GESTURE IN APHASIA: EFFECTS ON AUDITORY COMPREHENSION, SPEECH PRODUCTION AND COMMUNICATIVE EFFECTIVENESS

Isobel Angharad Chick

Department of Experimental Psychology
University College London

Submitted to University College London (UCL) in partial fulfilment of the requirements for the award of the degree of Doctor of Philosophy.

Primary supervisor: Professor Gabriella Vigliocco
Secondary supervisor: Professor Jeremy I. Skipper

Thesis submission date: 20th April 2025

Declaration

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

Abstract

Communication is multimodal. Humans dynamically use multiple channels seamlessly – speech, gesture, eye-gaze, facial expressions and more – to communicate. Yet research into disorders which disrupt communication, such as post-stroke aphasia, frequently focuses on just one of those channels, speech, to the exclusion of others. In this thesis, I investigate how people with post-stroke aphasia (PWA) understand, use and integrate gestures alongside speech in dyadic interactions, and I evaluate evidence for the functions gestures serve in aphasic communication. First, I explore whether gesture production and gesture observation facilitate speech comprehension and production in (semi-)naturalistic dyadic interactions between PWA and neurologically healthy interlocuters, and I identify cognitive markers that predict gesture (dis)benefit. Next, I investigate whether gestures produced by PWA contribute to communicative success from the perspective of neurologically healthy adults. Finally, I examine whether gestures can be used within a dynamic, dyadic therapy context to facilitate lexical retrieval in a core group of PWA with significant naming impairment. I discuss these findings in relation to two influential models of speech-gesture integration: the Lexical Retrieval Hypothesis and the Information Packaging Hypothesis.

I demonstrate that individual differences in working memory/executive function, gesture recognition and (visual) attention predict gesture observation and production benefit for speech comprehension and production in PWA. Further, I demonstrate that gestures serve a vital compensatory role in successful communication in PWA by encoding and transmitting semantic content when speech production is impaired. Lastly, I show that an intensive, self-paced, online therapy programme incorporating iconic gesture observation effectively improves naming accuracy in PWA, at least when gestures are closely semantically-related to target words.

Impact statement

Aphasia, an acquired communication disorder which affects both understanding and use of language, is a potentially life-changing consequence of stroke (and other neurological conditions) with significant societal impacts. As the population ages, aphasia prevalence is expected to increase. Although communication is a multimodal phenomenon, it was not until relatively recently that researchers started investigating how production and comprehension of visual aspects of language, such as co-speech gestures, were affected by aphasia. Certain models of speech-gesture integration predict that speech comprehension and production become easier when we see or use gestures, through processes of cross-modal priming and enhanced semantic activation of lexical items. If this is indeed the case, gestures may offer a pathway to improving speech comprehension and production, at least for some PWA. The work I present in this thesis aims to clarify which PWA may benefit from observing or producing gestures for communicative success, and which may not. Crucially, it aims to do so in an ecologically valid manner, with experimental tasks embedded within dyadic interactions.

First, in Chapter 2 I show that within (semi-)naturalistic interactions, discrete sub-groups of PWA benefit from (a) observing and (b) producing gestures in a single-word picture naming and speech comprehension task. I demonstrate that those with poorer working memory/executive function benefit from *observing* gestures to facilitate speech production, while those with better working memory/executive function benefit from *producing* gestures to facilitate speech production. Second, I also demonstrate that observing gestures facilitates speech comprehension in some PWA, but only those with intact gesture recognition abilities who can effectively extract semantic information from gestures and integrate it with speech. Third, in Chapter 3 I show that (visual) attention capacity moderates the extent to which PWA are able to integrate information from gestures in connected speech contexts. Fourth, in Chapters 3 and 4 I show that when PWA use gestures alongside connected speech their speech content is *less* informative but their overall

communication is *more* informative, suggesting that PWA use gestures as a compensatory means primarily to convey information to their interlocuters. I also show that the communicative effectiveness of gestures produced by PWA is moderated by upper limb apraxia severity. Finally, in Chapter 5 I demonstrate that an intensive self-paced online therapy programme combining speech with gesture observation improves speech production in a group of PWA with significant naming deficits.

These findings have clear implications for clinical practice and for personalised rehabilitation approaches for PWA. They also shed light on current theoretical debates on speech-gesture integration, particularly in terms of how wider cognitive factors, including attention, working memory and limb apraxia, interact with speech and gestures in the process of communication.

Dedication

For everyone living with aphasia

Acknowledgements

First and foremost, my heartfelt thanks to my supervisor and mentor, Professor Gabriella Vigliocco, whose support and guidance has been of immeasurable importance to me over the past few years. I am deeply grateful for your patience, resilience, creativity, and for teaching me what good research should look like. Thank you also for introducing me to some fantastic restaurants!

I would also like to thank the members of my thesis committee for their support and advice: my second supervisor Professor Jeremy I. Skipper, Professor Rosemary Varley, Professor Jenny Crinion, Dr Laurel J. Buxbaum and Dr Anna Krason. I have been very fortunate to work alongside many wonderful collaborators over the course of this project to whom I am extremely grateful, particularly Dr Amy Lebkuecher and Professor H. Branch Coslett. An immense thank you also to Dr Lotte Meteyard, who first encouraged me to consider this path. I am extremely grateful also to my funder, Brain Research UK, who saw the potential and value in this work.

Thank you to the members of the Language and Cognition Lab, past and present, who have helped keep me sane and well supplied with tea, chocolate and pot plants: Claudia, Gal, Ed, Chris, Viktor, Francesco, Marine, Dora, Guoda, Michi, Elena, Reena, Ruohan, Jenna, and many more. A special mention in particular to Norina, who has been the most wonderful friend and colleague I could ever have asked for.

My family, Joe and Ben, have been hugely supportive throughout the past few years. Ben trialled my various experimental tasks and gave useful (and at times robust!) feedback and encouragement, and Joe has patiently borne with me all the way through. Thank you, both of you.

Finally, my most grateful thanks to all the people with aphasia and their loved ones who gave so generously of their time to take part in this project. It goes

without saying, but I could not have done it without you. This thesis is dedicated to you, and to everyone living with aphasia.

Statement of publications

Chapter 1 of this thesis is partially based on the following:

Chick, I., Garrard, P., Buxbaum, L.J., Vigliocco, G. (2023). Co-speech Gesture Production in Spoken Discourse Among Speakers with Acquired Language Disorders. In: Kong, A.PH. (eds) Spoken Discourse Impairments in the Neurogenic Populations. Springer, Cham.

Reproduced with permission.

I, Isobel Chick, declare that I was fully involved through the entire process of producing the book chapter cited above.

UCL Research Paper Declaration Form

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

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?**

Co-speech Gesture Production in Spoken Discourse Among Speakers with Acquired Language Disorders.

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

https://link.springer.com/chapter/10.1007/978-3-031-45190-4_10

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

In: Kong, A.PH. (eds) Spoken Discourse Impairments in the Neurogenic Populations.

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

Springer, Cham

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

2023

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

Chick, I., Garrard, P., Buxbaum, L.J., Vigliocco, G.

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

Yes

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

Yes

   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)

IC: Conceptualisation, Funding Acquisition, Preparation of original draft, Approval of final manuscript for submission. PG: Provided revisions for draft, Approval of final manuscript for submission. LJB: Provided revisions for draft, Approval of final manuscript for submission. GV: Conceptualisation, Funding acquisition, Supervision, Provided revisions for draft, Approval of final manuscript for submission.

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

Chapter 1

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

[redacted]

Date:

20/03/2025

Supervisor/ Senior Author (where appropriate)

[redacted]

Date

20/03/2025

Table of contents

Declaration 2
Abstract 3
Impact statement 4
Dedication 6
Acknowledgements 7
Statement of publications 9
Table of contents 12
List of figures 17
List of tables 20
Abbreviations 22
Introduction 23
Chapter 1: Literature Review 25
1.0 Unimodal language, multimodal language, and rehabilitation in aphasia 25
1.1 Unimodal communication (speech) deficits in aphasia 26
1.2 Towards an understanding of multimodal communication deficits in aphasia 32
1.3 Approaches to language rehabilitation 58
1.4 Implications and the way forward 71
Chapter 2: “Go Fish!” 73
2.0 Abstract 73
2.1 Introduction 74
2.1.1 Gestures (may) facilitate speech comprehension 74
2.1.2 Gestures (may) facilitate speech production 76
2.1.3 Decontextualised language environments 78
2.2 The present study 79
2.3 Study aim and hypotheses 80
2.4 Methods 81
2.4.1 Participant sample 81
2.4.2 Background test battery 85
2.4.3 Neuroimaging data 86
2.4.4 Materials 86
2.4.5 Procedure 87

2.5 Data Analysis 91
2.5.1 Gesture coding 91
2.5.2 Speech response transcription and error coding 92
2.5.3 Statistical modelling strategy 93
2.5.4 Comprehension accuracy 94
2.5.5 Effects of gesture observation on naming accuracy 95
2.5.6 Effects of gesture production on naming accuracy 96
2.6 Results 97
2.6.1 Gesture observation behavioural analyses 97
2.6.2 Gesture production behavioural analyses 106
2.6.3 Descriptive lesion analyses 114
2.7 Discussion 120
2.7.1 Theoretical implications 121
2.7.2 Clinical implications 124
2.7.3 Limitations 125
2.8 Conclusion 126
Chapter 3: “Heads Up!” 128
3.0 Abstract 128
3.1 Introduction 129
3.1.1 Gesture and connected speech production in PWA 129
3.2 The present study 132
3.3 Study aim and hypotheses 133
3.4 Methods 135
3.4.2 Participant sample 135
3.4.3 Set up 135
3.4.4 Materials 135
3.4.5 Procedure 136
3.4.6 Naming error transcription and error typing 138
3.5 Data Analysis 139
3.5.1 Naming accuracy 139
3.5.2 Lexical diversity (Moving Average Type-Token Ratio) 140
3.5.3 Mean Length of Utterance (MLU) 141
3.5.4 Concreteness 141
3.6 Results 142
3.6.1 Gesture observation behavioural results 142

3.6.2 Speech analyses 148
3.7 Discussion 152
3.7.1 Theoretical implications 153
3.7.2 Clinical implications 156
3.7.3 Limitations 157
3.8 Conclusion 158
Chapter 4: Communicative Effectiveness 159
4.0 Abstract 159
4.1 Introduction 160
4.2 Study aim and hypotheses 162
4.3 Methods 165
4.3.1 The present studies 165
4.3.2 Participant sample 165
4.3.4 Materials 166
4.3.5 Procedure 167
4.4 Data analysis 168
4.4.1 Pre-processing of cosine similarity data 169
4.4.2 Primary analyses 170
4.4.3 Secondary analyses 170
4.5 Results 171
4.5.1 Cosine similarity scores 171
4.5.2 Communicative effectiveness ratings 175
4.5.3 Secondary analyses 178
4.6 Discussion 183
4.6.1 Theoretical implications 186
4.6.2 Clinical implications 188
4.6.3 Limitations 188
4.7 Conclusion 190
Chapter 5: Therapy 191
5.0 Abstract 191
5.1 Introduction 192
5.1.1 Therapy delivery in aphasia 192
5.1.2 Gesture therapy for word retrieval in aphasia 193
5.2 Study aim and hypotheses 196
5.3 Methods 198

5.3.1 The present study 198
5.3.2 Participant sample 199
5.3.4 Gesture norming 201
5.3.5 Online therapy design 201
5.3.6 Procedure 202
5.3.7 Therapy 204
5.3.8 Post-therapy naming assessments 206
5.4 Data analysis 207
5.4.1 Power analysis 207
5.4.2 Primary analyses 207
5.4.3 Secondary analyses 209
5.5 Results 210
5.5.1 Predictions 1a and 1b: Intervention effect on naming accuracy post-therapy and maintenance of treatment effects 210
5.5 2 Predictions 2a and 2b: Gesture effect on naming accuracy post-therapy 211
5.5.3 Predictions 3a and 3b: Effects of working memory and gesture norm scores 214
5.5.4 Prediction 4: Generalisation of treatment effects to untreated items 216
5.5.5 Secondary analyses 217
5.6 Discussion 218
5.6.1 Theoretical implications 219
5.6.2 Clinical implications 220
5.6.3 Limitations 223
5.7 Conclusion 224
Chapter 6: General Discussion 225
6.1 Summary of goals and findings 225
6.2 Theoretical implications 227
6.2.1 Moderating factors for gesture benefit 227
6.2.2 Models of speech-gesture integration 231
6.3 Clinical implications 232
6.4 Limitations and future directions 235
6.5 Conclusion 238
References 240
Appendix 1: PWA Background Battery Assessment Scores 271

Appendix 2: Chapter 2 “Go Fish!” Gesture Norming Scores 275
Appendix 3: PWA Gesture and Speech Transcription Coding Manual 284
Appendix 4: Chapter 2 “Go Fish!” Power Analyses 311
Appendix 5: Chapter 3 “Heads Up!” Power Analysis 316
Appendix 6: Chapter 3 MATTR Python Code 321
Appendix 7: Chapter 5 Gesture Norming Scores (Therapy) 322
Appendix 8: Chapter 5 PWA Therapy Instruction Sheet 323
Appendix 9: Data Analysis Models 327

List of figures

Chapter 2

Fig. 1. Sample stimulus cards: diving (action), dripping (action), pyramid (object)	87
Fig. 2. Go Fish gameplay.	90
Fig. 3. Go Fish gameplay.	90
Fig. 4. Effect plot: mean comprehension scores by gesture condition.	98
Fig. 5. Comprehension accuracy: interaction between comprehension accuracy scores in each gesture condition and PWA gesture recognition score.	99
Fig. 6. Comprehension accuracy: interaction between comprehension accuracy scores in each gesture condition and target word frequency.	99
Fig. 7. Effect plot: mean naming accuracy scores by gesture condition.	101
Fig. 8. Naming accuracy: interaction between naming accuracy scores in each gesture condition and target word length.	102
Fig. 9. Naming accuracy: interaction between naming accuracy scores in each gesture condition and working memory/executive function score.	102
Fig. 10. Proportion of error types by gesture condition (all PWA).	103
Fig. 11. Proportion of error types by gesture condition (Moderate PWA).	104
Fig. 12. Boxplot: effect of PWA gesture observation condition on semantic similarity (cosine similarity) score of PWA naming errors.	105
Fig. 13. Boxplot: effect of PWA gesture observation condition on phonemic similarity (Levenshtein) score of PWA naming errors.	106
Fig. 14. Frequency count for each gesture type produced by PWA.	107
Fig. 15. Effect plot: mean naming accuracy scores by PWA gesture production condition.	109
Fig. 16. Naming accuracy: interaction between naming accuracy scores in each PWA gesture production condition and working memory/executive function score.	109
Fig. 17. Proportion of naming error types by gesture condition (all PWA).	111
Fig. 18. Proportion of naming error types by gesture condition (Moderate PWA).	112
Fig. 19. Boxplot: effect of PWA gesture production condition on semantic similarity (cosine similarity) score of PWA naming errors.	113
Fig. 20. Boxplot: effect of PWA gesture production condition on phonemic similarity (Levenshtein) score of PWA naming errors.	114
Fig. 21. Positive GOBI coverage maps.	116
Fig. 22. Negative GOBI coverage maps.	116
Fig. 23. GOBI Benefit – Deficit subtraction maps.	116
Fig. 24. GOBI Deficit – Benefit subtraction maps.	117
Fig. 25. Positive GPBI coverage maps.	118

Fig. 26. Negatve GPBI coverage maps.	119
Fig. 27. GPBI Benefit – Deficit subtraction maps.	119
Fig. 28. GPBI Deficit – Benefit subtraction maps.	119

Chapter 3

Fig. 1. “Heads Up!” setup. Researcher providing description.	138
Fig. 2. “Heads Up!” setup. “It goes boom!”	138
Fig. 3. Naming accuracy: interaction between naming accuracy scores in each gesture condition and PWA visual attention capacity score.	144
Fig. 4. Proportion of naming error types produced by all PWA.	145
Fig. 5. Proportion of naming error types produced by PWA with moderate aphasia.	146
Fig. 6. Boxplot: effect of PWA gesture observation condition on semantic similarity (cosine similarity) score of PWA naming errors.	147
Fig. 7. Boxplot: effect of PWA gesture observation condition on phonemic similarity (Levenshtein) score of PWA naming errors.	148
Fig. 8. Effect plot: PWA MLU by gesture condition.	151
Fig. 9. Effect plot: PWA mean utterance concreteness by gesture condition.	152

Chapter 4

Fig. 1. PWA describing the target word “devil” in the gesture permitted condition.	167
Fig. 2. Effect plot: Participant cosine similarity scores by PWA gesture production condition (audio-visual version).	173
Fig. 3. Boxplot: Distribution of participant cosine similarity scores by PWA upper limb apraxia severity (Gesture To Sight score) within the gesture present stimuli (audio-visual version).	174
Fig. 4. Effect plot: Participant cosine similarity scores by PWA gesture production condition (audio-only version).	175
Fig. 5. Effect plot: Participant communicative effectiveness ratings by PWA gesture production condition (audio-visual version).	175
Fig. 6. Boxplot: Distribution of participant communicative effectiveness ratings by PWA upper limb apraxia severity (Gesture To Sight score) within the gesture present stimuli (audio-visual version).	177
Fig. 7. Effect plot: Participant communicative effectiveness ratings by PWA gesture production condition (audio-only version).	178
Fig. 8. Boxplot: Distribution of participant cosine similarity scores for errors by PWA upper limb apraxia severity (Gesture To Sight score) within the gesture present stimuli (audio-visual version).	180
Fig. 9. Boxplot: Distribution of participant cosine similarity scores for errors by PWA upper limb apraxia severity (Gesture To Sight score) within the gesture present stimuli (audio-only version).	180
Fig. 10. Boxplot: Distribution of participant cosine similarity scores for errors by PWA gesture production condition (audio-only version).	181

Fig. 11. Mean participant reaction time for assigning communicative effectiveness ratings by gesture condition and study version.	182

Chapter 5

Fig. 1. Indicative timescales for intervention.	199
Fig. 2. Intervention design.	206
Fig. 3. Change in naming accuracy for treated items by participant at each testing point.	211
Fig. 4. Effect plot: Mean naming accuracy for treated items by gesture condition at each testing point.	213
Fig. 5. Naming accuracy: effect of Aphasia Quotient on naming accuracy for treated items at each testing point.	214
Fig. 6. Naming accuracy: effect of gesture norm score on naming accuracy for treated items at each testing point.	215
Fig. 7. Change in naming accuracy for untreated items by participant at each testing point.	217

List of tables

Chapter 1

Table 1. Summary of identified gesture types.	36
Table 2. Summary of aphasic sub-types and gesture production characteristics.	39

Chapter 2

Table 1. Participant demographic data.	83
Table 2. Background battery assessments.	85
Table 3. Auditory comprehension model results.	98
Table 4. Naming accuracy model results.	101
Table 5. Cosine similarity of naming errors model results.	104
Table 6. Phonemic similarity of naming errors model results.	106
Table 7. Frequency count for each gesture type.	107
Table 8. Naming accuracy model results.	108
Table 9. Proportion of naming error types by gesture condition (all PWA).	110
Table 10. Proportion of naming error types by gesture condition (Moderate PWA).	111
Table 11. Cosine similarity of naming errors model results.	112
Table 12. Phonological Levenshtein distance model results.	113
Table 13. Participant GOBI scores.	115
Table 14. Participant GPBI scores.	118

Chapter 3

Table 1. Naming accuracy model output.	143
Table 2. Naming accuracy model output (UK subset data).	143
Table 3. Naming error types analysis output.	145
Table 4. Cosine similarity of naming errors model output.	146
Table 5. Phonemic similarity of naming errors model output.	148
Table 6. MLU model output.	151
Table 7. Concreteness model output.	152

Chapter 4

Table 1. Audio-visual study results (cosine similarity).	173
Table 2. Audio-only study results (cosine similarity).	174
Table 3. Audio-visual study results (communicative effectiveness).	176
Table 4. Audio-only study results (communicative effectiveness).	178
Table 5. Audio-visual study results (cosine similarity scores of errors).	179

Table 6. Audio-only study results (cosine similarity scores of errors).	179
Table 7. ANOVA output.	182
Table 8. Tukey's HSD output.	182
Table 9. Tukey's HSD pairwise contrasts.	182

Chapter 5

Table 1. Participant demographics.	200
Table 2. Error categorisation.	203
Table 3. Examples of target words and matched untreated words.	204
Table 4. Paired t-test results for naming accuracy of treated items at 1-week and 1-month post-intervention.	211
Table 5. Naming accuracy at 1-week post-intervention.	212
Table 6. Naming accuracy at 1-month post-intervention.	212
Table 7. Effects of working memory/executive function and gesture norm score on naming accuracy at 1-week post-intervention.	214
Table 8. Effects of working memory/executive function and gesture norm score on naming accuracy at 1-month post-intervention.	214
Table 9. Paired t-test results for naming accuracy of untreated items at 1-week and 1-month post-intervention.	216
Table 10. Comparison of number of untreated items named correctly at 1-week and 1-month post-intervention.	216
Table 11. Cosine similarity of PWA responses at 1-week post-intervention.	217
Table 12. Cosine similarity of PWA responses at 1-month post-intervention.	217
Table 13. Phonemic similarity of PWA responses at 1-week post-intervention.	218
Table 14. Phonemic similarity of PWA responses at 1-mont post-intervention.	218

Abbreviations

AoS	Apraxia of Speech
AQ	Aphasia Quotient
ATL	Anterior temporal lobe
CCT	Camel and Cactus Test
CVA	Cerebrovascular accident, stroke
GOBI	Gesture observation benefit index
GPBI	Gesture production benefit index
IFG	Inferior frontal gyrus
KDT	Kissing and Dancing Test
LAG	Language-action game
LRH	Lexical Retrieval Hypothesis
MATTR	Moving-average type-token ratio
MFG	Middle frontal gyrus
MLU	Mean length of utterance
MNI	Montreal Neurological Institute
MRI	Magnetic resonance imaging
MRRI	Moss Rehabilitation Research Institute
MTG	Middle temporal gyrus
OCS	Oxford Cognitive Screen
PWA	Person with aphasia
ST	Scenario Test
STM	Short-term memory
STT	Synonymy Triplets Test
TT	Token Test
WAB-R	Western Aphasia Battery-Revised
WM/EF	Working memory and executive function

Introduction

Human communication is inherently multimodal and interactive. In conversation, we dynamically and seamlessly integrate and use cues from a range of channels – speech, gesture, eye-gaze, facial expressions, body position, tone of voice – to express ourselves, and we rapidly decode cues produced by our interlocuters to understand *their* meaning. Yet in the face of neurological damage, processing of communication input and output across channels may be disrupted. This is aphasia, a primary communication disorder.

In this dissertation I investigate how people with aphasia (PWA) understand and use gestures alongside speech in naturalistic dyadic interactions, and explore what the function(s) of gesture may be in aphasic communication. In Chapter 1, I review the literature on unimodal (speech) and multimodal (speech and gesture) production and comprehension. I discuss the neurobiological bases of multimodal communication, and I explore theories of speech-gesture integration and function (the Lexical Retrieval Hypothesis and Information Packaging Hypothesis) and review evidence for and against these. I examine aphasia therapy approaches that incorporate gesture, and I consider how ecological validity concerns have impacted clinical findings. Finally, I consider how individual differences in cognitive functioning moderate gesture effects in aphasia

Chapters 2-5 are experimental chapters; in Chapter 2 I report findings from a novel dyadic language-action game (LAG) study, “Go Fish!”, which aimed to explore the extent to which gesture observation and production facilitated speech comprehension and single-word naming in PWA, and whether gesture effects were predictable based on cognitive profiles. Chapter 3 extends the findings from Chapter 2, to consider gesture observation and production effects in a connected speech context. In Chapter 4 I take an alternative approach and consider gesture effects from the perspective of neurologically healthy interlocuters. Specifically, I ask whether gestures produced by PWA contribute to communicative effectiveness by conveying

meaning to interlocuters, and whether effects are moderated by factors such as limb apraxia. In Chapter 5, I trial and report findings from a novel speech and language therapy intervention for naming based on combing gesture observation with speech production. In Chapter 6, I draw together insights from Chapters 2-5 and evaluate how my findings shed light on theories of speech-gesture integration and the functions performed by gestures in aphasic communication. I also consider experimental limitations and propose directions for future study.

Chapter 1: Literature Review

1.0 Unimodal language, multimodal language, and rehabilitation in aphasia

Aphasia is an acquired communication disorder characterised by disruption in receptive and expressive language processing, which is commonly caused by cerebrovascular accident (CVA, or stroke), usually to the left cerebral hemisphere. Approximately one third of people who survive a stroke experience aphasia (to a greater or lesser extent) afterwards; in around 30% of cases, aphasic symptoms are lifelong. In the UK annual aphasia incidence is estimated at around 50,600 people, with an estimated prevalence of 367,000 people (Brady et al., 2022). In the first two weeks after stroke (i.e. the acute phase), language recovery tends to be rapid, particularly within the first 48 hours, with speed of recovery slowing markedly after 15 days. Recovery in the acute phase is thought to occur primarily due to two mechanisms: reperfusion of damaged neural tissue resulting in return of functionality, and neural plasticity (i.e. neural reorganisation and restoration of damaged neural connections). In the second phase (the sub-acute phase), language recovery proceeds less rapidly with ongoing synaptic remodelling and axonal sprouting occurring alongside recovery from distal diaschisis (Marsh & Hillis, 2006). Improvement in functioning can continue, albeit more slowly, well into the chronic phase of recovery (i.e. more than 12 months post-stroke) with progressive improvements in speech production, speech comprehension, word repetition and naming observed in some people with aphasia (PWA) up to 25 years post-stroke (Smania et al., 2010).

Although stroke causes well-known disruptions to speech comprehension and production, communication is multimodal. Interlocuters flexibly use and comprehend information from a variety of communicative channels, including gesture, facial expressions and eye-gaze, which may vary depending on the needs of the audience (Levinson & Holler, 2014). Yet the majority of research into communication impairment after stroke has focused primarily on a single modality, speech, and multimodal communication impairments remain

significantly understudied. In this chapter, I first review how production and comprehension of (single) spoken words, which I term unimodal communication, is affected in aphasia (Section 1.1). In Section 1.2, I turn to multimodal communication (in this thesis referring to speech and co-speech gestures specifically) and review how co-speech gesture production is affected by aphasia and whether gesture production/observation facilitates speech production. I then turn to comprehension and review models and evidence for whether gesture observation facilitates speech comprehension in PWA. Finally, in Section 1.3 I review therapeutic approaches for speech production and comprehension for PWA that utilise (or constrain) gesture production/observation and consider factors which may moderate gesture effects on speech production and comprehension.

1.1 Unimodal communication (speech) deficits in aphasia

1.1.1 Anomia

Although the effects of aphasia vary considerably between individuals, anomia, or word-finding difficulties, is an almost universal symptom in the acute phase of aphasia and tends to persist (to some degree) in at least 30-40% of PWA even when other areas of language impairment have improved (Bonilha et al., 2016). Anomia is not restricted to aphasia following stroke; it may also occur in people with many other neurological conditions including dementia (Henry et al., 2008), Multiple Sclerosis (Cordova Luna, 2021), Huntington's (Clarke et al., 2023), Parkinson's (Kristensson et al., 2024) and epilepsy (Trebuchon et al., 2018).

Word-retrieval has been hypothesised in some models (though not all, see for example Levelt, 1989) to be an interactive, multi-stage process with feedback and feedforward between stages. For example, Dell et al. proposed that the lexical network consists of three levels: the semantic level (conceptualisation), where semantic features are selected; the lexical level, where semantic features are mapped to words; and the phonemic level, where phonological features are selected. Activation spreads between all

three levels until a state is reached where the word can be articulated. The process is bidirectional, with spreading activation occurring up and down each level, although the early stages of the process are thought to be dominated more by semantic selection and later stages by phonemic selection (Dell et al., 1999).

According to the Dell Model, anomia may occur when activation between different levels of word representation is disrupted or insufficient during the process of retrieval. As a result, anomia is highly variable in presentation and may present differently depending on the specific loci of impairment(s). There are two main naming error types produced by PWA, both of which may be further sub-divided: omission errors and commission errors. Omission errors occur when a PWA is unable to retrieve a word at all (complete omission), or when an explanation/description of the target word is produced instead (circumlocution error). Omission errors are assumed to arise due to loss of semantic representations altogether, or to disconnections between semantic representations and lexical items. Omissions are associated with lesions to a wide range of neural regions, including the left middle frontal gyrus, pars triangularis, middle temporal gyrus (MTG) and the temporal pole (McKinnon et al., 2018; Mirman et al., 2013).

In contrast, commission errors occur when a word is produced, but either the entire word, or part of it, is incorrect. Commission errors include semantic errors (i.e. meaning errors, producing “dog” instead of “cat”), phonemic errors (i.e. sound errors, producing “tat” [tæt] instead of “cat” [kæt]), mixed errors (i.e. the produced word has both semantic and phonemic elements in common with the target word, producing “rat” instead of “cat”), and unrelated errors (i.e. the produced word has no discernible semantic or phonemic connection to the target word). Semantic errors are thought to occur when semantic representations are degraded and have been associated with lesions to anterior and mid-temporal regions extending into parieto-occipital regions, or to anterior-posterior disconnection (Mirman et al., 2013). Phonemic errors are hypothesised to occur primarily due to impaired access to phonological forms or errors in phonological assembly, and have been

linked to damage to the left precentral gyrus, premotor cortex and supramarginal gyrus (Halai et al., 2018). Mixed errors are thought to occur due to breakdowns in cascading activation or connectivity between the semantic level and phonological forms of words (Dell et al., 2007). Unrelated errors, which are less common than other error types, are somewhat understudied but may occur due to damage in multiple regions spanning the anterior to posterior temporal cortex (i.e. the ventral stream of language processing) (Meier et al., 2021). Overall, omission and unrelated errors may be generally associated with more severe anomia; semantic, phonemic and mixed errors with less severe anomia (Schwartz & Brecher, 2000).

In terms of recovery, anomia severity often improves over time, particularly within the first few months' post-onset, although multiple factors moderate degree of improvement. In a longitudinal single case study of "HH", who initially presented with severe anomia, Raymer et al. (2000) found steady improvements in naming accuracy for single words and a corresponding reduction in the frequency of omission and unrelated errors up to six months' post-stroke, although improvements plateaued between six and 15 months post-stroke. Notably, HH did not engage in speech and language therapy. In contrast, in a study of 33 PWA who all received the same anomia treatment from six months post-stroke onwards (comprising at least 10 sessions of picture naming therapy with progressive cueing), Lambon Ralph et al. (2010) found ongoing and significant improvements in naming accuracy both during the course of the intervention and at follow-up (5 weeks after the end of therapy), although degree of gain was moderated by cognitive impairment. This suggests that while improvements in anomia severity may occur spontaneously without targeted intervention especially during the acute and sub-acute phases of stroke recovery, therapy may extend gain into the chronic period, although as Lambon Ralph et al. did not carry out longitudinal follow up (i.e. beyond 12 months' post-stroke) effects cannot be assumed to be maintained.

1.1.2 Speech comprehension

Although speech comprehension deficits in aphasia occur at both single word and connected speech levels, within this section I focus on single word comprehension only. Speech comprehension requires transformation of the acoustic waveform into discrete units (e.g. phonemes, syllables, words) which are then rapidly mapped to meaning(s). However, the speech signal may be distorted (e.g. by background noise), and there is no direct connection between speech sounds and phonological categories (i.e. the same phoneme may be produced differently by speakers with different accents, or even by the same speaker in different contexts) (Eisner & McQueen, 2018).

Speech comprehension is hypothesised to involve three main processing steps: auditory pre-processing, where the acoustic waveform is extracted from other noise; pre-lexical processing, where the extracted acoustic cues are categorised into units (e.g. phonemes, syllables); and lexical processing, where the categorised units are combined and mapped to semantic representations in the lexicon (Eisner & McQueen, 2018). Speech comprehension is assumed to be closely linked to speech production (Arjmandi & Behroozmand, 2024), although this idea is complicated by the fact that in PWA there may be dissociation in the severity of impairment in each area. For example, in a single case study Stasenko et al. (2015) found that a PWA with a severe naming impairment was nonetheless able to accurately distinguish between minimal pairs that differed by a single phoneme (e.g. AGA and ADA). Indeed, it has long been recognised that while speech comprehension and production deficits may co-occur in PWA, they may also dissociate and therefore are commonly assessed separately by clinicians; the classical classification of aphasia subtypes, which I discuss in more detail in Section 1.2, proposes a major dissociation between impairments in speech production and comprehension (see Berthier, 1995, for in-depth discussion).

As with the Dell Model of speech production described above, some models of (single word) speech comprehension assume that the process is interactive and bidirectional, with spreading activation between levels representing phonological features (e.g. voicing), phonemes, lexical items and semantic representations, until sufficient activation is reached and the listener comprehends the word. For example, the TRACE II model proposes that auditory and visual features (e.g. mouth movements) of a word are processed simultaneously and interact at all levels of word perception and comprehension bidirectionally, which allows for the listener to cope with distorted speech input and accurately map the word to its meaning in the lexicon (McClelland & Elman, 1986). Serial models, by contrast, assume that word comprehension is a step-by-step bottom-up process without inter-level interactions. One of the earliest serial models of speech comprehension, the Cohort Model, proposes that listeners activate multiple possible words in the lexicon (i.e. the 'cohort') as soon as acoustic speech input is detected, which are then sequentially narrowed down as additional acoustic information is added. Words within the activated cohort which mismatch the incoming acoustic input are gradually eliminated until the correct lexical item is identified and mapped to its meaning in the lexicon (Marslen-Wilson & Welsh, 1978). Importantly, both serial and interactive models of comprehension may explain (aspects of) impairments seen in PWA.

In aphasia, phoneme discrimination, phonological sequencing and semantic mapping impairments have all been identified as possible causes of comprehension breakdown (and indeed all may occur in the same individual) (Munro & Siyambalapitiya, 2017). In a study of 11 PWA and age-matched healthy adult controls, Mirman et al. (2011) found that PWA with primarily phonological deficits in speech production demonstrated strong rhyme competition effects in single word comprehension tasks (e.g. greater difficulty comprehending 'carrot' vs 'parrot'), whereas comprehension accuracy for semantic competitors did not vary from performance by neurologically healthy controls. The authors suggested that either difficulty discriminating between minimal pairs or difficulty suppressing phonemic competitors yielded these effects, while mapping between lexical items and semantic

representations was (assumed to be) unaffected. According to the Cohort Model, this effect could be due to impairments in early auditory pre-processing resulting in an inaccurate initial cohort being selected; according to the TRACE II model, this effect could be due to failure to suppress phonemic competitors throughout the process of comprehension.

In terms of identifying neural substrates of comprehension and lesion effects in PWA, Lwi et al. (2021) used univariate and multivariate lesion-symptom mapping to identify loci of comprehension impairments in 168 PWA with chronic aphasia; the authors found that single word comprehension deficits were associated with lesions in the left middle to posterior MTG, which has been shown to be a crucial hub for mapping lexical items to semantic representations (Binder, 2002). Similar results from lesion-symptom mapping analyses were also found by Fridriksson et al. (2018) and Turken & Dronkers (2011), although other studies (including studies of people with progressive comprehension deficits such as Primary Progressive Aphasia) have found that anterior temporal lesions are more closely associated with single word comprehension deficit. For example, in a study of 72 people with Primary Progressive Aphasia, single word comprehension deficits were (only) found to be associated with lesions to the anterior temporal pole and anterior temporal cortex; MTG and Wernicke's area lesions, by contrast, were associated with comprehension impairments at the level of connected speech (Mesulam et al., 2015). Indeed, it is likely that lesions to both areas contribute to comprehension impairments in PWA; lesions to the MTG may be more associated with deficits in accurately mapping the word to its meaning(s), while lesions to the ATL may be more associated with loss of semantic representations altogether.

1.1.3 Interim summary

In summary, aphasia has wide-ranging and long-lasting effects on both speech production and comprehension. Interactive models of word production such as the Dell Model account for speech errors made by PWA by hypothesising disconnection or insufficient activation between semantic,

phonemic and lexical levels of representations in the lexicon, or by degradation or loss of semantic, phonemic and/or lexical representations altogether (Dell et al., 2007). Serial and interactive models of (single word) comprehension may also account for patterns of comprehension disruption seen in PWA; for example, either due to initial perceptual deficits resulting in incorrect cohort selection (i.e. the Cohort Model), or by failure to supress lexical competitors effectively during the ongoing process of word comprehension (e.g. the TRACE II model). Comprehension deficits may also be caused by insufficient activation or loss of semantic representations in the lexicon (Mesulam et al., 2015). Speech production and comprehension deficits in PWA may be caused be lesions in a range of neural regions, although left temporal (anterior temporal pole, ATL, MTG) and inferior frontal regions (middle and IFG) are consistently implicated across multiple studies; importantly, research into (dis)connection between brain regions after stroke also suggests that intact connectivity between white matter tracts (particularly in the MTG and ATL) is an important predictor for language recovery (Bonilha et al. 2016; Roth et al., 2022). Spontaneous improvements in both anomia and speech comprehension deficits are observed in the early stages after stroke, although speed of recovery may plateau after 6-12 months (Raymer et al., 2000; Lambon Ralph et al. 2010). Beyond this point effective language therapy may extend and maintain speech production and comprehension gains, although precisely what comprises effective therapy is not well understood and factors mediating recovery trajectories are unclear; I return to this point in more detail in Section 1.3.

1.2 Towards an understanding of multimodal communication deficits in aphasia

Although co-speech gestures (i.e. gestures that occur alongside speech) are ubiquitous in communication, changes to how PWA produce and comprehend gestures remains an understudied area, particularly in terms of how and whether gestures contribute to language recovery by facilitating (or not) speech production and comprehension. In this section I first discuss how gestures are classified and review how gesture production may be affected

by stroke and progressive neurological conditions (Alzheimer's Disease, Primary Progressive Aphasia and Parkinson's). In Section 1.2.4 I then consider two key models describing how speech and gestures interact in the process of speech production, the Lexical Retrieval Hypothesis and the Information Packaging Hypothesis. In Section 1.2.5 I move onto multimodal comprehension and consider two further models, accounting for how gesture observation may facilitate speech comprehension, the Integrated-Systems Hypothesis and the Compensation Account.

1.2.1 Co-speech gesture classification

Although multiple gesture classification systems exist, one of the earliest and most influential attempts at distinguishing co-speech gestures from other communicative hand and body movements was proposed by Kendon (Kendon, 1980) and elaborated by McNeill (McNeill, 1992), who suggested that such movements comprise a 'continuum'. This ranges from gesticulations at one end – which are interactional in nature and require the presence of speech to be understood – to sign languages, which are full linguistic systems independent of speech. Co-speech gestures, which Kendon originally termed 'language-like gestures,' fall somewhere in the middle: they carry semantic content and are fully integrated into utterances, but like gesticulations their interpretation relies upon the presence of speech. Co-speech gestures differ from *emblems*, which express a conventionalised meaning within a community (such as the 'OK' or thumbs up symbols), and full *pantomimes* (for example, imitating using a hammer), both of which do not require speech to be understood.

While Kendon's Continuum is useful in providing a global overview of manual communication, it should be noted that there is often no clear boundary between different gesture categories in everyday communication: a speaker may move their index and middle finger back and forth while talking about cutting a piece of paper (i.e. a language-like gesture), and may then use the same gesture when mimicking the movement of the blades of scissors without speaking (a pantomime gesture).

McNeill subsequently divided Kendon's 'language-like' gesture category into four sub-types, termed 'speech accompanying' (or co-speech) gestures. These subtypes are iconic, metaphoric, deictic, and beat gestures (McNeill, 1992), and have largely been studied in typical populations. *Iconic gestures* represent a concrete referent imagistically by manner of movement or shape, such as depicting "ball" by moving the hands in a circle to mimic its shape or showing the swinging movement of a hammer. *Metaphoric gestures* typically represent abstract concepts, such as moving the index finger upwards while talking about the concept of inflation. Other metaphoric gestures include *number* gestures, which indicate a number on the fingers, and *time* gestures, which use space to differentiate between the past and future, such as using the front of the body to indicate the future and the back to indicate the past (Sekine & Rose, 2013). Iconic and metaphoric gestures may be depicted from a character viewpoint, where the speaker depicts an event as though performing it themselves (such as moving the arms through the air as though swimming), or from an observer viewpoint, where the speaker depicts an event as though seeing it from afar (e.g., moving the fingers back and forth to mimic a person running) (McNeill, 1992; Sekine & Rose, 2013).

In neurologically healthy adults, production of iconic gestures alongside speech facilitates speech comprehension in listeners, and this effect persists across the lifespan (Dargue et al., 2019). In connected speech, iconic and metaphoric gestures tend to precede the words they refer to, allowing for prediction of the following word (Zhang et al., 2021). Iconic and metaphoric gestures may also support speakers in word production; for example, there is evidence that neurologically healthy adults learn novel words faster when they are paired with meaningful gestures, and gestures may also support more effective schematising of information, particularly spatial information (Hupp & Gingras, 2016; Kita et al., 2017).

Deictic gestures are typically points (which includes both points produced with the index finger and points produced using the thumb, the whole hand, pointing using an object etc.) and may refer to concrete visible referents, such as pointing to a house while talking about it or pointing to the self;

concrete displaced or absent referents, such as pointing towards the door to indicate where someone has gone; or abstract referents, such as pointing backwards to refer to some time period in the past. Deictic gestures are used to disambiguate speech and to direct the listener's attention to a specific referent (Cochet & Vauclair, 2014).

Beats are small, often repetitive up-and-down hand movements that follow the rhythm of speech. Unlike the former three categories of gesture, beats have traditionally been assumed to carry no semantic content, however, more recent studies have suggested that beats may carry spatial information; Yap et al., for example, found that the upwards and downwards timing of beats correlated with spatial semantics in some utterances, such as saying "The ship *sank*" with a time-locked downward beat, or "The balloon went *up"* with a upward beat (Yap & Casasanto, 2018). The communicative purpose of beats is unclear: they may be used for emphasis or to introduce new information (McNeill, 1992), as visual markers of lexical stress (Bosker & Peeters, 2021), or to manage the flow of discourse (Bavelas et al., 1992). In people with acquired language impairments, there is some evidence that beats function to regulate the pace of speech and generate rhythm, similar to a metronome (Ferré, 2021; Leonard & Cummins, 2011).

