62630-1.00020-2

Ruksenaite, J., Volkmer, A., Jiang, J., Johnson, J. C., Marshall, C. R., Warren, J. D., & Hardy, C. J. (2021). Primary Progressive Aphasia: Toward a Pathophysiological Synthesis. *Current Neurology and Neuroscience Reports, 21*(3). https://doi.org/10.1007/s11910-021-01097-z

Russell, L. L. (2022). International collaboration and the FTD Prevention Initiative. *Alzheimer's & Dementia, 18*(S6), e063263. https://doi.org/https://doi.org/10.1002/alz.063263

Saija, J. D., Akyürek, E. G., Andringa, T. C., & Başkent, D. (2014, Feb). Perceptual restoration of degraded speech is preserved with advancing age. *J Assoc Res Otolaryngol, 15*(1), 139-148. https://doi.org/10.1007/s10162-013-0422-z

Sammler, D., Grosbras, M.-H., Anwander, A., Bestelmeyer, Patricia E. G., & Belin, P. (2015, 2015/12/07/). Dorsal and Ventral Pathways for Prosody. *Current Biology, 25*(23), 3079-3085. https://doi.org/https://doi.org/10.1016/j.cub.2015.10.009

Samuel, A. G. (1981, Dec). Phonemic restoration: insights from a new methodology. *J Exp Psychol Gen, 110*(4), 474-494. https://doi.org/10.1037//0096-3445.110.4.474

Samuel, A. G. (1981). The role of bottom-up confirmation in the phonemic restoration illusion. *Journal of Experimental Psychology: Human Perception and Performance, 7*, 1124-1131. https://doi.org/10.1037/0096-1523.7.5.1124

Samuel, A. G. (1991, 1991/08/01). A further examination of attentional effects in the phonemic restoration illusion. *The Quarterly Journal of Experimental Psychology Section A, 43*(3), 679-699. https://doi.org/10.1080/14640749108400992

Samuel, A. G. (1997, 1997/03/01/). Lexical Activation Produces Potent Phonemic Percepts. *Cognitive Psychology, 32*(2), 97-127. https://doi.org/https://doi.org/10.1006/cogp.1997.0646

Samuel, A. G. (2011). Speech Perception. *Annual Review of Psychology, 62*(1), 49-72. https://doi.org/10.1146/annurev.psych.121208.131643

Samuel, A. G., & Kraljic, T. (2009, Aug). Perceptual learning for speech. *Atten Percept Psychophys, 71*(6), 1207-1218. https://doi.org/10.3758/app.71.6.1207

Samuel, A. G., & Ressler, W. H. (1986). Attention within auditory word perception: Insights from the phonemic restoration illusion. *Journal of Experimental Psychology: Human Perception and Performance, 12*, 70-79. https://doi.org/10.1037/0096-1523.12.1.70

Sauseng, P., Klimesch, W., Schabus, M., & Doppelmayr, M. (2005, 2005/08/01/). Fronto-parietal EEG coherence in theta and upper alpha reflect central executive functions of working memory. *International Journal of Psychophysiology, 57*(2), 97-103. https://doi.org/https://doi.org/10.1016/j.ijpsycho.2005.03.018

Sauter, D. (2006). *An Investigation into Vocal Expressions of Emotions: The Roles of Valence, Culture, and Acoustic Factors* University College London].

Schirmer, A., & Kotz, S. A. (2006, 2006/01/01/). Beyond the right hemisphere: brain mechanisms mediating vocal emotional processing. *Trends in Cognitive Sciences, 10*(1), 24-30. https://doi.org/https://doi.org/10.1016/j.tics.2005.11.009

Schön, D., Magne, C., & Besson, M. (2004). The music of speech: Music training facilitates pitch processing in both music and language. *Psychophysiology, 41*(3), 341-349. https://doi.org/10.1111/1469-8986.00172.x