In the context of co-speech gesture production in people with acquired language disorders specifically, it has been argued that additional gesture types exist, used either to compensate for impaired speech production or to facilitate word retrieval (de Beer et al., 2019). A prime example of this is *letter gestures*, where the speaker traces single letters or strings of letters in the air or on a flat surface and then 'reads' the traced word aloud. Indeed, de Beer et al. found that letter gestures were exclusively used by people with post-stroke language impairment (de Beer et al., 2019). A second example is *homophone gestures*, in which the speaker produces a gesture indicating a homophone of the word they want to say, for example by indicating their wedding ring when trying to retrieve the word "ring", as in "I need to ring [telephone] the doctor." (Cocks et al., 2011)

In total, at least 13 different co-speech gesture types have been identified: referential; concrete deictic; points to self; iconic observer viewpoint; iconic character viewpoint; pantomime; metaphoric; time; beats; emblems; letter gestures; number gestures; and homophone gestures. It should be noted, however, that there is little agreement on how reliably these categories can be identified and coded, and researchers vary in how they classify gestures.

Table 1. *Summary of identified gesture types.* Adapted from: Sekine K, Rose ML. The relationship of aphasia type and gesture production in people with aphasia. Am J Speech Lang Pathol. 2013 Nov 1;22(4):662–72.

Gesture type	Description
Referential	Indicates an absent/displaced concrete referent (Gullberg, 2006)
Concrete deictic	Indicates a present concrete referent (McNeill, 1992)
Points to self	Indicates the speaker, typically with a point (Sekine & Rose, 2013)
Iconic observer viewpoint	Speaker depicts an event as though observing it from a distance, e.g. using fingers to indicate someone running from left to right (McNeill, 1992; Sekine & Rose, 2013)
Iconic character viewpoint	Speaker depicts an event as though they are performing it themself, such as moving arms as though swimming (McNeill, 1992; Sekine & Rose, 2013)
Pantomime	Use of the hand/body to enact a part in a narrative; may not require speech to be understood (McNeill, 1992)
Metaphoric	Represents a (typically abstract) referent imagistically (McNeill, 1992)
Time	Uses space to differentiate the past and future (Sekine & Rose, 2013)
Beats	Small repetitive movements, may carry spatial content (Ferré, 2021; Leonard & Cummins, 2011; McNeill, 1992)
Emblems	Conventionalised meaning within a community or language group; does not require speech to be understood, e.g. thumbs up or ‘OK’ symbols (Kendon, 1980)
Letter gestures	Traces individual letters or strings (Cicone et al., 1979)
Number gestures	Indicates a number with the fingers (Cicone et al., 1979)
Homophone gestures	Indicates a spoken word that is a homophone to the gestured meaning, e.g. pointing to a wedding ring to

	indicate a “ring” on the telephone (Cocks et al., 2011).

1.2.2 Co-speech gesture production in stroke

1.2.2.1 Post-stroke aphasia

The majority of research into co-speech gesture production in neurogenic populations is drawn from PWA. As described in Section 1.1, severity of language dysfunction in aphasia often involves both expressive and receptive language abilities (Thye & Mirman, 2018). The classical model categorises aphasia into sub-types depending on the degree of language impairment across three dimensions: fluency of speech, auditory comprehension, and repetition (Geschwind, 1970; Goodglass & Kaplan, 1972). This yields eight aphasic variants: anomic, conduction, transcortical motor, Broca’s, transcortical sensory, Wernicke’s, transcortical mixed (also known as isolation aphasia) and global aphasia. The boundaries between variants are not clear cut and indeed it is estimated that up to a quarter of PWA may present with an aphasia type that is unclassifiable, although the classical model remains the most commonly-used system in English speaking countries (Kasselimis et al., 2017).

Most research into co-speech gesture production in PWA is drawn from studies of individuals with (relatively) unimpaired auditory comprehension (i.e. those with anomic, conduction, transcortical motor and Broca’s aphasias under the classical model), or with Wernicke’s aphasia. Research into gesture use in PWA with global, isolation/transcortical mixed and transcortical sensory aphasias is sparse.

Evidence suggests that gestures produced by PWA may vary as a function of aphasia sub-type. In a study of narrative production in PWA with non-fluent Broca’s and fluent Wernicke’s aphasias, Carlomagno & Cristilli found that just as speaking rate differed significantly between groups with Broca’s aphasics producing fewer words per minute, gesture rate also varied, although in the opposite direction: Broca’s aphasics produced more gestures per word than

Wernicke's aphasics (Carlomagno & Cristilli, 2006). However, contrasting results were found by Pritchard et al., who also included PWA with anomic and conduction aphasias as well as Wernicke's and Broca's in their sample, with no difference in gesture rate found between groups (Pritchard et al., 2015). The reason for these differing findings is unclear, although as both studies included very small numbers of Broca's aphasic participants (five and three, respectively), further research is required to determine if there is indeed a significant difference in gesture rate between fluent and non-fluent aphasic groups.

Beyond gesture rate, there is some evidence that the relative frequency of gesture types used by PWA differs both from neurologically healthy adults and indeed within people with different aphasic variants. PWA with preserved auditory comprehension and semantic processing (i.e. anomic, conduction, transcortical motor and Broca's aphasias) have been found to produce more iconic and pantomime gestures than do PWA with other aphasic sub-types. By contrast, PWA with Wernicke's aphasia produce relatively more referential, metaphoric, deictic, and beat gestures. For example, in an early study of gesture production in unstructured discourse in Wernicke's and Broca's aphasics, le May et al. found that PWA with Wernicke's aphasia used significantly more metaphoric gestures (referred to as 'kinetographs') in unstructured conversation and picture description tasks than Broca's aphasics, who used significantly more iconic gestures (referred to as 'ideographs') (le May et al., 1988). Similar results were found by Sekine & Rose in a study of 46 PWA in a story re-tell task, with Wernicke's aphasics found to use significantly more metaphoric gestures than PWA with other aphasic sub-types, in particular PWA with Broca's aphasia who used more iconic gestures (Sekine & Rose, 2013). Wernicke's aphasics have also been found to produce more gestures per minute (though not necessarily per word) than other aphasic groups, but their gestures have been described as vague and difficult for interlocuters to interpret, mirroring characteristics of their speech output (le May et al., 1988; Sekine et al., 2013; Sekine & Rose, 2013).

Iconic gesture use in particular has been found to be more common in PWA with anomic and Broca's aphasias during episodes of word-finding difficulty than during fluent connected speech. This could be because gestures may have a facilitatory effect on lexical retrieval (which I discuss in more detail below), or as a compensatory measure to replace speech (Behrmann & Penn, 1984). Akhavan et al. demonstrated that in PWA with anomic and Broca's aphasias of varying severity, iconic gesture use was strongly associated with successful word-retrieval in a story retell task: 88% of episodes of successful word-retrieval were associated with production of an iconic gesture (Akhavan et al., 2017). However, the vast majority of word-finding episodes in the sample remained unresolved, suggesting that even if iconic gestures were used, they were not always facilitatory of lexical retrieval and were sometimes used to replace speech. Indeed, this alternative role for gesture use is supported by evidence that PWA with Broca's and transcortical motor aphasias use co-speech gestures significantly more frequently alongside production of non-words and word truncations, indicating that in at least some cases, gesture use is a compensatory strategy for impaired speech production (Ferré, 2021).

The broad gesture production patterns for different aphasic variants described in the previous paragraphs are summarised in the table below.

Table 2. *Summary of aphasic sub-types and gesture production characteristics.*

Aphasia sub-type	Key features of gesture production
Anomic	- Gesture production profile broadly similar to neurologically healthy adults but may produce more iconic gestures and beat gestures during episodes of word-finding difficulty (Akhavan et al., 2017; Ferré, 2021; Sekine & Rose, 2013).
Broca's; Conduction	- Produce significantly more iconic gestures than neurologically healthy adults and people with fluent aphasia (Carlomagno & Cristilli, 2006; le May et al., 1988). - Produce relatively more iconic, pantomime and deictic gestures than other gesture types (Sekine et al., 2013; Sekine & Rose, 2013).

	- May produce more iconic gestures and beat gestures during episodes of word-finding difficulty (Akhavan et al., 2017; Ferré, 2021). - May produce gestures slowly and haltingly in line with speech production (McNeill D & Pedelty LL, 1995).
Global; transcortical mixed/isolation	- Limited data available, may use emblems more frequently than other gesture types (de Ruiter, 2000).
Transcortical motor	- Gesture production profile broadly similar to neurologically healthy adults (Sekine & Rose, 2013). - Produce a high number of semantically rich gestures (Sekine & Rose, 2013). - May use beat gestures to get the pace of speech and generate rhythm, particularly during episodes of word-finding difficulty (Ferré, 2021).
Transcortical sensory	- Limited data available.
Wernicke's	- Use a more restricted range of gesture types than other aphasic groups (Sekine & Rose, 2013). - Gestures may be vague and difficult for interlocuters to interpret (Sekine & Rose, 2013). - Use relatively more metaphoric, referential, deictic, and beat gestures and relatively fewer iconic gestures (le May et al., 1988; Sekine & Rose, 2013). - Produce fewer iconic gestures per word than PWA with Broca's aphasia (Carlomagno & Cristilli, 2006). - May produce gestures relatively rapidly in line with speech production (McNeill D & Pedelty LL, 1995).

Unsurprisingly, because the speech modality is typically impaired in aphasia, evidence suggests that PWA transmit relatively more information via the gesture channel compared with neurologically healthy adults. Rose et al. demonstrated that listener comprehension was greater when PWA with conduction, Broca's, transcortical motor and anomic aphasias combined speech and gesture to transmit messages, compared to when messages were transmitted via speech or gesture alone (Rose et al., 2017). For PWA, gestures may carry more of the intended meaning relative to speech than they do for neurologically healthy adults; Pritchard et al. found that in a procedural discourse task, PWA with anomic, conduction, Wernicke's and Broca's aphasias relied on gestures to convey information absent from

speech to a significantly greater degree than neurologically healthy adults, presumably because the speech channel was less available to them, although it was not clear whether gesture usefulness varied between aphasic groups (Pritchard et al., 2015). In turn, interlocuters also rate gestures produced by PWA to be essential to the successful transmission of the message more frequently than they do for neurologically healthy adults (de Ruiter, 2000; de Ruiter & de Beer, 2013; Van Nispen et al., 2017). For instance, van Nispen et al. found that on average 22% of co-speech gestures (specifically emblems, deictics and iconic gestures) produced by PWA with conduction, transcortical motor, Wernicke's, anomic and Broca's aphasias of varying severity were deemed to be essential for successful communication compared with around 5% for neurologically healthy adults, although sample size for each sub-type was small (5-10 per group) (Van Nispen et al., 2017). Effective use of the gesture channel does however appear to be of considerable importance in successful communication for at least some PWA.

1.2.2.3 Right hemisphere damage (RHD)

As the language and praxic control centres in the brain are usually left lateralised, right hemisphere stroke does not generally cause a primary language or gesture production impairment, but typically results in prosodic and discourse processing deficits in at least 50% of patients (Côte et al., 2010). Deficits include monotone speech production, impaired turn-taking, atypical eye-gaze behaviours and poor comprehension of non-literal language and emotional prosody (Blake et al., 2010; Lehman Blake, 2018). Visuo-spatial processing deficits are also common (Clough & Duff, 2020). Compared with the literature available on co-speech gesture production in aphasia, far less is known about co-speech gesture production deficits following right hemisphere stroke.

Early case studies suggested that RHD resulted in an 'agestural' presentation in which patients were unable to express emotional states through gesture, although imitative pantomime gesture production was

spared (Ross & Mesulam, 1979). More recent studies have concluded that people with RHD produce fewer iconic gestures than neurologically healthy controls, especially in contexts of high emotional content (Cocks et al., 2007; Hadar et al., 2010; Hogrefe et al., 2016). Akbıyık et al. found no difference in overall gesture rates between people with RHD and neurologically healthy controls and indeed found that gesture rate was correlated with overall narrative competence in this group (Akbıyık et al., 2018). Interestingly, some studies have found that people with RHD show increased rates of non-communicative manual behaviours during interaction, such as self-touching, scratching, or grooming (Blonder et al., 1995; Cocks et al., 2007). To summarise, patients with RHD demonstrate reduced production of iconic gestures, particularly alongside speech with higher emotional content, while their gesture rate is comparable to that of neurologically healthy adults. Patients with RHD may also produce a greater number of non-communicative hand movements during speech.

1.2.3 Co-speech gesture production in progressive neurogenic populations

1.2.3.1 Dementia

Communication impairment occurs across a range of dementia sub-types. I have chosen here to focus on two in particular: Alzheimer's Disease, the most common form of dementia, and Primary Progressive Aphasia, a presentation of fronto-temporal dementia which is characterised (in the early stages) by selective deficits in language processing.

Alzheimer's Disease (AD) is a progressive neurological condition in which impairments in memory and learning are the dominant clinical features. Neural degeneration in AD is diffuse but usually starts in mesial temporal regions, accounting for the early memory deficits typical of the condition (Mueller et al., 2018). In later stages, however, global deficits develop as neural degeneration becomes more widespread. Speech production impairments have been well studied in AD and in the initial stages significant

word retrieval deficits, circumlocutions, and semantically 'empty' speech (i.e. Wernicke's aphasia) are often evident (Faber-Langendoen et al., 1988; Nicholas et al., 1985). In late-stage AD generalised semantic deficits appear and speech production decreases and/or becomes distorted; in the very late stages people with AD may have minimal or no speech output at all (Caramelli et al., 1998).

In contrast to the relatively well studied speech deficits that occur in AD, gesture production deficits have received less attention. Upper limb apraxia is known to occur early in AD but is particularly evident during late-stage disease (Parakh et al., 2004; Rousseaux et al., 2012). Research into the efficiency of co-speech gesture production in AD is mixed. Overall rates of gesture production have been found to be comparable between people with moderate AD and neurologically healthy adults, but gesture informativeness may be variable (Geladó et al., 2022; Glosser et al., 2010). In people with AD, the evidence suggests that gesture production breaks down in parallel with speech, with gestures devoid of semantic content mirroring similarly 'empty' speech production (Carlomagno et al., 2005; Rousseaux et al., 2010) . In summary, while overall gesture production frequency is comparable to that of neurologically healthy adults in at least the early to moderate stages of AD, production of semantically-rich gestures may be impaired alongside the lexico-semantic deficits seen in speech production in this population.

Compared to AD, there is significantly less known about gesture production in PPA. PPA is a presentation of fronto-temporal dementia in which impairment of language processing is the primary deficit. Currently, three sub-types are recognised: semantic (svPPA), logopenic (lvPPA) and non-fluent (nfvPPA)[1] (Gorno-Tempini et al., 2011), with a fourth type, mixed PPA, encompassing those whose profiles do not fit readily into any of the aforementioned categories (Ruksenaite et al., 2021). Because neural degeneration is (initially) localised to the frontal and temporal lobes in PPA,

[1] Note that in the literature multiple terms are used to refer to each sub-type of PPA. svPPA is also known as semantic dementia and PPA-S; lvPPA is also known as logopenic progressive aphasia and PPA-L; nfvPPA is also known as progressive non-fluent aphasia, non-fluent/agrammatic variant PPA, agrammatic PPA and PPA-G.

patients may experience minimal or no other symptoms for several years. However, disease progression eventually results in global dementia and in some cases of lvPPA and nfvPPA, degenerative syndromes involving the motor system such as corticobasal degeneration (Harciarek et al., 2014).

There is to date no comprehensive analysis of co-speech gesture production in any sub-type of PPA. In a single case study, Macoir et al. demonstrated that a verb cueing technique utilising gesture observation improved action naming in an individual with svPPA, with the treatment effect lasting for at least 4 weeks after the intervention. However, the effect of gesture *production* on action naming was not assessed (Macoir et al., 2015).

Limb apraxia (discussed in more detail in Section 1.3) is common in PPA, occurring in all three variants, and is associated with decreased volume of the left anterior inferior parietal cortex extending into the posterior superior temporal gyrus (Nelissen et al., 2010; Yliranta & Jehkonen, 2020). Upper limb apraxia, rigidity, and extrapyramidal features like those seen in corticobasal syndrome and progressive supranuclear palsy have been found in patients with more advanced nfvPPA. In svPPA, loss of semantic knowledge results in progressive lexico-semantic deficits in speech production. A single case study of an individual with svPPA showed evidence of significantly reduced iconic gesture production with predominant use of deictics indicating location, suggesting a gestural presentation similar to that seen in post-stroke Wernicke's aphasia (Kindell et al., 2013). People with svPPA also show increased production of conceptual gesture errors (e.g., vague or unrecognisable gestures) and a recent case series demonstrated that individuals with svPPA – unlike lvPPA and nfvPPA – showed reduced benefits of meaning in gesture imitation accuracy, consistent with conceptual apraxia (Baumard et al., 2019; Dresang et al., n.d.). In lvPPA, decreased speech output with phonemic paraphasias is common, with co-occurring apraxia and other clinical features reminiscent of AD (Harciarek et al., 2014). Finally, two further case studies showed that two people with lvPPA and mixed PPA respectively were able to use gestures as replacements for

impaired speech, compensating for word-finding difficulty (Volkmer et al., 2023).

Taken together, it may be hypothesised that co-speech gesture production in PPA is impaired along the lines of the deficits seen in speech production, as in some studies of people with AD. However, further research is needed in this population to confirm this.

1.2.3.2 Parkinson's Disease

Parkinson's Disease (PD) is a progressive neurological condition which is regarded as primarily a disorder of movement, due to the dominance of its three major clinical features: rigidity, tremor and brady-hypokinesia (slowness and reduced amplitude of movement). Impairments in co-speech gesture production in this population have therefore been assumed to relate to motor deficits. More recently, however, it has been recognised that cognitive deficits are also present in PD, including impairments in memory, executive function, attention, visuo-spatial processing, and language processing (Raskin et al., 1992; Verbaan et al., 2007). In particular, processing of action words is impaired in PD, including in early stages of the disease before significant cognitive impairment is present, one explanation for which is that people with PD may have a specific action semantics deficit that contributes to production difficulties for action-related language (Bocanegra et al., 2015).

There is some evidence that the putative action semantics deficits in patients with PD may differ from that seen in posterior temporal stroke or svPPA, in that PD is associated with abnormalities in motor *simulation* rather than visual semantic feature processing (Klostermann et al., 2022). In one of the few systematic studies of gesture production in PD, Humphries et al. found a difference in the specific type of iconic gestures produced by people with PD compared to neurologically healthy controls: although overall rates of iconic (and metaphoric and deictic) gesture production did not differ between groups, people with PD produced more gestures from an character viewpoint (i.e. in which the speaker depicts an event as though they are performing it

themselves), compared to healthy controls who described the same event from an observer viewpoint (in which the speaker depicts an event as though observing it from afar) (Humphries et al., 2016). A possible reason for this finding is that patients with PD have difficulty mentally representing performing actions themselves and need to rely instead on visual information (which largely relates to extra-personal space) to activate motor patterns for gestures (Humphries et al., 2016, 2021). In terms of the rate of gesture production in PD evidence is mixed, with some studies finding reduced gesture rate per 100 words (taking into account the markedly slower speech production rate in PD) compared with neurologically healthy controls, and others finding no difference (Cleary et al., 2011; Poliakoff et al., 2021; Saatci Yurtsever et al., 2021).

To summarise, specific gesture production impairments found in people with PD cannot solely be attributed to motor symptom severity and may instead reflect a more subtle deficit in implicit simulation of motor actions. Reduced gesture rate has been found in this population, along with a tendency towards producing action gestures from a character viewpoint perspective. However, the frequency of production of iconic, metaphoric and deictic gestures mirror those produced by neurologically healthy adults overall.

1.2.4 Gesture facilitation effects on speech production

Two highly influential psycholinguistic models have been proposed to account for the gesture production characteristics described above: the Lexical Retrieval Hypothesis (LRH) (Rauscher et al., 1996) and the Information Packaging Hypothesis (Kita, 2000). Both models are predicated on the idea that co-speech gestures have a speaker-oriented function: for the LRH, this is related to lexical retrieval, and for the Information Packaging Hypothesis to efficient chunking of information. However, the models differ in the point in time at which they conceive speech and gestures to interact: according to the LRH, speech and gestures interact at later stages of word production, while the Information Packaging Hypothesis assumes interaction from the earliest, conceptual stage.

1.2.4.1 The Lexical Retrieval Hypothesis (LRH)

The Lexical Retrieval Hypothesis proposes that the function of co-speech gestures (referred to in this model as 'lexical gestures') is to facilitate the speaker's lexical retrieval and word production. As outlined in Section 1.1, word production is hypothesised to be a non-linear process in which semantic units (concepts), lexical units (words) and phonological structure (phonemes, syllables etc.) interact multi-directionally. Activation spreads from the semantic level (conceptualisation), to the word level (lexical selection), to the phonological level (phoneme selection), resulting in a state that enables the word to be articulated (Dell et al., 1999). According to the LRH, gestures interact with speech after the conceptualisation stage, and accompany activity at the level of either lexical selection or phoneme selection. Gestures encode spatial representations that may prime word retrieval, at least when the content of speech is spatial in nature. For example, when describing a trip to the seaside, the speaker activates imagistic mental representations of the paths taken during the trip. These mental representations lead to the retrieval and production of gestures that correspond to features of the motion along these paths. In turn, during lexical and phoneme selection, these gestures support the retrieval and articulation of the corresponding words. Constraining a speaker from gesturing should therefore result in dysfluent and anomic speech production, particularly for words with spatial content, because the facilitatory effect of gestures on lexical or phoneme selection is absent (Rauscher et al., 1996).

Evidence for the facilitation of word retrieval by gestures in both neurologically healthy adults and adults with acquired language impairment is mixed. In studies of 'tip of the tongue' states, where lexical retrieval failure is typically phonological, preventing neurologically healthy adults from gesturing significantly reduces the frequency of successful lexical retrieval (Frick-Horbury & Guttentag, 1998); neurologically healthy adults have also been found to produce speech significantly more fluently when they are permitted to gesture, compared with when they are constrained from doing so (Rauscher et al., 1996). However, a more recent study by Kisa et al. in

which neurologically healthy adults were prevented from gesturing found no effect of gesture on fluency or word retrieval in a story re-tell task, either for words with concrete (e.g. “The scuba diver *went down*”) or metaphoric (e.g. “My opinion of him *went up*”) spatial content (Kısa et al., 2021).

In populations with acquired language impairment, the picture is equally unclear. In broad terms those with primarily phonological deficits have been found to benefit from production of gestures during word retrieval; for those with language impairment linked to conceptual or lexico-semantic deficits, gesture effects on word retrieval are either absent or inconsistent (Bonifazi et al., 2013; Boo & Rose, 2010; Raymer et al., 2006; Raymer et al., 2012; Rose et al., 2010; M. Rose & Sussmilch, 2008). For example, in a novel word learning paradigm Kroenke et al. found that successful word retrieval was associated with iconic gesture production in PWA, but only those with phonological deficits benefited. PWA with impaired semantic processing did not (Kroenke et al., 2013). Comparable results were found by Akhavan et al. and Kistner et al., with iconic gestures facilitating word retrieval for PWA with intact semantic processing in both studies, in a story retell task and in spontaneous conversation respectively (Akhavan et al., 2017; Kistner et al., 2019). This suggests that where gesture is effective in facilitating lexical retrieval, it does so across communicative contexts. In contrast, Kong et al. reported an association between increased gesture use and episodes of word-finding difficulties in some aphasic Cantonese speakers, but a minimal facilitatory effect of iconic gestures on lexical retrieval. In fact, iconic gestures were found to play a greater role reinforcing content already present in speech than in facilitating its production (Kong et al., 2017; Pak-Hin Kong et al., 2015). Indeed, where gestures do accompany successful lexical retrieval, this may not be due to the facilitatory effect of the gestures themselves: an alternative hypothesis is that lexical items with stronger conceptual representations simply give rise to both speech and gestures (Akhavan et al., 2017).

Interestingly, the hypothesised cross-modal priming effect for gestures may also be generated by the interlocuter, when their gestures are seen by the

speaker (i.e. resulting in a gesture observation effect for lexical retrieval). For example, in a study of the effects of gesture observation on verb retrieval in 6 PWA, Bonifazi et al. (2013) found that only PWA with intact action semantic processing benefited; PWA with impaired action semantic processing showed no change in verb retrieval accuracy. However, Routhier et al. (2015) found no benefit for iconic gesture observation on action naming accuracy in 2 PWA, regardless of whether the primary area of impairment was semantic or phonological processing. As with gesture production effects discussed above the specific factors which predict whether PWA will benefit from gesture observation remain uncertain.

In summary, it is unclear to what extent gestures may facilitate lexical retrieval. While it is indeed the case that successful word retrieval may be accompanied by gestures, this is not universal. Particularly in neurogenic populations alternative functions of gesture, such as its compensatory role in replacing or enhancing verbal information, should be considered alongside the possibility that gestures facilitate word retrieval.

1.2.4.2 Information Packaging Hypothesis

In contrast with the LRH, the Information Packaging Hypothesis assumes interaction between speech and gestures occurs from the earliest phase of conceptualisation of the message. According to this hypothesis, gestures help speakers to organise or ‘package’ spatial information via ‘spatio-motoric’ thinking, a method of information processing distinct from analytic thinking, in which information is organised hierarchically into words and phrases. Analytic thinking is the mode by which speech is produced, and its templates for speech comprise linguistically-specified information, such as semantic and pragmatic features of words, how words can be combined into phrases, and so on. In contrast, spatio-motoric thinking, which underlies gesture production, relies on information organised according to action schemas, specifying how the body interacts with the environment, both physical and imaginary. When producing speech and gestures simultaneously, the speaker therefore has access to two modes of information processing, which

continuously influence each other during production of the utterance and enable more effective ‘packaging’ or chunking of complex information during communication (Kita, 2000).

Unlike the LRH, the Information Packaging Hypothesis assumes that speech and gesture are separate and independent systems which continually interact in the service of a common communicative goal. The model is compatible with the production of gestures in the absence of speech, and with incongruities between speech and co-speech gestures, both of which occur in people with acquired communication disorders. For example, in studies of narrative production in people with post-stroke language impairment, Dipper et al. and Pritchard et al. both found evidence for a particular gesture-language incongruence in discourse production in PWA: the pairing of ‘empty’ or ‘light’ verbs (i.e. those that contain minimal semantic information) with semantically rich gestures, such as pairing the verb “go” with gestures indicating the manner and direction of movement (Dipper et al., 2015; Pritchard et al., 2015).

To summarise, co-speech gestures are hypothesised to perform speaker-oriented functions during discourse, including aiding lexical retrieval and facilitating effective packaging of communicative information. Both the LRH and the Information Packaging Hypothesis may account for some findings from the neurogenic communication disorder literature. However, the models differ in their respective conceptions of the starting point of gesture integration with speech: in the LRH, in line with gesture’s role in facilitating speech production, gestures interact with speech at the level of lexical or phoneme selection; whereas according to the Information Packaging Hypothesis gestures and speech bidirectionally interact from conceptualisation of the message. Further research is required to refine our analyses of the timepoint at which speech and gestures interact, and indeed to determine whether and how this differs between neurologically healthy speakers and people with acquired communication impairments.

1.2.5 Multimodal comprehension

Compared with studies exploring whether gesture observation and/or production facilitates speech production in PWA, there are considerably fewer studies focusing on whether gestures produced by interlocuters facilitate PWA comprehension. In neurologically healthy adults, observation of (iconic) gestures during interaction is known to facilitate speech comprehension (Dargue et al., 2019; Zhang et al., 2021). However, in PWA findings are less consistent. Importantly, research in this area has focused predominantly on the effect of pantomimes and representational gestures (i.e. iconic and metaphoric gestures) on speech comprehension. Whether other types of gesture, such as deictic, beat or pragmatic gestures, may also have a facilitative effect on PWA comprehension remains unclear, as do the mechanisms of action by which facilitation may occur.

Experimental paradigms vary considerably between studies, particularly in terms of whether PWA are presented with single words or connected speech stimuli alongside gestures, making generalisations difficult. Records (1994) presented 12 PWA with single word stimuli in three conditions (speech only, speech and iconic gesture, and iconic gesture only); PWA were asked to point to a picture, from a choice of two, that best fit the word they had heard/seen gestured. PWA with poorer speech comprehension relied more on gestural information than PWA with better comprehension, who relied more on speech. By contrast, Cocks and colleagues found that speech comprehension in PWA was hindered by speakers combining short two or three word phrases with additive iconic gestures (i.e. gestures that added information absent from speech, such as saying “I washed it” alongside a “scrubbing hair” gesture) (Cocks et al., 2018), whereas Eggenberger and colleagues found that speakers using speech (short phrases, e.g. “time to brush your teeth”) and redundant iconic gestures (i.e. gestures that included information also contained in speech, such as a “brushing teeth” gesture) facilitated comprehension in some PWA, but not in others (Eggenberger et al., 2016). Finally, in an eye-gaze tracking study, Preisig and colleagues found that PWA attended to meaningful co-speech gestures in conversation

(i.e. connected speech of varying lengths) more frequently than to meaningless hand gestures, although as comprehension of the speech content was not subsequently assessed this does not necessarily mean that attention to meaningful gestures produced by the speaker resulted in enhanced comprehension (Preisig et al., 2018). Taken together, these studies suggest that PWA attend to gestures and may use information encoded in the interlocuter's gestures to understand messages when speech comprehension is impaired, but effects may be moderated by complexity of the speech input and whether gestures are additive or redundant. In the latter case, redundant iconic gestures may facilitate comprehension by highlighting key semantic features of the target word(s), making them more salient.

In neurologically healthy adults, comprehension of both co-speech iconic and silent pantomime gestures is associated with activity in a highly distributed fronto-temporo-parietal network that is primarily (but not entirely) left lateralised. Unsurprisingly, the brain regions recruited for multimodal comprehension are similar to those used in speech comprehension and include the left IFG, areas of the left temporal lobe and the premotor cortex, with some studies also finding bilateral activation in homologous right hemisphere areas (Straube et al., 2011). Kircher et al. (2009) found that the IFG was more strongly activated when neurologically healthy adults heard speech alongside seeing gestures, than when they heard speech alone. Increased left IFG activation has also been associated with processing of gestures incongruent with speech, or when gestures relate to more abstract concepts. For example, in an fMRI study of 17 neurologically healthy native German speakers, Straube et al. (2011) found that IFG activation was significantly increased when participants saw (iconic) gestures produced alongside abstract speech content, such as "John produced the work to a *very high standard"*, compared with when they saw the same gesture produced alongside concrete speech content, such as "John lived at the top of a *very high hill*". In the concrete speech condition, by contrast, activation was primarily found in bilateral middle temporal regions. As the IFG is known to play a role in resolving ambiguities and competing interpretations in

speech comprehension (Rodd et al., 2012), these results suggest that the IFG may play a similar disambiguating role in multimodal comprehension as well (see Emmorey & Ozyurek, 2014, for an in-depth discussion).

Kalénine et al. (2010) has also suggested that the left posterior MTG (which is also implicated in mapping lexical items to semantic representations, e.g. Binder, 2002) may function as a 'hub' for the association of actions (including gestures) with meaning(s). In a study of 43 PWA, participants were asked to complete a gesture comprehension task in which they watched pairs of videos of a researcher performing two semantically-similar actions (e.g. brushing teeth and combing hair, or hammering and sawing). They were then asked to choose which video best matched a given target word. Lesion-symptom mapping analysis was then used to determine whether lesions in the posterior temporal lobe, IFG or inferior parietal lobule best predicted performance on the behavioural task. Lesions to the posterior temporal lobe were found to be the best predictor of poor performance, suggesting that integrity of the posterior temporal lobe is also key to action semantics processing.

1.2.6 Gesture facilitation effects on speech comprehension

Although listeners make (at least some) use of gestural information produced by speakers for message comprehension (Krason, 2023), models of speech-gesture integration vary in the extent to which they assume this information is relied upon by listeners, i.e. whether it is equivalent to, or subordinate to, speech information. The Integrated-Systems Hypothesis (Kelly et al., 2010), which proposes that speech and gestures mutually influence each other during comprehension, takes a "hand-in-hand" approach where information from both modalities is assumed to have comparable impacts on comprehension. By contrast, the Compensation Account (de Ruiter & de Beer, 2013) takes a "trade off" approach, where information from gestures is assumed to be relevant only in cases where speech comprehension is otherwise challenging, such as in speech-in-noise conditions. Broadly, these

models differ in the extent to which they assume listeners integrate gestural information with speech for message comprehension.

1.2.6.1 Integrated-Systems Hypothesis

Similarly to the Information Packaging Hypothesis for speech and gesture production reviewed above, the Integrated-Systems Hypothesis assumes that words and gestures form a common semantic and semiotic system and interact at the level of semantic processing. Speech and (iconic) gestures are assumed to bidirectionally influence each other (equally) in the process of comprehension, and interlocuters obligatorily derive information from both channels. That is, the Integrated-Systems Hypothesis assumes that interlocuters cannot help but attend to and (attempt to) integrate meaning from both speech and gestures during comprehension (Kelly et al., 2010). This yields two predictions: firstly, comprehension will be disrupted when gestures and speech are incongruent in meaning; and secondly, even when a task does not *require* attention to gestures (e.g. listening to someone reading a story aloud), any gestural information present will nevertheless influence how interlocuters comprehend speech content.

Some evidence for the Integrated-Systems Hypothesis comes from speech-gesture mismatch paradigm studies. For example, Kelly et al. (2010) found that (neurologically healthy) adults matched action prime words (e.g. ‘chopping’) more rapidly and accurately to videos with congruent speech and gestures (e.g. saying ‘chopping’ and performing a ‘chopping’ gesture), than to videos containing incongruent speech and gestures (e.g. saying ‘chopping’ but performing a ‘cutting’ gesture, or saying ‘cutting’ but performing a ‘chopping’ gesture). Interestingly, this effect was uninfluenced by target modality; that is, comprehension speed and accuracy did not significantly differ depending on whether the gesture or the speech was incongruent with the prime, suggesting that each modality comparably influenced comprehension of the other. In a follow-up study using the same stimuli, participants were directed to attend to the speech content of each video only; despite this, participant response accuracy still declined when speech and

gestures were incongruent, suggesting that they were unable to avoid integrating gestural information even when specifically instructed to do so (Kelly et al., 2010). The authors concluded that participants were unable to suppress integration of information from the gesture channel. Similar results were found by Perniss and colleagues (2020), who asked native British Sign Language (BSL) signers and English speakers (16 Deaf and 10 hearing) to watch videos of a BSL signer producing (single) target signs with congruent/incongruent mouth movements, or an English speaker producing (single) target words with congruent/incongruent gestures, then choose a picture which best matched the target word. Even when instructed to focus on the primary communication channel (i.e. BSL for signers, speech for speakers), comprehension accuracy declined in the presence of incongruent visual cues. These results further suggest that speakers automatically attempt to integrate information from gestures (or indeed other visual information) to facilitate comprehension.

Speech-gesture congruency effects for comprehension have also been found in PWA (Vigliocco et al., 2020), although whether attention to both modalities is mandatory as proposed by the Integrated-Systems Hypothesis remains unclear. For example, as described in Section 1.2.5 above, Eggenberger et al. (2016) found that congruent gestures facilitated speech comprehension in PWA using a similar paradigm to Kelly et al. (2010). However, they also found that PWA with more severe speech comprehension deficits fixated on the interlocuter's gesturing hands for significantly longer than PWA with less severe comprehension deficits suggesting that while PWA may attend to both speech and gestures, they may selectively attend to gestural information more intently to (attempt to) compensate for impaired speech comprehension. Further, these results also suggest that gestural information may be insufficient to significantly facilitate speech comprehension when the deficit is severe.

Importantly, the Integrated-Systems Hypothesis relates only to the role of representational (i.e. iconic and metaphoric) gestures in comprehension; other gesture types are not included and it remains unclear whether, for

example, beat gestures, which are not typically assumed to carry semantic content, are predicted to facilitate comprehension according to the Integrated-Systems Hypothesis.

1.2.6.2 Compensation Account

Unlike the Integrated-Systems Hypothesis, which is predicated on the idea that speech and gestures comparably influence comprehension of the other modality, the Compensation Account proposes that information encoded in gestures is supplementary in speech comprehension, i.e. it becomes relevant for comprehension only when speech is too difficult to decode, such as in noisy conditions or when auditory processing is impaired. The Compensation Account derives from the Asymmetric-Redundancy Sketch Model (de Ruiter & de Beer, 2013), a parallel model of speech and gesture production which proposes that information encoded in (iconic) gestures is redundant and that speech is the primary (and preferred) modality by which information is conveyed (de Beer et al., 2020).

Some evidence for the Compensation Account comes from speech-in-noise experiments; for example, Obermeier et al. (2012) presented hearing-impaired adults with videos of an actor in background noise producing spoken homonyms which were disambiguated by a co-occurring (iconic) gesture or a subsequent word. The hearing-impaired adults were found to preferentially rely on the gesture to disambiguate the homonym over the word, suggesting that the gesture contained semantic information which was facilitatory of speech comprehension. Further evidence comes from L2 learners (especially lower level L2 learners), who have been found to rely more on gestures than native or fluent L2 speakers for comprehension. For example, Sueyoshi & Hardison (2005) presented low level and advanced/fluent English L2 speakers with recorded lectures by a native speaker, presented in three conditions: as audio-only, as audio and face only, and as audio, face and gestures. Comprehension accuracy for all learners was higher in response to both visual conditions compared to the audio-only condition, but low level learners performed significantly better in the audio,

face and gesture condition compared to the speech and face condition. By contrast, comprehension accuracy in advanced/fluent L2 speakers did not vary significantly between the visual conditions. This suggests that when listeners find speech difficult to decode, they may preferentially use gestural information for comprehension even when other visual cues (e.g. facial expressions, mouth movements) are available; but when speech is easier to decode, information included in gestures is less relevant for comprehension. Similar results were also found by Zhang et al. (2023), who found that L2 learners relied more on meaningful gestures for comprehension than L1 speakers, who relied on multimodal cues (including gestures, but also mouth movements and prosody) less for comprehension overall. For PWA then, representational gestures produced by interlocuters may play a more substantial role in comprehension particularly when speech comprehension is more significantly impaired, although the evidence in this area remains sparse, particularly in terms of comparing gesture effects in PWA with varying degrees of comprehension impairment.

1.2.7 Interim summary

To summarise, after stroke, impairments in co-speech gesture production may affect gesture production frequency, rate and type, and some studies suggest that speech and gestures may break down in parallel in at least some PWA, although evidence is mixed (Rousseaux et al., 2012). How gestures interact with speech to facilitate (or not) speech production and comprehension remains an area of debate; while some studies of PWA do show gesture facilitation effects both for speech production (e.g. Marangolo et al. 2010; Bonifazi et al. 2013) and for comprehension (e.g. Eggenberger et al., 2016), others do not. Importantly, models for gesture facilitation effects primarily focus on representational gestures, especially iconic gestures; whether gestures such as beats, which are not assumed to carry semantic content, are also assumed to facilitate speech production or comprehension remains unclear. Importantly, the studies outlined above primarily considered co-speech gesture effects within experimental contexts only. In the next

section (Section 1.3), I consider how gestures have been used within structured therapy contexts as part of language rehabilitation.

1.3 Approaches to language rehabilitation

1.3.1 Impact of aphasia

The impact of aphasia on quality of life can be significant, both for PWA and for those they interact with. Indeed, in a population-based study of 66,000 adults with a range of health conditions, Lam & Wodchis (2010) found that aphasia was rated as having the greatest negative impact on quality of life of all conditions reported, surpassing both cancer and Alzheimer's. Aphasia is also significantly correlated with incidence of depression after stroke, particularly in the early stages of recovery (Kauhanen et al., 2000). Overall, up to 70% of PWA are expected to experience depression after stroke, with depressive symptoms continuing beyond 12 months in up to 60% of PWA (Strong, 2023). Aphasia is also a predictive factor in destination following post-stroke hospital discharge; Gonzalez-Fernandez et al. (2012) found that PWA with significant ongoing deficits in auditory comprehension, reading comprehension and naming were significantly more likely to be discharged to settings other than home (i.e. nursing/care homes), although access to timely and effective rehabilitation was shown to reduce this risk.

Le Dorze & Brassard (1995) carried out in-depth structured interviews with 18 PWA about the impacts of aphasia and revealed significant negative effects on interpersonal relationships, careers, loss of autonomy and significant feelings of loneliness and stigmatisation. Carers of PWA have also been found to be at significant risk of developing poor mental health, experiencing relationship breakdown, loneliness and physical and mental exhaustion (Winkler et al., 2014). Moreover, while the working age population of PWA is predicted to increase in future, aphasia has been identified as a key risk factor in being unable to successfully return to work following stroke. This effect was found to be exacerbated when PWA were unable to access timely and effective speech and language therapy (Matérne et al., 2019). In

sum, aphasia has been identified as having key negative impacts on quality of life for both PWA and carers across a range of domains, including depression and mental health, inter-personal relationships, the home setting, and employment.

The symptom burden of aphasia is predicted to increase in future, both as the population ages resulting in more people experiencing strokes, and as medical science advances to keep stroke survivors alive for longer (King et al., 2020). As a consequence, developing and delivering effective rehabilitation approaches which facilitate communication recovery and community participation are key healthcare priorities in the UK. In particular, PWA and their carers have identified developing treatments for moderate/severe aphasia as particularly important (Franklin et al., 2018).

1.3.2 Unimodal and multimodal therapy approaches

1.3.2.1 Anomia

Anomia therapy typically falls into two main categories: multimodal therapy approaches that encourage the free use of communicative channels: speech, gesture, writing, drawing, facial expressions etc.; and unimodal therapy approaches that restrict the use of non-speech communicative channels except under certain circumstances.

Broadly, these two categories of intervention are referred to as Multimodal Aphasia Therapy (M-MAT) and Constraint-Induced Aphasia Therapy (CIAT). These approaches rely on distinctly different hypothesised mechanisms of recovery: in M-MAT, a structured combination of multimodal cues (speech, gestures, drawings, reading aloud, writing etc.) aims to leverage all remaining neural networks involved in language processing, facilitating communication through rich inter-systemic and cross-modal cueing which is argued to more closely represent ‘real world’ multimodal communication (Rose, 2013). By contrast, in CIAT PWA are trained to (try to) use speech alone, and are constrained from using alternative communicative modalities

through explicit instruction and modelling, on the basis that through processes of experience-dependent neural plasticity, forced use will gradually result in neural reorganisation and/or regeneration and eventually improved language processing (Difrancesco et al., 2012). In early versions of CIAT, use of non-speech channels by PWA to communicate was expressly forbidden. In more recent versions (which are sometimes known as CIATplus), PWA are typically seated behind barriers, which block the interlocuter's view of any gestures they make. PWA may therefore attempt to gesture (or write), but communication with the interlocuter is only judged as successful based on the content of the speech output (Kurland et al., 2016).

Although M-MAT is primarily concerned with gesture production effects on naming, a very small number of studies have explored whether gesture observation therapy also improves word retrieval. For example, Oppici et al. (2023) found that gesture observation was as effective as gesture production in facilitating word retrieval in a meta-analysis including 7 studies (n = 309 brain damaged adults with left and right hemisphere lesions). However, the authors were primarily interested in (L2) word learning, rather than re-learning of words following stroke. In a small study of 6 PWA, who took part in a structured naming therapy programme combining words with gesture observation (one therapy session daily for two weeks) Bonifazi et al. (2013) found that only PWA with primarily phonological processing showed improvements in word retrieval; PWA with impaired action semantic processing showed no improvements in verb retrieval accuracy, suggesting that activation of underlying semantic representations was needed for successful retrieval.

Studies comparing the efficacy of M-MAT and CIAT/CIATplus suggest that they are similarly effective in treating chronic anomia across a broad cohort of PWA; it is currently unclear what factors may predict benefit or make one approach more appropriate for a given PWA. Attard et al. (2013) found that two PWA with severe Broca's aphasia, who received both types of therapy under an alternating treatment design for 32 hours over two weeks, showed equal improvement in naming accuracy regardless of intervention type.

However, no generalisation to untreated items was detected. Similar results were found by Rose et al. (2013) in a larger study of 11 PWA with a variety of aphasia subtypes (anomic, Broca's, conduction) and severities (mild, moderate and severe). Interestingly, at follow-up a majority of participants expressed a strong preference for M-MAT over CIAT, suggesting that the M-MAT approach may be more acceptable to a larger number of PWA (Rose et al., 2013). Finally, Rose et al. (2022) carried out a randomised multi-site trial of 201 PWA, comparing the effectiveness of M-MAT, CIAT and usual care on word retrieval, functional communication measures and communication-based quality of life. In both the M-MAT and CIAT groups, all three outcomes improved significantly compared to the usual care group, and improvements in word retrieval were also maintained at 12 weeks post-intervention. However, while there is evidence that M-MAT and CIAT may be equally effective across cohorts of PWA, there is also evidence that individual PWA differ in how they respond to each type of intervention. Analysis conducted by Rose et al. (2022) indicated that baseline aphasia severity may moderate treatment response, with PWA with mild and severe aphasia benefiting more from M-MAT and PWA with moderate aphasia benefiting more from CIAT in improving lexical retrieval. Additional behavioural and brain-based markers (lesion site, lesion volume) which may predict gesture response and allow for identification of specific rehabilitation pathways for PWA, remain understudied.

1.3.2.2 Speech comprehension

Compared with multimodal therapy approaches for anomia, there are very few structured speech comprehension therapies for PWA utilising (or indeed constraining) co-speech gestures. *Total Communication* approaches, in which the PWA's communication partner is trained to use all multimodal channels alongside speech to facilitate comprehension (including the use of pictures and communication books), have shown some success in facilitating speech comprehension in PWA with severe aphasia, although the effects appear to be contingent on the continued use of multimodal strategies by communication partners (Simmons-Mackie, 2009). That is, comprehension

improves when the communication partner uses, and the PWA observes, multimodal cues including gestures, but effects may not carry over beyond the interaction. Additionally, it is unclear whether certain types of visual cues (e.g. gestures or pictures) are more beneficial than other types, or whether certain types of gestures are more facilitatory of comprehension.

1.3.3 Generalisation of treatment gains and ecological validity

A key feature in aphasia rehabilitation literature is the failure of gains made in therapy to generalise beyond the treatment room (Castro et al., 2021). That is, while clinically significant change may occur in items treated during intervention, there is not necessarily corresponding improvement in untreated items or change in overall functioning or quality of life measures. Similarly, experimental effects observed in the clinic or in controlled laboratory studies may not be subsequently observed in more naturalistic environments. For example, in a study of 13 PWA with chronic mild-moderate aphasia who engaged in 30 hours of M-MAT and 30 hours of CIAT, generalisation of gains in word retrieval to untreated items was observed in some, but not all participants. Discourse samples were collected at three time points (pre-treatment, post-M-MAT and post-CIAT) to track change in four outcome measures: total number of Correct Information Units (CIUs) produced, number of CIUs produced per minute, total number of words produced and number of words produced per minute. No consistent effects were found across measures regardless of intervention type, with significant inter-participant variability in response including in both positive and negative directions (Rose et al., 2016).

1.3.3.1 Language-action games (LAGs)

It is possible that one reason for failure of generalisation of treatment gains and failure to replicate laboratory effects in naturalistic interactions is a lack of ecological validity in the contexts in which these effects were observed or elicited. A potentially useful method of maintaining experimental control over key variables while allowing for more spontaneity of communication between

interlocuters, is to utilise game-based tasks. These gamified tasks, often known as language-action games (LAGs), have a long history in aphasia studies (Romani et al., 2018). LAGs can also be embedded into therapy approaches, and both M-MAT and CIAT make extensive use of LAGs. LAGs were initially developed to provide a space for PWA to practice different communicative forms in a structured interactive setting, such as storytelling or giving map directions to interlocuters (Pulvermüller, 1990).

Typically, LAGs consist of turn-taking activities between two to four players with a shared communicative goal, such as requesting or passing a card/object between players. Clinicians facilitate LAGs but may or may not act as players themselves, depending on the structure of the game. Each round of turn-taking in a LAG involves at least two phases: the request phase, in which a player communicates a request to other players, and the action phase, in which the responding player(s) make the appropriate answer (e.g. responding yes/no, passing over the requested item, etc.). Additional phases, such as a clarifying phase where a responding player may ask the requesting player questions, may also be included. Other game-like elements such as barriers, dice, playing cards, counters and game boards may also be used in LAGs. Depending on the rules of the LAG, players may be restricted to communicating through a single channel (e.g. speech), or may be free to communicate via a range of channels (e.g. speech, gesture, writing, drawing, nodding etc.) (Difrancesco et al., 2012).

Crucially, although LAGs are structured and goal-based, each turn in a LAG is not fully predictable by players. That is, the responding players cannot predict precisely what will be requested of them, and requesting players cannot predict the response they will receive. Therefore, as in naturalistic everyday interactions, turns in a LAG may be considered more akin to action trees than action chains, in that a single act by the requesting player (e.g. “Do you have…?”) may be followed by a (contextually constrained) range of responses from the other players (e.g. “Yes”, “No”, “Here it is”, “Please repeat”, head nod/shake, point etc.). Additionally, turns in a LAG typically also enable players to practise both language comprehension and language

production skills, mirroring naturalistic interactions (Difrancesco et al., 2012).

Unlike fully unstructured and spontaneous interactions, some degree of control over variables of interest is possible because the game rules and stimuli are fully manipulable by the researcher. For example, researchers and confederates may choose to use (or not) specific multimodal cues to observe their effect, or they may tailor tasks and stimuli to elicit specific responses from players. For example, a LAG may allow for manipulation of the gesture use/non-use, by permitting or constraining use of gestures by the interlocuter (focusing on the effect of gesture observation) and/or the PWA (focusing on the effect of gesture production). It is also possible to control the order of presentation of stimuli, control lexical variables such as word frequency and length, and record a range of outcome measures (e.g. response latency, naming accuracy etc.) through a LAG.

Finally, due to their goal-based nature LAGs may enhance motivation more than other structured tasks, because they are more engaging and communication partners can also be trained to act as players, resulting in greater compliance in experimental tasks and therapy, as well as allowing for more intensive input (Meinzer et al., 2007). LAGs may therefore be more acceptable to PWA than other lab or therapy tasks. For example, in a study of six PWA who engaged in an intensive therapy programme (54 hours over six weeks) based around a picture naming LAG, PWA reported that they preferred the LAG format to standard 1:1 speech and language therapy; furthermore, four out of six PWA made clinically significant gains of at least 20% improvement in naming of trained items, and improvements were maintained at least six months after conclusion of the intervention (Romani et al., 2018).

1.3.3.2 Factors moderating gesture benefit for production and comprehension

Beyond the design of therapy activities, individual differences in cognition are key moderating factors in developing effective rehabilitation approaches for PWA (Brady, 2022). In considering whether gesture observation and/or production facilitate speech comprehension or production in PWA, semantic processing, working memory (WM) and upper limb apraxia have been identified as of particular relevance in moderating gesture effects.

1.3.3.2.1 Semantic processing

Some studies have identified a connection between deficits in (conceptual) semantic processing in PWA and speech and gesture comprehension and production impairments. In an unimpaired semantic processing system, multimodal conceptual features are activated to comprehend verbal (e.g. words) and non-verbal (e.g. pictures, gestures) stimuli (Jefferies & Lambon Ralph, 2006). In PWA, access to lexical semantics and phonological forms may be impaired, hindering word production and speech comprehension, but underlying conceptual representations may remain intact (Howard & Gatehouse, 2006). However, if underlying conceptual representations are degraded, PWA may be unable to use semantic information encoded in gestures either to comprehend speech or facilitate speech production.

In an early study of 128 adults with neurological damage (both left and right hemisphere strokes), deficits in symbolic (picture) understanding were significantly associated with deficits in gesture comprehension, suggesting that conceptual semantic processing deficits may underlie both impairments, although speech comprehension was not assessed (Gainotti & Lemmo, 1976). In terms of gesture production, in a study exploring spontaneous use of gesture by PWA in a story narration task, severity of conceptual semantic processing deficits was found to predict the diversity of gesture types produced in 24 PWA with severe (primarily global) aphasia, in that PWA with the most severely impaired semantic processing skills produced significantly fewer meaningful gestures than those with less impaired semantic processing abilities (Hogrefe et al., 2012). The authors suggested that the gestures produced by PWA with the most severely impaired conceptual

semantic processing abilities were underspecified in terms of semantic features, presumably due to correspondingly underspecified conceptual semantic representations. Whether the gestures produced by PWA facilitated their speech production was not assessed, but it may be hypothesised that gestures which were underspecified in terms of semantic features may less strongly activate the semantic features of corresponding words, and therefore be less likely to resolve word-finding difficulties.

Kong et al. (2015) also found that conceptual semantic processing abilities predicted gesture production frequency in a study of 48 Cantonese-speaking PWA. In two narrative story retell tasks, PWA with less severe semantic processing abilities produced a significantly higher frequency of gestures (of all types) than PWA with more severe semantic processing abilities, but this did not correlate with improved lexical retrieval suggesting that gestures were used as a compensatory measure for impaired speech production. Contrasting results were found by Lanyon et al. (2009), who found that 'meaning-laden' gesture production (i.e. iconic, emblem and pantomime gestures) in PWA with relatively intact semantic processing was significantly correlated with successful lexical retrieval, suggesting that the semantic content of the gestures may have enhanced activation sufficient to retrieve the target words. It is possible that the contrasting results in these studies derive from the different types of gestures considered in each; Kong et al. (2015) considered gestures of all types, including pragmatic and beat gestures without semantic content, which would not be assumed to facilitate semantic activation in the same way as the 'meaning-laden' gestures considered by Lanyon and colleagues.

Taken together, these studies suggest that underlying semantic processing abilities may moderate whether (at least some) gestures facilitate comprehension and production in PWA, on the basis that gestures may enhance semantic activation and facilitate access to the lexicon in cases where semantic representations remain intact. However, effects on speech comprehension remain particularly understudied and it is not clear whether

beat gestures, which do not carry semantic content, could also facilitate speech and comprehension and production in this way.

1.3.3.2.2 Working memory (WM)

Working memory (WM) may also moderate gesture facilitation effects in PWA. In neurologically healthy adults, visual and auditory information has been found to be retained in WM for up to 6 seconds if people are distracted in the intervening period between stimulus presentation and testing (Ricker & Cowan, 2010), and up to 10 seconds without distractions (Zhang & Luck, 2009). In PWA, the speed of trace decay of information in WM is known to be faster than in neurologically healthy adults. For example, PWA with both mild and severe aphasia were found to retain information at 1 second following stimulus presentation, but by 5 seconds the information had been lost without any distractions (Silkes et al., 2021). However, in both PWA and neurologically healthy adults, the length of time information is retained in WM can be extended through rehearsal and re-activation of content, which delays trace decay. Indeed, it has been suggested that gestures reduce cognitive load by maintaining activation of visuo-spatial and spatial-motoric information in WM. For example, young neurologically healthy adults with poorer WM were found to gesture more frequently in story retell tasks than young adults with better WM scores, which may be a compensatory measure to maintain information in WM more effectively by using gestures alongside speech to extend trace activation time (Hostetter & Alibali, 2007). Similarly in PWA, Morsella & Krauss (2004) found that pairing producing gestures with spatial concept words facilitated subsequent spatial word retrieval more than presenting the words alone, which they hypothesised was caused by the gestures acting to sustain activation of semantic features of the target words in WM.

Fewer studies have been conducted into the effects of observing (as opposed to producing) gestures in sustaining activation of information in WM, especially in clinical populations, but Thompson & Guzman (1999) found that although young neurologically healthy adults benefited from seeing iconic

gestures alongside speech in a comprehension task in speech-in-noise conditions, older neurologically healthy adults did not. In contrast, Schubotz et al. (2021) found that older neurologically healthy adults did benefit from observing gestures alongside speech in a similar task, but the gain was significantly smaller than for neurologically healthy younger adults and was moderated by WM capacity, with older adults with poorer WM capacity benefiting less from gestures than those with better WM. Interestingly, this suggests that while gesture production may facilitate trace activation in WM, gesture observation may not, or at least not in neurologically healthy older adults.

1.3.3.2.3 Limb apraxia

Apraxia commonly co-occurs with post-stroke aphasia and is a major confounding factor in studies of co-speech gesture production and comprehension after stroke. Apraxia is characterised by deficits in skilled movements which cannot otherwise be explained by motor or cognitive deficits (Buxbaum & Randerath, 2018). There are, broadly speaking, two major types of limb apraxia characterised by differing underlying mechanisms, lesion profiles and performance patterns, though the literature is inconsistent. The first type, ideational or conceptual apraxia, reflects loss of conceptual information about actions and manipulable objects. In stroke as well as in the semantic variant of Primary Progressive Aphasia, the disorder most often occurs following damage that includes the left posterior temporal cortex and temporo-parietal junction, although it may also occur following damage to the IFG. The temporo-parietal region is critical for knowledge of action meaning (‘action semantics’) and the appearance of gestures associated with manipulable objects (e.g., what a ‘hammering’ gesture looks and feels like, or how the hand looks and feels when using scissors).

The second type, production or ideomotor apraxia, reflects a deficit in transforming relatively intact action semantic knowledge into a motor plan for positioning the body in space over time. It is most often associated with

damage that includes the fronto-parietal cortex but spares the temporal lobe (Buxbaum & Kalénine, 2020). Both subtypes are associated with impaired performance on gesture production tasks and are thus difficult to distinguish behaviourally with most clinical assessment tools, but ideational apraxia is additionally characterised by poor performance on gesture recognition tasks and poor production of both meaningful and meaningless gestures (i.e., meaning does not benefit gesture production). A functional-neuroanatomic framework has not been applied in prior studies of communicative gesture in PWA, but may be helpful in making sense of at least some of the contradictions in the literature.

Early studies of gesture production in PWA identified impairments in pantomime gesture imitation, meaningless gesture imitation and tool use, leading to suggestions that gesture production deficits in PWA may be attributable (at least partly) to apraxia (Goodglass & Kaplan, 1963; L. Wang & Goodglass, 1992). However, these studies did not test the informativeness of gestures produced by individuals with and without apraxia, leaving unresolved the question of whether apraxia influences gestural communication. Other studies have concluded that apraxia may indeed affect the comprehensibility of gestures produced by PWA (Völsch et al., 2021). For example, Feyereisen et al. reported that although patients with more severe apraxia used the gestural modality to communicate more frequently than patients with less severe apraxia in a referential communication task, their gestures were more difficult for interlocuters to comprehend (Feyereisen et al., 1988). Similar results were found by Borod et al., who asked nurses and speech and language therapists to rate the spontaneous gestural communication of patients with apraxia: the more severe the patient's apraxic deficit, the less effective their use of gestures to communicate (Borod et al., 1989). More recently, van Nispen et al. found in a group of PWA that impairments in the production of manipulable object use gestures on an apraxia test were associated with substantial reductions in the informativeness of gestures in conveying information to interlocuters (van Nispen et al., 2018). Hogrefe et el. also demonstrated that impairment on a pantomime to command test strongly predicted comprehensibility of gestures

and weakly predicted gesture diversity in a story retell task in individuals with severe aphasia (Hogrefe et al., 2012).

Taken together, the literature suggests that limb apraxia may be associated with reductions in the communicative content that can be conveyed through the gesture channel. Of interest is whether apraxia profiles (i.e., and in particular, the status of semantic knowledge about actions and manipulable objects) mediate this effect. Based on the prior studies in PWA, it may be the case that patients whose aphasia is associated with damage to the middle and posterior temporal lobe (e.g., Wernicke's) will exhibit conceptual (ideational) apraxia and reduced comprehensibility of the communicative gestures produced, whereas patients with non-fluent aphasias (e.g., Broca's) and relative preservation of the posterior temporal lobe will exhibit production (ideomotor) apraxia and relatively preserved gesture comprehensibility. Additionally of interest is assessment of whether apraxia profile predicts the ratio of iconic to metaphoric and other gesture types employed by PWA; based on their loss of conceptual knowledge of actions and manipulable objects, it might be predicted that individuals with ideational apraxia may exhibit reductions in at least some types of iconic gestures.

1.3.4 Interim summary

Aphasia, regardless of which communicative channels are impaired, can have a significant negative impact on quality of life. Furthermore, the symptom burden of aphasia in the UK is predicted to increase in future as the population ages and frequency of strokes increase (King et al., 2020). Two main therapeutic approaches involving gestures exist for anomia: M-MAT, which encourages the use of gestures (and other channels) to facilitate word-retrieval; and CIAT, which restricts the use of channels other than speech. Although both approaches have been found to be effective in (at least some) PWA, it is unclear which PWA benefit more from each approach. Few structured therapy programmes for speech comprehension utilising gestures exist, although Records (1994) found that (at least some) PWA whose communication partners used a *Total Communication* approach which

incorporated gestures demonstrated better comprehension, at least while the interaction was ongoing.

Considerations of ecological validity and individual moderating factors may go some way towards accounting for varying gesture effects between studies. Although semantic processing, WM and limb apraxia have been identified as of particular relevance for gesture-based therapies, none have been studied in depth and crucially, interactions between factors (and other variables such as lesion site) may not be accounted for. Taken together, it remains unclear whether gesture therapies facilitate speech production and comprehension in PWA (particularly in spontaneous interactions), and which moderating factors predict benefit. Moving forwards, the ability to accurately predict which PWA benefit from gesture for comprehension and production, and which do not, could significantly open up possibilities for tailored rehabilitation (Varkanitsa & Kiran, 2022).

1.4 Implications and the way forward

Drawing on the literature reviewed above, in this thesis I focus on the following elements. Firstly, communication is multimodal, and evidence suggests that speech and gestures are subserved by an underlying fronto-temporo-parietal neural network that is (at least partially) shared (Andric & Smal, 2012). Aphasia therefore has consequences for both unimodal and multimodal communication. Secondly, co-speech gestures are hypothesised to have speaker-oriented effects in facilitating speech production, perhaps through cross-modal priming (the LRH) or enabling efficient chunking of information (the Information Packaging Hypothesis); and for listeners, gestures may facilitate rapid speech comprehension, either “hand-in-hand” with speech (the Integrated-Systems Hypothesis) or as an adjunctive measure (the Compensation Account). Yet how gesture processing and speech-gesture integration is affected by aphasia remains unclear. Thirdly, for (some) PWA, gestures may offer a path to personalised language rehabilitation pathways (Varkanitsa & Kiran, 2022). Factors which affect how and predict whether gestures facilitate speech production and

comprehension in PWA, such as WM, semantic processing and limb apraxia effects are not well understood, nor are the extent to which they interact. Indeed, this may go some way to explaining why therapies which encourage the use of gestures (such as M-MAT and *Total Communication*) appear to yield comparable outcomes in PWA to therapies which constrain the use of gestures (such as CIAT/CIATplus); without identification of relevant predictors, therapy is as likely to have no effect as to produce improvement, with inevitable consequences for the healthcare system and PWA (Brady, 2022; Rose et al., 2013). Fourthly, an additional factor which should not be overlooked is ecological validity, and an urgent need to extend generalisation of treatment gains beyond the clinic or laboratory by positioning experimental studies and therapy into a more naturalistic context (Romani et al., 2018).

Chapter 2: “Go Fish!”

2.0 Abstract

As described in Chapter 1, previous studies of the effects of gesture on auditory comprehension and naming have typically elicited or shown gestures and speech outside of dyadic interactions, limiting ecological validity. To address this, and to more closely mimic naturalistic interaction while still maintaining a degree of experimental control, I asked 36 PWA to play a dyadic interactive picture naming LAG, “Go Fish!”, with a neurologically healthy interlocuter (the researcher) to test the predictions of the LRH and to ascertain if gestures produced by the interlocuter facilitate auditory comprehension in PWA. In half of game trials, players were permitted to gesture, while in the other half of trials gesture use was constrained. I investigated whether gesture observation (i.e. PWA observing gestures produced by the researcher) facilitated PWA speech comprehension or subsequent naming accuracy, and whether gesture production (i.e. PWA producing gestures during naming attempts) facilitated naming accuracy. Contrary to predictions, I found no clear facilitatory effect of gesture for speech comprehension or naming accuracy across PWA. However, I found that gesture observation/production benefit for naming did occur in a minority of PWA and analysis of neuropsychological variables revealed that gesture observation benefit was associated with poorer working memory/executive function (WM/EF), while gesture production benefit was associated with better WM/EF. Furthermore, I found that gesture observation benefit for speech comprehension was moderated by gesture recognition abilities. Analysis of lesion profiles also indicated that PWA with relatively preserved frontal lobes benefited more from gesture production in naming, while PWA with relatively preserved temporal lobes benefited more from gesture observation. I discuss these findings in relation to theories of speech-gesture integration.

2.1 Introduction

Co-speech gestures are ubiquitous in human communication and occur in both neurologically healthy and clinical populations. However, their effects in facilitating auditory comprehension and speech production are not well understood in PWA, particularly within dynamic multi-person interactions, where most communication takes place.

2.1.1 Gestures (may) facilitate speech comprehension

From the perspective of the (neurologically healthy) interlocuter, gestures produced by the speaker facilitate comprehension of speech, at both single word and connected speech levels (Dargue et al., 2019). Neurologically healthy speakers typically produce gestures before the word(s) they refer to, so the listener is able to use gestures to constrain lexical competitors and predict upcoming words, facilitating rapid comprehension (Donnellan et al., 2022). Furthermore, seeing gestures that encode semantic information incongruent with speech content may result in comprehension failure, especially when gestures are strongly incongruent. For example, Kelly et al. (2010) asked neurologically healthy adults to watch videos of an actor producing iconic gestures which were semantically congruent, weakly incongruent or strongly incongruent with speech content, then answer comprehension questions about the videos; participants were significantly more likely to make comprehension errors when gestures were strongly incongruent (e.g. actor saying “cut” and gesturing “twist”) than weakly incongruent (e.g. actor saying “cut” and gesturing “chop”), indicating that interlocuters attended to gestures and attempted to integrate semantic information from both speech and gesture channels for comprehension, even when the content did not fully cohere.

However, in PWA findings are inconsistent and studies have found varying effects of gesture observation on comprehension. In an eye-movement tracking study of 20 PWA and 30 healthy controls in which participants were presented with short video clips of a person speaking and producing

congruent or incongruent representational (i.e. iconic or metaphoric) gestures, Eggenberger et al. (2016) found that PWA fixated on the speaker's gesturing hands significantly more often than the control group (although both groups fixated most frequently on the speaker's face). Notably, PWA with the most severe aphasia fixated on the hands more often than PWA with less severe aphasia. However, extent of fixation on the gesturing hands was not found to significantly improve PWA comprehension of the speaker's message; while some PWA did show evidence of multimodal gain for comprehension, this occurred only when gestures were congruent with speech. Therefore, while PWA may attend to gesturing hands as a means of compensating for impaired speech comprehension, this strategy may not be sufficient in the face of significant comprehension deficits or when gestures encode semantic information incongruent with speech.

Furthermore, in a single case study Cocks et al. (2009) demonstrated that "SR", a PWA with severe naming and speech comprehension deficits but intact gesture comprehension, showed significantly lower multimodal gain for comprehension than neurologically healthy controls. SR and 20 age-matched neurologically healthy controls were asked to watch 21 videos of an actor (with mouth covered to hide lip movements) producing short sentences alongside iconic gestures (e.g. "I cut it" with a "cutting" gesture). Randomised clips were presented in one of three conditions: speech and gesture, gesture only with speech muted, and speech only with the video replaced by a still photo of the actor. After watching each clip, participants were asked to point to a picture (from a choice of four) which best depicted the message conveyed. Although SR performed better in the combined speech and gesture trials than either of the single modality trials, the effect was minimal. SR's multimodal gain score for comprehension was calculated to be 11.6%, compared to an average of 30.0% for neurologically healthy adults. This suggests that (iconic) gestures do not inevitably facilitate speech comprehension in PWA, even in cases where gesture comprehension remains intact. These results were extended by Cocks et al. (2018), who used the same paradigm to test the effect of gesture observation on speech comprehension in 31 PWA with mild-moderate aphasia and 30 age-matched

neurologically healthy controls. While healthy controls scored at ceiling in comprehending the messages regardless of condition, PWA performed significantly worse in response to speech and gesture stimuli, than in either of the single modality conditions. Cocks and colleagues suggested that for (at least some) PWA, presentation of information through multiple channels results in 'multimodal loss' for comprehension accuracy, perhaps due to difficulty allocating attention sufficiently to multiple sources of input or integrating information successfully from multiple channels.

2.1.2 Gestures (may) facilitate speech production

Considering the purpose of gestures from the perspective of the speaker, one proposed function is to facilitate speech production through a cross-modal priming effect at the level of lexical retrieval or phoneme selection (i.e. the LRH) (Krauss et al., 2000) (Rauscher et al., 1996). In a within-subjects study in which neurologically healthy adults were either permitted to, or constrained from, gesturing while describing short cartoon stories to listeners, Rauscher et al. (1996) found that speech with spatial content was markedly more disfluent in the gesture constrained condition. Non-spatial speech content was not affected by gesture constraint or production. The authors concluded that gestures facilitated speech production, at least when speech content was spatial in nature. Pine et al. (2007) found similar results with neurotypical children, except that the beneficial effects of gesture for word retrieval were not limited to speech with spatial content. The authors suggested that gestures functioned to reduce cognitive effort by exploiting a visuo-spatial route to the lexicon, facilitating retrieval of phonological forms.

This cross-modal priming effect for gestures may also be generated by the interlocuter, when their gestures are seen by the speaker (i.e. a gesture observation effect). In a meta-analysis of 7 studies with both neurologically healthy and brain damaged participants (n = 309 adults), Oppici et al. (2023) found that gesture observation was as effective as gesture production in facilitating word retrieval, although the authors were primarily interested in (L2) word learning rather than production of already known words. Murteira

et al. (2019) found that observation of (pantomime) gestures facilitated verb naming in 26 neurologically healthy adults, who were shown video clips of an actor performing a congruent or incongruent gesture, then immediately afterwards asked to name an action picture. Naming latencies were significantly faster and more accurate when the video clips depicted congruent gestures, although the authors were unable to determine whether facilitation occurred due to activation of shared semantic features between the gesture and the target word, or due to more efficient access to the word's phonological or lexical form.

Previous studies of PWA have shown that (at least some) PWA do benefit from both (iconic) gesture production (e.g. Marangalo et al., 2010) and/or observation (e.g. Marangolo et al., 2012; Bonifazi et al., 2013) in facilitating word retrieval. However, the extent to which PWA benefit varies. For example, in a study of the effects of gesture observation on verb retrieval in 6 PWA, Bonifazi et al. (2013) found that only PWA with intact action semantic processing benefited; PWA with impaired action semantic processing showed no change in verb retrieval accuracy. Rose & Douglas (2003) found similar results for the effects of gesture production on noun retrieval in 6 PWA, with only PWA with primarily phonological (but not semantic) processing deficits benefiting from gestures. Routhier et al. (2015) found no benefit for iconic gesture observation on action naming accuracy in 2 PWA, regardless of the locus of impairment. Analogous results for gesture production were found by Völsch et al. (2021), who further compared the effects of iconic vs. meaningless gesture production on verb naming in 5 PWA with primarily phonological, semantic or mixed impairments; no improvement in naming accuracy was found for any PWA, regardless of the type of gesture produced or area of impairment.

To summarise, the effects of both gesture observation and gesture production on word retrieval in PWA remain unclear, although PWA with primarily (action) semantic processing impairments may be more likely to show no effect of gesture compared to PWA with primarily phonological processing impairments, perhaps due to difficulty maintaining activation of

semantic and visuo-spatial features while the target word is being retrieved (Pine et al., 2007; Murteira & Nickels, 2020).

2.1.3 Decontextualised language environments

A possible reason for differing experimental results across both multimodal comprehension and production studies with PWA may be the context in which they occur. Studies of speech and gesture integration in PWA often occur in decontextualised language environments where speech and/or gestures are presented or elicited outside dyadic interactions, such as through story-retells or procedural narratives (Clough & Duff, 2020). This has the benefit of being highly controlled experimentally, but strips speech and gestures from the interactive environment in which they evolved and in which they usually appear. Speech comprehension and elicitation tasks are treated similarly in aphasia assessments, with single word confrontation (picture) naming of decontextualised items a typical way of assessing the severity of naming deficits, and yes/no responses to decontextualised questions (e.g. “Do you cut the grass with an axe?”) a typical way of assessing auditory comprehension impairments. While useful in establishing a clinical baseline, these approaches do not necessarily capture the effectiveness of PWA communication in everyday situations, and indeed PWA may perform very differently in informal interactions than during formal assessments (Doedens & Meteyard, 2018).

To improve the ecological validity of studies while maintaining a degree of experimental control, language-action games (LAGs) have been used to study more naturalistic communication in PWA and to deliver therapy (Romani et al., 2018). Typically, LAGs consist of multiple rounds where players have a shared communicative goal which is understood by all participants (e.g., collecting sets of cards, tracing a route on a map, etc.). LAGs emphasise turn-taking activities between multiple players and may incorporate inter-player competition to increase engagement. Experimental manipulations can be built into LAG “rules” by researchers (e.g. permitting or constraining gesture use, requiring single word or connected speech

production, etc.), while still providing motivation for players. While not fully naturalistic and arguably more formalised than everyday interactions, LAGs nevertheless allow for more spontaneity of PWA responses than lab tasks or structured therapy activities (Davis & Wilcox, 1985; Difrancesco et al., 2012). Barrier games, card games, guessing games and story-telling games have all been utilised to elicit speech and gesture production in PWA, and both Multi-Modality Aphasia Therapy (M-MAT) and Constraint-Induced Aphasia Therapy (CIAT) approaches have made use of LAGs (Kurland et al., 2016; Pierce et al., 2017).

2.2 The present study

The goal of the present study was to investigate whether gestures (both observation of gestures by PWA and production of gestures by PWA) facilitated single-word speech comprehension and single-word picture naming accuracy in the context of a short, semi-structured dyadic interaction. To maintain PWA engagement and more closely mimic naturalistic interactions, I developed a novel turn-based LAG based on the card game "Go Fish!". A within-subjects design was used with the primary experimental manipulation being whether game players (comprising dyads of the researcher and a PWA) were permitted to use, or were constrained from using, gestures alongside speech in each game turn (i.e. each trial). I measured whether PWA responded correctly to the researcher's requests for specific playing cards (comprehension measure), and whether PWA produced the correct names of their own playing cards (naming accuracy measure). In order to maximise the number of participants, data collection was carried out at two sites: UCL in London, UK and Moss Rehabilitation Research Institute (MRRI) in Philadelphia, USA. To ensure consistency for PWA, all interactions were conducted with the same researcher either at UCL or at MRRI[2]. Ethical approval to carry out this study was received from the High Risk Research Ethics Committee at UCL (21765/001) and the MRRI Institutional Review Board (IRB-2023-1046).

[2] Isobel Chick and Dr. Amy Lebkuecher at UCL and MRRI respectively.

2.3 Study aim and hypotheses

The aims of this study were: (1) to investigate whether speech comprehension in PWA was facilitated by (iconic) gestures produced by the interlocuter (the researcher); (2) to investigate whether lexical retrieval for target words improved when PWA observed (iconic) gestures produced by the interlocuter (i.e. gesture observation effect); and (3) to investigate whether lexical retrieval for target words improved when PWA produced gestures themselves alongside attempting to say the target word (i.e. gesture production effect).

Following Eggenberger et al. (2016), I firstly predicted that PWA would score significantly higher for comprehension in trials where they could see the researcher's gestures than in trials where the researcher's gesture use was constrained. As the researcher used only representational (iconic) gestures which were closely semantically-aligned to the target word, I hypothesised that PWA (at least those with intact action semantic processing who could extract information from gestures) should be able to use the semantic content of the gesture to comprehend the target word, even when speech comprehension was impaired.

Secondly, following the LRH, I predicted that PWA naming accuracy would be significantly higher in gesture permitted condition trials where PWA would have recently heard the researcher produce the word and *observed* the associated gesture, than in gesture constrained condition trials where the researcher produced the word only. I hypothesised that if the LRH was correct, the researcher's gestures would facilitate PWA lexical retrieval through a process of cross-modal priming, making the word more activated in the lexicon and therefore easier for PWA to successfully retrieve.

Thirdly, I also predicted that PWA naming accuracy would be significantly higher in gesture permitted trials where PWA themselves *produced* a gesture, as there would also be a cross-modal priming effect in these cases which would be absent in the gesture constrained condition trials.

Finally, although lesion data was not available for all PWA, it was available for a subset (see Section 2.4.3 for details of neuroimaging data acquisition). Lesion analyses must be considered exploratory given the small amount of data available and therefore I did not make *a priori* predictions for correlations between lesion site and gesture observation/production effects in PWA.

Prediction 1: PWA will perform significantly better on speech comprehension measures when they are able to observe the researcher's gestures alongside speech, than when the researcher does not use gestures.

Prediction 2a: PWA will perform significantly better on naming accuracy measures when they are able to observe the researcher's gestures alongside speech, than when the researcher does not use gestures.

Prediction 2b: PWA will perform significantly better on naming accuracy measures when they are able to produce gestures alongside speech, than when their gesture use is constrained.

2.4 Methods

2.4.1 Participant sample

I recruited 36 adult participants (29 males, 7 females) from the MRRI Research Registry, the UCL Communication Clinic, and London-based stroke/aphasia-support organisations. All participants had a clinical diagnosis of aphasia made by a neurologist and/or Speech-Language Pathologist/Therapist. Neuroimaging data was available for 19 participants (see Section 2.4.3 for further information). Table 1 presents demographic information for participants. Age ranged from 42 to 83 years (mean = 57.51, SD = 11.09), education level from 10 to 19 years (mean = 15.0, SD = 2.0), and participants were 12 to 272 months post-stroke (mean = 102.0, SD = 58.0).

Inclusion criteria for PWA were: (i) at least one left hemisphere lesion caused by stroke; (ii) a diagnosis of aphasia on standardised clinical tests made by a qualified speech and language therapist; (iii) at least 12 months post-stroke; (iv) fluent English speaker; (v) no history of other neurological conditions; (vi) no visual and/or hearing deficits that could not be corrected to normal levels; (vii) no visual field or sensory perceptual deficits; (viii) no severe acquired apraxia of speech or severe upper limb apraxia as established via standardised clinical tests; (ix) no other significant cognitive deficits.

Table 1. *Participant demographic data.*

Participant	Sex	Age at stroke (y;m)	Months post-stroke at testing	Age at testing (y;m)	Handedness (pre-stroke)	Hemiplegia	Years of education	Aphasia sub-type	WAB-R Aphasia Quotient
2	M	36;10	65	42;4	Right	Right	17	Anomic	97.0
3	M	59;6	27	61;9	Right	Right	18	Conduction	82.0
4	M	49;8	51	54;0	Right	None	13	Conduction	81.5
5	F	64;10	44	68;6	Left	Right	13	Conduction	80.3
6	M	62;4	57	67;1	Left	Right	17	Anomic	84.3
8	M	50;2	130	61;0	Right	Right	16	Conduction	57.0
9	M	69;7	83	76;1	Right	None	16	Anomic	95.8
10	M	62;4	139	73;11	Right	None	17	Anomic	84.2
11	M	65;2	216	83;3	Right	None	17	Anomic	62.1
12	M	53;6	85	60;7	Right	None	16	Anomic	94.0
13	M	64;0	113	73;8	Right	Right	13	Anomic	78.6
14	M	73;3	48	77;3	Right	None	16	Conduction	69.0
17	M	35;10	103	44;5	Right	Right	16	Anomic	85.7
18	M	68;2	47	72;1	Right	None	13	Anomic	95.2
19	M	34;3	43;3	120	Right	Right	17	Anomic	88.9
20	F	62;10	75	69;2	Right	Right	16	Conduction	59.0
22	F	48;4	132	59;11	Right	Right	17	Anomic	73.1
25	M	56;8	142	68;6	Right	None	13	Broca’s	51.5
27	F	51;4	101	53;4	Right	Right	13	Anomic	86.3
28	M	48;4	117	57;3	Right	Right	17	Broca’s	68.9
29	F	54;0	43	57;7	Right	None	17	Conduction	78.4
30	M	31;6	152	44;2	Right	Right	16	Anomic	91.2
32	M	54;6	87	61;9	Right	None	16	Conduction	77.5
33	M	53;7	105	62;4	Right	Right	18	Anomic	93.4

34	F	39;0	92	46;8	Left	Right	13	Anomic	80.0
S01	M	57;2	60	62;2	Right	Right	16	Conduction	79.2
S02	M	51;3	135	62;6	Right	None	18	Mixed	87.6
S03	M	58;5	25	60;6	Right	Right	16	Anomic	82.9
S04	M	54;4	272	77;1	Right	Right	18	Broca's	51.6
S05	M	64;9	12	65;9	Right	None	16	Anomic	94.6
S06	M	65;11	161	79;4	Right	None	16	Wernicke's	51.8
S07	M	48;4	131	59;3	Right	Right	10	Anomic	81.3
S08	M	51;5	18	67;1	Right	None	12	Anomic	83.0
S09	M	47;9	217	65;9	Right	Right	14	Conduction	78.1
S10	F	53;6	52	57;10	Right	Right	19	Wernicke's	77.1
S11	M	73;7	84	80;7	Right	Right	18	Anomic	92.2
S12	M	65;9	55	70;4	Right	Right	12	Anomic	88.7

2.4.2 Background test battery

To obtain a baseline overview of cognitive and linguistic functioning, a comprehensive background test battery was completed for each PWA. This battery measured broad aphasia severity; confrontation naming abilities; phonological processing; object, action and lexical semantic processing; short-term and working memory/executive function; spatial relationships, functional communication; and the presence and severity of upper limb apraxia. Full details of PWA performance on these tasks may be found in Appendix 1.

Table 2. *Background battery assessments.*

Assessment	Focus
Western Aphasia Battery-Revised (WAB-R) (Kertesz, 2006)	Single word, sentence and discourse comprehension and production, aphasia severity.
Camel & Cactus Test (Bozeat et al., 2000)	Object semantics
Kissing & Dancing Test (Bak & Hodges, 2003)	Action semantics
Synonymy Triplets Test (Saffran et al., 1989)	Lexical semantics
Object and Action Naming Battery (OANB) (Masterson & Druks, 1998)	Object and action naming
Middleton Lab Naming Battery (Middleton et al., 2016)	Object naming
Oxford Cognitive Screen (Demeyere et al., 2015)	Cognitive functioning
Digit span tests (Wechsler, 1997)	Short-term memory, working memory/executive function
The Token Test (De Renzi & Vignolo, 1962)	Spatial relationships
The Scenario Test UK (Hilari et al., 2018)	Functional communication
MRRI Apraxia Battery (Tarhan et al., 2015)	Gesture recognition, gesture production (gesture-to-sight, praxis)

2.4.3 Neuroimaging data

Where available, structural MRI scans were acquired for a subset of participants (n = 19). This included 7 PWA recruited from the US and 12 PWA from the UK. Scan data for participants in the UK cohort was taken from the Predicting Language Outcome and Recovery After Stroke (PLORAS) database (Seghier et al., 2016); scan data for participants from the US was taken from the MRRI Research Registry neuroimaging archive. In both cases T1-weighted anatomical whole-brain scans were acquired for all participants. For the UK participants T1-weighted images were analysed using an automated lesion identification toolbox developed by Seghier and colleagues (2008) which generated a binary lesion image for each PWA. For the US-cohort a similar process was followed; lesions were automatically segmented from T1-weighted images using LINDA (Pustina et al., 2016). Following acquisition, all binary lesion images were normalised in MNI space for comparison (Evans et al., 1993). To ensure accuracy of the automated lesion segmentation process all binary lesion images were subsequently reviewed by a neurologist for correctness.[3]

2.4.4 Materials

Materials consisted of:

- Two sets of stimulus playing cards depicting objects or actions, individualised for each PWA (Fig. 1). Each playing card set consisted of at least 4 cards showing objects/actions that the participant had correctly named during baseline testing and 16-20 cards that they had been unable to name correctly. Cards were individualised for each participant based on his/her performance on the OANB (Masterson & Druks, 1998) and Middleton Lab Naming Battery (Middleton et al., 2016). Correctly named cards were included to establish a baseline for PWA performance

[3] Professor Alex Leff (UCL Queen Square Institute of Neurology) and Professor H. Branch Coslett (Penn Medicine and University of Pennsylvania) reviewed binary lesion images for the UK and US participants respectively.

and to reduce frustration on the part of the participant. To the extent possible, cards were balanced across participants for the number of object and actions words and for the number of semantic and phonological errors made during background testing. Cards were randomly ordered in each set and participants were free to shuffle the cards.