Schütt, H. H., Harmeling, S., Macke, J. H., & Wichmann, F. A. (2016, 2016/05/01/). Painfree and accurate Bayesian estimation of psychometric functions for (potentially) overdispersed data. *Vision Research, 122*, 105-123. https://doi.org/https://doi.org/10.1016/j.visres.2016.02.002

Scott, S. K. (2019). From speech and talkers to the social world: The neural processing of human spoken language. *Science, 366*(6461), 58-62. https://doi.org/doi:10.1126/science.aax0288

Scott, S. K., Blank, C. C., Rosen, S., & Wise, R. J. (2000, Dec). Identification of a pathway for intelligible speech in the left temporal lobe. *Brain, 123 Pt 12*(Pt 12), 2400-2406. https://doi.org/10.1093/brain/123.12.2400

Scott, S. K., & McGettigan, C. (2013, 2013/09/01/). The neural processing of masked speech. *Hearing Research, 303*, 58-66. https://doi.org/https://doi.org/10.1016/j.heares.2013.05.001

Scott, S. K., McGettigan, C., & Eisner, F. (2009, 2009/04/01). A little more conversation, a little less action — candidate roles for the motor cortex in speech perception. *Nature Reviews Neuroscience, 10*(4), 295-302. https://doi.org/10.1038/nrn2603

Scott, S. K., Rosen, S., Lang, H., & Wise, R. J. S. (2006). Neural correlates of intelligibility in speech investigated with noise vocoded speech—A positron emission tomography study. *The Journal of the Acoustical Society of America, 120*(2), 1075-1083. https://doi.org/10.1121/1.2216725

Sedley, W., Gander, P. E., Kumar, S., Kovach, C. K., Oya, H., Kawasaki, H., Howard, M. A., & Griffiths, T. D. (2016). Neural signatures of perceptual inference. *eLife, 5*. https://doi.org/10.7554/elife.11476

Sehm, B., Schnitzler, T., Obleser, J., Groba, A., Ragert, P., Villringer, A., & Obrig, H. (2013, Oct 2). Facilitation of inferior frontal cortex by transcranial direct current stimulation induces perceptual learning of severely degraded speech. *J Neurosci, 33*(40), 15868-15878. https://doi.org/10.1523/jneurosci.5466-12.2013

Shafiro, V., Sheft, S., Gygi, B., & Ho, K. T. (2012, Jun). The influence of environmental sound training on the perception of spectrally degraded speech and environmental sounds. *Trends Amplif, 16*(2), 83-101. https://doi.org/10.1177/1084713812454225

Shahin, A. J., Bishop, C. W., & Miller, L. M. (2009). Neural mechanisms for illusory filling-in of degraded speech. *NeuroImage, 44*(3), 1133-1143. https://doi.org/10.1016/j.neuroimage.2008.09.045

Shannon, R. V., Zeng, F. G., Kamath, V., Wygonski, J., & Ekelid, M. (1995). Speech Recognition with Primarily Temporal Cues. *Science, 270*(5234), 303-304. https://doi.org/10.1126/science.270.5234.303

Shany-Ur, T., & Rankin, K. P. (2011, Dec). Personality and social cognition in neurodegenerative disease. *Curr Opin Neurol, 24*(6), 550-555. https://doi.org/10.1097/WCO.0b013e32834cd42a

Shaw, A. D., Hughes, L. E., Moran, R., Coyle-Gilchrist, I., Rittman, T., & Rowe, J. B. (2021, Feb 5). In Vivo Assay of Cortical Microcircuitry in Frontotemporal Dementia: A Platform for Experimental Medicine Studies. *Cereb Cortex, 31*(3), 1837-1847. https://doi.org/10.1093/cercor/bhz024

Sheffert, S. M., Pisoni, D. B., Fellowes, J. M., & Remez, R. E. (2002). Learning to recognize talkers from natural, sinewave, and reversed speech samples. *Journal of Experimental Psychology: Human Perception and Performance, 28*(6), 1447-1469. https://doi.org/https://doi.org/10.1037/0096-1523.28.6.1447