- Additional playing cards (as many as required) for teaching gameplay. These cards were drawn at random from the remaining pictures in the OANB (Masterson & Druks, 1998) and Middleton Lab Naming Battery (Middleton et al., 2016). Familiarisation cards were identical in format to those used during the experimental task.

Fig. 1. *Sample stimulus cards: diving (action), dripping (action), pyramid (object)*

2.4.5 Procedure

2.4.5.1 Room layout

Depending on PWA preferences and mobility requirements experimental testing took place at participants' homes or in accessible testing suites at UCL or MRRI. Prior to starting the task, the researcher and PWA were seated at right angles to each other in a quiet, well-lit room, with a second researcher positioned to the side of the PWA and able to observe the interaction. Two video cameras filmed interactions, with one camera facing the participant and the other positioned to capture the full interaction space. Both the researcher and the participant wore clip-on microphones to facilitate

audio recording. In line with UCL and MRRI's COVID-19 guidelines researchers wore clear face masks or clear visors throughout interactions; participants were not required to wear masks although could choose to do so if they wished. To counteract fatigue, participants were permitted to take breaks at any point during the task, as often as required.

2.4.5.2 Gameplay

Participants completed (up to) 90 rounds of the gamified "Go Fish!" task with the researcher. Prior to starting the experimental task participants were taught how to play the game using practice cards, for at least three practice trials. There was no upper limit set on the number of practice trials and participants could choose to complete additional practice trials if needed to ensure they were fully familiar with gameplay. Experimental testing commenced at the point that the PWA indicated that s/he was happy to begin. Participants were informed that the aim of the game was: "To collect matching pairs of cards." To motivate participation, participants were also informed that they would be playing against the researcher and the player with the largest collection of matching pairs at the end of the game would be the winner.

At the start of the game, the researcher and participant were dealt a 'starter hand' of 7 random playing cards, with the remainder of the cards placed face down in a pile in the centre of the table, within easy reach of each player. The researcher and participant proceeded to take turns requesting additional cards from each other to collect pairs of matching cards (using the carrier phrase "Do you have [*target*]?"). If the researcher or participant did not have the requested card, s/he was instructed to say "Go fish!" and take a card from the centre pile to add to her/his hand. Participants were informed that the game would formally end when all cards from the centre pile had been removed and played.

Barriers were set up between the researcher and participant to ensure that each could not see the other's cards. Participants were encouraged to lay

their cards flat on the table behind the barrier during the game, to reduce any negative impact on participants with hemiplegia who may have struggled to hold the cards and to enable participants to keep their hands free to gesture. Barriers were approximately 20 cm high and were deliberately constructed to be high enough to keep playing cards hidden and low enough to enable the researcher and participants to see any gestures produced over the top.

In half of trials both the researcher and PWA were free to use gestures while requesting a card (gesture permitted condition). In this condition the researcher consistently produced a meaningful iconic gesture semantically aligned to the target word in each trial; participants were not specifically asked to gesture and were free to choose in each trial. In the remaining half of trials gesture use was constrained. Gesture use/constraint was blocked and counterbalanced across participants. Following gesture-constraint paradigms developed by Pulvermüller and colleagues, physical restraints to discourage gesturing were not used; participants were constrained from gesturing through explicit instruction and reinforcement between trials (Pulvermüller et al., 2001). In the gesture constrained condition, the researcher and participant were required to keep their hands flat on the table in front of them. The second researcher provided reminders of this instruction as required to the participant. To minimise participant fatigue, a maximum of 30 minutes were spent playing the game in each condition with a minimum 10 minute break between conditions. Participants were free to have additional/longer breaks as needed.

Fig. 2. *Go Fish gameplay. Researcher is describing "pyramid" and gesturing the shape.*

Fig. 3. *Go Fish gameplay. Participant is describing "snake" and gesturing how it moves.*

2.4.5.3 Gesture norming

To ensure consistency in data collection between UK and US sites, UK and US researchers agreed on a single set of (iconic) gestures that were used consistently during game trials. Gesture norming to establish iconicity of the gestures was carried out online using the Prolific platform with 30

neurologically healthy adults who viewed short video clips of the researchers producing each gesture; participants were then asked to rate on a 1-7 Likert scale how well each gesture matched a given target word. Norming scores for each gesture produced may be found in Appendix 2. Norming scores were also entered as control variables in each model.

2.5 Data Analysis

Logistic and linear mixed-effects regression analyses were performed using the *lme4* package (Bates et al., 2015) in RStudio (RStudio Team, 2020). The primary dependent variables were PWA comprehension accuracy and naming accuracy, with gesture condition as the main experimental manipulation and variable of interest. Logistic models were fit with the Laplace approximation with a binomial distribution and the *logit* link function. The *bobyqa* or *Nelder-Mead* optimisers were used to ensure model convergence. Working memory/executive function score (i.e. WAIS-III Backwards Digit Span Wechsler, 1997) was log transformed for model stability, to reduce skew and ensure model assumptions were met. Target word lengths were automatically calculated in Python and target word frequencies were drawn from Subtlex-UK (van Heuven et al., 2014) and Subtlex-US (for US particiants) (Brysbaert & New, 2009). Continuous variables were scaled and centred and categorical variables were orthogonally coded to test for effects. We determined that there was no significant multicollinearity in the models by calculating the variance inflation factor of all predictors using the *vif* function in the R *car* package (all variance inflation factors < 2.50) (Fox & Weisberg, 2019). Where additional packages were used for analyses these are indicated with the model details below. Full R code for all models below may be found in Appendix 9.

2.5.1 Gesture coding

Detailed coding of researcher gestures was not carried out as part of this study, although all gestures used by the researcher were iconic and subject to norming as described above. Coding of gestures produced by PWA was

carried out in *ELAN* v6.7 (Lausberg & Sloetjes, 2009) (https://archive.mpi.nl/tla/elan) by the researcher and five trained research assistants. Coders used images from two camera angles (Camera A directly facing the PWA and Camera B capturing the entire interaction space) to identify and code gestures. Broadly, the following gesture types were coded: iconic and metaphoric gestures (representational gestures), points/deictic gestures, emblems, pragmatic gestures and other gestures. Coders used an adapted version of the *MultiModal MultiDimensional Labeling System* (M3D) developed by Rohrer (2022) to arrange coding tiers. Descriptions of each gesture type and full details of the coding scheme may be found in the Coding Manual in Appendix 3. Where coders could not decide on the coding of a specific gesture, consensus meetings were held to come to agreement.

To ensure reliability of PWA gesture type coding, a proportion of video files were selected at random for recoding by a trained researcher unrelated to the project (equating to approximately 11.0% of the total data collected). Videos were fully recoded for gesture presence and type. I then calculated Cohen's kappa to determine inter-rater reliability for gesture type coding (McHugh, 2012). I found that there was strong agreement between raters' judgements, $\kappa = 0.89$, indicating that gesture type coding was robust.

2.5.2 Speech response transcription and error coding

PWA verbal responses were transcribed into International Phonetic Alphabet (IPA) form for scoring. Responses were scored as 1 (correct) or 0 (incorrect). Participant responses were scored as correct if they produced the target word correctly, produced a British English synonym of the target word (e.g. "spectacles" instead of "glasses", "cockerel" instead of "rooster"), produced an informal/slang name for the target word (e.g. "brolly" instead of "umbrella", "fridge" instead of "refrigerator"), or produced a non-British English dialectal equivalent name (e.g. "eggplant" instead of "aubergine", "ladybug" instead of "ladybird"). Responses with incorrect 0 scores were also allocated to an error category: semantic (meaning error), phonemic (sound error with at least 50% of phonemes in common with the target word), mixed (semantic and

phonemic error), omission (no name produced or explicit indication that the participant was unable to produce the name, e.g. “I can’t say that one”, “I don’t know it”, etc.) or unrelated (no clear semantic or phonemic relation to the target word).

2.5.3 Statistical modelling strategy

2.5.3.1 Data preparation

Distribution of all continuous variables was visually examined by the researcher via histograms prior to model inclusion. Variables (raw scores) were scaled and centred prior to model inclusion to ensure they were of comparable magnitude. In order to ensure there was a good spread of scores across participants (i.e. capturing the full range of individual variability in responses), outliers were not removed but where variables were not normally distributed and there was significant right or leftwards skew, variables were log-transformed to improve model stability and meet model assumptions.

2.5.3.2 Fixed variable selection and random effects selection

Fixed variables were selected based on the literature and previous research reviewed in Chapter 1 and in line with the hypotheses being tested. Therefore, in each chapter of this thesis the models vary depending on which variables are hypothesised to be relevant. Random effects were also included in models where appropriate to account for individual baseline differences (random intercepts, e.g. 1 | Participant) and to allow the effect of variables to vary across participants (random slopes, e.g. Participant | Condition). The contribution of random effects to model fit was assessed using adjusted R^2 . Random effects were removed if there was no improvement in adjusted R^2.

2.5.3.3 Dealing with multicollinearity

Multicollinearity of variables in the model was assessed by examining the Variance Inflation Factor (VIF) using the *vif* function in R. Where variables were found to be highly multicollinear (i.e. VIF > 3.0) they were either:

(1) Removed from the model where an alternative variable was available and explained more variance, calculated by examining the adjusted R^2 of the model with variables systematically included and excluded.
(2) Combined to produce a composite variable following a Principal Component Analysis approach (e.g. combining lexical semantic score, object semantic score and action semantic score into a single semantic access score).

This approach effectively reduced multicollinearity in models to acceptable levels (i.e. VIF < 3.0) and maximised the extent to which variables captured independent elements within the data. This approach was also used to reduce model complexity and manage instability, as model stability improved (as shown through iterative testing with multiple optimisers as below) when (orthogonal) variables were included with low levels of multicollinearity.

2.5.3.4 Convergence and model stability

To identify the most stable parameter estimates and check convergence, models were iteratively fitted with multiple optimisers using the *allFit* function in R. Where there were no convergence errors and parameter estimates did not vary significantly between optimisers, the default optimiser was selected (i.e. *bobyqa/Nelder-Mead*). [4]

2.5.4 Comprehension accuracy

I conducted logistic mixed effects regression analysis with comprehension accuracy (1/0) as the dependent variable. I entered two-way interactions

[4] Note that in this thesis the default optimisers were sufficient as models successfully converged and parameter estimates were essentially stable. Had this not been the case, the optimiser which achieved the highest log-likelihood (within numerical tolerance) would have been selected.

between gesture condition and other predictors. I entered gesture recognition score, target word length, target word frequency, WAB-R *Aphasia Quotient*, research site (US/UK), working memory/executive function, and semantic access as predictors into the model. I also included a random intercept of participant to account for variability in participant response to gestures.

2.5.5 Effects of gesture observation on naming accuracy

I limited analysis of the effects of gesture observation on PWA naming accuracy to a maximum of five game turns following presentation of the gesture by the researcher, to ensure that short-term effects of gesture observation were successfully captured. I entered naming accuracy as the dependent variable and gesture condition as the primary independent variable. I entered target word length, target word frequency, short-term memory, semantic access score (composite score comprising the *Synonymy Triplets Test, Camel and Cactus Test* and *Kissing and Dancing Test*), WAB-R *Spontaneous Speech* score, research site (US/UK), working memory/executive function, gesture recognition score, first experimental condition (i.e. gesture permitted/gesture constrained), back count (i.e. number of turns previously the researcher produced the gesture) and semantic access as predictors into the model, including two-way interactions between these predictors and gesture condition. I also included a random intercept of participant to account for variability in participant response to gestures.

2.5.5.1 Semantic and phonemic similarity of naming errors

For more fine-grained analysis of the effects of gesture observation on PWA speech output, I also analysed whether gesture observation differentially affected cosine similarity (i.e. semantic similarity) or phonemic similarity (phonological Levenshtein distance) of errors produced by PWA. Correct responses were excluded from this analysis. For these analyses I conducted linear mixed effects regressions with cosine similarity or phonemic similarity as the dependent variables. I included gesture condition as the independent

variable of interest and Gesture to Sight (apraxia praxis severity, part of the MRRI Apraxia Battery), first gesture condition (i.e. gesture permitted/gesture constrained), target word length and target word frequency as control variables. I also included a random intercept of participant.

2.5.6 Effects of gesture production on naming accuracy

I conducted a logistic mixed effects regression analysis with naming accuracy as the dependent variable and PWA gesture production as the primary independent variable. I entered target word length, target word frequency, WAB-R *Spontaneous Speech* score, short-term memory (forwards digit span) working memory/executive function score, first experimental condition (gesture permitted/gesture constrained), apraxia praxis severity (Gesture to Sight score) and semantic access score as predictors into the model. Note that as PWA were not specifically instructed to gesture in the gesture permitted condition, not all PWA chose to do so and therefore gesture production analyses were restricted to 32 PWA (out of a total of 36 PWA) who used gestures. I also included random intercepts by-participant and by-research site to account for individual variability in response to gestures.

2.5.6.1 Semantic and phonemic similarity of naming errors

For more fine-grained analysis of the effects of PWA gesture production on PWA speech output, I also analysed whether gesture production differentially affected cosine similarity (i.e. semantic similarity) or phonemic similarity (phonological Levenshtein distance) of naming errors produced by PWA. For these analyses we conducted linear mixed effects regressions with cosine similarity or phonemic similarity as the dependent variables and broadly the same control predictors as in the previous model above, with the addition of research site and trial order as fixed effects as required for model convergence.

2.6 Results

2.6.1 Gesture observation behavioural analyses

2.6.1.1 Comprehension

Post-hoc power analyses were conducted with the collected data to determine if it was sufficient to detect a small, moderate or large effect. Following the approach of Cabiddu et al. (2022), simulations were conducted (500 iterations) with the observed data for a range of effect sizes (range: 0.0 - 2.0); the effect of the independent variable (gesture condition) was varied across iterations while all other fixed and random effects were kept consistent. The full analysis and code adapted from Cabiddu et al. (2022) may be found in Appendix 4. This analysis indicated that the data was sufficient to detect a medium effect (effect size = 0.50; $\alpha = 0.05$; power = 0.80).

In both gesture conditions, comprehension accuracy was high (gesture constrained condition mean accuracy: 0.97; gesture present condition mean accuracy: 0.98), reflective of the fact that PWA within the sample were selected primarily for anomia rather than comprehension deficits. As shown in Table 3 and Fig. 4, contrary to Prediction 1 overall there was no significant effect of gesture observation on PWA comprehension. Although PWA comprehension scores were indeed marginally higher when the researcher produced gestures, the difference was non-significant ($\beta = 0.505$, SE = 0.394, $z = 1.283$, $p = 0.200$). However, there was a significant interaction between gesture recognition score and gesture condition, with PWA with higher gesture recognition scores deriving significantly more benefit from gestures for auditory comprehension accuracy, than PWA with lower gesture recognition scores ($\beta = 0.572$, SE = 0.292, $p = 0.049$) (Fig. 5). The control variables target word frequency and WAB-R Aphasia Quotient also significantly predicted auditory comprehension accuracy (both $p < 0.05$). Finally, there was a significant interaction between gesture condition and

target word frequency (Fig. 6), with less frequent words comprehended significantly more accurately when the researcher produced gestures.

Table 3. *Auditory comprehension model results.*

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	4.706	0.578	8.145	0.000
Gesture condition (present)	0.505	0.394	1.283	0.200
Working memory	0.195	0.394	0.495	0.620
Semantic access score	0.389	0.348	1.120	0.263
Gesture recognition test score	-0.262	0.300	-0.873	0.383
Target word frequency	0.499	0.191	2.611	0.009
Target word length	0.190	0.182	1.045	0.296
Aphasia Quotient	0.808	0.326	2.480	0.013
Site	-0.093	0.646	-0.144	0.886
Gesture condition (present) * Working memory	-0.078	0.319	-0.245	0.807
Gesture condition (present) * Semantic access score	0.482	0.355	1.357	0.175
Gesture condition (present) * Gesture recognition test	0.572	0.292	1.961	0.049
Gesture condition (present) * Target word frequency	-0.916	0.290	-3.161	0.002
Gesture condition (present) * Target word length	-0.328	0.281	-1.166	0.243
Gesture condition (present) * Aphasia Quotient	-0.340	0.312	-1.089	0.276

Fig. 4. *Effect plot: mean comprehension scores by gesture condition.*
In this figure the Y axis represents comprehension scores (0.0-1.0 scale) with the score range indicated by bars and the mean score represented by dots. Scores in each gesture condition (i.e. constrained, present) are presented separately.

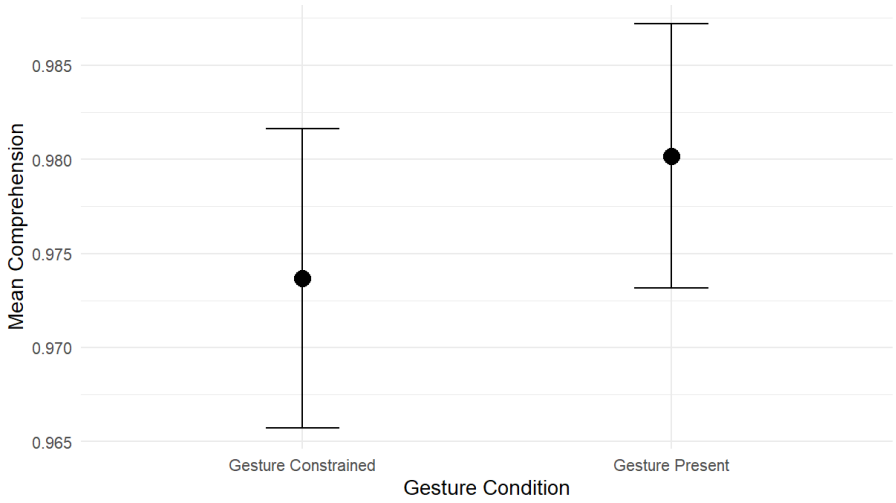


Fig. 5. *Comprehension accuracy: interaction between comprehension accuracy scores in each gesture condition and PWA gesture recognition score.*

In this figure the Y axis represents scaled comprehension accuracy and the X axis indicates scaled gesture recognition score. The red line indicates change in comprehension accuracy in the gesture present condition and the blue line indicates change in comprehension accuracy in the gesture constrained condition. The shaded areas represent confidence intervals.

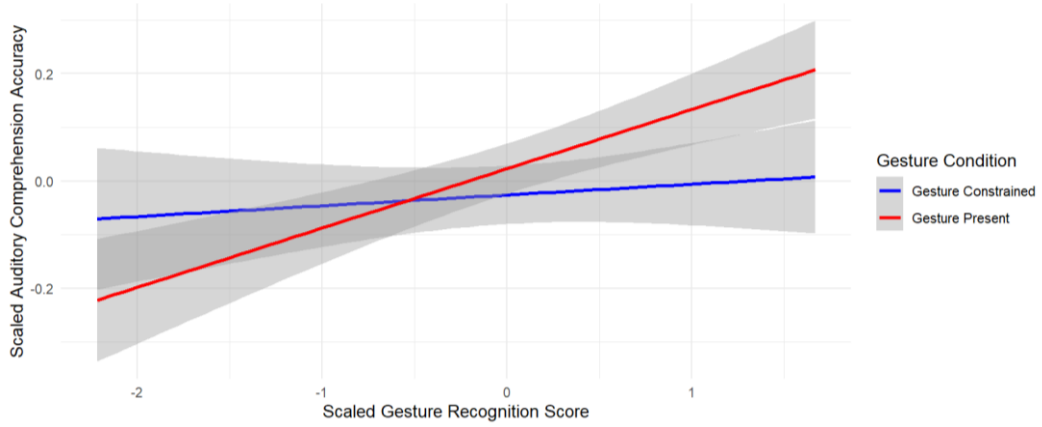
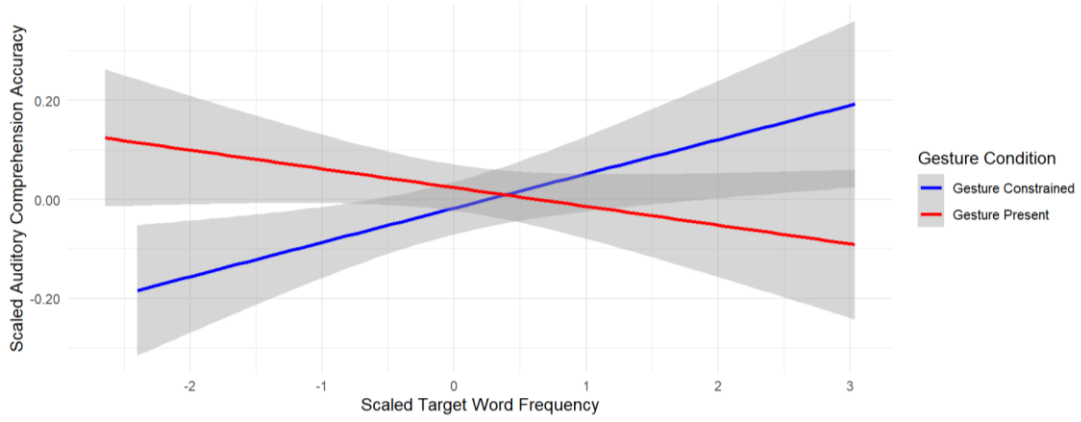


Fig. 6. *Comprehension accuracy: interaction between comprehension accuracy scores in each gesture condition and target word frequency.*

In this figure the Y axis represents scaled comprehension accuracy and the X axis indicates target word frequency. The red line indicates change in comprehension accuracy in the gesture present condition and the blue line indicates change in comprehension accuracy in the gesture constrained condition. The shaded areas represent confidence intervals.



2.6.1.2 Naming accuracy

As for comprehension, post-hoc power analyses following Cabiddu et al. (2022) were conducted with the collected data to determine if it was sufficient to detect a small, moderate or large effect. This analysis indicated that the data was sufficient to detect a medium effect (effect size = 0.45; α = 0.05; power = 0.80). See Appendix 4 for further information.

In the gesture constrained condition, the mean naming accuracy was 0.76; in the gesture present condition the mean naming accuracy was 0.80. As shown in Table 4 and Fig. 7 below, contrary to Prediction 2a, although PWA naming accuracy was marginally higher in the gesture present condition, the effect was non-significant (β = 0.053, SE = 0.277, z = 0.191, *p* = 0.848). However, I found a significant simple positive effect of working memory/executive function on naming accuracy, regardless of gesture condition (β = 0.694, SE = 0.337, z = 2.060, *p* = 0.039). In addition, shorter words were named significantly more accurately by PWA than longer words, regardless of gesture condition (β = -0.591, SE = 0.206, z = -2.872, *p* = 0.004).

I also found significant interactions between gesture observation and target word length, and between gesture observation and working memory/executive function measures. I found that gesture observation significantly moderated the word length effect in PWA, with longer words produced significantly more accurately when PWA had seen the researcher produce the gesture alongside the target word in a previous turn, than when the researcher had produced the word alone (β = 0.652, SE = 0.284, z = 2.298, *p* = 0.022). Furthermore, I found that PWA with poorer working memory/executive function produced target words significantly more accurately when they had previously seen the researcher produce a semantically-aligned gesture within 5 previous game turns, than when no gestures were used (β = -0.963, SE = 0.385, z = -2.499, *p* = 0.012), indicating that gesture observation effect on naming accuracy was

moderated by individual differences in PWA working memory/executive functioning abilities.

Table 4. *Naming accuracy model results.*

Variable	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	1.817	0.450	4.041	0.000
Gesture condition (present)	0.053	0.277	0.191	0.848
Working memory	0.694	0.337	2.060	0.039
Short term memory	0.092	0.277	0.331	0.741
Semantic access score	-0.373	0.272	-1.375	0.169
Gesture recognition score	-0.044	0.268	-0.162	0.871
Spontaneous speech score	0.063	0.259	0.244	0.807
Back count (no. turns since card presented)	-0.020	0.132	-0.150	0.880
Target word frequency	0.420	0.189	2.226	0.026
Target word length	-0.591	0.206	-2.872	0.004
First gesture condition	0.230	0.223	1.031	0.302
Site	-0.381	0.500	-0.761	0.446
Gesture condition (present) * Working memory	-0.963	0.385	-2.499	0.012
Gesture condition (present) * Short term memory	0.415	0.357	1.163	0.245
Gesture condition (present) * Semantic access score	0.252	0.317	0.795	0.426
Gesture condition (present) * Gesture recognition score	0.201	0.303	0.661	0.509
Gesture condition (present) * Spontaneous speech score	-0.253	0.311	-0.812	0.417
Gesture condition (present) * Target word frequency	-0.194	0.297	-0.653	0.514
Gesture condition (present) * Target word length)	0.652	0.284	2.298	0.022

Fig. 7. *Effect plot: mean naming accuracy scores by gesture condition.*
In this figure the Y axis represents naming accuracy (0.0-1.0 scale) with the score range indicated by bars and the mean score represented by dots. Scores in each gesture condition (i.e. constrained, present) are presented separately.

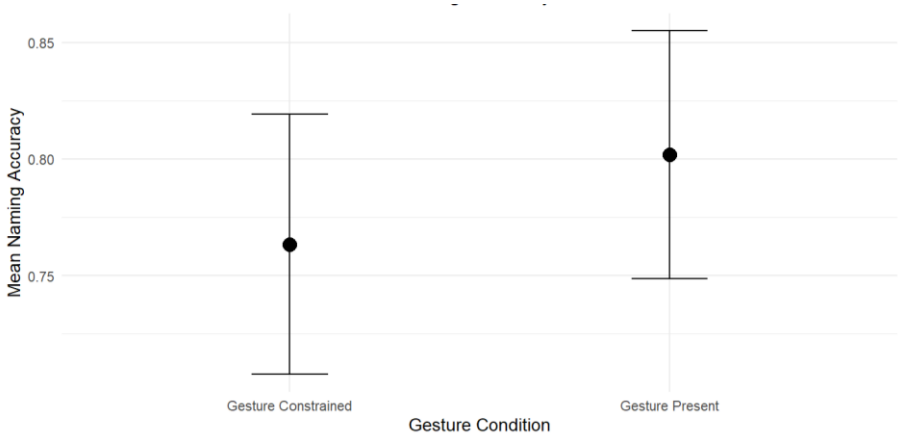


Fig. 8. *Naming accuracy: interaction between naming accuracy scores in each gesture condition and target word length.*
In this figure the Y axis indicates scaled naming accuracy score and the X access indicates target word length. The red line indicates change in naming accuracy in the gesture permitted condition, while the blue line represents change in naming accuracy in the gesture constrained condition.

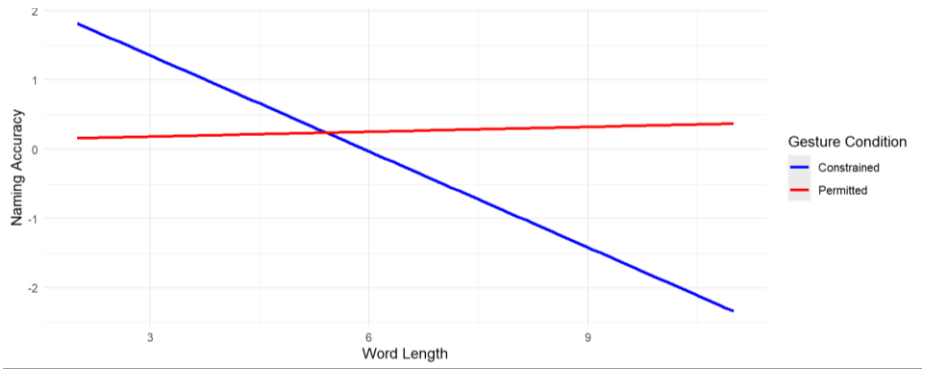


Fig. 9. *Naming accuracy: interaction between naming accuracy scores in each gesture condition and working memory/executive function score.*
In this figure the Y axis indicates scaled naming accuracy score and the X access indicates WM/EF. The red line indicates change in naming accuracy in the gesture permitted condition, while the blue line represents change in naming accuracy in the gesture constrained condition.

2.6.1.3 Gesture observation secondary analyses

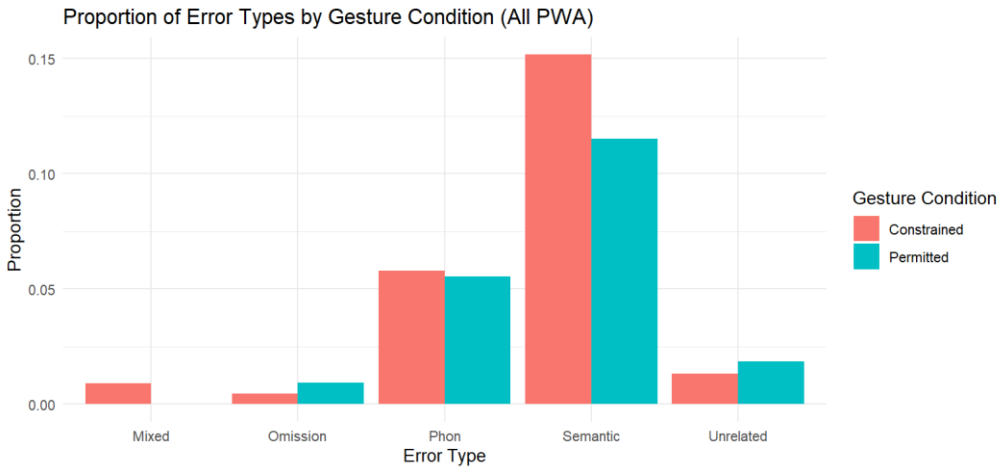
In addition to the primary analyses outlined above, I also ran several secondary analyses exploring the effects of gesture observation on naming

error type, cosine similarity of naming errors, and phonemic similarity (phonological Levenshtein distance) of naming errors. Results from these analyses are presented below.

2.6.1.3.1 Naming error type

Firstly, I explored whether the proportion of naming error types varied depending on whether PWA did, or did not, observe the researcher producing gestures alongside the target word in the preceding 5 trials. In this analysis only incorrectly named items produced by PWA were included; all correctly named items were excluded. Following Kim (2017), given the relatively small number of observations available for analysis I initially performed the Chi-Square test of approximation and then used Fisher's Exact Test to confirm the result. No significant association between error type and gesture observation was found in the Chi-square analysis ($\chi^2 = 2.88$, $p = 0.66$) or Fisher's Exact Test analysis ($p = 0.68$). Proportion of error types produced by gesture condition is shown in Fig. 10 below.

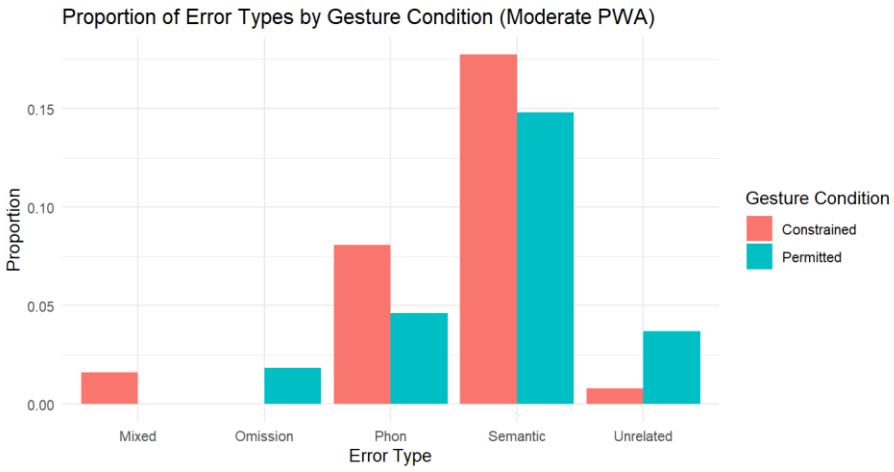
Fig. 10. *Proportion of error types by gesture condition (all PWA).*



Next, in order to explore whether this finding held true for all PWA in the sample or was driven by the relatively few naming errors produced by PWA with mild aphasia, I repeated the analysis including only trials of PWA with moderate aphasia. As before no significant association between error type

and gesture observation was found in the Chi-square analysis ($\chi^2 = 7.51$, $p = 0.08$) or Fisher's Exact Test analysis ($p = 0.11$). However, there was a trend towards significance in both analyses with relatively more phonemic naming errors produced by PWA in the gesture constrained condition and relatively more omission errors produced by PWA in the gesture permitted condition; these trends were not observed in PWA with mild aphasia. Proportion of error types produced by PWA with moderate aphasia in each gesture condition is shown in Fig. 11 below.

Fig. 11. *Proportion of error types by gesture condition (Moderate PWA).*



2.6.1.3.2 Cosine similarity of naming errors

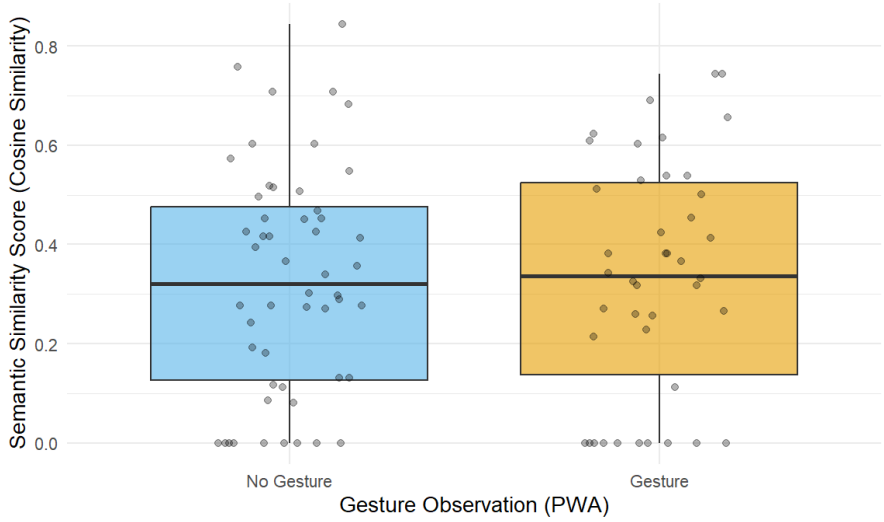
Next, I explored whether gesture observation predicted the cosine similarity (i.e. semantic distance from the target word) of naming errors produced by PWA, depending on whether PWA did, or did not, observe the researcher producing gestures alongside the target word in the preceding 5 trials. This analysis included naming errors only and correct trials were excluded from the analysis. Across 92 observations there was no significant effect of gesture observation on cosine similarity score of naming errors ($\beta = -0.004$, SE = 0.051, t(84) = -0.071, $p = 0.943$) (Table 5 and Fig. 12).

Table 5. *Cosine similarity of naming errors model results.*

	Estimate	Std. Error	df	t value	Pr(>|t|)

(Intercept)	0.340	0.036	35.963	9.413	0.000
Gesture condition (present)	-0.004	0.051	84.109	-0.072	0.943
Working memory	-0.057	0.029	20.766	-1.967	0.063
Gesture to sight test score	0.025	0.028	31.012	0.911	0.369
First gesture condition	0.014	0.028	21.191	0.509	0.616
Target word length	-0.032	0.026	84.862	-1.237	0.219
Target word frequency	0.021	0.025	84.307	0.862	0.391

Fig. 12. *Boxplot*: *effect of PWA gesture observation condition on semantic similarity (cosine similarity) score of PWA naming errors.*
In this figure the Y axis represents semantic similarity score (cosine similarity, 0.0-1.0 scale). Boxplots indicate the range of scores with the mean score shown as a line in each box. Scores in each condition are presented separately (i.e. no gesture condition, gesture condition).



2.6.1.3.3 Phonemic similarity of naming errors

Next, I explored whether gesture observation predicted the phonological Levenshtein distance (i.e. phonemic distance from the target word) of naming errors produced by PWA, depending on whether PWA did, or did not, observe the researcher producing gestures alongside the target word in the preceding 5 trials. As for the cosine similarity analysis above, this analysis included naming errors only and correct trials were excluded from the analysis. Across 90 observations I found no significant effect of gesture

observation on phonological Levenshtein distance of naming errors (β = 0.005, SE = 0.058, t(81) = -0.0923, *p* = 0.927) (Table 6 and Fig. 13).

Table 6. *Phonemic similarity of naming errors model results.*

	Estimate	Std. Error	Df	t value	Pr(>|t|)
(Intercept)	0.364	0.044	28.671	8.192	0.000
Gesture condition (present)	0.005	0.058	81.391	0.092	0.927
Working memory	0.045	0.037	16.609	1.212	0.243
Gesture to sight score	0.005	0.034	23.346	0.152	0.880
First gesture condition	0.012	0.036	17.416	0.326	0.748
Target word length	0.073	0.030	82.992	2.412	0.018
Target word frequency	0.011	0.028	78.518	0.396	0.693

Fig. 13. *Boxplot*: *effect of PWA gesture observation condition on phonemic similarity (Levenshtein) score of PWA naming errors.*
In this figure the Y axis represents phonological similarity score (Levenshtein distance, 0.0-1.0 scale). Boxplots indicate the range of scores with the mean score shown as a line in each box. Scores in each gesture condition are presented separately (i.e. no gesture condition, gesture condition).

2.6.2 Gesture production behavioural analyses

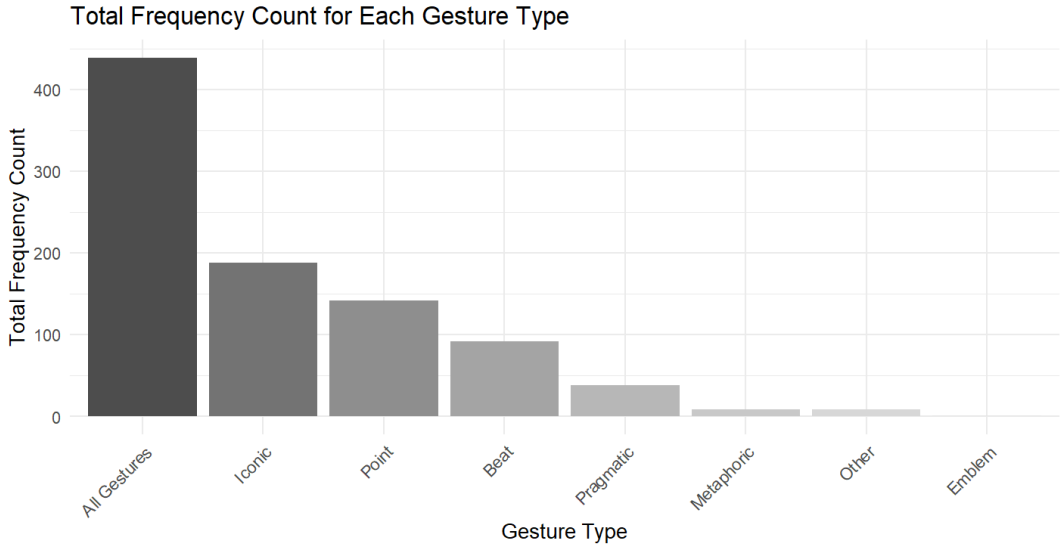
I initially calculated the frequency distribution of gestures of different types (iconic, metaphoric, emblem, pragmatic, deictic/points, other) produced by PWA during game trials in the gesture permitted condition. In total, PWA

produced 439 gestures across game turns, although there was significant individual variability (mean no. gestures produced per PWA: 12.19; range: 0.0 – 40.0). The most frequently produced gesture type was iconic, accounting for 42.84% of the total gestures produced. Points accounted for 32.34%, beats 20.95%, and all other gesture types 3.87%. See Table 7 and Fig. 14 below for frequency counts for each gesture type. To more closely represent naturalistic communication, where PWA (and neurologically healthy adults) typically use a range of gesture types alongside speech, subsequent analyses considered the effects of *all* gesture types on lexical retrieval.

Table 7. *Frequency count for each gesture type.*

Gesture type	Frequency
All gestures	*439*
Iconic	188
Deictic/point	142
Beat	92
Metaphoric	8
Other	8
Emblem	1

Fig. 14. *Frequency count for each gesture type produced by PWA.*



2.6.2.1 Naming accuracy

As for the previous analyses, post-hoc power analyses following Cabiddu et al. (2022) were conducted with the collected data to determine if it was sufficient to detect a small, moderate or large effect. This analysis indicated that the data was sufficient to detect a medium to large effect (effect size = 0.60; α = 0.05; power = 0.80). See Appendix 4 for further information.

As shown in Table 8 below, contrary to the LRH and Prediction 1b, I found a significant negative effect of gesture production on PWA naming accuracy. Overall, across 1326 observations from 32 PWA, I found that PWA produced target words significantly less accurately when they also produced a gesture (regardless of type) within the same game turn (β =−0.704, SE = 0.177, z =−3, $p < 0.001$), indicating that gesture production did not facilitate naming accuracy in this group of PWA. I also found a significant interaction between gesture condition and working memory/executive function abilities; while the cohort of PWA as a whole performed significantly worse in naming accuracy when they produced gestures, the effect was particularly pronounced for PWA with poorer working memory/executive function performing (β = −0.828, SE = 0.300, z =−2.761, p =0.006).