Shenhav, A., Botvinick, Matthew M., & Cohen, Jonathan D. (2013, 2013/07/24/). The Expected Value of Control: An Integrative Theory of Anterior Cingulate Cortex Function. *Neuron, 79*(2), 217-240. https://doi.org/https://doi.org/10.1016/j.neuron.2013.07.007

Sidaras, S. K., Alexander, J. E., & Nygaard, L. C. (2009, May). Perceptual learning of systematic variation in Spanish-accented speech. *J Acoust Soc Am, 125*(5), 3306-3316. https://doi.org/10.1121/1.3101452

Sivasathiaseelan, H., Marshall, C. R., Agustus, J. L., Benhamou, E., Bond, R. L., van Leeuwen, J. E. P., Hardy, C. J. D., Rohrer, J. D., & Warren, J. D. (2019, Apr). Frontotemporal Dementia: A Clinical Review. *Semin Neurol, 39*(2), 251-263. https://doi.org/10.1055/s-0039-1683379

Skoe, E., & Karayanidi, K. (2019, Feb). Bilingualism and Speech Understanding in Noise: Auditory and Linguistic Factors. *J Am Acad Audiol, 30*(2), 115-130. https://doi.org/10.3766/jaaa.17082

Snowden, J. S., Pickering-Brown, S. M., Mackenzie, I. R., Richardson, A. M., Varma, A., Neary, D., & Mann, D. M. (2006, Nov). Progranulin gene mutations associated with frontotemporal dementia and progressive non-fluent aphasia. *Brain, 129*(Pt 11), 3091-3102. https://doi.org/10.1093/brain/awl267

Snowden, J. S., Stopford, C. L., Julien, C. L., Thompson, J. C., Davidson, Y., Gibbons, L., Pritchard, A., Lendon, C. L., Richardson, A. M., Varma, A., Neary, D., & Mann, D. M. A. (2007, 2007/01/01/). Cognitive Phenotypes in Alzheimer's Disease and Genetic Risk. *Cortex, 43*(7), 835-845. https://doi.org/https://doi.org/10.1016/S0010-9452(08)70683-X

Snowden, J. S., Thompson, J. C., & Neary, D. (2004, Apr). Knowledge of famous faces and names in semantic dementia. *Brain, 127*(Pt 4), 860-872. https://doi.org/10.1093/brain/awh099

Snyder, M. (1974). Self-monitoring of expressive behavior. *Journal of Personality and Social Psychology, 30*, 526-537.

Sohoglu, E., & Davis, M. H. (2016, Mar 22). Perceptual learning of degraded speech by minimizing prediction error. *Proc Natl Acad Sci U S A, 113*(12), E1747-1756. https://doi.org/10.1073/pnas.1523266113

Sohoglu, E., & Davis, M. H. (2020). Rapid computations of spectrotemporal prediction error support perception of degraded speech. *eLife*. https://doi.org/https://doi.org/10.7554/eLife.58077

Sohoglu, E., Peelle, J. E., Carlyon, R. P., & Davis, M. H. (2014, Feb). Top-down influences of written text on perceived clarity of degraded speech. *J Exp Psychol Hum Percept Perform, 40*(1), 186-199. https://doi.org/10.1037/a0033206

Souza, P., & Arehart, K. (2015). Robust relationship between reading span and speech recognition in noise. *International Journal of Audiology, 54*(10), 705-713. https://doi.org/10.3109/14992027.2015.1043062

Spinelli, E. G., Mandelli, M. L., Miller, Z. A., Santos-Santos, M. A., Wilson, S. M., Agosta, F., Grinberg, L. T., Huang, E. J., Trojanowski, J. Q., Meyer, M., Henry, M. L., Comi, G., Rabinovici, G., Rosen, H. J., Filippi, M., Miller, B. L., Seeley, W. W., & Gorno-Tempini, M. L. (2017, Mar). Typical and atypical pathology in primary progressive aphasia variants. *Ann Neurol, 81*(3), 430-443. https://doi.org/10.1002/ana.24885