Table 8. *Naming accuracy model results.*

Effect	Est.	Std. Error	z-value	Pr(>|z|)
(Intercept)	1.847	0.619	2.981	0.003
Gesture produced by PWA	- 0.704	0.177	- 3.986	< 0.0001
Working memory	0.393	0.246	1.601	0.109
Short term memory	0.144	0.211	0.683	0.495
Semantic access score	- 0.103	0.194	- 0.531	0.595
Gesture to sight score	0.107	0.173	0.615	0.538
Spontaneous Speech score	- 0.002	0.168	- 0.015	0.988
First gesture condition	- 0.002	0.179	- 0.010	0.992
Target word length	- 0.233	0.092	- 2.541	0.011
Target word frequency	0.318	0.101	3.145	0.002
Gesture produced by PWA * Working memory	- 0.828	0.300	- 2.761	0.006
Gesture produced by PWA * Short term memory	0.398	0.227	1.752	0.079
Gesture produced by PWA * Semantic access	0.284	0.245	1.157	0.247
Gesture produced by PWA * Gesture to sight score	0.165	0.217	0.762	0.446
Gesture produced by PWA * Spontaneous speech score	- 0.378	0.222	- 1.707	0.088
Gesture produced by PWA * Target word length)	0.019	0.182	0.103	0.918
Gesture produced by PWA * Target word frequency)	- 0.062	0.170	- 0.363	0.717

Fig. 15. *Effect plot: mean naming accuracy scores by PWA gesture production condition.*

In this figure the Y axis represents naming accuracy (0-1.0 scale) with the score range indicated by bars and the mean score represented by dots. Scores in each condition (i.e. gesture not produced, gesture produced) are presented separately.

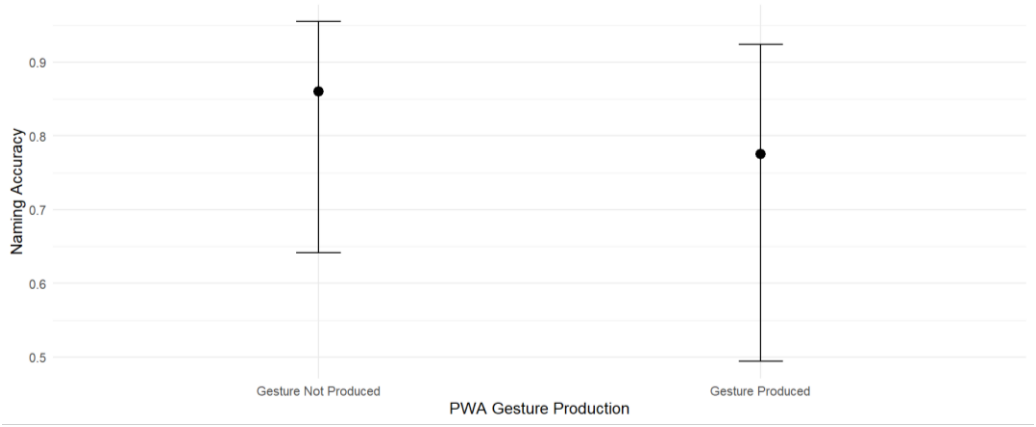
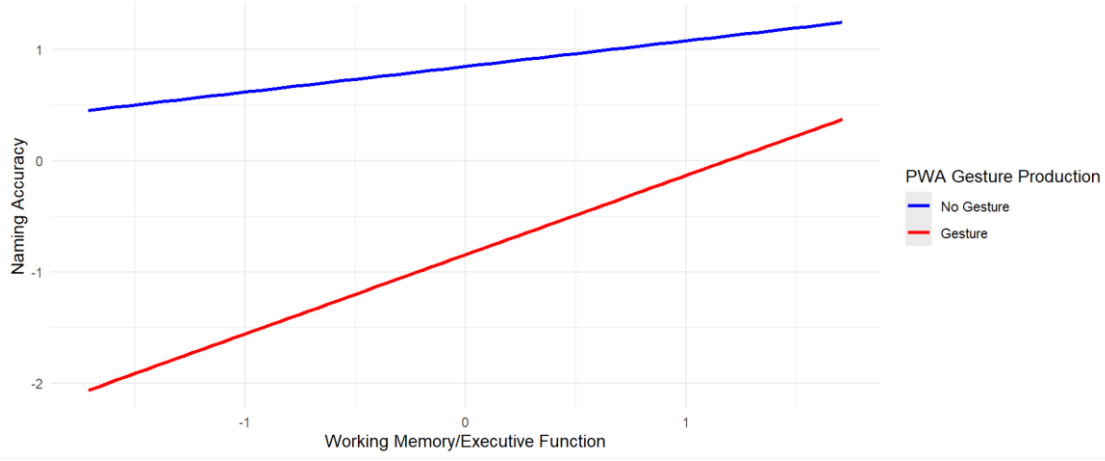


Fig. 16. *Naming accuracy: interaction between naming accuracy scores in each PWA gesture production condition and working memory/executive function score.*

In this figure the Y axis indicates scaled naming accuracy score and the X access indicates scaled WM/EF score. The red line indicates change in naming accuracy in the gesture produced condition, while the blue line represents change in naming accuracy in the no gesture produced condition.



2.6.2.2 Gesture production secondary analyses

In addition to the primary analyses outlined above, I also ran secondary analyses exploring the effects of gesture production on naming error type, cosine similarity of naming errors, and phonemic similarity (phonological Levenshtein distance) of naming errors.

2.6.2.2.1 Naming error type

Firstly, I explored whether the proportion of naming error types varied depending on whether PWA did, or did not, produce gestures alongside the target word. In this analysis only incorrectly named items produced by PWA were included; all correctly named items were excluded. As for the gesture observation analyses and following Kim (2017), given the relatively small number of observations available for analysis I initially performed the Chi-Square test of approximation and then used Fisher's Exact Test to confirm the result. I found a significant association between error type and gesture production in both the Chi-square analysis ($\chi^2 = 17.18$, $p = 0.002$) and Fisher's Exact Test analysis ($p = 0.001$).

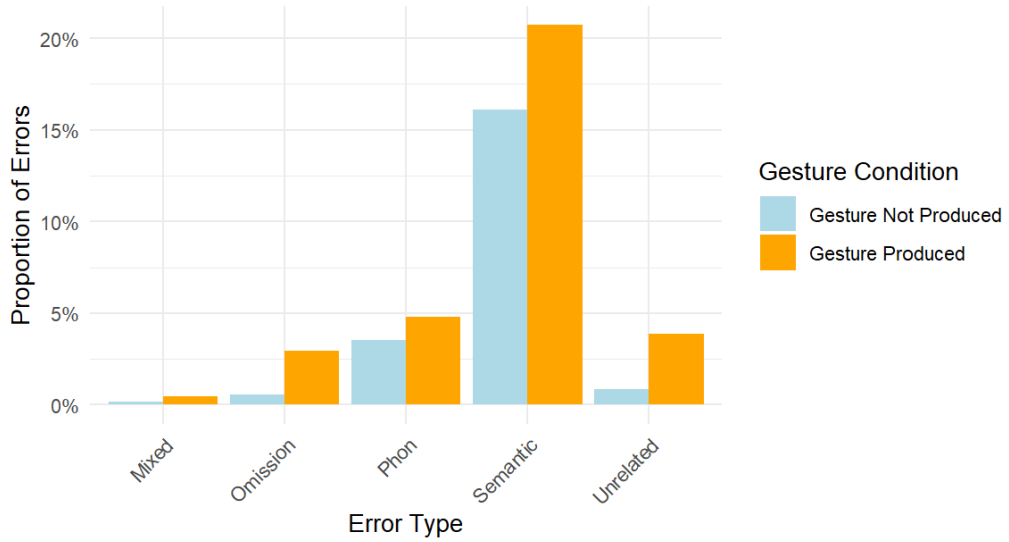
In order to determine whether specific error types were driving this result, I next performed pairwise analysis of error types by gesture condition (gesture produced vs. gesture not produced) using a two-sample test for equality of proportions. I also applied the Bonferroni correction for multiple comparisons.

Table 9. *Proportion of naming error types by gesture condition (all PWA).*

Error type	Mixed	Omission	Phonemic	Semantic	Unrelated
p value	1.0	0.002	1.0	0.19	0.0005

These results indicate that the frequency of omission and unrelated errors varied significantly across conditions, with significantly more omission and unrelated naming errors produced by PWA when they also produced gestures within the same game turn.

Fig. 17. *Proportion of naming error types by gesture condition (all PWA).*



I next explored whether this finding varied for PWA with moderate aphasia. I found a significant association between error type and gesture production in both the Chi-square analysis ($\chi^2 = 17.6$, $p = 0.001$) and Fisher's Exact Test analysis ($p = 0.001$).

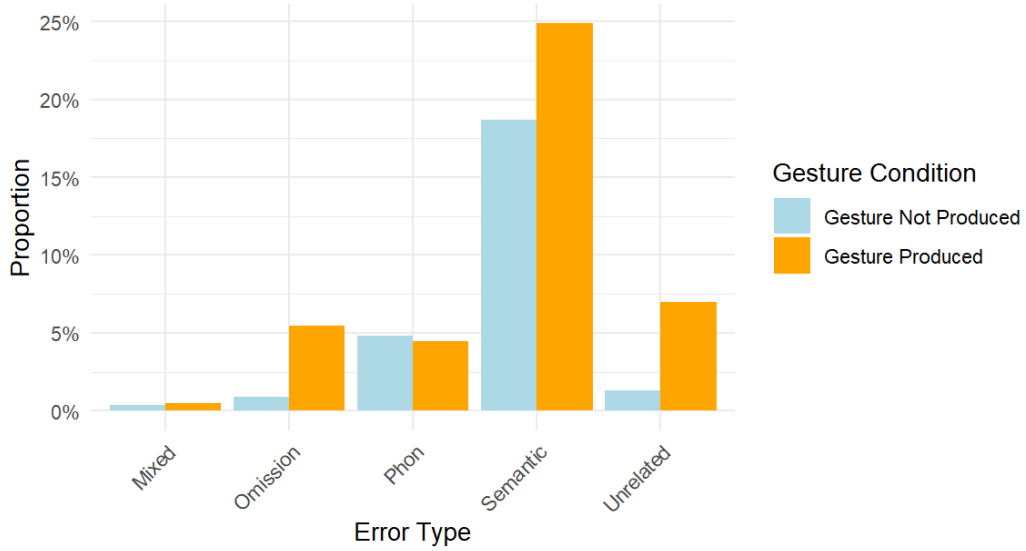
In order to determine whether specific error types were driving this result and whether results for PWA with moderate aphasia varied from results for the whole cohort, I again performed pairwise analysis of error types by gesture condition (gesture produced vs. gesture not produced) using a two-sample test for equality of proportions, with the Bonferroni correction for multiple comparisons.

Table 10. *Proportion of naming error types by gesture condition (Moderate PWA).*

Error type	Mixed	Omission	Phonemic	Semantic	Unrelated
p value	1.0	0.002	1.0	0.39	0.0005

As in the previous analysis these results indicate that the frequency of omission and unrelated errors varied significantly across conditions, with significantly more omission and unrelated naming errors produced by PWA with moderate aphasia when also produced gestures.

Fig. 18. *Proportion of naming error types by gesture condition (Moderate PWA).*



2.6.2.2.2 Cosine similarity of naming errors

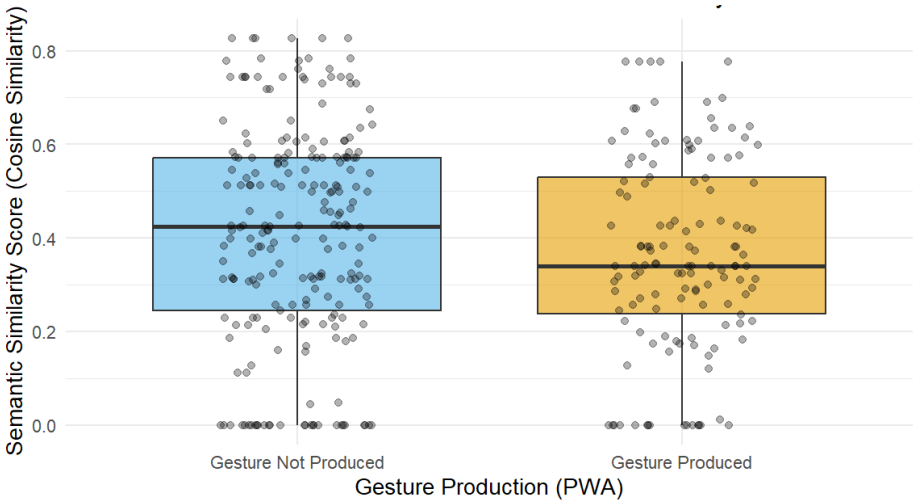
Next, I explored whether PWA gesture production predicted the cosine similarity (i.e. semantic distance from the target word) of naming errors produced by PWA in the same game turn. This analysis included naming errors only and correct trials were excluded from the analysis. As Table 11 and Fig 19. show, across 330 observations I found no significant effect of gesture production on cosine similarity score of naming errors, although there was a trend towards lower cosine similarity scores in trials where PWA produced gestures (p = 0.077).

Table 11. *Cosine similarity of naming errors model results.*

	Estimate	Std.Error	df	t_value	Pr(>|t|)
(Intercept)	0.410	0.028	23.810	14.633	> 0.0001
Gesture produced by PWA	-0.047	0.026	315.100	-1.774	0.077
Working memory	0.039	0.033	25.420	1.161	0.257
Short term memory	0.000	0.040	23.440	0.003	0.997
Aphasia Quotient	-0.048	0.057	21.000	-0.850	0.405
Semantic access	-0.054	0.032	25.340	-1.681	0.105
Gesture to sight score	0.027	0.031	19.210	0.881	0.389
Spontaneous speech score	0.032	0.042	23.560	0.753	0.459
First gesture condition	0.018	0.028	20.900	0.629	0.536
Target word length)	-0.017	0.014	316.900	-1.211	0.227

Target word frequency)	-0.017	0.014	318.700	-1.226	0.221

Fig. 19. *Boxplot*: *effect of PWA gesture production condition on semantic similarity (cosine similarity) score of PWA naming errors.*
In this figure the Y axis represents semantic similarity score (cosine similarity, 0.0-1.0 scale). Boxplots indicate the range of scores with the mean score shown as a line in each box. Scores in each gesture condition are presented separately (i.e. gesture produced, no gesture produced).



2.6.2.2.3 Phonemic similarity of naming errors

Next, I explored whether PWA gesture production predicted the phonological Levenshtein distance (i.e. phonemic distance from the target word) of naming errors produced by PWA in the same game turn. As for the cosine similarity analysis above, this analysis included naming errors only and correct trials were excluded from the analysis. As Table 12 and Fig. 20 show, across 316 observations I found no significant effect of gesture production on phonological Levenshtein distance of naming errors (p = 0.602).

Table 12. *Phonological Levenshtein distance model results.*

	Estimate	Std.Error	df	t_value	Pr(>|t|)
(Intercept)	0.332	0.032	27.287	10.521	> 0.0001
Gesture produced by PWA	-0.013	0.026	297.466	-0.522	0.602
Working memory	-0.028	0.037	29.059	-0.738	0.466
Short term memory	0.027	0.045	26.355	0.597	0.556

Aphasia Quotient	0.005	0.064	24.460	0.083	0.934
Semantic access	0.030	0.036	28.388	0.839	0.408
Gesture to sight score	-0.002	0.035	22.853	-0.062	0.951
Spontaneous speech score	0.056	0.048	26.891	1.188	0.245
First gesture condition	-0.009	0.032	24.054	-0.267	0.791
Target word length)	0.055	0.013	298.729	4.111	> 0.0001
Target word frequency)	0.012	0.013	302.778	0.889	0.375

Fig. 20. *Boxplot: effect of PWA gesture production condition on phonemic similarity (Levenshtein) score of PWA naming errors.*

In this figure the Y axis represents phonemic similarity score (Levenshtein distance, 0.0-1.0 scale). Boxplots indicate the range of scores with the mean score shown as a line in each box. Scores in each gesture condition are presented separately (i.e. gesture produced, no gesture produced).

2.6.3 Descriptive lesion analyses

Below are presented results from lesion analyses for a subset of PWA for whom structural scans were available. Lesion overlap maps for PWA are presented to demonstrate lesion locations, followed by lesion subtraction maps indicating areas of lesion differentiation between PWA who showed gesture observation benefit/detriment for naming accuracy, and PWA who showed gesture production benefit/detriment for naming accuracy.

2.6.3.1 Gesture Observation Benefit Index for Naming (GOBI)

To explore which brain regions were associated with researcher gesture production (i.e. PWA gesture observation) benefit for naming accuracy, I first generated a Gesture Observation Benefit Index (GOBI) score for participants for whom imaging data was available. The GOBI was generated by subtracting the % naming accuracy in trials where PWA observed gestures (i.e. all gesture permitted trials) from the % naming accuracy in trials where PWA did not observe gestures (i.e. all gesture constrained trials), yielding a positive GOBI score where PWA naming accuracy was higher in gesture permitted trials, or a negative GOBI score where PWA naming accuracy was higher in gesture constrained trials. PWA with a GOBI score of 0 (i.e. indicating no difference in naming accuracy between gesture conditions) were removed from subsequent analyses, resulting in a final set of 14 PWA available for analysis.

Table 13. *Participant GOBI scores.*

Participant ID	Go Fish GOBI Score
P10	4
P11	-11
P13	17
P25	27
P27	25
P3	-25
P30	-14
P32	-25
P34	-29
S01	-100
S02	33
S04	7
S08	33
S11	100

Coverage maps were generated to visualise lesion location for all PWA included (see Figs. 21 and 22) using MRICron (Rorden & Brett, 2000). Next, subtraction maps were generated to identify lesioned voxels in PWA

correlated with: (1) positive GOBI scores; and (2) negative GOBI scores (see Figs. 23 and 24). Colours below each lesion map indicate the number of PWA with lesions in each area.

Fig. 21. *Positive GOBI coverage maps. Indicating lesion locations of PWA with positive GOBI scores.*

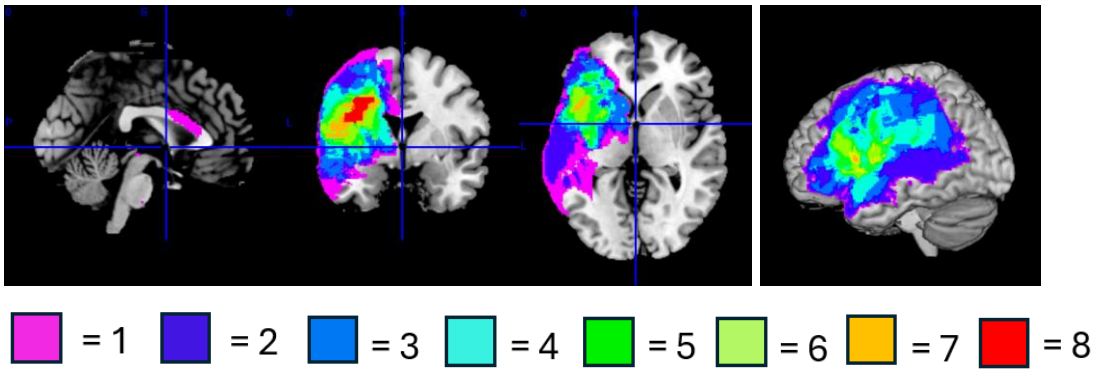


Fig. 22. *Negative GOBI coverage maps. Indicating lesion locations of PWA with negative GOBI scores.*

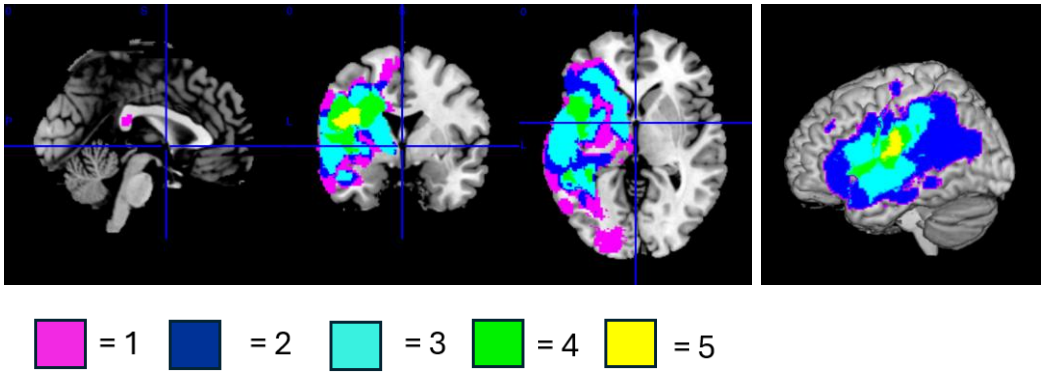


Fig. 23. *GOBI Benefit – Deficit subtraction maps.*

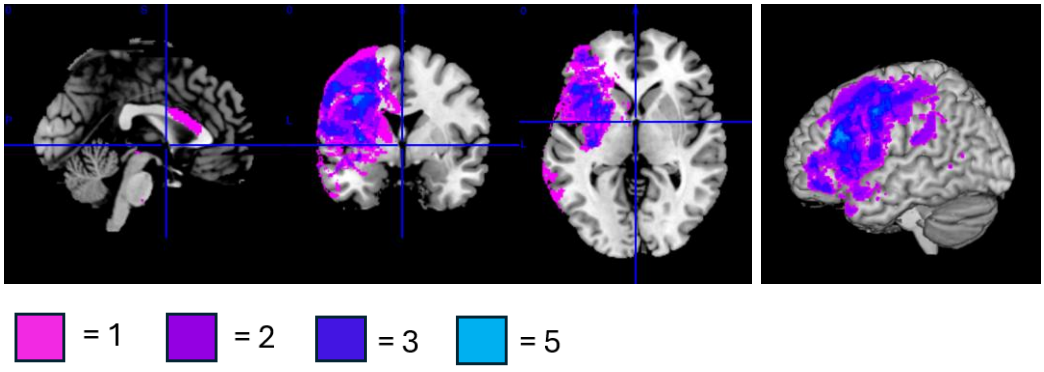
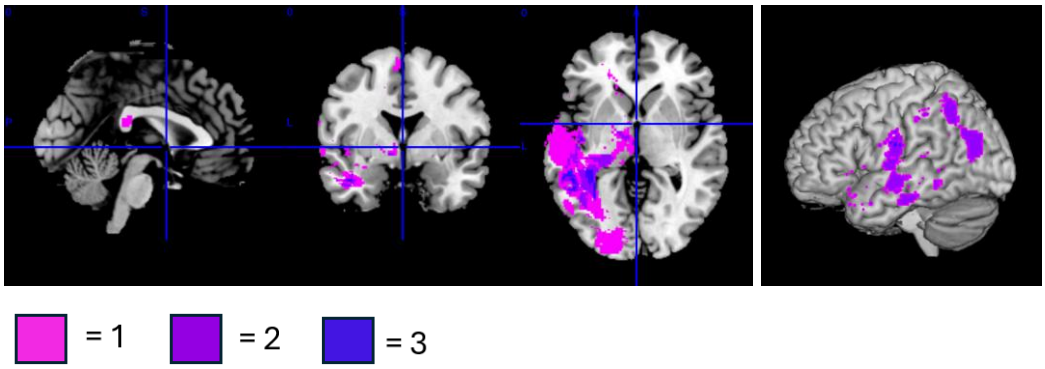


Fig. 24. *GOBI Deficit – Benefit subtraction maps.*



Due to the small amount of scan data statistical analysis was not possible, however, by observing broad patterns of lesions across PWA I found indications that PWA with positive GOBI scores (indicating gesture observation benefit for naming) presented with lesions clustered anteriorly along the dorsal pathway for language processing (including the IFG and anterior portions of the superior temporal lobe), with lesions also to the dorsolateral prefrontal cortex. In contrast, PWA with negative GOBI scores (indicating gesture observation detriment for naming) presented with heterogenous lesions clustered posteriorly, overlapping the ventral pathway for language processing (including the angular gyrus, middle and inferior temporal gyri).

2.6.3.2 Gesture Production Benefit Index for Naming (GPBI)

To explore whether PWA gesture production facilitated naming accuracy in PWA I also generated a Gesture Production Benefit Index (GPBI) score for participants. The GPBI was generated by subtracting the % naming accuracy in trials where PWA produced gestures (i.e. gesture permitted trials) from the % naming accuracy in trials where PWA did not produce gestures (i.e. gesture constrained trials), yielding a positive GPBI score where PWA naming accuracy was higher in gesture permitted trials, or a negative GPBI score where PWA naming accuracy was higher in gesture constrained trials Note that as some participants did not produce any gestures at all during gesture permitted trials, these PWA were not allocated a GPBI score and were excluded from further analyses. PWA with a GPBI score of 0 (i.e.

indicating no difference in naming accuracy between gesture conditions) were also removed from subsequent analyses, resulting in a final set of 15 PWA available for analysis.

Table 14. *Participant GPBI scores.*

Participant ID	Go Fish GPBI score
P10	-24
P11	-23
P13	-2
P2	23
P25	-55
P27	-34
P3	-8
P30	20
P32	-17
P33	5
P34	-6
S01	2
S04	-44
S07	-11
S10	6

As before, coverage maps were generated to visualise lesion location for all PWA included (see Figs 25 and 26). Next, subtraction maps were generated to identify lesioned voxels in PWA correlated with (1) positive GPBI scores; and (2) negative GPBI scores (see Figs. 27. and 28.).

Fig. 25. *Positive GPBI coverage maps. Indicating lesion locations of PWA with positive GPBI scores.*

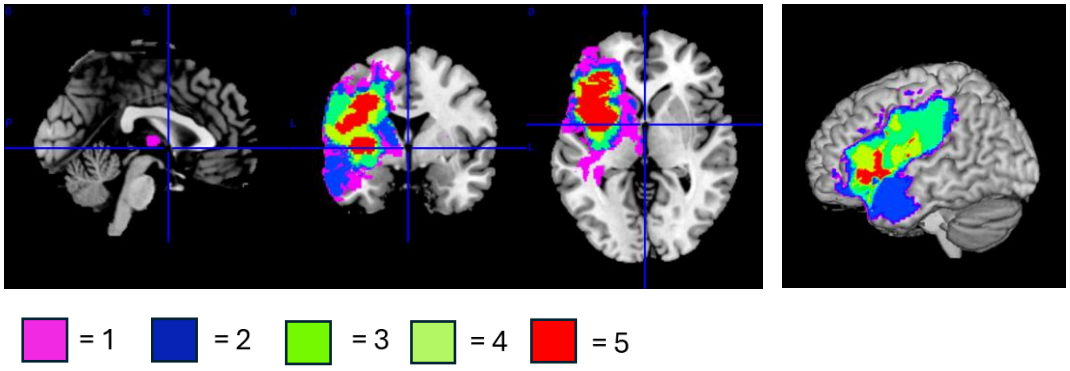


Fig. 26. *Negative GPBI coverage maps. Indicating lesion locations of PWA with negative GPBI scores.*

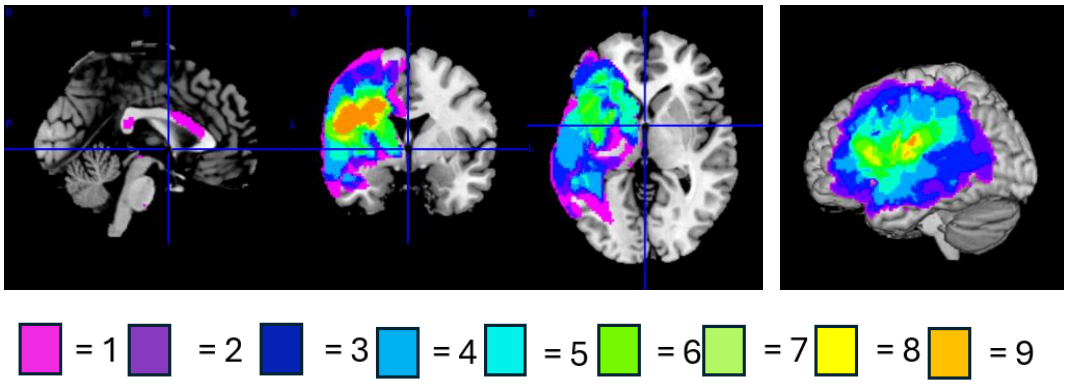


Fig. 27. *GPBI Benefit – Deficit subtraction maps.*

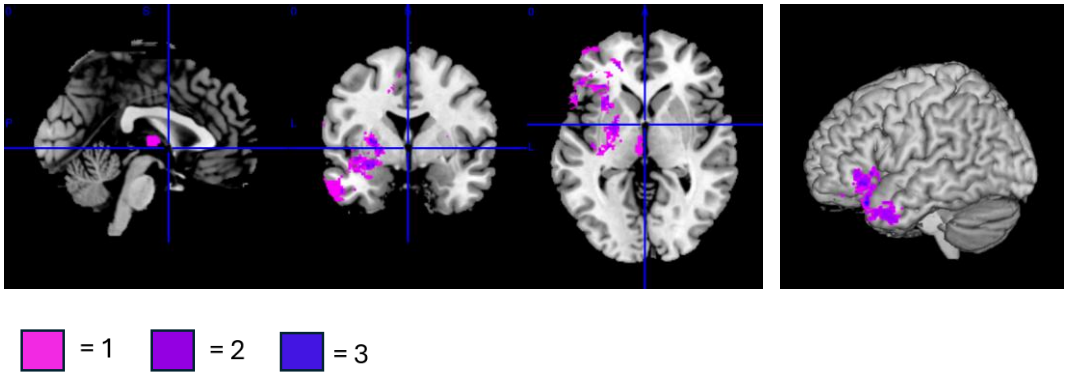
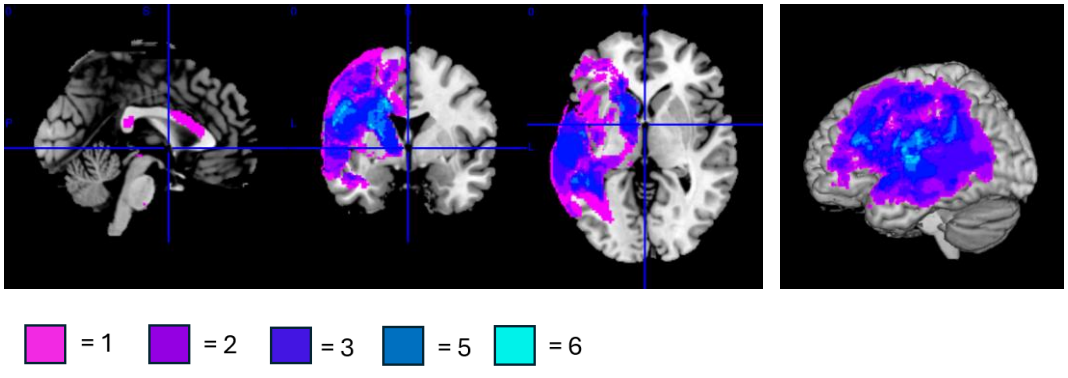


Fig. 28. *GPBI Deficit – Benefit subtraction maps.*



I found indications that PWA with negative GPBI scores (indicating gesture production detriment for naming) presented with diffuse lesions broadly distributed across both the dorsal and ventral language processing pathways and the dorsolateral prefrontal cortex but with the IFG largely spared. In

contrast, PWA with positive GPBI scores (indicating gesture production benefit for naming) presented with lesions clustered at the IFG and temporal pole.

2.7 Discussion

In this study, I asked 36 PWA to play an interactive picture naming language-action game, “Go Fish!”, with researchers at MRRI (US) or UCL (UK). Players took turns to collect matching pairs of cards by naming cards in their hand and asking the researcher for cards they needed to make pairs (naming accuracy dependent variable). Participants also needed to respond accurately to the researcher’s requests for cards in turn, by handing over the requested cards (speech comprehension dependent variable). In half of game turns, PWA and the researcher were free to use gestures alongside speech when asking for cards; in the other half, gesture use for both players was constrained.

I first investigated whether PWA gesture observation facilitated speech comprehension accuracy; i.e. when the researcher produced iconic gestures when making requests for cards, did this result in improved comprehension for PWA? I found no significant main effect of gesture observation on comprehension accuracy, although there was a significant interaction between gesture condition and gesture recognition score, with PWA with higher gesture recognition scores deriving more benefit from gestures than PWA with lower gesture recognition scores.

I next investigated whether PWA gesture observation facilitated naming accuracy, as proposed by the LRH. Although there was no significant effect of gesture observation alone, I found a significant interaction between gesture observation and target word length, indicating that when PWA heard longer words (> 5 phonemes) presented alongside semantically-aligned gestures, they produced these words more accurately in subsequent turns of the game (up to a maximum of 5 game turns), than when these words were presented without gestures. In other words, gesture presence moderated the

word length effect in these participants. I also explored whether any neurocognitive variables predicted gesture observation benefit for naming accuracy. I identified that PWA with poorer WM/EF function benefited from gesture observation for naming, but PWA with better WM/EF scores showed no such benefit.

Finally, I explored whether gesture production by PWA facilitated their own naming accuracy, as proposed by the LRH. Overall, I found a significant negative effect of gesture production on naming accuracy; PWA who produced gestures during the same game turn were significantly less likely to name the target word accurately. I also identified a dissociation with gesture observation effects on naming accuracy, in that PWA with poorer WM/EF showed a particular *disbenefit* of gesture production for naming accuracy. In addition, I found that the proportion of naming error types varied as an effect of gesture production, with PWA producing proportionally more omission and unrelated naming errors when they produced gestures, than when they did not produce gestures.

2.7.1 Theoretical implications

Results from this study shed light on the functions of gesture in interactions between PWA and neurologically healthy adults. Firstly, although (iconic) gestures produced by researchers did not facilitate PWA speech comprehension across the board, PWA who were better able to recognise the researcher's gestures did benefit. Although gesture comprehension in PWA is historically understudied, iconic gestures have previously been shown to benefit comprehension in both neurologically healthy and clinical populations (e.g. Wang et al., 2004). However, the mechanism by which this occurs is not fully understood and indeed may differ depending on the population. For example, although Wang and colleagues demonstrated that children with ADHD showed improved comprehension when teachers used speech and iconic gestures to explain a task, compared to when teachers used speech alone, it was unclear whether this was because the semantic content of the gestures enhanced processing of the (spoken) message, or

simply acted as an attentional device which helped children concentrate on key parts of the teacher's instructions (Wang et al., 2004). The results outlined above suggest that, given the interaction between PWA gesture recognition abilities, which relate to comprehending the semantic content of gestures, and improved speech comprehension, in this case gestures facilitated comprehension by conveying semantic information and those PWA who were better able to extract (and integrate) that information were the ones to benefit.

Secondly, contrary to the predictions of the LRH, I did not find clear evidence that PWA gesture production facilitated lexical retrieval. Furthermore, while I found evidence that gesture observation may facilitate lexical retrieval in some PWA, this effect was not universal and appeared to be restricted to PWA with poorer WM/EF abilities. It is possible that gesture observation facilitated naming through a priming process in which the gesture enhanced the target word's activation in working memory for a short time after it was produced by the researcher, delaying trace decay; this then made it more salient and easier for PWA with poorer working memory to retrieve during subsequent game turns. For PWA with better working memory, no such facilitatory effect would have been required, as the target word(s) would already have been sufficiently retained in memory for retrieval to occur without needing gesture priming.

In contrast to gesture observation, gesture *production* appeared to be particularly taxing for PWA with poorer WM/EF and resulted in significantly reduced naming accuracy, indicating a dissociation between gesture observation effects and gesture production effects in these PWA. The pronounced negative effects of gesture production for these participants may be due to a dual-task penalty, in which producing the gesture and the target word in close temporal proximity was too cognitively challenging. PWA are known to be affected by dual-task demands, perhaps due to attentional deficits; for example, Harmon et al. (2019) demonstrated that speech production in a story retell task was significantly more impaired in PWA when they were simultaneously asked to distinguish between high and low tonal

sounds playing in the background. Similar results were also found by Hula et al. (2007) for PWA completing a picture naming task and an auditory processing task simultaneously. Notably both these studies involved one active (i.e. speech output) and one passive (i.e. auditory processing) task. In contrast and similar to the study above which involved two active tasks (i.e. speech production and gesture production), Plummer et al. (2020) showed that producing speech while walking negatively affected gait speed in PWA (with mild to moderate aphasia); PWA walked significantly faster when not speaking. While no studies to my knowledge have specifically explored dual-task effects of naturalistic gesture production on speech output in PWA, it is possible that the same dual-task penalty would arise. An alternative interpretation of this finding is that these PWA may simply have produced a gesture instead of a word because it was easier (i.e. less cognitive taxing) to do so; in this case, PWA would be producing a gesture to compensate for impaired speech production.

Thirdly, I also found indications of varying lesion patterns for PWA who benefited (or not) from gesture observation/production for naming accuracy. For gesture observation benefit, associated with poorer working memory/executive function behaviourally, PWA presented with lesions clustered anteriorly along the dorsal language processing pathway (including the IFG and anterior portions of the superior temporal lobe) and the dorsolateral prefrontal cortex. For gesture production benefit, associated with better working memory/executive function behaviourally, the small minority of PWA who showed an effect presented with lesions to the IFG and temporal pole only, without additional fronto-temporo-parietal involvement. The dorsolateral prefrontal cortex is crucial for maintaining and manipulating information in working memory (Barbey et al., 2013); PWA with lesions to this region may therefore be expected to present with working memory deficits, in line with the behavioural findings. In contrast, for gesture observation detriment for naming, PWA presented with lesions clustered posteriorly along the ventral pathway for language processing (including the angular gyrus, middle and inferior temporal gyri). The ATL (bilaterally) have been found to function as hubs for integration of multimodal information (Ralph et al.,

2016). PWA with ATL lesions may be unable to integrate information from speech and gestures simultaneously, resulting in poorer naming accuracy. Gesture production detriment for naming was associated with diffuse and heterogenous lesions to both the ventral and dorsal language processing pathways, reflecting the fact that the majority of PWA in the sample showed gesture production detriment for naming.

I also found that the proportion of naming error types varied depending on whether PWA produced gestures. Specifically, the number of omission and unrelated errors significantly increased when PWA produced gestures. This may be because (as briefly discussed above) PWA used gesture production as a compensatory measure for impaired lexical retrieval. That is, when they were unable to retrieve the correct target word, they used gesture(s) to effectively convey the message instead. This is consistent with previous research, where PWA have been shown to use gestures to convey information that they cannot convey through speech (e.g. Akhavan et al., 2017 ; de Beer et al., 2019). This possibility is explored further in Chapter 4 of this thesis.

2.7.2 Clinical implications

Although I found no significant main effect of gesture observation on PWA speech comprehension accuracy, the interaction between gesture condition and gesture recognition score is suggestive of the fact that PWA may differentially derive benefit from gesture observation as a consequence of upper limb apraxia. Indeed, Ferro et al. (1980) suggested that ideomotor apraxia (i.e. gesture praxis) and ideational apraxia (i.e. gesture recognition deficits) are functionally distinct and can be dissociated, in the same way that expressive and receptive language deficits may be dissociated in the same individual. As a result, clinicians should assess both gesture recognition and gesture production (praxis) to determine if they are likely to benefit from gesture observation for comprehension.

Furthermore, results from this study indicate that PWA do not (necessarily) use gestures as a means of self-cueing for word retrieval. Indeed, the use of gestures was negatively correlated with naming accuracy. Instead, it appears that PWA used gestures as a compensatory measure for impaired speech production; these results indicate that when they were unable to produce speech (i.e. an omission error), PWA used gestures instead. Clinicians should therefore consider PWA goals when recommending therapy; if the goal is to facilitate speech production and naming accuracy, gesture production may not be beneficial; in contrast, if the goal is to enhance overall communicative effectiveness, gesture production may be encouraged.

Results from this study also suggest that gesture (dis)benefit for single word naming accuracy in PWA is significantly moderated by individual differences in WM/EF deficits, with PWA with poorer WM/EF benefitting significantly more from gesture observation for naming accuracy. In contrast, PWA with better WM/EF showed limited benefit from gesture observation. Differential benefits in gesture observation and production suggests that a “gesture pathway” for anomia may be unhelpful; instead, clinicians may need to consider whether individual PWA are predicted to benefit, or not, from gestures produced by interlocuters (or themselves) for naming accuracy. The backwards digit span task used in this study to assess working memory/executive function could be incorporated into aphasia assessment batteries to achieve this.

2.7.3 Limitations

An important limitation of this study is that although there was a significant interaction effect between gesture observation and gesture recognition score for predicting PWA speech comprehension accuracy, the effect size was relatively small and was likely impacted by the fact that PWA in the study as a cohort generally scored highly for single word comprehension accuracy. Indeed, power to observe higher-level interactions has not been established and is likely to be limited. As a result, a beneficial effect of gesture observation for PWA with more severe comprehension deficits cannot be

assumed. This remains an important yet under-researched area in aphasiology, as auditory comprehension is associated with significant reductions in quality-of-life measures in PWA (Davidson et al., 2008) (Moineau et al., 2005). Future studies should therefore incorporate PWA with a wider range of severity of comprehension deficits, to shed light on the question of whether gesture observation facilitates comprehension, as has previously been found in studies of neurologically healthy adults (Zhang et al., 2021). Additionally, while the use/non-use of gestures was manipulated in this study, PWA had access to other visual cues which could have facilitated comprehension for target words, i.e. images on the playing cards and the researcher's mouth movements. In future studies, these aspects could also be manipulated to more clearly quantify how gestures (as a sole visual cue, rather than one among many) impact comprehension.

Additionally, although the experimental design was based around a LAG and was intended to be as naturalistic as possible while maintaining a degree of control, it was not fully naturalistic. In particular, the tasks were designed to elicit single words (and to a lesser extent, gestures) and therefore may not represent spontaneous interactions where connected speech production is more common. This limitation is addressed further in Chapter 3. Finally, although lesion data was available for a subset of participants, larger sample sizes are needed for reliable results and statistical analysis which were unavailable here. Therefore, while the lesion observations align relatively well with the behavioural findings, they should be taken as tentative.

2.8 Conclusion

This study represents the first attempt to utilise a novel LAG based on the card game "Go Fish!", to investigate the effects of gesture observation and production in facilitating comprehension and single word naming accuracy within a dyadic interaction between a neurologically healthy adult and a PWA. Although observing iconic gestures did not improve speech comprehension in all PWA, PWA who were best able to extract semantic information from the researcher's gestures showed the greatest benefit. I

also found a significant interaction between gesture observation and working memory/executive functioning for naming, as PWA with poorer working memory/executive functioning abilities benefited more from observing iconic gestures in subsequent naming accuracy. These results suggest that iconic gestures may function to maintain trace activation of recently presented words in working memory in PWA, as previously suggested by Krauss & Morsella (2004). By contrast and contrary to the predictions of the LRH, gesture production was significantly associated with lower naming accuracy, especially in PWA with poorer working memory/executive function, perhaps due to a dual-task penalty. These findings illustrate the critical importance of comprehensively assessing individual differences in PWA cognitive profiles and shed light on the functions of gesture in interactions between PWA and neurologically healthy adults.

Importantly, findings from this study, while indicative of gesture benefit for naming accuracy in (some) PWA, relate only to single-word picture naming. However, most dyadic interactions involve far more than the cooperative exchange of single words paired with single iconic gestures, and the role of gestures (including non-representational gestures) in facilitating communication during connected speech exchanges remain unclear. In the following chapter, I therefore turn to the effects of gesture in a context of connected speech production.

Chapter 3: “Heads Up!”

3.0 Abstract

In Chapter 2 I demonstrated that individual differences in WM/EF predicted differential gesture observation and production benefits for single word naming in PWA. Yet communication is not limited to single words. To extend the findings reported in Chapter 2, and to determine if individual differences predict gesture benefit for PWA communication in a connected speech context, I developed a second novel dyadic LAG, based around the card game “Heads Up!” In this study, 36 PWA played “Heads Up!” with a neurologically healthy interlocuter (the researcher). To elicit connected speech in a naturalistic manner, players were asked to produce descriptions of everyday objects and actions in response to picture stimuli. In half of game trials both players were permitted to gesture alongside speaking, while in the other half of trials gesture use was constrained. I investigated whether gesture observation (i.e. PWA observing gestures produced by the researcher) and gesture production (i.e. PWA producing gestures during speech production) affected PWA communication. Contrary to predictions, there was no significant main effect of gesture observation on PWA communication, although for a subset of PWA (visual) attention capacity moderated gesture benefit; PWA with better attentional capacity benefited from gesture observation for comprehension, while PWA with poorer attention did not. Further, gesture production significantly affected PWA connected speech measures: Mean Length of Utterance (MLU) was significantly higher when PWA were constrained from gesturing, while mean concreteness score was significantly higher when PWA were permitted to gesture. Moving Average Type-Token Ratio (MATTR) did not significantly differ between conditions. I discuss these findings in relation to the predictions of the LRH and the Information Packaging Hypothesis.

3.1 Introduction

While the previous chapter focused on single word comprehension and production in PWA, during everyday interactions interlocuters typically flexibly use multimodal cues (gestures, facial expressions, tone of voice and so on) alongside *connected speech* output. Even in PWA with few deficits in single word speech production or comprehension, connected speech production (i.e. production of a continuous sequence of words to form utterances) or comprehension may be impaired, and there appears to be little correlation in severity of deficits in single word and connected speech processing in PWA (Alyahya, 2018). Therefore, the same patterns of gesture facilitation (or not) for single word production and comprehension described in Chapter 2 cannot be assumed to hold for connected speech. In this chapter, I investigated whether gesture observation facilitation for naming accuracy in PWA occurred in a connected speech context, rather than a single word task, and whether PWA gesture production measurably affected aspects of PWA connected speech output. I again used an interactive LAG based on a simple card game, “Heads Up!”, where players (the researcher and PWA) produced descriptions of everyday objects and actions; in half of game trials players were permitted to gesture, and in the other half gesture use for both players was constrained.

3.1.1 Gesture and connected speech production in PWA

The primary focus of the LRH is speech-gesture interaction at the level of *word* retrieval (Rauscher et al., 1996), while the Information Packaging Hypothesis is concerned with how speech and gestures interact during *message* conceptualisation and encoding – which may occur at the level of single words, but may also extend across units comprising multiple words (e.g., utterances), depending on the complexity of the message (Kita, 2000). Therefore, although both models are concerned with when speech and gestures interact in the process of speech production, they may be concerned with different levels of lexical representation (i.e. word production for the LRH; word *and* utterance production for the Information Packaging

Hypothesis). Because of this, while the Information Packaging Hypothesis makes several predictions about how gesture may affect aspects of connected speech production, the LRH largely does not (with the exception of fluency, discussed briefly below). However, taking the basic premise of the LRH as gesture production facilitating word retrieval through cross-modal priming, it is possible to extend it to generate novel predictions for connected speech production. In the following sections I discuss how the predictions of both an extended LRH and the Information Packaging Hypothesis may be investigated in PWA connected speech production using three measures: lexical diversity (operationalised through MATTR), utterance length (MLU), and utterance concreteness.

3.1.1.1 Gesture and lexical diversity (MATTR)

The LRH proposes that gesture production serves to facilitate lexical retrieval, and while this relates (primarily) to retrieval of single words, there is evidence that retrieval may be affected by gesture production in connected speech contexts as well. For example, Rauscher et al. (1996) found that constraining gesture production in neurologically healthy adults resulted in a significant observed increase in dysfluencies and word-finding difficulties in connected speech production; they suggested that the increase in dysfluencies was a direct result of impaired word-finding due to gesture constraint, i.e. that constraining gesture use led to impaired lexical retrieval, which in turn caused speakers to interrupt their utterances with repetitions, empty pauses or fillers. Similar results were also found by Beattie & Coughlan (1999), who observed that neurologically healthy speakers experienced fewer ‘tip-of-the-tongue’ word-retrieval difficulties during connected speech production when allowed to gesture.

Lexical access is closely related to lexical diversity; where efficient word retrieval is compromised (as in PWA, or indeed in neurologically healthy adult L2 learners) lexical diversity is typically reduced, because speakers fall back on repeatedly producing words they find easier to access (Cunningham & Haley, 2020; Schmid & Jarvis, 2014). In a connected speech context,

therefore, lexical diversity can be a useful additional measure of ease of lexical retrieval. Following the LRH, lexical diversity is predicted to be higher for utterances where PWA use gestures than where they do not, because the process of producing gestures facilitates word retrieval, giving rise to more lexically diverse utterances. In contrast, the Information Packaging Hypothesis is silent on the issue of lexical diversity; more efficient packaging of information does not necessarily assume greater diversity of lexical output, because some of the information conceptualised within the message is assumed to be conveyed through gestures (rather than speech) in any case.

3.1.1.2 Gesture and mean length of utterance (MLU)

MLU has been used in multiple populations and provides a general measure of the number of morphemes produced in spontaneous utterances (Rice et al., 2010). MLU is known from multiple studies to be reduced in PWA compared to healthy controls, partly as a function of increased word-finding difficulties (e.g. Fromm et al., 2017; Bryant et al., 2013; Webster & Morris, 2019). Following the LRH, MLU may be *longer* when speakers use gestures, because word-retrieval will be easier. By contrast, the Information Packaging Hypothesis predicts that MLU will be *shorter* when speakers gesture, because they have access to spatio-motoric thinking which allows for (some) information to be preferentially conveyed via gestures. This effect was found by Özer et al. (2017) in a study of speech output in a route description task: participants completed the task successfully despite producing significantly less speech content when permitted to gesture, which the authors concluded was because they communicated some relevant information through gestures instead. Results are not consistent across all studies, however: for example, in a narrative retell task in which 12 Italian speakers were asked to describe the content of short comic strips to listeners, Cravotta et al. (2021) found that encouraging speakers to gesture while speaking resulted in a significant *increase* in overall discourse length (i.e. in line with the LRH) compared to when speakers did not gesture. Importantly, in no study has the effect of gesture on MLU been investigated in relation to PWA.

3.1.1.3 Utterance concreteness

Finally, utterance concreteness may also be used as a measure for investigating gesture effects on connected speech production, although it has rarely been applied to narratives produced by PWA. Concreteness relates to the extent to which a word/utterance refers to a sensorily perceptible entity in the world (Paivio et al., 1968); more concrete (or less abstract) words relate to more perceptible entities, less concrete (or more abstract) words to less perceptible entities. In neurologically healthy adults, gestures have been found to co-occur more frequently with more concrete/less abstract words, presumably because more concrete words typically encode sensorimotor or spatial content related to our direct interaction with the world; sensorimotor simulations are assumed to more readily give rise to gestures (Hader & Krauss, 1999; Kok, 2016; Tong, 2023). Although neither the LRH nor the Information Packaging Hypothesis speak directly to word or utterance concreteness, it may be hypothesised that if speech and gestures interact from the earliest stages of message conceptualisation (i.e. in line with Information Packaging Hypothesis), speakers may be expected to adjust their speech content when gesture production is manipulated, resulting in increased concreteness when permitted to gesture, and increased abstractness when constrained from gesturing. By contrast, because the LRH assumes the message is conceptualised *before* gestures interact with speech, gesture use/constraint would not be predicted to influence speech concreteness content.

3.2 The present study

In this study, I elicited single words and connected speech output from English-speaking PWA within the context of a dyadic interaction between participants and a neurologically healthy interlocuter (the researcher). The interaction was based on a turn-taking naming-to-description card game, “Heads Up!”. A within-subjects design was used with the primary experimental manipulation being whether game players were permitted to use, or constrained from using, gestures alongside speech. Players took

turns describing objects/actions; the opposing player had to guess what object/action was being described. I measured whether PWA retrieved the correct target word in response to the researcher's descriptions (naming accuracy measure) and I also recorded and transcribed PWA connected speech production to carry out analyses to investigate gesture production effects on connected speech output measures. Data collection was carried out at two sites: UCL in London, UK and Moss Rehabilitation Research Institute (MRRI) in Philadelphia, USA. To ensure consistency for PWA, as in the previous study all interactions were conducted with the same researcher either at UCL or at MRRI. Ethical approval to carry out this study was received from the High Risk Research Ethics Committee at UCL (21765/001) and the MRRI Institutional Review Board (IRB-2023-1046).

3.3 Study aim and hypotheses

In this study, I firstly aimed to extend findings from Chapter 2 and investigate whether gestures produced by the interlocuter (gesture observation) facilitated PWA naming accuracy in a naming-to-description task. Secondly, I aimed to investigate whether gestures produced by PWA affected connected speech production measures.

Firstly, I predicted that PWA naming accuracy would be significantly higher in the gesture permitted condition of the "Heads Up!" task (i.e. when PWA were able to observe researcher gestures alongside speech), as gestures produced by the researcher would enhance semantic activation and enable PWA to more effectively retrieve the correct target word. I also predicted an interaction effect between gesture comprehension measures (gesture recognition ability) and gesture condition; PWA with poorer gesture comprehension were not predicted to benefit to the same degree as PWA with more intact gesture comprehension, as they would be unable to extract semantic information from the gestures produced by the researcher.

Prediction 1a: PWA naming accuracy will be higher in the gesture permitted condition.

Prediction 1b: Gesture comprehension abilities (gesture recognition score) will moderate gesture benefit.

Beyond the effect of gesture observation on naming accuracy, I also predicted that PWA gesture production would affect PWA speech output and that PWA connected speech measures would vary significantly depending on whether PWA were permitted to gesture or constrained from gesturing. In each case I generated differing predictions for the LRH and the Information Packaging Hypothesis. Firstly, for MATTR, the Information Packaging Hypothesis is silent about lexical diversity and therefore no difference between conditions is predicted; however, if the LRH is correct lexical diversity will be significantly higher in the gesture permitted condition. Secondly, for PWA MLU, if the Information Packaging Hypothesis is correct, MLU is predicted to be shorter in the gesture permitted condition; if the LRH is correct, MLU is predicted to be longer in the gesture permitted condition. Finally, if speech and gestures interact from the earliest stages of conceptualisation, mean utterance concreteness is predicted to be higher in the gesture permitted condition; if not, no difference in utterance concreteness is expected regardless of gesture use/constraint.

Prediction 2a: PWA MATTR will not differ significantly between conditions, in line with the Information Packaging Hypothesis.

Prediction 2b: PWA MATTR will be significantly higher in the gesture permitted condition, in line with the LRH.

Prediction 3a: PWA MLU will be significantly shorter in the gesture permitted condition, in line with the Information Packaging Hypothesis.

Prediction 3b: PWA MLU will be significantly longer in the gesture permitted condition, in line with the LRH.

Prediction 4a: If speech and gestures interact from conceptualisation, PWA mean utterance concreteness will be significantly higher in the gesture permitted condition.

Prediction 4b: If speech and gestures do not interact from conceptualisation, PWA mean utterance concreteness will not differ significantly between conditions.

3.4 Methods

3.4.2 Participant sample

I utilised the same group of participants for this task as for the task described in Chapter 2, "Go Fish!". Full details of participant demographic may be found in Section 2.4.1. For information about participant performance on background battery assessments refer to Appendix 1.

3.4.3 Set up

The room layout used for this task was the same as for the "Go Fish!" task described in Chapter 2. For full details please refer to Section 2.4.5.

3.4.4 Materials

Materials consisted of:

- Two sets of stimulus cards depicting objects or actions for each task (as described in Chapter 2). Each individualised set consisted of at least 4 cards showing objects/actions that the participant had correctly named during baseline testing and 16-20 cards that they had been unable to name correctly. For each PWA, one card set was randomly designated the "gesture set" and one was designated the "no gesture set". Cards in both sets were individualised for each participant based on his/her performance on the *Object and Action Naming Battery* and

MRRI Middleton Lab Naming Battery. Correctly named cards were included to establish a baseline for PWA performance and to reduce frustration on the part of the participant. To the extent possible, cards were balanced for the number of objects and actions and for the number of semantic and phonological errors made during background testing. Cards were randomly ordered in each set and participants were free to shuffle the cards if they wanted.

- Additional cards (as many as required) for teaching gameplay and familiarising participants with the card format. These cards were drawn at random from the remaining pictures in the *Object and Action Naming Battery* and *MRRI Middleton Lab Naming Battery.* They were identical in format to those used during the experimental task.
- Large sand-timer to indicate time left for each game turn.

3.4.5 Procedure

Participants completed (up to) 24 rounds of “Heads Up!”, an interactive naming-to-description LAG, with the researcher. Prior to starting the experimental task participants were taught how to play the game using practice cards. Participants were able to play as many practice turns as required to become comfortable with gameplay before starting the experimental task.

The participant and researcher were each dealt a set of cards, face down on the table. Participants were instructed to hold out a single card, facing away from them towards the researcher, such that the researcher could see the picture on the face of the card but the participant could not. The researcher described the picture on the card to the participant without using its name (up to a maximum time of two minutes per game turn, indicated to participants with a large sand timer) after which the participant was invited to guess the name of the object or action depicted on the card.

Descriptions consisted of a range of information about the target word, including its form, function, origin etc. Although broad content was planned

beforehand, the descriptions were produced by the researcher without referring to scripts during the interaction. Participants were encouraged but not required to wait until the description was complete before making a guess. Participants were also permitted to request additional information from the researcher. In each trial, the participant's first clear attempt at naming was scored; the same scoring procedure was carried out as for "Go Fish!" as outlined in Chapter 2. Gameplay then transferred to the other player, with the researcher holding out a card and the participant invited to describe the picture on its face for the researcher to guess, eliciting connected speech samples from the participant.

In half of trials the researcher and participant were free to use gestures while producing descriptions. In the remaining half of trials gesture use was constrained. Gesture use/constraint was blocked and counterbalanced across participants. Following gesture-constraint paradigms developed by Pulvermüller and colleagues, physical restraints to discourage gesturing were not used; participants were constrained from gesturing through explicit instruction and reinforcement between trials (Pulvermüller et al., 2001).

In the gesture constrained condition, the researcher and participant were required to keep their hands flat on the table in front of them. The second researcher provided reminders of this instruction as required to the participant. Trials in the gesture constrained condition where participants produced gestures were discarded and not used in analyses. Trials were counter-balanced with half of participants completing the gesture permitted condition first and half completing the gesture constrained condition first. In the gesture permitted condition the researcher always used gesture in each trial when producing descriptions; participants were not specifically instructed to gesture and were free to choose whether or not to gesture in each trial. To reduce participant fatigue, time spent on each condition was limited to a maximum of 30 minutes (excluding breaks).

Fig. 1. *“Heads Up!” setup. Researcher providing description.*

Fig. 2. *“Heads Up!” setup. “It goes boom!”*

3.4.6 Naming error transcription and error typing

The same process for speech transcription and error typing was followed as outlined in Chapter 2. See Section 2.5.2 for further details.

3.5 Data Analysis

The same statistical modelling strategy was adopted in this chapter as described in Section 2.5.3. To evaluate the predictions outlined the previous section, the following analyses were conducted. Analyses were performed in *R* (version 4.2.1) (https://www.r-project.org/) using the *lme4* package (Bates et al., 2015) . Logistic models were fit with the Laplace approximation with a binomial distribution and the *logit* link function. The *bobyqa* or *Nelder-Mead* optimisers were used to ensure model convergence. Working memory/executive function score was log transformed for model stability, to reduce skew and ensure model assumptions were met. Continuous variables were scaled and centred and categorical variables were orthogonally coded to test for effects. We determined that there was no significant multicollinearity in the models by calculating the variance inflation factor of all predictors using the *vif* function in the R *car* package (all variance inflation factors < 3.00) (Fox & Weisberg, 2019). Where additional packages were used for analyses these are indicated with the model details below. Full R code for all models may be found in Appendix 9.

3.5.1 Naming accuracy

I conducted logistic mixed effects regression analyses with naming accuracy as the dependent variable. Target word length, target word frequency, spontaneous speech production (WAB-R Spontaneous Speech score), research site (US/UK), working memory/executive function, short-term memory, first experimental condition, apraxia gesture recognition score, trial order, whether the target word was an object/action and semantic access score were entered as predictors into the model. I also included a random slope of gesture condition with a random intercept by-participant to account for variability in participant response to gestures.

3.4.1.1 Semantic and phonemic similarity of naming errors

For more fine-grained analysis of the effects of gesture observation on PWA speech output, I also analysed whether gesture observation differentially affected cosine similarity (i.e. semantic similarity) or phonemic similarity (phonological Levenshtein distance) of errors produced by PWA. Correct responses were excluded from this analysis. For these analyses we conducted linear mixed effects regressions with cosine similarity or phonemic similarity as the dependent variables.

3.5.2 Lexical diversity (Moving Average Type-Token Ratio)

Lexical diversity is typically calculated using Type-Token Ratio (TTR): dividing the number of different words produced in a speech sample (types) by the total number of words produced in the whole sample (tokens). As a stable TTR requires utterances of relatively equal length for direct comparison, MATTR may be used as an alternative length invariant measure, where the TTR is calculated for a moving window of tokens (Fergadiotis et al., 2013). Following Covington & McFall (2010) I calculated a MATTR for a moving window of 50 items independent of utterance length, allowing utterances of different lengths to be directly compared. In each utterance, the number of tokens equated to the number of items (regardless of repetition); the number of types equated to the number of unique items.

MATTR was calculated in Python. PWA transcribed utterances were imported into Python (v3.9.7, 2021) and initially lemmatized to their stem forms using the *lemmatizer* function of the *Natural Language Toolkit* package (Loper & Bird, 2002). We wrote a function to iterate over all utterances produced by PWA and calculate the number of types and tokens for a moving window of 50 items. The number of types was then divided by the number of tokens to yield the type-token ratio. This process was repeated for both gesture conditions. Python code for calculating MATTR may be found in Appendix 6.

3.5.3 Mean Length of Utterance (MLU)

MLU is a measure of linguistic productivity and is calculated by dividing the number of morphemes in an utterance by the total number of utterances (Brown, 1973). Although MLU is typically calculated based on the number of morphemes in 100 utterances, I chose to calculate MLU in this instance based on all analysable utterances in the dataset to account for individual variability in expressive language output across PWA. As for MATTR PWA transcribed utterances were imported into Python and lemmatized. To calculate MLU I used the standard formula:

MLU = Total Number of Utterances / Total Number of Morphemes

To determine whether the difference in MLU between conditions was significant, and whether gesture condition was predictive of MLU, I defined a linear mixed effects regression model with MLU as the dependent variable and gesture condition as the primary independent variable of interest. Aphasia severity (WAB-R Aphasia Quotient), working memory/executive function and semantic access were entered as predictors in the model with random intercepts for participant and research site to account for inter-PWA variability.

3.5.4 Concreteness

I also investigated whether mean utterance concreteness varied across gesture conditions, as has been found in neurologically healthy adults (see Kok, 2016 for a review). Individual concreteness ratings were generated for all words in each utterance in the analysable dataset. Utterances were imported into Python and lemmatized, then lemmatized forms were checked against concreteness ratings for single words developed by Brysbaert et al. (2009). I then generated utterance-level mean (average) concreteness ratings by dividing the concreteness rating for each item in the utterance by the total number of words in the utterance, yielding a single mean concreteness rating for each utterance.

To determine whether mean concreteness ratings for each utterance varied by gesture condition, I defined a linear mixed effects regression model with mean concreteness rating as the dependent variable and gesture condition as the primary independent variable of interest. Aphasia severity (WAB-R Aphasia Quotient), recruitment site, working memory/executive function, research site and semantic access were entered as predictors in the model with a random intercept for participant.

3.6 Results

3.6.1 Gesture observation behavioural results

3.6.1.1 Naming accuracy

Post-hoc power analyses were conducted with the collected data to determine if it was sufficient to detect a small, moderate or large effect. Following the approach of Cabiddu et al. (2022), simulations were conducted (500 iterations) with the observed data for a range of effect sizes (range: 0.0 - 2.0); the effect of the independent variable (gesture condition) was varied across iterations while all other fixed and random effects were kept consistent. The full analysis and code adapted from Cabiddu et al. (2022) may be found in Appendix 5. This analysis indicated that the data was sufficient to detect a small to medium effect (effect size = 0.35; $\alpha = 0.05$; power = 0.80).

In the gesture constrained condition the mean naming accuracy was 0.668; in the gesture permitted condition it was 0.665. As shown in Table 1 below, contrary to Prediction 1a above, across 1240 observations overall there was no significant effect of gesture observation on PWA naming accuracy ($p > 0.05$). Additionally, in this analysis I found no significant interactions between cognitive variables and gesture condition, although the control variables expressive aphasia severity (WAB-R Spontaneous Speech) and STM significantly predicted naming accuracy ($p < 0.001$ and $p = 0.002$ respectively).

Table 1. *Naming accuracy model output.*

Effect	Estimate	Std.Error	z value	Pr(>|z|)
(Intercept)	1.125	0.330	3.406	0.001
Gesture condition (present)	0.029	0.227	0.128	0.898
Working memory	0.278	0.151	1.837	0.066
Short term memory	0.462	0.150	3.075	0.002
Semantic access score	-0.043	0.145	-0.294	0.769
Gesture recognition score	0.138	0.126	1.097	0.272
Spontaneous speech score	0.390	0.116	3.362	0.001
Trial order	0.017	0.070	0.249	0.803
Site	-0.448	0.276	-1.626	0.104
First gesture condition	-0.193	0.246	-0.783	0.434
Object word	0.277	0.211	1.310	0.190
Target word length)	-0.148	0.103	-1.430	0.153
Target word frequency	0.068	0.105	0.643	0.520
Gesture condition (present) * Working memory	0.115	0.218	0.526	0.599
Gesture condition (present) * Short term memory	-0.026	0.227	-0.114	0.909
Gesture condition (present) * Semantic access	0.200	0.213	0.938	0.348
Gesture condition (present) * Gesture recognition score	-0.120	0.192	-0.625	0.532
Gesture condition (present) * Spontaneous speech	-0.192	0.183	-1.050	0.294
Gesture condition (present) * Object word	-0.070	0.298	-0.234	0.815
Gesture condition (present) * Target word length	0.134	0.149	0.896	0.370
Gesture condition (present) * Target word frequency	0.172	0.149	1.155	0.248

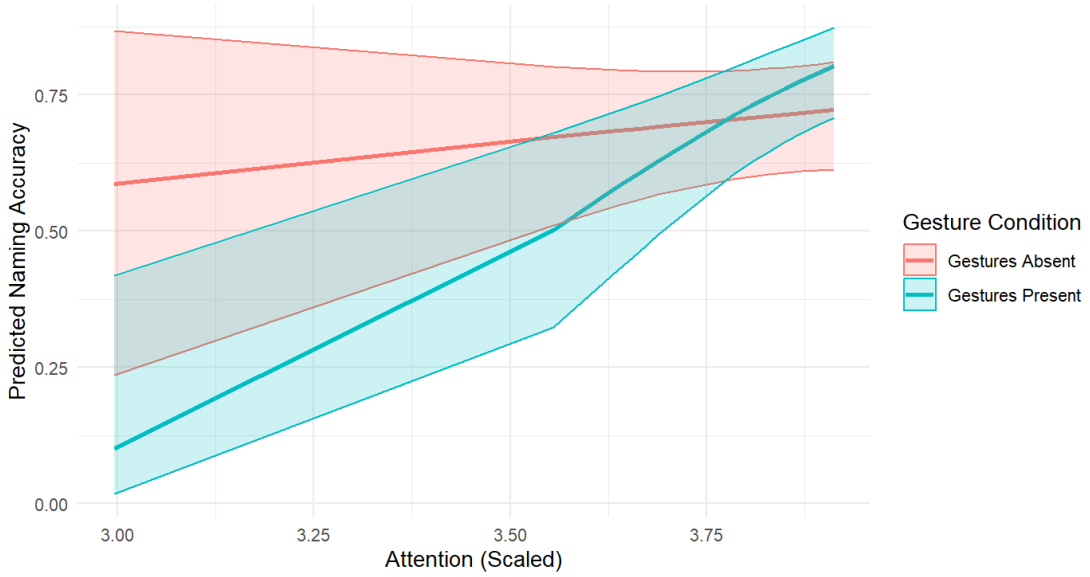
However, using a subset of participants (UK-based, n = 24 PWA) for whom there was additional behavioural data derived from the *Oxford Cognitive Screen* assessment, across 984 observations I found a significant interaction between (visual) attention capacity and gesture condition (p = 0.004), as shown in Table 2 and Fig. 3 below. The control predictor WM/EF also significantly predicted naming accuracy (p = 0.001).

Table 2. *Naming accuracy model output (UK subset data).*

Effect	Estimate	Std.Error	z value	Pr(>|z|)
(Intercept)	0.882	0.244	3.616	< 0.001
Gesture condition (present)	0.097	0.173	0.559	0.576
Working memory	0.587	0.178	3.293	0.001
Semantic access score	-0.014	0.174	-0.080	0.936
Gesture recognition score	0.316	0.171	1.853	0.064
Trial order	0.032	0.075	0.424	0.671
First gesture condition	-0.193	0.348	-0.555	0.579

Visual attention	0.107	0.143	0.751	0.453
Target word length)	-0.157	0.111	-1.422	0.155
Target word frequency	0.149	0.111	1.346	0.178
Gesture condition (present) * Working memory	0.120	0.185	0.652	0.514
Gesture condition (present) * Semantic access	-0.017	0.181	-0.094	0.925
Gesture condition (present) * Gesture recognition score	-0.140	0.192	-0.730	0.465
Gesture condition (present) * Visual attention	0.527	0.182	2.900	0.004
Gesture condition (present) * Target word length	0.077	0.160	0.481	0.630
Gesture condition (present) * Target word frequency	0.102	0.161	0.634	0.526

Fig. 3. *Naming accuracy: interaction between naming accuracy scores in each gesture condition and PWA visual attention capacity score.*
In this figure the Y axis represents naming accuracy (0.0-1.0 scale) and the X axis indicates scaled visual attention score. The red line indicates change in naming accuracy in the gesture absent condition and the blue line indicates change in naming accuracy in the gesture present condition. Shaded areas represent confidence intervals.

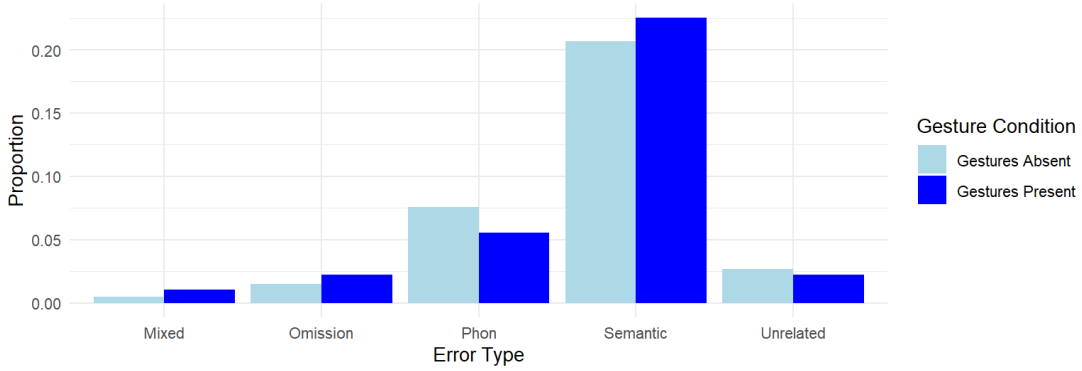


3.6.1.2 Naming error type

Next, I explored whether the proportion of naming error types varied depending on whether PWA did, or did not, observe the researcher producing gestures alongside producing the description for each trial. In this analysis only incorrectly named items produced by PWA were included; all correctly named items were excluded. The conservative Bonferroni

correction for multiple comparisons was applied to reduce the risk of a Type I error. Given the relatively small number of observations available for analysis (n = 395) I initially performed the Chi-Square test of approximation and then used Fisher's Exact Test to confirm the result. No significant association between error type and gesture observation was found in the Chi-square analysis ($\chi^2 = 5.94$, $p = 0.29$) or Fisher's Exact Test analysis ($p = 0.30$). Fig. 4 below shows these results.

Fig. 4. *Proportion of naming error types produced by all PWA.*



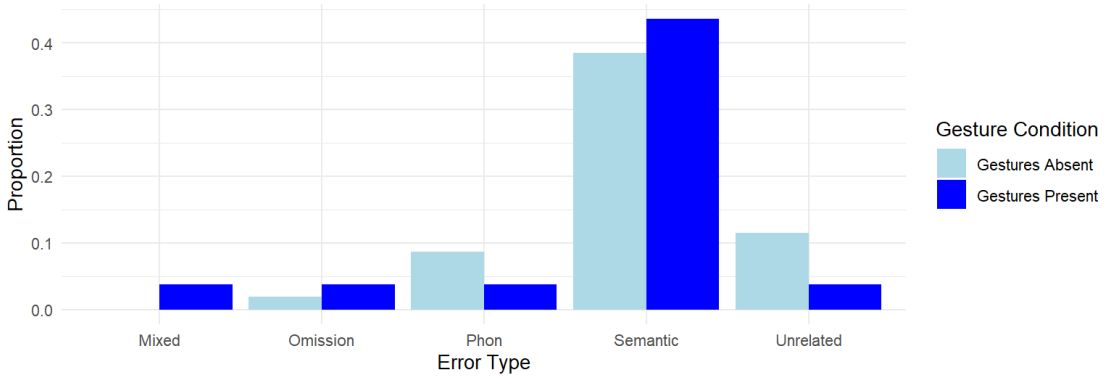
Next, in order to explore whether this finding held true for all PWA in the sample or was driven by the relatively few naming errors produced by PWA with mild aphasia, I repeated the analysis including only incorrect responses from PWA with moderate aphasia (n = 142). A significant association between naming error type and gesture observation was found in the Chi-square analysis ($\chi^2 = 13.36$, $p = 0.012$) and Fisher's Exact Test analysis ($p = 0.013$). Pairwise comparisons initially indicated that the proportion of unrelated errors varied by gesture condition ($p = 0.046$), with more unrelated naming errors produced by PWA when the researcher did not produce gestures; however, when the Bonferroni correction for multiple comparisons was applied, the results were no longer statistically significant ($p = 0.231$) (Table 3 and Fig. 5).

Table 3. *Naming error types analysis output.*

Error Type	Original *p*-value	Bonferroni Adjusted *p*-value
Mixed	0.109	0.543

Omission	0.609	1.0
Phon	0.215	1.0
Semantic	0.181	0.904
Unrelated	0.046	0.231

Fig. 5. *Proportion of naming error types produced by PWA with moderate aphasia.*



3.6.1.3 Cosine similarity of naming errors

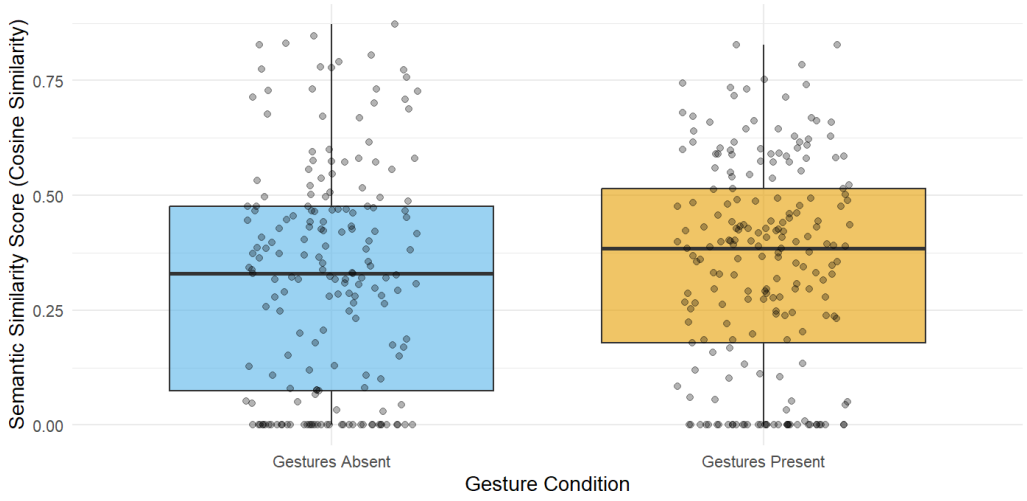
On the basis that (iconic) gestures produced by the researcher may enhance semantic activation of target words and semantic neighbourhood competitors, I next explored whether the cosine similarity (semantic similarity) of naming errors produced by PWA was affected by gesture presence or absence. Only naming errors were included in this analysis; correct responses were excluded. As shown in Table 4 and Fig. 6 below, across 395 trials where PWA produced naming errors in response to researcher descriptions I found no significant effect of gesture observation (β=0.016, SE=0.022, t(383.41)=0.74, p=0.463). However, the control variable target word frequency significantly predicated cosine similarity of naming errors, with more frequent target words resulting in PWA guesses that were semantically closer to the target word (p < 0.001). The control variable sematic access score also approached significance (p = 0.098).

Table 4. *Cosine similarity of naming errors model output.*

	Estimate	Std. Error	df	t value	Pr(>|t|)
(Intercept)	0.340	0.032	34.330	10.751	0.000

Gesture condition (present)	0.016	0.022	383.409	0.735	0.463
Working memory	-0.022	0.018	32.397	-1.218	0.232
Aphasia Quotient	0.028	0.020	26.450	1.442	0.161
Semantic access score	-0.032	0.019	32.777	-1.701	0.098
First gesture condition (present)	0.004	0.039	28.775	0.100	0.921
Target word length	-0.010	0.012	374.798	-0.882	0.378
Target word frequency	0.053	0.012	380.134	4.556	< 0.001

Fig. 6. *Boxplot: effect of PWA gesture observation condition on semantic similarity (cosine similarity) score of PWA naming errors.*
In this figure the Y axis represents semantic similarity score (cosine similarity, 0.0-1.0 scale). Boxplots indicate the range of scores with the mean score shown as a line in each box. Scores in each gesture condition are presented separately (i.e. gestures absent, gestures present).



3.6.1.4 Phonemic similarity of naming errors

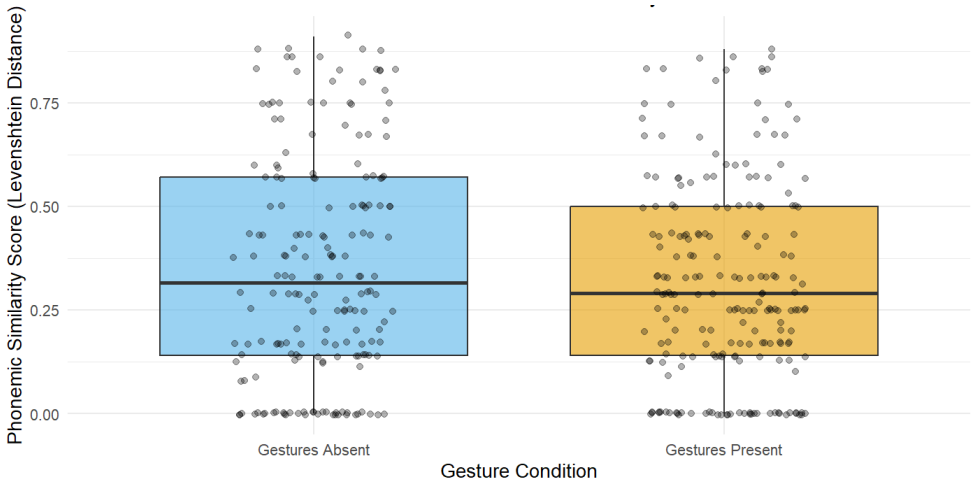
To test the proposal of the LRH that gestures function to facilitate phoneme selection and lexical retrieval, I also investigated whether phonemic similarity (phonological Levenshtein distance) of naming errors produced by PWA was affected by gesture condition. As for cosine similarity results reported above, there was no significant effect of gesture observation on phonemic distance of naming errors produced by PWA (=−0.027, SE=0.025, t(387.44)=−1.09, *p* = 0.277) (Table 5 and Fig. 7). However, the control variable target word length significantly predicated cosine similarity of naming errors, with longer

target words resulting in PWA guesses that were phonemically closer to the target word ($p < 0.001$).

Table 5. *Phonemic similarity of naming errors model output.*

	Estimate	Std. Error	df	t value	Pr(>|t|)
(Intercept)	0.345	0.032	32.311	10.823	0.000
Gesture condition (present)	-0.027	0.025	387.443	-1.088	0.277
Working memory	0.005	0.019	29.883	0.288	0.775
Aphasia Quotient	0.011	0.020	24.233	0.538	0.595
Semantic access score	0.032	0.019	29.947	1.696	0.100
First gesture condition (present)	-0.006	0.039	26.274	-0.163	0.872
Target word length	0.042	0.013	378.923	3.293	< 0.001
Target word frequency	-0.015	0.013	384.845	-1.197	0.232

Fig. 7. *Boxplot: effect of PWA gesture observation condition on phonemic similarity (Levenshtein) score of PWA naming errors.*
In this figure the Y axis represents phonemic similarity score (Levenshtein distance, 0.0-1.0 scale). Boxplots indicate the range of scores with the mean score shown as a line in each box. Scores in each gesture condition are presented separately (i.e. gestures absent, gestures present).



3.6.2 Speech analyses

I next analysed the effects of PWA gesture production on their speech output using three output measures: MATTR, MLU and mean utterance concreteness.

3.6.2.1 Transcription of speech data

Speech data from PWA was extracted from the video recordings and converted to MP3 audio format for analysis using VLC Media Player (VideoLan, 2006). MP3 audio files for each PWA were then imported into Praat (Boersma, 2002) for manual transcription. Automatic transcription was unavailable due to the unreliability of off-the-shelf automated systems at accurately transcribing aphasic speech (Mahmoud et al., 2023). Speech transcription was therefore carried out manually in Praat by the researcher and three trained transcribers. The unit of transcription was the *utterance*, understood as a unit of speech of variable length bounded by silence or a change of speaker (Bakhtin, 1986). See Appendix 3 for full details of the Coding Manual.

In total, 15,612 utterances were initially recorded. 1080 utterances (consisting of 6.9% of the total) were subsequently discarded and removed from the analysis; of these, 685 were discarded due to poor quality audio recording rendering transcription impossible, and 395 were discarded where at least two trained transcribers were unable to come to agreement about the correct transcription of the utterance. Therefore, in total 14,532 transcribed utterances were available for analysis. 7168 utterances were produced by PWA in the gesture constrained condition of the “Heads Up!” task and 7364 utterances were produced in the gesture permitted condition.

3.6.2.2 MATTR

For utterances in the gesture constrained condition, the MATTR was 0.609 (number of tokens: 25,199; number of types: 1831). For utterances in the gesture permitted condition, the MATTR was 0.615 (number of tokens: 22,982; number of types: 1671). In both conditions the MATTR was indicative of a moderate level of lexical diversity. MATTR was relatively low within the sample regardless of gesture condition; Cunningham & Haley (2020) found that 225 PWA with anomic aphasia had a MATTR of 0.90, although the chosen window size was 5 items compared to a window size of 50 items in

the current analysis. While small MATTR window size is effective for capturing very short-term lexical diversity, given the number of tokens in the analysis and the variability in utterance output between PWA a larger window size was chosen for this analysis for a more robust and reliable measure, following Covington (2007).

To determine if MATTR varied by gesture condition, I used a bootstrapping method to resample the data. I resampled all tokens from each condition 1000 times with replacement and recalculated MATTR for each resample to create a distribution of MATTR scores for each condition. Comparing distributions, I found no significant difference in MATTR by gesture condition (observed difference in MATTR between conditions = 0.0055; mean difference of –0.0008 across resampled pairs; p = 0.37), indicating that lexical diversity was comparable in utterances produced by PWA regardless of whether they did or did not use gestures.

3.6.2.3 MLU

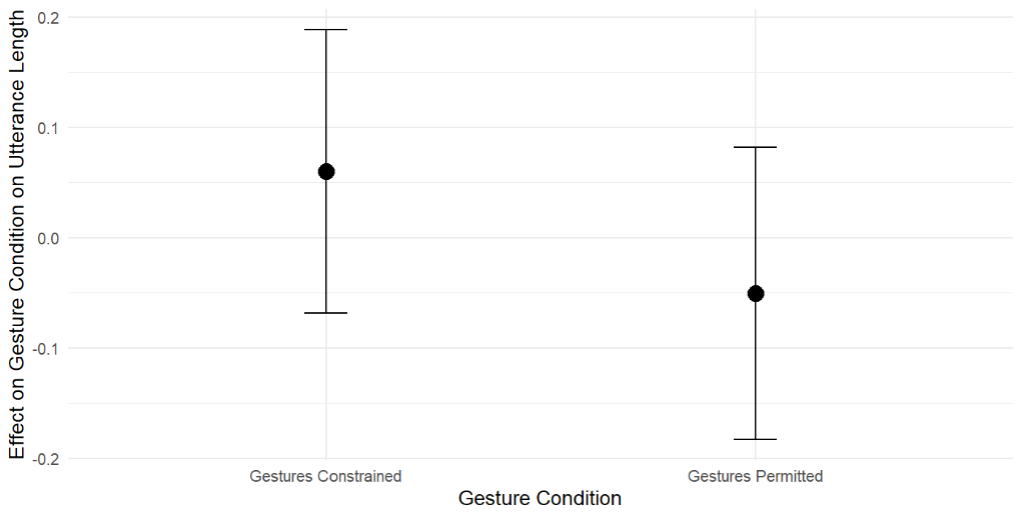
In the gesture constrained condition, the MLU for lemmatized forms of utterances was 3.52 (standard deviation 3.78); in the gesture permitted condition, the MLU for lemmatized forms of utterances was 3.12 (standard deviation 3.01). In both conditions MLU was relatively low; in a discourse-elicitation study of 92 PWA with anomic aphasia, Fromm et al. (2017) found an average MLU of 7.0. However, it should be noted that PWA in the current study presented with significantly more severe aphasia, with a reported average WAB-R Aphasia Quotient of 85.0 in Fromm et al. (2017) compared to 61.9 in the current study.