Stacey, P. C., & Summerfield, A. Q. (2007, May). Effectiveness of computer-based auditory training in improving the perception of noise-vocoded speech. *J Acoust Soc Am, 121*(5 Pt1), 2923-2935. https://doi.org/10.1121/1.2713668

Stalpaert, J., Miatton, M., Sieben, A., Langenhove, T. V., Mierlo, P. v., & Letter, M. D. (2021). The Electrophysiological Correlates of Phoneme Perception in Primary Progressive Aphasia: A Preliminary Case Series. *Frontiers in Human Neuroscience*. https://doi.org/https://doi.org/10.3389/fnhum.2021.618549

Stanislaw, H., & Todorov, N. (1999, 1999/03/01). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, & Computers, 31*(1), 137-149. https://doi.org/10.3758/BF03207704

Stasenko, A., Bonn, C., Teghipco, A., Garcea, F. E., Sweet, C., Dombovy, M., McDonough, J., & Mahon, B. Z. (2015, 2015/02/17). A causal test of the motor theory of speech perception: a case of impaired speech production and spared speech perception. *Cognitive Neuropsychology, 32*(2), 38-57. https://doi.org/10.1080/02643294.2015.1035702

Stevenson, J. S., Clifton, L., Kuźma, E., & Littlejohns, T. J. (2022). Speech-in-noise hearing impairment is associated with an increased risk of incident dementia in 82,039 UK Biobank participants. *Alzheimer's & Dementia, 18*(3), 445-456. https://doi.org/https://doi.org/10.1002/alz.12416

Strange, B. A., Otten, L. J., Josephs, O., Rugg, M. D., & Dolan, R. J. (2002). Dissociable Human Perirhinal, Hippocampal, and Parahippocampal Roles during Verbal Encoding. *The Journal of Neuroscience, 22*(2), 523-528. https://doi.org/10.1523/jneurosci.22-02-00523.2002

Stuart, A., Kalinowski, J., Rastatter, M. P., & Lynch, K. (2002, May). Effect of delayed auditory feedback on normal speakers at two speech rates. *J Acoust Soc Am, 111*(5 Pt 1), 2237-2241. https://doi.org/10.1121/1.1466868

Sunami, K., Ishii, A., Takano, S., Yamamoto, H., Sakashita, T., Tanaka, M., Watanabe, Y., & Yamane, H. (2013, Nov 6). Neural mechanisms of phonemic restoration for speech comprehension revealed by magnetoencephalography. *Brain Res, 1537*, 164-173. https://doi.org/10.1016/j.brainres.2013.09.010

Swaminathan, J., Mason, C. R., Streeter, T. M., Best, V., Kidd, J. G., & Patel, A. D. (2015). Musical training, individual differences and the cocktail party problem. *Scientific Reports, 5*(1), 11628. https://doi.org/10.1038/srep11628

Takaso, H., Eisner, F., Wise, R. J., & Scott, S. K. (2010, Apr). The effect of delayed auditory feedback on activity in the temporal lobe while speaking: a positron emission tomography study. *J Speech Lang Hear Res, 53*(2), 226-236. https://doi.org/10.1044/1092-4388(2009/09-0009)

Taler, V., Baum, S. R., Chertkow, H., & Saumier, D. (2008, Mar). Comprehension of grammatical and emotional prosody is impaired in Alzheimer's disease. *Neuropsychology, 22*(2), 188-195. https://doi.org/10.1037/0894-4105.22.2.188

Taylor, B. (2003). Speech-in-noise tests: How and why to include them in your basic test battery. *The Hearing Journal, 56*(1), 40,42-46. https://doi.org/10.1097/01.HJ.0000293000.76300.ff

Testa, J. A., Beatty, W. W., Gleason, A. C., Orbelo, D. M., & Ross, E. D. (2001, Oct 23). Impaired affective prosody in AD: relationship to aphasic deficits and emotional behaviors. *Neurology, 57*(8), 1474-1481. https://doi.org/10.1212/wnl.57.8.1474

Tinnemore, A. R., Zion, D. J., Kulkarni, A. M., & Chatterjee, M. (2018, Sep/Oct). Children's Recognition of Emotional Prosody in Spectrally Degraded Speech Is Predicted by Their Age and Cognitive Status. *Ear Hear, 39*(5), 874-880. https://doi.org/10.1097/aud.0000000000000546

Tourville, J. A., Reilly, K. J., & Guenther, F. H. (2008, 2008/02/01/). Neural mechanisms underlying auditory feedback control of speech. *NeuroImage, 39*(3), 1429-1443. https://doi.org/https://doi.org/10.1016/j.neuroimage.2007.09.054

Tsodyks, M., & Gilbert, C. (2004, Oct 14). Neural networks and perceptual learning. *Nature, 431*(7010), 775-781. https://doi.org/10.1038/nature03013

Utianski, R. L., Duffy, J. R., Clark, H. M., Strand, E. A., Botha, H., Schwarz, C. G., Machulda, M. M., Senjem, M. L., Spychalla, A. J., Jack, C. R., Petersen, R. C., Lowe, V. J., Whitwell, J. L., & Josephs, K. A. (2018, 2018/09/01/). Prosodic and phonetic subtypes of primary progressive apraxia of speech. *Brain and Language, 184*, 54-65. https://doi.org/https://doi.org/10.1016/j.bandl.2018.06.004

Utoomprurkporn, N., Hardy, C. J. D., Stott, J., Costafreda, S. G., Warren, J., & Bamiou, D. E. (2020). “The Dichotic Digit Test” as an Index Indicator for Hearing Problem in Dementia: Systematic Review and Meta-Analysis. *Journal of the American Academy of Audiology, 31*(09), 646-655. https://doi.org/10.1055/s-0040-1718700

Utoomprurkporn, N., Stott, J., Costafreda, S. G., & Bamiou, D. E. (2021, Jun 20). Lack of Association between Audiogram and Hearing Disability Measures in Mild Cognitive Impairment and Dementia: What Audiogram Does Not Tell You. *Healthcare (Basel), 9*(6). https://doi.org/10.3390/healthcare9060769

Van't Hooft, J. J., Pijnenburg, Y. A. L., Sikkes, S. A. M., Scheltens, P., Spikman, J. M., Jaschke, A. C., Warren, J. D., & Tijms, B. M. (2021, Mar). Frontotemporal dementia, music perception and social cognition share neurobiological circuits: A meta-analysis. *Brain Cogn, 148*, 105660. https://doi.org/10.1016/j.bandc.2020.105660

van Dyck, C. H., Swanson, C. J., Aisen, P., Bateman, R. J., Chen, C., Gee, M., Kanekiyo, M., Li, D., Reyderman, L., Cohen, S., Froelich, L., Katayama, S., Sabbagh, M., Vellas, B., Watson, D., Dhadda, S., Irizarry, M., Kramer, L. D., & Iwatsubo, T. (2023, Jan 5). Lecanemab in Early Alzheimer's Disease. *N Engl J Med, 388*(1), 9-21. https://doi.org/10.1056/NEJMoa2212948

Vidal, M. M., Lousada, M., & Vigário, M. (2020). Music effects on phonological awareness development in 3-year-old children. *Applied Psycholinguistics, 41*(2), 299-318. https://doi.org/10.1017/S0142716419000535

Wang, C., Wang, Z., Xie, B., Shi, X., Yang, P., Liu, L., Qu, T., Qin, Q., Xing, Y., Zhu, W., Teipel, S. J., Jia, J., Zhao, G., Li, L., & Tang, Y. (2022). Binaural processing deficit and cognitive impairment in Alzheimer's disease. *Alzheimer's & Dementia, 18*(6), 1085-1099. https://doi.org/https://doi.org/10.1002/alz.12464

Warren, J. D., Fletcher, P. D., & Golden, H. L. (2012, 2012/08/01). The paradox of syndromic diversity in Alzheimer disease. *Nature Reviews Neurology, 8*(8), 451-464. https://doi.org/10.1038/nrneurol.2012.135