I used a linear mixed effects regression model to investigate if MLU varied significantly by gesture condition. There was a significant effect of gesture condition on MLU, with PWA utterances produced in the gesture permitted condition showing a significantly shorter MLU than PWA utterances produced in the gesture constrained condition ($\beta=-0.111$, $t(14532)=-6.789$, $p < 0.001$), as shown in Table 6 and Fig. 8.

Table 6. *MLU model output.*

	Estimate	Std. Error	df	t value	Pr(>|t|)
(Intercept)	0.060	0.065	0.350	0.919	0.675
Gesture condition (present)	-0.111	0.016	14532.000	-6.789	< 0.001
Aphasia Quotient	0.067	0.070	29.940	0.961	0.344
Working memory	0.012	0.049	30.080	0.247	0.807
Lexical semantics (STT)	0.029	0.066	5.433	0.444	0.675
Object semantics (CCT)	0.060	0.069	17.980	0.870	0.396
Action semantics (KDT)	0.007	0.072	27.950	0.097	0.924

Fig. 8. *Effect plot: PWA MLU by gesture condition.*
In this plot the Y axis indicates scaled utterance length. The bars represent range of utterance lengths and the dots represent mean utterance length. Utterance length for each gesture condition (i.e. constrained, permitted) is presented separately.



3.6.2.4 Concreteness

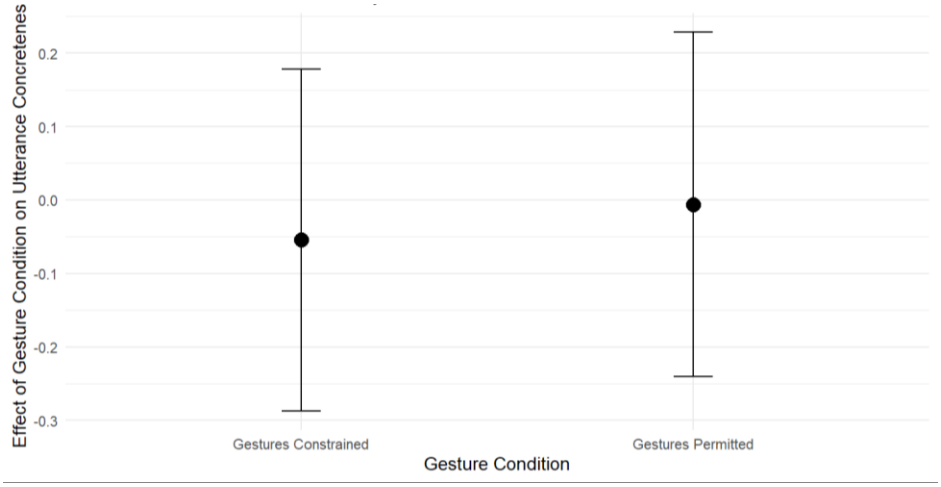
I next explored whether concreteness of utterances varied depending on whether they were produced in the gesture permitted or gesture constrained condition. Mean utterance concreteness in the gesture constrained condition was 2.092 (SD = 1.246); in the gesture permitted condition it was 2.138 (SD = 1.335). As shown in Table 7 and Fig. 9 this difference was significant (β = 0.049, t(14532) = 2.941, *p* = 0.003), with significantly higher mean

concreteness (less abstractness) of speech content found in the gesture permitted condition.

Table 7. *Concreteness model output.*

	Estimate	Std. Error	df	t value	Pr(>|t|)
(Intercept)	-0.055	0.119	29.650	-0.461	0.648
Gesture condition (present)	0.049	0.017	14532.000	2.941	0.003
Aphasia Quotient	-0.024	0.067	28.790	-0.356	0.725
Working memory	0.012	0.046	29.950	0.260	0.797
Lexical semantics (STT)	0.048	0.073	28.700	0.650	0.521
Object semantics (CCT)	0.006	0.070	28.720	0.092	0.928
Action semantics (KDT)	-0.015	0.071	29.160	-0.207	0.838
Site	0.070	0.139	29.320	0.508	0.615

Fig. 9. *Effect plot: PWA mean utterance concreteness by gesture condition.* In this plot the Y axis indicates scaled utterance concreteness. The bars represent range of utterance concreteness and the dots represent mean utterance concreteness. Utterance concreteness for each gesture condition (i.e. constrained, permitted) is presented separately.



3.7 Discussion

In this study, I asked 36 PWA to play a dyadic interactive naming-to-description game, “Heads Up!”, in which the researcher and the PWA took turns describing common objects and actions. Players were not permitted to say the name of the target word as part of the description. After hearing the

description, players were invited to guess what was being described. In half of trials, PWA and the researcher were constrained from gesturing alongside speech; in the other half of trials, gesture use was permitted. I investigated the extent to which gestures produced by the researcher and observed by the PWA facilitated naming accuracy in PWA; although the researcher's speech content in both conditions was broadly scripted, the researcher used gestures freely during gesture permitted trials with no restriction on gesture type to mimic gesture production in naturalistic interactions. Next, I investigated whether speech output measures (MATTR, MLU and utterance concreteness) varied depending on whether PWA were permitted to use, or constrained from using, gestures alongside speech.

Firstly, I found that trials in which PWA observed gestures produced by the researcher did not yield significantly more accurate naming responses than trials where the researcher's gesture use was constrained; nor did I find that PWA gesture recognition abilities or WM/EF moderated gesture observation benefit. However, when considering a subset of PWA for whom there was additional behavioural data, I found that (visual) attention capacity significantly moderated gesture observation benefit. While naming accuracy was relatively stable across PWA in the gesture constrained condition regardless of visual of attention capacity, PWA with poorer visual attention performed worse in the gesture present condition. Additionally, I found that lexical diversity of utterances as measured by MATTR did not differ significantly by gesture condition, but PWA gesture production affected both MLU (significantly shorter MLU when PWA used gestures) and mean concreteness of PWA utterances (utterances significantly more concrete when PWA used gestures); these findings may be (albeit indirectly) supportive of the Information Packaging Hypothesis rather than the LRH.

3.7.1 Theoretical implications

3.7.1.1 Gesture observation effects

As naming accuracy did not significantly differ by gesture condition, gesture observation did not facilitate naming accuracy for most PWA in this task, nor was this effect moderated by WM/EF or gesture recognition, contrary to predictions. However, (visual) attention capacity did moderate gesture observation benefit for naming, at least for a subset of PWA. Rather than gesture facilitating speech production in this case, I suggest that gesture observation effects targeted PWA *comprehension*. That is, observing the researcher's gestures did not necessarily facilitate PWA retrieval of the target words directly, but may instead have facilitated their understanding of the researcher's descriptions, by providing additional semantic information (i.e. through representational gestures) or highlighting key words in the speech stream (i.e. through beats and pragmatic gestures). Indeed, the fact that naming accuracy interacted with visual attention suggests that this is likely to be the case: where gestures were absent naming accuracy was (broadly) stable regardless of PWA visual attention capacity; when gestures were present, naming accuracy was lower in PWA with poorer visual attention capacity, suggesting they were unable to attend to and integrate gestural information with speech in subsequently retrieving the target word (i.e. naming accuracy). This moderating effect of visual attention capacity was (unsurprisingly) only evident in the gesture present condition, where visual information was presented alongside speech. This interpretation is supported by results from two further secondary analyses of PWA naming production, exploring whether gesture observation resulted in the production of incorrect words which were nevertheless semantically or phonemically close to the target word. Results for both these analyses were non-significant, suggesting that (most) PWA did not find the researcher's gestures helpful in "narrowing down" the correct target word, either phonemically or semantically.

3.7.1.2 Gesture production effects

Next, I evaluated the effects of PWA gesture production on their connected speech output. As discussed in Section 3.1, if gestures facilitated word retrieval, lexical diversity would presumably be higher when PWA produced gestures. However, this was not the case: lexical diversity did not differ

significantly by gesture condition ($p = 0.37$). This finding is therefore not supportive of the LRH, while the Information Packaging Hypothesis does not assume that lexical diversity is affected by gesture production (i.e. Prediction 2a).

Secondly, I compared MLU for PWA utterances produced in each gesture condition. MLU was significantly shorter for PWA utterances produced in the gesture permitted condition than for PWA utterances produced in the gesture constrained condition ($p < 0.001$). This result aligns with the predictions of the Information Packaging Hypothesis (i.e. Prediction 3a), and may also provide some evidence in support of the idea that gestures compensate for impaired speech production in PWA by conveying key parts of the message that they cannot easily convey through speech (Van Nispen et al., 2017). Therefore, (at least some) PWA were able to flexibly use less impaired communication channels to compensate for speech production deficits to manage communication efficiently (i.e. to maximise comprehensibility of the message while minimising effort for the sender and receiver), as proposed in Gibson et al.'s Efficiency Theory of communication (Gibson et al., 2019). I discuss the compensatory role of gestures in PWA communication in more detail in Chapter 4.

Finally, I also found that gesture condition predicted mean concreteness of PWA utterances, with concreteness ratings of PWA utterances significantly higher in the gesture permitted condition than in the gesture constrained condition. This finding aligns with previous research in neurologically healthy adults, e.g. Kok (2016). Importantly, findings from this study bear on the directionality of this relationship; it appears not only that gestures co-occur more frequently with more concrete lexical items, but that gesture production (or constraint) may *influence* (at least partly) the concreteness/abstractness of PWA speech content. While neither the LRH nor the Information Packaging Hypothesis speak directly to this effect, it suggests that gesture production/constraint integrates with speech production from earliest conceptualisation of the message (i.e. Prediction 4a).

3.7.2 Clinical implications

Findings from this study and the previous study “Go Fish!” underline the importance of individualised communication partner training for effective communication between PWA and neurologically healthy adults. Just as WM/EF and gesture recognition moderated gesture observation benefit for naming in a single-word context (as described in Chapter 2), visual attention capacity moderated gesture observation benefit/detriment for comprehension in a connected speech context. Therefore, clinicians should ensure that these aspects of cognition are comprehensively assessed prior to starting communication partner training; the use of gestures by communication partners should not be *assumed* either to be beneficial or detrimental a priori.

Furthermore, gestures (and other visual signals) may be more facilitatory of PWA speech comprehension in certain situations, regardless of the moderating effects of visual attention. For example, Krason (2023) found that observing both mouth movements and (iconic) gestures facilitated PWA comprehension in conditions of degraded speech, compared to when visual cues were absent, relative to normal speech conditions. In particular, mouth movements may act to facilitate phonological processing (because the shape of the mouth supports predictions about the auditory nature of the upcoming word), while gestures may facilitate semantic processing (because the gesture encodes semantic information about the upcoming word) (Krason, 2023). Further investigation into how visual communication signals (gesture, mouth movements, facial expressions, body movements etc.) interact to facilitate (or not) PWA speech comprehension, and exploration of the situations where this occurs, would be beneficial for developing personalised rehabilitation programmes.

Finally, findings from this study also indicate that where PWA use gestures, they do so to convey information as a compensation for impaired speech production. Therefore, communication partners should also be supported to attend to PWA gestures during interactions, without focusing on PWA speech output alone.

3.7.3 Limitations

An important limitation of this study is that although I argue above that gesture observation may facilitate PWA comprehension leading to downstream positive effects on speech production, the dependent variable used was naming accuracy (i.e. a measure of speech production, not comprehension). However, it is plausible that PWA may well have comprehended the description yet still been unable to retrieve the target word. A passive measure of comprehension without relying on speech production (such as pointing to a picture of the correct target word from a choice of four), may have been a more effective way of assessing PWA comprehension of the researcher's descriptions.

A second important limitation of this study is that in producing their descriptions of target words, PWA were required to inhibit production of the target word itself; any trials where PWA did produce target words were removed from analyses. Thus, PWA were required to devote cognitive resources both to inhibition *and* to word retrieval. Arguably, inhibition of target word production is non-naturalistic and the cognitive effort required to inhibit target word production may have affected how easily PWA were able to retrieve (other) words or plan their output. As discussed in section 3.6.2, MLU and MATTR for the transcribed samples were relatively low compared to samples derived from other studies; it is possible that the cognitive effort required to maintain inhibition of the target word negatively affected the extent to which PWA were able to produce connected speech effectively.

Finally, as with the higher-level interactions discussed in the previous chapter, power to observe higher-level interactions has not been established and may be limited. Results involving higher-level interactions should therefore be treated as tentative. Going forwards, these interactions could be more comprehensively evaluated via bootstrapping analyses.

3.8 Conclusion

To conclude, in a connected speech context gesture facilitation for PWA comprehension was closely related to (visual) attention capacity; only PWA who could attend to and integrate information from the researcher's gestures were successfully able to use this information to comprehend the message. Furthermore, PWA speech output varied as a function of gesture constraint or production; the concreteness of PWA utterances varied depending on whether they produced, or did not produce, gestures, suggesting that gesture use/constraint influenced PWA speech content from the point of conceptualisation of the message. Finally, the MLU of PWA descriptions was significantly shorter when PWA were able to use gestures alongside speech, suggesting that when permitted to gesture, PWA may use gestures in a compensatory manner to convey information to the interlocuter that they cannot convey through speech.

In the next chapter, I turn to this final point in further detail and consider PWA gestures from the perspective of the dyadic partner. Specifically, I ask how gestures produced by PWA contribute to communicative effectiveness from the perspective of the interlocuter, and whether any effects are moderated by individual differences in PWA cognitive profiles.

Chapter 4: Communicative Effectiveness

4.0 Abstract

In the previous chapters, I demonstrated how gesture observation and gesture production affect speech production in PWA, at single-word and connected speech levels. I evaluated these effects against the predictions of speaker-oriented theories of the functions of gesture. However, everyday interactions involve (at least) two people, and it is therefore vital to understand how the gestures produced by PWA in interactions affect communicative effectiveness from the perspective of their interlocuters. In this study, 168 neurologically healthy adults watched and/or listened to approximately 140 connected speech passages produced by PWA with moderate aphasia, drawn from the descriptions they produced in the “Heads Up!” game described in Chapter 3. Clips were presented either in audio-visual format, where participants could see and hear PWA, or as audio-only recordings, where participants could hear PWA but not see them. In both versions of the study, half of clips were drawn from descriptions elicited with gestures, and half without. Neurologically healthy participants were asked, firstly, to guess what PWA were describing in each clip, and secondly, when told the correct target word, to rate how well they felt each description matched the target word. Results indicated that neurologically healthy adults use information encoded in gestures produced by PWA to compensate for impaired speech production in successful communication. Participant guesses were significantly closer to target words (as determined through cosine similarity measures) in the gesture present condition of the audio-visual version of the study compared to the gesture absent condition, but this effect was not found in the audio-only version of the study. Communicative effectiveness ratings were also significantly higher in the gesture present condition of the audio-visual version of the study, an effect which was also not found in the audio-only version. Significant interactions with limb apraxia severity (PWA gesture praxis abilities) were also found in the audio-visual version of the study, but not the audio-only version. In sum, these results suggest that gestures, as a visual cue which encodes semantic information,

are vital in conveying meaning for interlocuters in the presence of impaired speech output.

4.1 Introduction

In the previous chapters, I have largely focused on how gestures (produced or observed) facilitate communication for PWA, such as their role in word-retrieval (i.e. the LRH; Krauss et al., 2000) and in 'chunking' information for verbalisation (i.e. the Information Packaging Hypothesis; Kita, 2000). Yet since interaction always requires (at least) two people, it is equally important to consider how gestures produced by PWA contribute to communication for their interlocuters. Interlocuter-oriented functions of gesture include emphasising perceptual or motor information (Goldin-Meadow & Alibali, 2013), drawing the interlocuter's attention to specific entities (Novack & Goldin-Meadow, 2017), and complementing speech content by adding extra information (Breckinridge Church et al., 2007). In some cases, information may only be conveyed to interlocuters by gestures: for example, in a picture description task, Melinger & Levelt (2005) found that neurologically healthy adults conveyed spatial information solely through gesture. In a similar narrative task, Cohen et al. (2011) found that around a third of the semantic information conveyed through gesture was not represented in speech and could not be inferred from participants' speech output. Finally, Kelly et al. (1999) demonstrated that listeners were significantly more likely to understand the meaning of an indirect request to close the door if it was accompanied by a gesture, e.g. "I feel cold" with a point towards the open door. These findings suggest that gestures may be an efficient way for speakers to convey semantic information to interlocuters, at least in some circumstances.

PWA are also able to use gestures to convey information to (neurologically healthy) interlocuters and may be more likely to do so when speech output is impaired (Ferré, 2021). In a single case study of "Chil", a PWA with severe expressive language deficits but relatively intact receptive language, Goodwin (2010) showed that Chil was able to use points and very limited

speech output (“yes”, “no”, “and”) to effectively convey meaning to his communication partners. Rose et al. (2017) asked neurologically healthy participants to write down messages conveyed by PWA through video clips in three conditions: speech-only, gesture-only, and speech and gesture combined. They found that the accuracy of messages was significantly higher in the combined condition than in either of the single modality conditions, suggesting that use of multiple channels by PWA was more effective for conveying information. In a similar study, Hermann et al. (1988) also found that PWA with severe aphasia produced informative gestures that could be understood by interlocuters even when speech output was severely impaired. Finally, de Beer et al. (2017) demonstrated that neurologically healthy interlocuters found narrations produced by PWA to be more comprehensible when they included both speech and gestures together, than when they were produced with speech alone. Taken together, these studies indicate that gestures are informative for interlocuters and can be used effectively by PWA to transmit information.

However, it has also been recognised that the degree to which PWA can successfully convey information via gesture can vary. Some studies have explored whether individual differences in PWA moderate any beneficial effect of gesture on communicative effectiveness. For example, Hogrefe and colleagues filmed 24 PWA with severe aphasia producing short story retells, then removed PWA speech content; naïve raters were asked to rate how well the gestural narrations depicted each story. They found that presence of semantic processing deficits and upper limb apraxia (gesture production; praxis) severity predicted the diversity and informativeness of gestures produced by PWA, with PWA with more severe deficits in each area respectively producing less diverse and less informative gestures from the perspective of the raters. Importantly however, as the gestures were divorced from speech content, researchers were unable to determine whether the gestures produced by PWA reinforced speech content, included information absent from speech, or facilitated lexical retrieval (Hogrefe et al., 2012).

Aphasia severity has also been explored in relation to gesture effectiveness. Mol et al. (2013) videoed 25 PWA (13 with moderate aphasia, 12 with severe aphasia) completing *The Scenario Test*, a communication assessment in which PWA role play communicative situations with the assessor, such as ordering food from a waiter in a restaurant. They found that both PWA with moderate and severe aphasia produced informative gestures from the perspective of the assessor, but that gestures produced by PWA with moderate aphasia were rated as more informative. Notably however, upper limb apraxia presence and severity was not assessed. In sum, while it appears to be the case that PWA use gestures to compensate for impaired speech production and convey information to interlocuters, the extent to which they are able to do this effectively may be moderated by individual differences in cognitive profiles.

4.2 Study aim and hypotheses

In the present study, I investigated the extent to which gestures produced by PWA with moderate aphasia (i.e. *Western Aphasia Battery-Revised* Aphasia Quotient score between 50.1 and 74.9) conveyed meaning to neurologically healthy interlocuters and whether any gesture benefit for interlocuters was moderated by individual differences in PWA, specifically limb apraxia (praxis) severity.

Unlike previous similar studies (e.g. de Beer et al., 2017) where interlocuters were presented with a relatively limited range of gestures produced by PWA, I placed no such restrictions on PWA gesture production and presented interlocuters with gestures of all types produced by PWA, including iconic/metaphoric, beat, emblem, points and pragmatic gestures, to more accurately reflect naturalistic interactions.

We took descriptions of everyday objects and actions produced by PWA derived from the “Heads Up!” task described in Chapter 3 and asked neurologically healthy adults to complete two tasks: firstly, I asked them to guess what the PWA was describing; and secondly, when given the correct

target word, I asked them to rate how well the description matched that target word. Stimuli were divided into two conditions: descriptions produced by PWA with gestures, and descriptions produced without gestures. I ran two concurrent versions of the study: an audio-visual version, where stimuli were presented through MP4 video clips and participants could see and hear PWA, and an audio-only version, where stimuli were presented through MP3 audio recordings so PWA voices could be heard but they were not visible.

First, following the findings of Cohen et al. (2011) I hypothesised that gestures produced by PWA could include key information that they could not convey through speech due to word-retrieval difficulties; as such I hypothesised that the gestures would contribute to successful transmission of PWA descriptions. However, gestures could only have such an effect if they could be seen. Therefore, I predicted that neurologically healthy participants would have a significantly greater understanding of the topic when gestures were visible (and therefore their guesses as to what PWA were describing would be more accurate), compared with when gestures were either not present (as in the no-gesture condition of the audio-visual study), or not visible (as in both conditions of the audio-only study). Similarly, communicative effectiveness ratings were also predicted to be significantly higher when gestures were visible.

However, in line with the proposals of the Lexical Retrieval Hypothesis on the role of gestures in facilitating lexical retrieval, I also predicted that gesture production would have dual effects both of conveying part of the message as outlined above, and also of facilitating speech output through cross-modal priming effects as described in Chapter 1. I therefore predicted that positive effects of gesture on word retrieval and speech production would be observed in the with-gesture conditions in both the audio-visual and audio-only versions of the study, regardless of whether gestures were visible to participants or not.

Prediction 1a: Cosine similarity scores of participant guesses and communicative effectiveness ratings will be significantly higher in the with-

gesture condition of the audio-visual version of the study, compared with the no-gesture condition of the audio-visual version and both conditions of the audio-only version.

Prediction 1b: Cosine similarity scores of participant guesses and communicative effectiveness ratings will be significantly higher in the with-gesture condition of the audio-only study, compared with the no-gesture condition of both study versions.

Finally, following Hogrefe and colleagues' finding that gesture informativeness was moderated by severity of upper limb apraxia in PWA (Hogrefe et al., 2012), I also predicted interaction effects between the communicative effectiveness of gestures and upper limb apraxia (praxis) severity. For descriptions produced by PWA with less severe/no upper limb apraxia presented in the audio-visual version of the study, I predicted that: (1) neurologically healthy participants' guesses would be closer to the correct target word(s); and (2) they would rate the descriptions as more effectively conveying the meaning of the target word(s). For descriptions produced by PWA with more severe upper limb apraxia, I predicted the opposite effects. No effects of limb apraxia were predicted either in the no-gesture condition of the audio-visual version of the study or in either condition of the audio-only version.

Prediction 2: Limb apraxia will moderate gesture benefit; in the with-gesture condition of the audio-visual version of the study, neurologically healthy adults' guesses will be closer to the correct target word(s) and descriptiveness ratings will be higher when descriptions are produced by PWA with no/less severe limb apraxia.

4.3 Methods

4.3.1 The present studies

I ran two inter-linked computer-based studies to assess the extent to which gestures produced by PWA facilitated communicative effectiveness from the perspective of neurologically healthy adults. In the first study, neurologically healthy adults viewed short video clips of PWA producing descriptions of common everyday objects and actions (e.g. *kitchen, comb, jumping, swimming*); videos included both auditory and visual information. In the second study, neurologically healthy adults listened to audio-only descriptions produced by PWA with no access to visual information. In both studies, in half of video/audio descriptions PWA were permitted to gesture alongside speaking; in the other half, PWA were not permitted to gesture.

A between-subjects design was used with neurologically healthy participants allocated a selection of approximately 35-40 video descriptions (audio-visual study) or audio descriptions (audio-only study). Video/audio descriptions were counterbalanced to ensure an equivalent proportion of with-gesture and without-gesture descriptions were presented to each participant. Video/audio descriptions were presented to participants in a random order. To ensure that participants were unfamiliar with the stimuli presented, participants were only permitted to take part in one version of the study (i.e. audio-visual or audio-only) and were excluded from participating in the other version. Ethical approval to carry out these studies was received from the High Risk Research Ethics Committee at UCL (Project ID 21765/001)

4.3.2 Participant sample

I recruited 168 neurologically healthy adult participants (101 females; 67 males; mean age: 37.4 years, SD: 13.9 years) through the Prolific recruitment platform (https://www.prolific.com/) who were paid for their participation in line with UCL participant payment guidelines. Of these, 85 participants completed the audio-only version of the study (50 females, 35

males; average age: 39.5 years) and 83 participants completed the audio-visual version of the study (51 females, 32 males; average age: 35.2 years).

Inclusion criteria for participants were: (i) native English speaker; (ii) no known history of neurological conditions; (iii) no visual and/or hearing deficits that could not be corrected to normal levels; (iv) access to a laptop or desktop computer with speakers and stable internet connection; (v) no significant previous experience of communicating with PWA. Participants were not permitted to complete either study on a mobile phone or tablet device. All participants provided informed consent in order to participate. Participants were advised to sit in a quiet room with no distractions for the duration of the task and were advised to wear headphones to maximise audibility.

4.3.4 Materials

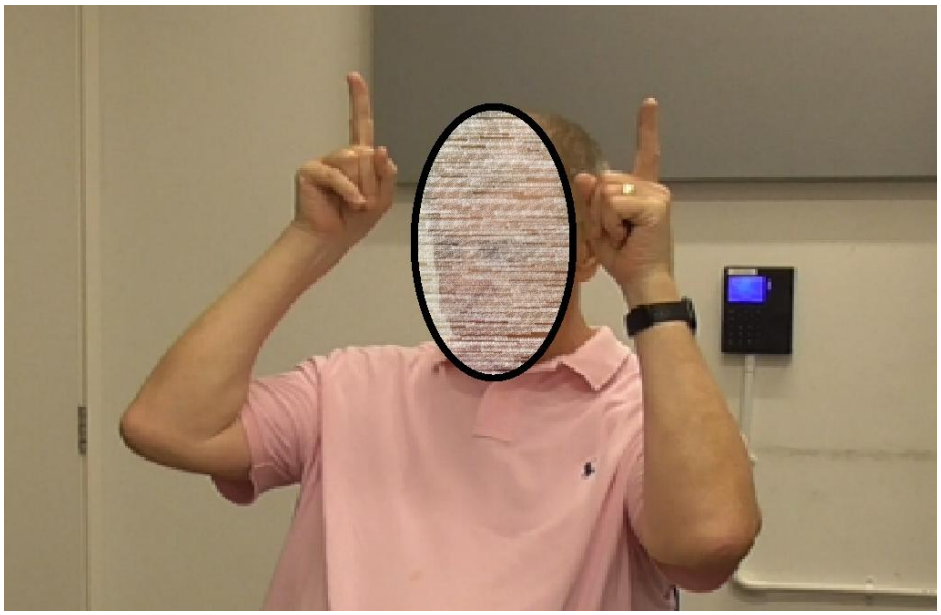
Audio-visual study: 140 approximately 1-minute MP4 video clips of PWA following left-hemisphere stroke describing an everyday action (e.g. eating, drinking, running, etc.) or object (e.g. banana, computer, hairbrush etc.) without using its name. The faces of all aphasic adults in the clips were blurred to avoid the risk of identification. Target words chosen for description by PWA were balanced for length and frequency. All aphasic adults were assessed as having moderate aphasia by a qualified speech and language therapist based on *Western Aphasia Battery-Revised* Aphasia Quotient scores of between 50.1 and 74.9. All PWA also previously completed assessments of upper limb apraxia severity (*Gesture to Sight test*, *MRRI Apraxia Battery*, Tarhan et al., 2015). For full details of PWA assessment performance, please refer to Appendix 1.

No videos containing aphasic adults with co-morbid severe apraxia of speech and/or severe dysarthria were included. In half the video clips, descriptions were produced by aphasic adults who were permitted to gesture while speaking; in the other half, descriptions were produced using speech only and PWA gesture use was constrained. Videos were drawn from a pre-

existing corpus developed by the researcher. Videos were split into four randomised lists of both with-gesture and without-gesture descriptions; lists were balanced for total length of stimuli.

Audio-only study: 140 approximately 1-minute MP3 audio clips which were stripped from the same video clips described above using VLC Media Player (https://www.videolan.org/). To maximise audibility of each clip, background noise was reduced by processing each clip in HandBrake, an open source video transcoder, using the *Denoise* functionality (https://handbrake.fr/). As above audio clips were split into four randomised lists of both with-gesture and without-gesture descriptions; lists were balanced for total length of stimuli.

Fig. 1. *PWA describing the target word "devil" in the gesture permitted condition.*



4.3.5 Procedure

In the audio-visual study, after providing informed consent, participants were randomly allocated to one of four experimental groups corresponding to one of the four video stimuli lists as described above. Participants were shown approximately 35-40 1-minute video clips of PWA describing an everyday

action or object as described in the *Materials* section above. Clips were presented on the Gorilla Experiment Builder software platform (https://gorilla.sc/).

Participants were informed prior to starting the study that clips would show adults with aphasia describing objects/actions and were given a brief outline of how aphasia affects communication. Before experimental trials began, participants were presented with a practice trial for familiarisation purposes. In each trial, participants were first shown a fixation cross for 250 milliseconds after which the video clip started playing automatically.

After presentation of each video clip, participants were asked: (1) to state what object or action they believed the PWA was describing, by typing their answer into a textbox; (2) following this, the correct target word was shown to participants who were then asked to rate, on a 7-point Likert scale, how effective in communicating the target word they felt the description was. Participants were required to watch each clip in full and each clip was only shown once. Participants were not permitted to skip, pause or rewind clips. Participants were offered an optional break after watching approximately 20 clips. Although the study was self-paced, participants were asked to complete the whole study within 75 minutes. To ensure participants attended throughout the task, 8 attention checks (forced choice yes/no questions) were included at random intervals.

For the audio-only version of the study, the same procedure was followed as for the audio-visual version, but audio clips were presented to participants instead of video clips. All other study aspects remained identical.

4.4 Data analysis

The same statistical modelling strategy was adopted in this chapter as described in Section 2.5.3. Linear mixed-effects regression analyses were performed using the *lme4* package (Bates et al., 2015) in RStudio (RStudio Team, 2020). Continuous variables were scaled and centred and categorical

variables were orthogonally coded to test for effects. To assess robustness of main effects (i.e. gesture condition), parametric bootstrapping analyses were also performed with 1000 simulations. In order to capture the diversity of gestures produced by PWA and to extend the findings of de Beer et al. (2017) which focused on a limited range of PWA gestures, all gesture types were included in the analyses. The *vif* function within the *car* package (Fox & Weisberg, 2019) was used to assess multicollinearity of predictors; in all models the variance inflation factor was below 2.0 indicating no significant multicollinearity between predictors. Full R code for all models below may be found in Appendix 9.

4.4.1 Pre-processing of cosine similarity data

Prior to analysing the cosine similarity data, I first extracted all participant text responses (i.e. participant guesses as to what PWA were describing in each clip). For trials where participants provided more than one response (e.g. “a rabbit or maybe a cat”) I took the first response. Trials where participants did not provide any response (e.g. “NA”) or indicated that they were unable to respond (e.g. “I can’t work this one out”) were removed from the analysis. In total 157 responses (3.14% of the total responses) were discarded.

I next lemmatized participant responses using the *lemmatizer* function of the *Natural Language Toolkit* package (Loper & Bird, 2002) in Python (v3.9.7, 2021) to reduce each response to its bare lemma in order to ensure direct comparability between participant responses. I then calculated the cosine similarity of each participant response in relation to the correct target word using pre-trained GloVe vector embeddings (100 dimensional trained on the 6B corpus) (Pennington et al., 2014) and the *cosine_similarity* function from the *PyTorch* package (Paszke et al., 2019). This yielded cosine similarity scores of between -1.0 and +1.0 for each participant response. Scores of –1.0 indicated the response was orthogonal in meaning to the target word; scores of +1.0 indicated the response was identical in meaning to the target word. These scores were then entered as the dependent variable for analysis.

4.4.2 Primary analyses

4.4.2.1 Cosine similarity analysis

To assess the effect of PWA gesture production on cosine similarity scores of participant guesses, for each version of the study I defined a linear mixed-effects regression model with cosine similarity as the dependent variable and gesture condition as the independent variable of interest. As I hypothesised that PWA aphasia severity and upper limb apraxia severity may affect gesture production and therefore how well participants were able to understand PWA descriptions, I also entered PWA aphasia severity (*Western Aphasia Battery-Revised* Aphasia Quotient) and upper limb apraxia severity (median PWA Gesture to Sight score) as variables into the model. Finally, I included target word frequency, target word length, and whether the target word was an object or action word as control variables. We also included a random intercept for PWA to capture variability in how each PWA produced descriptions.

4.4.2.2 Communicative effectiveness ratings analysis

To assess the effect of PWA gesture production on participants' communicative effectiveness ratings, for each version of the study I defined a linear mixed effects regression model with communicative effectiveness rating as the dependent variable. I entered the same predictors as for the previous analysis (see above).

4.4.3 Secondary analyses

I also carried out two exploratory secondary analyses to further investigate the role of gestures from the perspective of neurologically healthy adults.

4.4.3.1 Error analysis

First, on the basis that participants might benefit from semantic information conveyed via gestures produced by PWA but not sufficiently so to guess the

target word correctly, I investigated whether cosine similarity scores of *incorrect* guesses for target words made by participants varied by gesture condition and study version. For this analysis, I used the same model as the cosine similarity analysis in 4.4.2.1 above, but included only trials where guesses were incorrect in the analysis (n = 1889 incorrect guesses).

4.4.3.2 Reaction time analysis

I next investigated whether reaction time (response latency in milliseconds) for communicative effectiveness ratings varied by gesture condition and study version (n = 5037). Reaction time of participant responses was tracked automatically in Gorilla Experiment Builder; timing started when the participant response screen opened (timestamp 1) and finished when the participant entered a response (timestamp 2), with the reaction time being the difference between timestamps. I predicted that if observing gestures made PWA descriptions easier for participants to understand, participants would respond to the stimuli more rapidly, than for trials where gestures were not present or not visible. For this analysis I defined a Two-Way ANOVA with reaction time as the dependent variable, and gesture condition (gestures present/absent) and study version (audio-visual/audio-only) as the independent variables. I also included an interaction term between study version and gesture condition.

4.5 Results

4.5.1 Cosine similarity scores

In the audio-visual version of the study, the mean cosine similarity score across all target words and participant responses was 0.49 (SD: 0.35, total number of analysable responses: 2575). In the audio-only version of the study, the mean cosine similarity score across all target words and participant responses was 0.44 (SD: 0.34, total number of analysable responses: 2431).

In line with Prediction 1a, in the audio-visual version of the study I found a significant effect of gesture condition on cosine similarity scores. Cosine similarity scores of participant responses to PWA utilising both speech and gestures were significantly higher (indicating they were closer in meaning to the target word) than cosine similarity scores of responses to descriptions produced by PWA using speech alone, with an estimated effect size of 0.171 (SE = 0.043, t(2564.112) = 3.996, $p < 0.001$). To test robustness of the main effect of gesture condition on cosine similarity, I performed a parametric bootstrap with 1000 simulations using the *bootMer* function in the *lme4* package. The bootstrapped estimate of the effect size was consistent with the original model with minimal bias (β^=0.171; bias=0.001; 95% bootstrap confidence interval: 0.087 to 0.254).

In line with Prediction 2, I also found a significant interaction effect between gesture condition and apraxia severity. The cosine similarity scores of participant responses to descriptions of PWA with less severe upper limb apraxia were significantly higher than scores of responses to descriptions produced by PWA with more severe upper limb apraxia (β=−0.234, SE = 0.088, t(2566.452) = -2.646, $p = 0.008$).

However, contrary to Prediction 1b, in the audio-only version of the study no significant effect of gesture condition on cosine similarity score was found (β=−0.052, SE = 0.037, t(2422.397) = -1.408, $p = 0.159$). As before a parametric bootstrap with 1000 simulations was performed to test robustness of the (null) effect of gesture condition. Although the original estimate of the effect was negative (β=−0.052), the 95% bootstrap confidence interval (-0.125, 0.015) included zero, indicating that the effect was not statistically significant and gesture condition did not significantly predict cosine similarity. Furthermore, I found no interaction between gesture condition and upper limb apraxia severity in this version of the task (β=−0.022, SE = 0.038, t(2421.34) = -0.586, $p = 0.558$).

In both versions of the study, object words were guessed correctly significantly more frequently than action words, and more lexically frequent

words were guessed correctly significantly more often than less frequent words. Furthermore, aphasia severity (Aphasia Quotient) and target word length did not significantly predict cosine similarity scores in either version of the study.

Table 1. *Audio-visual study results (cosine similarity).*

	Estimate	Std. Error	df	t value	Pr(>|t|)
(Intercept)	-0.108	0.228	3.041	-0.471	0.669
Gesture condition (present)	0.171	0.043	2564.112	3.996	0.0001
Median gesture to sight score	-0.245	0.390	3.126	-0.628	0.573
Object word	0.227	0.041	2566.692	5.544	0.0000
Target word frequency	0.093	0.020	2564.317	4.773	0.0000
Target word length	0.014	0.021	2565.697	0.688	0.492
Aphasia Quotient	0.146	0.213	2.948	0.684	0.544
Gesture condition (present) * Median gesture to sight	-0.234	0.088	2566.452	-2.646	0.008

Fig. 2. *Effect plot: Participant cosine similarity scores by PWA gesture production condition (audio-visual version).*

In this plot the Y axis represents scaled cosine similarity scores. The bars represent range of cosine similarity scores and the dots represent mean cosine similarity scores in each gesture condition (i.e. gestures absent, gestures present).

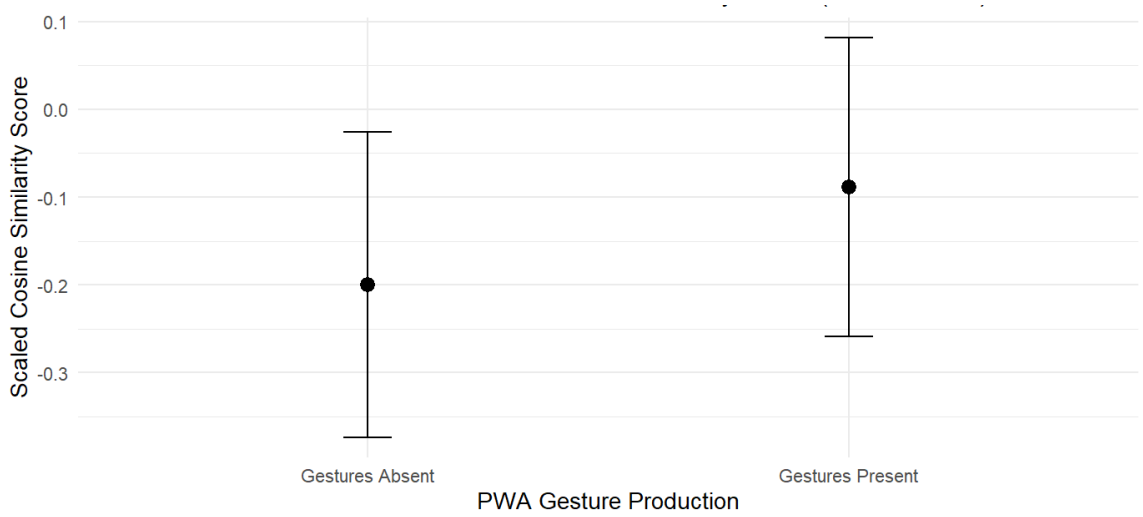


Fig. 3. *Boxplot: Distribution of participant cosine similarity scores by PWA upper limb apraxia severity (Gesture To Sight score) within the gesture present stimuli (audio-visual version).*

In this plot the Y axis represents scaled cosine similarity scores. Boxes indicate the range of scores with mean scores represented with bars.

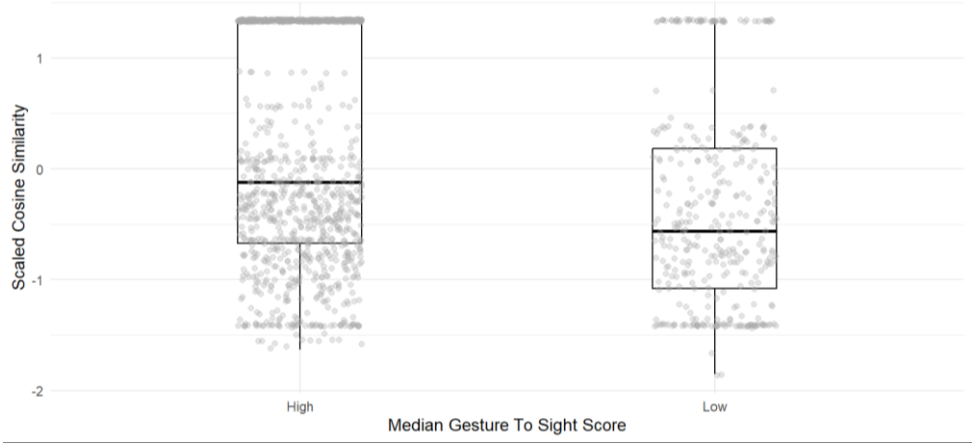
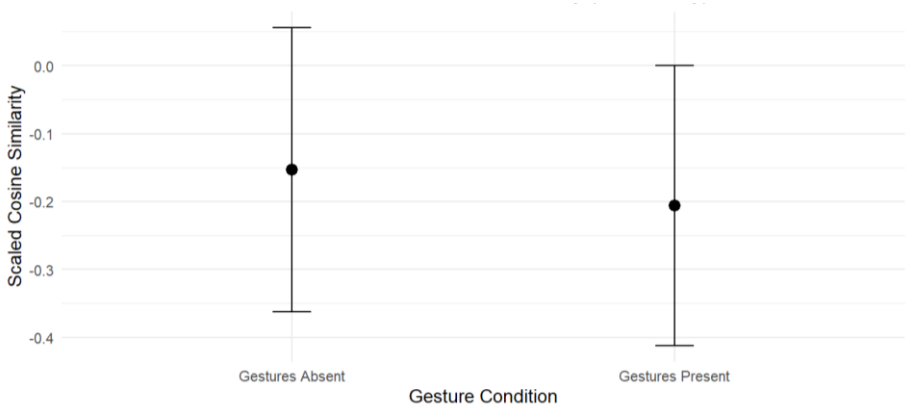


Table 2. *Audio-only study results (cosine similarity)*.