Warren, J. D., Rohrer, J. D., & Rossor, M. N. (2013, Aug 6). Clinical review. Frontotemporal dementia. *Bmj, 347*, f4827. https://doi.org/10.1136/bmj.f4827

Warren, J. D., Rohrer, J. D., Schott, J. M., Fox, N. C., Hardy, J., & Rossor, M. N. (2013, Oct). Molecular nexopathies: a new paradigm of neurodegenerative disease. *Trends Neurosci, 36*(10), 561-569. https://doi.org/10.1016/j.tins.2013.06.007

Warren, J. D., Scott, S. K., Price, C. J., & Griffiths, T. D. (2006, 2006/07/01/). Human brain mechanisms for the early analysis of voices. *NeuroImage, 31*(3), 1389-1397. https://doi.org/https://doi.org/10.1016/j.neuroimage.2006.01.034

Warren, J. E., Wise, R. J., & Warren, J. D. (2005, Dec). Sounds do-able: auditory-motor transformations and the posterior temporal plane. *Trends Neurosci, 28*(12), 636-643. https://doi.org/10.1016/j.tins.2005.09.010

Warren, R. M. (1970, Jan 23). Perceptual restoration of missing speech sounds. *Science, 167*(3917), 392-393. https://doi.org/10.1126/science.167.3917.392

Warren, R. M., & Obusek, C. J. (1971, 1971/05/01). Speech perception and phonemic restorations. *Perception & Psychophysics, 9*(3), 358-362. https://doi.org/10.3758/BF03212667

Warren, R. M., & Sherman, G. L. (1974, 1974/01/01). Phonemic restorations based on subsequent context. *Perception & Psychophysics, 16*(1), 150-156. https://doi.org/10.3758/BF03203268

Watson, C. L., Possin, K., Allen, I. E., Hubbard, H. I., Meyer, M., Welch, A. E., Rabinovici, G. D., Rosen, H., Rankin, K. P., Miller, Z., Santos-Santos, M. A., Kramer, J. H., Miller, B. L., & Gorno-Tempini, M. L. (2018). Visuospatial Functioning in the Primary Progressive Aphasias. *Journal of the International Neuropsychological Society, 24*(3), 259-268. https://doi.org/10.1017/S1355617717000984

Weinstein, J., Koenig, P., Gunawardena, D., McMillan, C., Bonner, M., & Grossman, M. (2011, Feb). Preserved musical semantic memory in semantic dementia. *Arch Neurol, 68*(2), 248-250. https://doi.org/10.1001/archneurol.2010.364

Weintraub, S., Wicklund, A. H., & Salmon, D. P. (2012). The Neuropsychological Profile of Alzheimer's Disease. *Cold Spring Harbor perspective in medicine, 2*(4).

White, B. E., & Langdon, C. (2021, 2021/10/15/). The cortical organization of listening effort: New insight from functional near-infrared spectroscopy. *NeuroImage, 240*, 118324. https://doi.org/https://doi.org/10.1016/j.neuroimage.2021.118324

Wild, C. J., Yusuf, A., Wilson, D. E., Peelle, J. E., Davis, M. H., & Johnsrude, I. S. (2012). Effortful Listening: The Processing of Degraded Speech Depends Critically on Attention. *Journal of Neuroscience, 32*(40), 14010-14021. https://doi.org/10.1523/jneurosci.1528-12.2012

Wildgruber, D., Ackermann, H., Krieifelts, B., & Ethofer, T. (2006). Cerebral processing of linguistic and emotional prosody: fMRI studies. *Progress in Brain Research, 156*, 19. https://doi.org/10.1016/S0079-6123(06)56013-3

Wilson, D., & Sperber, D. (2002). Relevance Theory. In G. Ward & L. Horn (Eds.), *Handbook of Pragmatics*. Blackwell.

Wilson, S. M., Saygin, A. P., Sereno, M. I., & Iacoboni, M. (2004, Jul). Listening to speech activates motor areas involved in speech production. *Nat Neurosci, 7*(7), 701-702. https://doi.org/10.1038/nn1263

Wise, R. J. S., Greene, J., Büchel, C., & Scott, S. K. (1999, 1999/03/27/). Brain regions involved in articulation. *The Lancet, 353*(9158), 1057-1061. https://doi.org/https://doi.org/10.1016/S0140-6736(98)07491-1

Wolpe, N., Ingram, J. N., Tsvetanov, K. A., Geerligs, L., Kievit, R. A., Henson, R. N., Wolpert, D. M., & Rowe, J. B. (2016, Oct 3). Ageing increases reliance on sensorimotor prediction through structural and functional differences in frontostriatal circuits. *Nat Commun, 7*, 13034. https://doi.org/10.1038/ncomms13034

Wöstmann, M., Erb, J., Kreitewolf, J., & Obleser, J. (2021). Personality captures dissociations of subject versus objective hearing in noise. *Royal Society Open Science, 8*. https://doi.org/https://doi.org/10.1098/rsos.210881

Yates, A. J. (1963). Delayed auditory feedback. *Psychological Bulletin, 60*(3), 213-232. https://doi.org/https://doi.org/10.1037/h0044155

Yoo, J., & Bidelman, G. M. (2019, Jun). Linguistic, perceptual, and cognitive factors underlying musicians' benefits in noise-degraded speech perception. *Hear Res, 377*, 189-195. https://doi.org/10.1016/j.heares.2019.03.021

Zatorre, R. J., & Baum, S. R. (2012). Musical Melody and Speech Intonation: Singing a Different Tune? *PLOS Biology, 10*(7).

Zatorre, R. J., Belin, P., & Penhune, V. B. (2002, 2002/01/01/). Structure and function of auditory cortex: music and speech. *Trends in Cognitive Sciences, 6*(1), 37-46. https://doi.org/https://doi.org/10.1016/S1364-6613(00)01816-7

Zatorre, R. J., & Gandour, J. T. (2008). Neural specializations for speech and pitch: moving beyond the dichotomies. *Philosophical Transactions of the Royal Society B: Biological Sciences, 363*(1493), 1087-1104. https://doi.org/doi:10.1098/rstb.2007.2161

Zendel, B. R., & Alain, C. (2009, Aug). Concurrent sound segregation is enhanced in musicians. *J Cogn Neurosci, 21*(8), 1488-1498. https://doi.org/10.1162/jocn.2009.21140

Zhang, L., Fu, X., Luo, D., Xing, L., & Du, Y. (2020, Aug 11). Musical Experience Offsets Age-Related Decline in Understanding Speech-in-Noise: Type of Training Does Not Matter, Working Memory Is the Key. *Ear Hear, 42*(2), 258-270. https://doi.org/10.1097/aud.0000000000000921

Zheng, Z. Z., Vicente-Grabovetsky, A., MacDonald, E. N., Munhall, K. G., Cusack, R., & Johnsrude, I. S. (2013). Multivoxel Patterns Reveal Functionally Differentiated Networks Underlying Auditory Feedback Processing of Speech. *The Journal of Neuroscience, 33*(10), 4339-4348. https://doi.org/https://doi.org/10.1523/JNEUROSCI.6319-11.2013

Zhou, J., Greicius, M. D., Gennatas, E. D., Growdon, M. E., Jang, J. Y., Rabinovici, G. D., Kramer, J. H., Weiner, M., Miller, B. L., & Seeley, W. W. (2010, May). Divergent network connectivity changes in behavioural variant frontotemporal dementia and Alzheimer's disease. *Brain, 133*(Pt 5), 1352-1367. https://doi.org/10.1093/brain/awq075

Zhu, J., Chen, X., Chen, F., & Weiner, S. (2022). Individuals With Congenital Amusia Show Degraded Speech Perception but Preserved Statistical Learning for Tone Languages. *Journal of Speech, Language, and Hearing Research, 65*(1), 53-69. https://doi.org/https://doi.org/10.1044/2021_JSLHR-21-00383