	Estimate	Std. Error	Df	t value	Pr(>|t|)
(Intercept)	-0.169	0.205	3.134	-0.825	0.468
Gesture condition (present)	-0.052	0.037	2422.397	-1.408	0.159
Median gesture to sight score	0.079	0.216	3.076	0.364	0.739
Object word	0.364	0.040	2421.381	9.118	0.000
Target word frequency	0.100	0.019	2422.261	5.190	0.000
Target word length	-0.027	0.021	2422.078	-1.312	0.189
Aphasia Quotient	0.284	0.267	2.977	1.067	0.365
Gesture condition (present) * Median gesture to sight	-0.022	0.038	2421.344	-0.586	0.558

Fig. 4. *Effect plot: Participant cosine similarity scores by PWA gesture production condition (audio-only version).*
In this plot the Y axis represents scaled cosine similarity scores. The bars represent range of cosine similarity scores and the dots represent mean cosine similarity scores in each gesture condition (i.e. gestures absent, gestures present).



4.5.2 Communicative effectiveness ratings

In the audio-visual version of the study, the mean communicative effectiveness rating provided by neurologically healthy adults was 3.58 (SD: 2.20; total number of responses: 2592). In the audio-only version, the mean communicative effectiveness rating provided by neurologically healthy adults was 3.48 (range: SD: 2.17, total number of responses: 2445).

In line with Prediction 1a, in the audio-visual version of the study I found a significant effect of gesture condition on communicative effectiveness ratings. Descriptions produced by PWA utilising both speech and gestures were rated as significantly more informative than descriptions produced by PWA using speech alone ($\beta = 0.267$, SE = 0.041, $t(2581.051) = 6.575$, $p < 0.001$). Bootstrapping was then carried out to test robustness of the main effect of PWA gesture production on communicative effectiveness ratings. The bootstrapped estimate of the effect size was consistent with the original model with minimal bias ($\beta\hat{}=0.266$, SE = 0.0402, bias = -0.0016, 95% bootstrap confidence interval: 0.1859 - 0.3423).

In line with Prediction 2 I also found a significant interaction between upper limb apraxia severity within the gesture present stimuli and communicative effectiveness rating, with significantly higher ratings for descriptions produced by PWA with less severe upper limb apraxia (β = -0.230, SE = 0.041), (t(2583.991) = 1.405, *p* = 0.001).

Contrary to Prediction 1b and in contrast with the audio-visual version of the task, no significant effect of gesture condition on communicative effectiveness rating was found in the audio-only task (β = -0.066, SE = 0.035, t(2435.121) = -1.857, *p* = 0.064). Although the result approached statistical significance, parametric bootstrapping confidence intervals included zero (95% bootstrap percentile confidence interval: -0.1324 - 0.0034) further suggesting the result was non-significant (β^=-0.066, SE = 0.035, bias = -0.0010). Furthermore, unlike in the audio-visual version I found no significant interaction between PWA upper limb apraxia severity and communicative effectiveness ratings (β = -0.018, SE = 0.039, t(2436.555) = -0.455, *p* = 0.649).

In both versions of the study, descriptions of object words were rated higher significantly more frequently than descriptions of action words. In the audio-only version, descriptions of more lexically frequent words were rated significantly higher than descriptions of less frequent words although this effect was not observed in the audio-visual version of the study. Aphasia severity (Aphasia Quotient) and target word length did not significantly predict ratings in either version of the study.

Table 3. *Audio-visual study results (communicative effectiveness).*

	Estimate	Std. Error	df	t value	Pr(>|t|)
(Intercept)	-0.032	0.275	3.016	-0.118	0.913
Gesture condition (present)	0.267	0.041	2581.051	6.575	0.000
Target word frequency	0.028	0.020	2583.991	1.405	0.160
Target word length	0.023	0.020	2582.478	1.188	0.235
Median gesture to sight score (GTS)	-0.513	0.468	3.061	-1.096	0.352
Aphasia Quotient	0.183	0.259	2.952	0.709	0.530
Object word	0.136	0.040	2583.980	3.366	0.001

Gesture condition (present) * Median GTS	-0.230	0.084	2582.884	-2.737	0.006

Fig. 5. *Effect plot: Participant communicative effectiveness ratings by PWA gesture production condition (audio-visual version).*
In this plot the Y axis represents scaled communicative effectiveness scores. The bars represent range of scores and the dots represent mean communicative effectiveness scores in each gesture condition (i.e. gestures absent, gestures present).

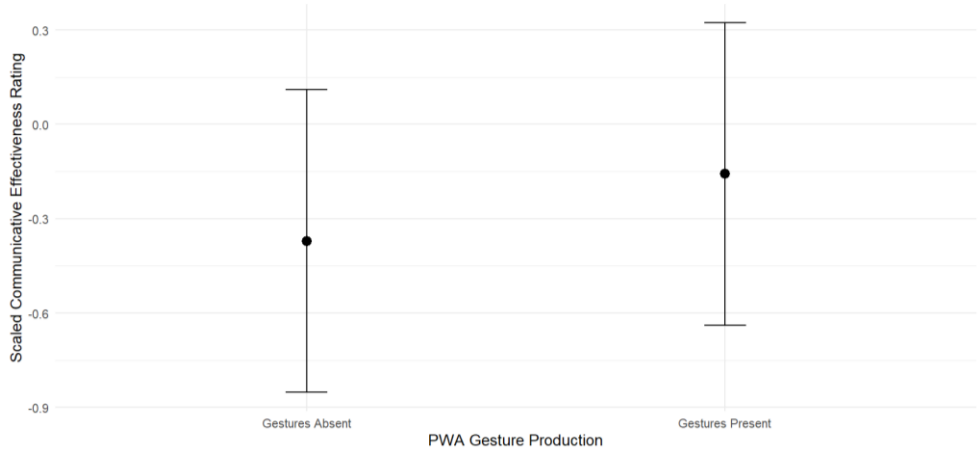


Fig. 6. *Boxplot: Distribution of participant communicative effectiveness ratings by PWA upper limb apraxia severity (Gesture To Sight score) within the gesture present stimuli (audio-visual version).*
In this plot the Y axis represents scaled communicative effectiveness scores. Boxes indicate the range of scores with mean scores represented with bars.

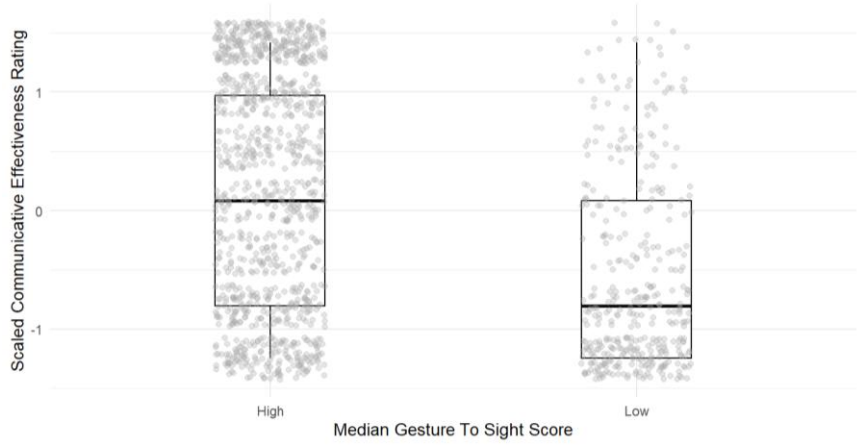
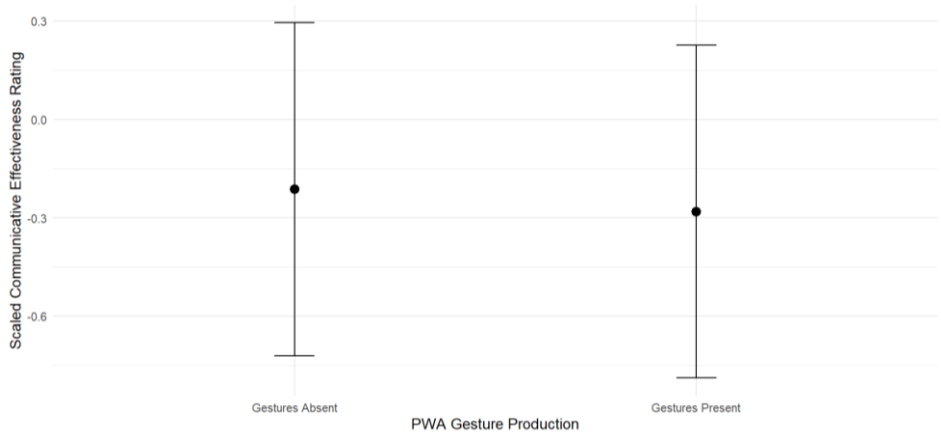


Table 4. *Audio-only study results (communicative effectiveness).*

	Estimate	Std.Error	df	t value	Pr(>|t|)
(Intercept)	0.001	0.217	3.104	0.005	0.996
Gesture condition (present)	-0.066	0.035	2435.121	-1.857	0.063
Target word frequency	0.091	0.020	2436.999	4.554	0.000
Target word length	-0.004	0.020	2436.078	-0.213	0.831
Gesture to sight score (GTS)	0.179	0.150	3.274	1.194	0.312
Aphasia Quotient	0.389	0.226	3.004	1.719	0.184
Object word	0.147	0.040	2436.911	3.706	0.000
Gesture_condition (present) * GTS	-0.018	0.039	2436.555	-0.455	0.649

Fig. 7. *Effect plot: Participant communicative effectiveness ratings by PWA gesture production condition (audio-only version).*
In this plot the Y axis represents scaled communicative effectiveness scores. The bars represent range of scores and the dots represent mean communicative effectiveness scores in each gesture condition (i.e. gestures absent, gestures present).



4.5.3 Secondary analyses

4.5.3.1 Secondary analysis 1: Cosine similarity scores of errors

Across both versions of the study, a total of 1889 incorrect guesses (i.e. words that were not the correct target word) were suggested by participants in responses to PWA descriptions. I lemmatized error words and generated cosine similarity scores using the procedure outlined in Section 4.4.1, resulting in a mean cosine similarity of 0.31 for error words in the audio-

visual version of the study (standard deviation: 0.20) and 0.29 for error words in the audio-only version of the study (standard deviation: 0.20).

In the audio-visual version of the study, I found no significant main effect of gesture condition on cosine similarity for errors, unlike in the primary analysis, but there was a significant interaction between gesture condition and apraxia praxis abilities, with lower gesture praxis abilities predicting lower cosine similarity scores in the gesture present condition (β = -0.031, SE = 0.053, t(1878.881) = -0.578, *p* = 564). In the audio-only version, no such interaction effect was found although interestingly there was a significant negative main effect of gesture condition in this analysis, with significantly lower cosine similarity scores for alternative words suggested in response to PWA descriptions in the gesture present condition (β = -0.233, SE = 0.111, t(1890.975) = -2.095, *p* = 0.036).

Table 5. *Audio-visual study results (cosine similarity scores of errors).*

	Estimate	Std.Error	df	t value	Pr(>|t|)
(Intercept)	-0.173	0.171	2.912	-1.015	0.387
Gesture condition (present)	-0.031	0.053	1878.881	-0.578	0.564
Gesture to sight score (GTS)	0.076	0.283	3.028	0.268	0.806
Object word	0.360	0.049	1880.245	7.296	0.0001
Target word frequency	0.097	0.023	1861.194	4.183	0.0001
Target word length)	0.029	0.025	1880.858	1.140	0.255
Aphasia Quotient	0.119	0.154	2.648	0.776	0.501
Gesture condition (present) * GTS	-0.218	0.101	1880.911	-2.153	0.031

Table 6. *Audio-only study results (cosine similarity scores of errors).*

	Estimate	Std.Error	df	t value	Pr(>|t|)
(Intercept)	-0.010	0.295	3.318	-0.032	0.976
Gesture condition (present)	-0.233	0.111	1890.975	-2.095	0.036
Gesture to sight score (GTS)	0.002	0.012	3.073	0.170	0.875
Object word	0.200	0.049	1890.929	4.095	0.0001
Target word frequency	0.170	0.023	1845.523	7.296	0.0001
Target word length)	0.029	0.025	1889.629	1.144	0.253
Aphasia Quotient	0.147	0.154	2.721	0.955	0.416
Gesture condition (present) * GTS	0.001	0.004	1890.688	0.165	0.869

Fig. 8. *Boxplot: Distribution of participant cosine similarity scores for errors by PWA upper limb apraxia severity (Gesture To Sight score) within the gesture present stimuli (audio-visual version).*
In this plot the Y axis represents scaled cosine similarity scores. Boxes indicate the range of scores with mean scores represented with bars. Scores are presented split by median gesture to sight score (high, low).

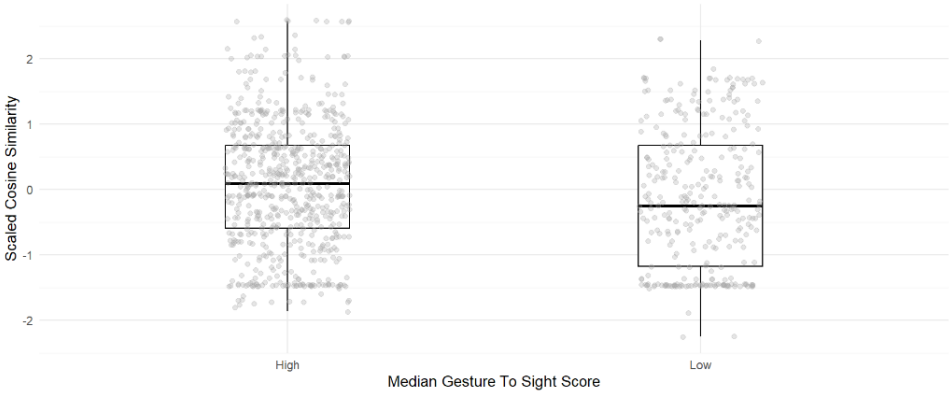


Fig. 9. *Boxplot: Distribution of participant cosine similarity scores for errors by PWA upper limb apraxia severity (Gesture To Sight score) within the gesture present stimuli (audio-only version).*
In this plot the Y axis represents scaled cosine similarity scores. Boxes indicate the range of scores with mean scores represented with bars. Scores are presented split by median gesture to sight score (high, low).

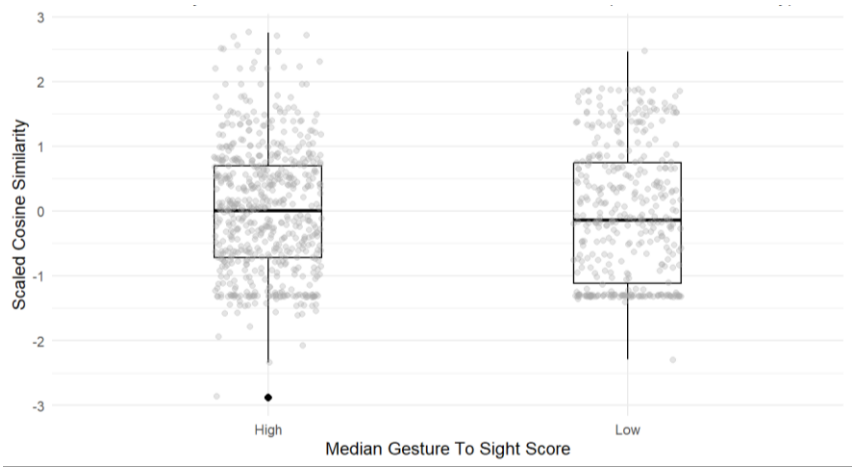
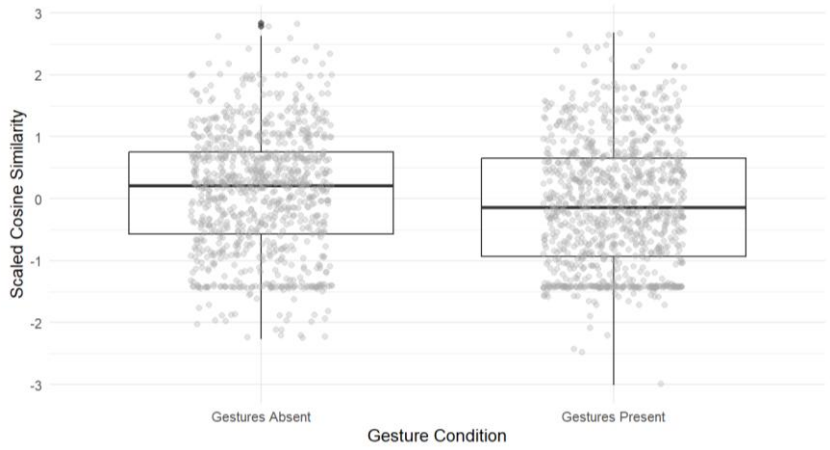


Fig. 10. *Boxplot: Distribution of participant cosine similarity scores for errors by PWA gesture production condition (audio-only version).*
In this plot the Y axis represents scaled cosine similarity scores. Boxes indicate the range of scores with mean scores represented with bars. Scores are presented split by gesture condition (i.e. gestures absent, gestures present).



4.5.3.2 Secondary analysis 2: Reaction time

In the audio-visual version of the study, the mean reaction time for assigning a rating to the descriptions provided by PWA in the gesture present condition was 3403.39 milliseconds; in the gesture absent condition it was 3401.47 milliseconds. In the audio-only version of the study the mean reaction times were 4240.56 milliseconds and 4343.57 milliseconds respectively, substantially longer than in the audio-visual version. To determine if differences in reaction times were significant by gesture condition or study version, I performed a Two-Way ANOVA with gesture condition, study version and interaction between them as independent variables and reaction time as the dependent variable. I then carried out post-hoc tests using Tukey's HSD to determine which levels of the variables showed significant differences in reaction time, with pairwise contrasts.

Overall, although I found that response times were significantly faster (regardless of gesture condition) in the audio-visual version of the study compared to the audio-only version ($p = 0.02$), I found no significant

difference in response times comparing gesture conditions within different versions of the same study (all $p > 0.05$).

Table 7. *ANOVA output.*

Factor	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Gesture_condition	1	2.993	2.993	2.998	0.0834
Study_Version	1	5.238	5.238	5.246	0.0220
Gesture_condition * Study_Version	1	3.011	3.011	3.016	0.0825
Residuals	5033	5025	0.998		

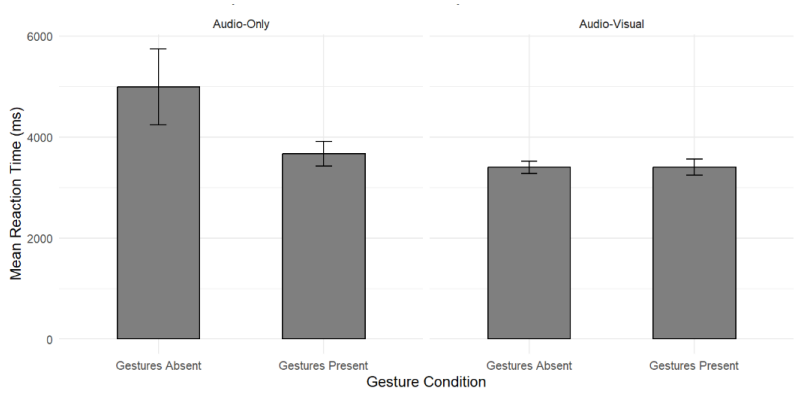
Table 8. *Tukey's HSD output.*

Gesture Condition	Study Version	Estimated Mean	SE	95% CI (Lower - Upper)
-1 (Absent)	Audio	0.0856	0.0295	[0.0278, 0.1433]
1 (Present)	Audio	-0.0124	0.0278	[-0.0668, 0.0420]
-1 (Absent)	AV	-0.0319	0.0292	[-0.0891, 0.0254]
1 (Present)	AV	-0.0317	0.0265	[-0.0836, 0.0202]

Table 9. *Tukey's HSD pairwise contrasts.*

Contrast	Estimate	SE	z-ratio	Pr(>z)
(Gesture Absent Audio) - (Gesture Present Audio)	0.0980	0.0405	2.421	0.0732
(Gesture Absent Audio) - (Gesture Absent AV)	0.1174	0.0415	2.830	0.0241
(Gesture Absent Audio) - (Gesture Present AV)	0.1173	0.0396	2.961	0.0162
(Gesture Present Audio) - (Gesture Absent AV)	0.0194	0.0403	0.482	0.9631
(Gesture Present Audio) - (Gesture Present AV)	0.0193	0.0384	0.503	0.9585
(Gesture Absent AV) - (Gesture Present AV)	-0.0001	0.0394	-0.004	1.0000

Fig. 11. *Mean participant reaction time for assigning communicative effectiveness ratings by gesture condition and study version.*



4.6 Discussion

In these studies, I showed 168 neurologically healthy adults video/audio clips of descriptions of everyday objects or actions produced by PWA. In half of the descriptions, PWA gestured when speaking; in the other half, gesture use was constrained. For each clip, I asked participants to: (1) guess the object/action described by the PWA, described either with gestures and speech or with speech alone; and (2) rate how well each description produced by the PWA matched a given target object/action word. Stimuli were presented either as audio-visual or as audio-only recordings. All neurologically healthy adult participants were naive to the purpose of the studies and had limited previous knowledge or experience of aphasia.

I first investigated whether gesture presence/absence predicted how well neurologically healthy adults were able to guess the identity of the object/action described by PWA, and whether this varied by study version. In the audio-visual version of the study, where participants were able to see the gestures produced by PWA, I found that cosine similarity scores of guesses from trials presented in the gesture present condition were significantly higher than cosine similarity scores of guesses from trials in the gesture absent condition, suggesting that PWA gestures were assisting participants to understand the spoken descriptions and resulting in guesses that were closer to the actual target word. However, as expected I found no effect of

gesture condition in the audio-only version of the study. This reinforces the suggestion that participants drew benefit from gestures when they were visible, but PWA gesture production did not impact speech and therefore they did not derive comprehension benefit from any effects of gesture production on speech output, as instead predicted by the LRH.

Next, I investigated whether gesture presence/absence predicted how well neurologically healthy adults rated the overall communicative effectiveness of the description. Again, I found that communicative effectiveness ratings were significantly higher in the gesture present condition compared to the gesture absent condition in the audio-visual version of the study, but this effect was not replicated in the audio-only version of the study. This further lends support to the suggestion that gestures functioned primarily to convey information key to understanding the message.

In the audio-visual version of the study, I also found significant interaction effects between severity of upper limb apraxia and gesture condition, in that upper limb apraxia moderated the extent to which gestures produced by PWA were informative (or not). In the audio-visual version of the study, guesses elicited in response to descriptions produced by PWA with less severe upper limb apraxia received higher cosine similarity scores than guesses elicited in response to descriptions produced by PWA with more severe upper limb apraxia. This suggests that for neurologically healthy adults, gestures were more informative (and presumably more intelligible) when produced by PWA with less severe upper limb apraxia. Unsurprisingly, this interaction effect was absent in the audio-only version of the study where participants were unable to see the PWA gestures produced. The same effects were observed in the analysis of communicative effectiveness ratings, with significantly higher ratings for descriptions produced by PWA with less severe limb apraxia in the gesture present condition of the audio-visual version of the study, compared with the gesture absent condition. As before, this effect was not seen in the audio-only version of the study. These findings replicate those by Hogrefe et al. (2012), who found that although naive communication partners generally benefited from gestures produced by PWA

in understanding their (spoken) messages, limb apraxia severity predicted how well communication partners were able to comprehend the gestures produced.

I next explored whether gesture presence/absence or study version predicted the cosine similarity scores of incorrect guesses of target words elicited from participants. In the audio-visual version of the study, unsurprisingly, I replicated the significant interaction effect between gesture condition and upper limb apraxia severity as described above. However, in the audio-only version of the study I found that there was a significant negative main effect of gesture presence. Cosine similarity scores of participant guesses were significantly lower when PWA produced gestures in this version of the study. This likely reflects PWA attempting to use gestures to replace speech in this condition (i.e. as compensation), which they were unable to do in the gesture constrained condition. As neurologically healthy adults could not see PWA gestures in the audio-only version, they could not use them to comprehend the message and so had to rely on PWA impaired speech output instead, resulting in guesses that were less semantically-aligned to the target word.

Finally, I explored whether response times for rating descriptions differed significantly between versions of the study and in interaction with gesture conditions. I found that neurologically healthy adults produced ratings significantly faster in the audio-visual version of the study compared to the audio-only version, although response latency difference between gesture conditions was non-significant. Response latencies in the audio-only gesture-absent condition were slowest of all, but difference between conditions was non-significant. The significantly faster responses in the audio-visual study version compared to the audio-only version suggest that neurologically healthy adults found it easier to comprehend descriptions when they were presented multimodally through videos rather than through audio recordings only. These findings are similar to those by Schwarz et al. (2022), who found that neurologically healthy adults (and children) were more accurate in their responses and responded more quickly to stimuli presented multimodally

than through speech only stimuli, even when the speech content was the same in each case.

The lack of gesture effect on reaction time is somewhat surprising, as it may be predicted that if gestures facilitated participant comprehension by conveying (part of) the message, neurologically healthy adults would respond significantly more quickly in the gesture present condition of the audio-visual study version, than in the gesture absent version. However, this pattern was not observed and indeed this finding replicates results by Perniss et al. (2020) in a study comparing response times for congruent/incongruent speech and gestures in comprehension tasks with neurologically healthy adults; gestures congruent with speech did not result in faster reaction times, although (as above) congruency did correlate with improved comprehension accuracy. The reasons for this pattern remain unclear, but it is possible that participants derived benefit from other visual cues (e.g. body positioning, facial expressions, mouth movements) in the audio-visual version of the study even when gesture use was constrained, which enabled them to more readily comprehend the descriptions and produce ratings more quickly.

4.6.1 Theoretical implications

These findings suggest that for at least some PWA with moderate aphasia, from the perspective of neurologically healthy interlocuters gestures function effectively as a means of conveying information absent from, or not clearly explicated in, speech. If gestures had facilitated PWA lexical retrieval, as proposed by the LRH, I should have seen significant effects of gesture condition across both primary analyses in the auditory-visual and audio-only versions of the study, as even if the gestures were not visible to interlocuters speech content should have been enhanced by the production of gestures, and therefore should have elicited higher communicative effectiveness ratings and cosine similarity scores for participant guesses.

Instead, these results lend further weight to the idea that gestures are meaningful and carry (part of) the message, in particular those aspects that PWA struggle to convey through speech due to impaired word-finding abilities. As proposed by Kita (2000), gestures may therefore facilitate organisation or ‘packaging’ of information for PWA prior to verbalisation. Thus, when verbalisation is unavailable (i.e. when ‘tip-of-the-tongue' states or word-finding difficulties occur), gestures may function as an alternative means to convey the intended information. This account is supported by van Nispen et al. (2017), who found that around a fifth of gestures produced by PWA were determined to be essential for successful interpretation of the message by neurologically healthy interlocuters, compared with approximately 5% of gestures produced by neurologically healthy adults. This study therefore provides additional evidence that gestures function as a compensatory measure for PWA with impaired speech production.

These results also shed light on the suggestion that speech output and gesture production break down in parallel in PWA (e.g., Buck & Duffy, 1980; Cicone et al., 1979). Indeed, this does not necessarily seem to be the case. PWA with moderate aphasia were able to produce informative gestures and convey semantic information successfully even in cases of very significant speech output deficits, suggesting that at least some aspects of semantic representations remained accessible to the gesture production system, even if unavailable or inaccessible to the speech production system. Furthermore, although upper limb apraxia severity moderated gesture benefit for conveying information, PWA gestures were somewhat informative across the board, even in cases of more severe limb apraxia. These findings support more recent studies by de Beer et al. (2017) and Kong et al. (2019), who also found evidence of differential severity of breakdowns in speech and gesture production in PWA. Indeed, it may be the case that while connections between semantic representations, phonological forms and the motor speech system is disrupted in (at least some) PWA, connections between semantic representations and motor systems for gesture production may remain (relatively) intact.

4.6.2 Clinical implications

These findings suggest that while gesture production in PWA may facilitate overall communicative effectiveness in interactions with neurologically healthy adults, gestures do not necessarily result in improved word-finding abilities. Gesture production by PWA may therefore more appropriately be considered as a compensatory measure for enabling increased participation in communication, rather than as a means of improving word-finding difficulties. Furthermore, since ability to convey meaning through gestures was moderated by upper limb apraxia severity, with gestures produced by PWA with more severe limb apraxia contributing less to message comprehension, clinicians should assess the presence and severity of upper limb apraxia as part of management. Communication partners may also need training to support them to interpret PWA gestures, especially in cases where gestures are less informative; programmes such as the *SPPARC Conversation Training Programme* (Lock et al., 2001) or its adaptation *Better Conversations with Aphasia* (Beeke et al., 2013), which support communication partners to attend to and interpret non-verbal cues, may be usefully incorporated into therapy sessions to achieve this.

4.6.3 Limitations

There were several limitations of the studies described above. Firstly, all PWA whose descriptions were used in this study were assessed as having relatively unimpaired lexical semantic, object semantic and action semantic processing abilities, all with diagnoses of conduction, anomic or Broca's aphasia. As discussed in Chapter 1, PWA with these aphasic sub-types typically produce significantly more informative representational gestures than PWA with other aphasic sub-types. The PWA whose descriptions featured in this study may therefore not be representative of PWA generally, especially not those with more impaired semantic processing. It is possible that gestures produced by PWA with more significant semantic processing impairments may be less informative to interlocuters; indeed, Kindell et al. (2013) demonstrated that people with semantic dementia produced

significantly fewer representational (iconic) gestures as disease progressed with corresponding increase in deictic gesture production, which may be attributed to progressive loss of semantic representations. The same patterns have been observed in PWA with Wernicke's aphasia (Sekine & Rose, 2013). Future studies should incorporate PWA with a wider range of semantic processing impairments in order to further investigate the extent to which semantic processing deficits moderate gesture informativeness.

Secondly, unlike in naturalistic interactions, neurologically healthy participants had no "common ground" established from which to understand PWA descriptions and no way to request clarifications or repair misunderstandings. Common ground consists of experiences and knowledge that are shared between two (or more) people (Clark, 2015). Common ground is known to make communication more efficient; familiar speakers may produce shorter utterances, while listeners can "fill in the gaps" using shared knowledge (Brown-Schmidt, 2009). Participants in the study had no prior knowledge of the objects/actions PWA would be describing in each clip (or indeed, any prior knowledge of aphasic communication) and therefore had no way to "fill in the gaps" in PWA speech production, perhaps meaning they needed to rely more on PWA gestures than familiar communication partners would need to. In future studies, both naïve and familiar raters could be used to assess the usefulness of PWA gestures, to determine if interlocuter familiarity moderates the extent to which gestures convey information separate from speech.

Thirdly, the reaction time analysis results should be treated with caution as they may not adequately capture speed of comprehension. In both versions of the study, participants were required the watch the videos/listen to the audio recordings in full *before* the response screen appeared and therefore may have understood what was being described well before they had the opportunity to indicate this. In future studies, participants could be permitted to respond as soon as they comprehend the message (i.e. even while the videos/audio recordings are still being played) to more accurately capture speed of comprehension.

4.7 Conclusion

To conclude, these results indicate that gestures produced by PWA successfully conveyed (part of) message content from the perspective of neurologically healthy adults, but only when they were visible. Gestures produced by PWA that were not visible to interlocuters did not aid comprehension, such as through facilitating lexical retrieval or aiding speech production. In sum, gestures provided semantic information to interlocuters which compensated for impaired speech production, significantly aiding communication. Importantly, I also found that gesture effect was moderated by upper limb apraxia severity, as previously suggested by Hogrefe et al. (2012), and that gestures produced by PWA with more severe limb apraxia contributed less to communicative effectiveness than gestures produced by PWA with less severe limb apraxia.

In the preceding chapters, I have demonstrated that individual differences in working memory/executive functioning predict gesture observation and production benefit for naming in PWA in semi-naturalistic interaction contexts, at least for single words, and I have also demonstrated that gestures produced by PWA effectively convey semantic information to interlocuters and can be critical for communicative success. In my next chapter, I turn to the question of whether gesture observation can therefore be used as part of an intensive speech and language therapy programme to support communication (specifically word-finding) in PWA.

Chapter 5: Therapy

5.0 Abstract

Implementation of effective aphasia therapy is a key healthcare priority and personalised therapy approaches that deliver sufficient dosage to effect change are increasingly sought after. Previous studies have utilised gestures (both production and observation) alongside speech stimuli to improve naming accuracy in PWA but effects have been variable and delivering sufficient dosage to be effective has proven challenging (Marshall et al., 2013). Building on the work in previous chapters, in this study I developed and trialled a novel computer-based, self-paced naming intervention utilising gestures for PWA. I recruited 4 PWA with moderate aphasia who had previously taken part in studies outlined in Chapters 2 and 3. I asked PWA to complete 20 sessions of self-paced computer-based therapy based around repeated naming of a personalised set of treatment items. Target words were presented via video clips, half of which were presented with the researcher using speech and paired iconic gestures, and half of which were presented with the researcher using speech only. I found that naming accuracy for treated items significantly improved one week after completing the intervention and effects were maintained one month later, although there was no significant difference in performance overall between items presented with gestures and items presented without. However, within the set of items presented with gestures, I found that the more informative the iconic gesture produced by the researcher the more accurately the word was named by PWA at follow-up testing. These findings suggest that gestures, at least those that are highly imagistic, may facilitate naming in PWA through a cross-modal priming process where sensory-motor information activates conceptual semantic representations, aiding retrieval. These findings further suggest that computer-based self-paced therapy is a viable means of delivering high-dose anomia therapy to (at least some) PWA.

5.1 Introduction

5.1.1 Therapy delivery in aphasia

Speech and language therapy interventions for anomia have been shown to be effective in improving naming accuracy in PWA, but dosage of interventions remains variable (Brady et al., 2016). A meta-analysis of 25 treatment studies for aphasia has revealed that the minimum dosage for therapy to be effective varies depending on the treatment target and acuteness of aphasia; for example, improvements in auditory comprehension were found to require at least 9 hours of therapy per week for up to 20 weeks, with intervention provided for at least 4 days per week for PWA with chronic aphasia. In contrast, PWA less than 3 months' post-stroke may require only 3-4 hours of intervention per week for the same timescale to show measurable changes in functional communication. Additionally, personalised and tailored therapy with regular home-based practice was found to yield the best results (Brady et al., 2022).

In many English-speaking countries, PWA do not receive sufficient therapy to achieve measurable or long-lasting improvements in communication (Richard et al., 2009). In the UK (excluding Scotland), data from the Stroke Sentinel National Audit Programme (SSNAP) revealed that the average length of a speech and language therapy session nationally was 20.9 – 32.0 minutes and 94% of PWA received intervention for no more than 3 months post-stroke (Intercollegiate Stroke Working Party, 2015; 2019). Palmer et al. (2018) reported that the median amount of therapy received by PWA living in the community was 6.3 hours across three months, with around one one-hour session delivered every two weeks. Notably, 90% of sessions were delivered in-person, one to one. Finally, a recent study of 227 speech and language therapists delivering aphasia therapy across the UK revealed significant variation in treatment dosage regionally, with PWA in Northern Ireland receiving approximately three times as much therapy per week (median 180 minutes) as PWA living in Southwest England (median 60 minutes) (Monnelly et al., 2023). Evidence of similar research-practice gaps

has been found in the delivery of aphasia therapy by services in the US (Cavanaugh et al., 2021), New Zealand (Gibson, 2023), Australia (Shrubsole et al., 2018) and Ireland (Manning et al., 2020).

A possible reason for this gap is lack of staff to deliver therapy at sufficient dosages, and online self-paced therapies have been in development to meet this need. Therapies generally consist either of structured programmes, where PWA move sequentially through a pre-defined set of activities, or free-flow programmes, where PWA choose their own activities and the order in which they complete them. One example of an effective self-paced structured programme is *iTalkBetter*, an adaptive anomia therapy app; a recent clinical trial of *iTalkBetter* involving 27 PWA over 6 weeks showed that PWA completed an average of 45 hours of therapy with a 13% improvement in naming accuracy for trained items at post-intervention testing. Importantly, PWA required minimal or no technical support from the research team over the course of the trial, reducing time pressure on clinicians (Upton et al., 2024).

5.1.2 Gesture therapy for word retrieval in aphasia

Therapy for anomia that includes a gestural element (either production or observation) has been found to be effective in improving naming of both object and action words in (some) PWA, although size of treatment effects is variable. The hypothesised mechanism of action is assumed to be cross-modal priming effects, where gestures activate either conceptual semantic features, or phonemic features/lexical selection of target words, following the Information Packaging Hypothesis or Lexical Retrieval Hypothesis respectively (outlined in detail in Chapter 1).

In a small study of 6 PWA, Rose & Douglas (2001) found that iconic gesture production facilitated immediate naming of a target object word. Notably, production of points did not facilitate naming, and the authors suggest that (for at least some PWA) production of the iconic gesture specifically added or reinforced conceptual semantic information that strengthened activation of

the target word at the level of the lemma, before phoneme selection. Similar results were found by Rose et al. (2002) who used a single case study design with multiple baselines to compare the effectiveness of iconic gesture, points, cued articulation and visualisation alongside speech as treatments for object picture naming. As in the previous study, significant improvements in naming were found for the iconic gesture production treatment only. For treatment of action word naming, Rose & Sussmilch (2008) found that semantic feature analysis treatment combined with gestures was effective in improving action naming accuracy, although the size of treatment effects varied and participants with the most severe semantic processing deficits did not show a strong treatment response. Finally, in a study of 4 PWA with moderate-severe word retrieval deficits, Rodriguez et al. (2006) found that one PWA showed significant improvement in action naming following speech and gesture treatment, while three did not. However, two of the three remaining PWA showed significant increases in meaningful gesture production following treatment, despite moderate-severe limb apraxia (praxis) deficits. The authors suggested that even where lexical retrieval does not improve with therapy, functional communication may improve with gestural treatments in PWA with severe aphasia.

Gesture observation has also been found to improve naming in at least some PWA. Kroenke et al. (2013) investigated whether pairing (novel) target words with gestures improved retrieval, compared to a verbal treatment alone. Interestingly, researchers found that PWA with better preserved semantic processing skills benefited more from multimodal training in retrieving the novel object words, while PWA with better preserved phonological processing skills benefitted more from purely verbal training strategies. As in the previous study by Rose & Douglas (2001), it is possible that for PWA with better preserved semantic processing skills, gestures facilitated word retrieval by priming and strengthening conceptual semantic activation, leading to more efficient retrieval despite the presence of phonological access impairments. This account is supported by Bonifazi et al. (2013), who found that observation of meaningful gestures paired with action words facilitated action naming in PWA without semantic processing deficits, while

observation of meaningless gestures paired with action words did not. Bonifazi and colleagues proposed that as concepts are represented multimodally in the lexicon, activating sensory-motor features of target words through gesture observation strengthened semantic activation, whereas observation of meaningless gestures had no such facilitatory effect (Bonifazi et al., 2013). Taken together, these results suggest that observation of (iconic, meaningful) gestures alongside speech may be an effective way of delivering anomia therapy, even in the absence of PWA producing gestures themselves.

Roper et al. (2016) demonstrated that anomia therapy incorporating gestures (and speech) could be effectively delivered through a self-administered computer-based programme. 20 PWA with severe chronic aphasia completed an average of 14 hours of an online gesture production-based therapy programme, with approximately 5 hours of direct support provided from a speech and language therapist. Participants worked through a series of videos depicting gestures produced alongside target words and were required to repeat the gesture (but not necessarily the target word) in each case; PWA gesture production accuracy was automatically monitored through vision-based gesture recognition software and an ”applause” track was played as feedback when PWA produced the correct gestures. Although the target of therapy was not lexical retrieval, researchers found that naming accuracy improved at post-intervention testing. Researchers additionally found that in comparison to in-person interventions which typically experience a relatively high dropout rate (Brady et al., 2016), only one participant failed to complete the full programme and treatment fidelity was high. This suggests that self-paced speech and language therapy interventions incorporating gesture delivered via computer can be accessible to PWA even in cases of severe aphasia.

In this study, I adopted a similar self-paced computer-based therapy approach to Roper et al. (2016) and trialled a gesture observation-based intervention targeting word-retrieval difficulties.

5.2 Study aim and hypotheses

In this study I aimed to develop and trial a novel computer-based self-paced intervention for word retrieval difficulties based on gesture observation alongside speech production. I manipulated whether PWA observed, or did not observe, iconic gestures produced by the researcher alongside hearing target words spoken aloud, which PWA were then asked to repeat. PWA were then re-tested (at one week and one-month post-completion of the intervention) for naming accuracy for the target words. Following the findings outlined in Chapter 2 and the literature reviewed above, I hypothesised that iconic gestures produced by the researcher and observed by PWA would facilitate naming in (at least some) PWA, through a process of cross-modal priming resulting in semantic activation of target words prior to phoneme selection, making the words more highly activated in the lexicon and therefore easier for PWA to retrieve.

Following the findings of Marshall et al. (2013) that both speech-only and speech and gesture therapy improved retrieval of treated items, I predicted that all participants would show a treatment effect of significantly improved naming accuracy for trained words regardless of gesture presence/absence. However, as I hypothesised that gesture observation would enhance semantic activation and make words easier to retrieve, I also predicted that naming accuracy for words presented alongside gestures would be significantly higher than for words presented without gestures. Additionally, I predicted that for treated items in both conditions, gains in naming accuracy would be maintained at 1-week and 1-month post-intervention.

Prediction 1a: Participants will show significantly increased naming accuracy for treated items, regardless of treatment condition, at testing 1-week post-intervention.

Prediction 1b: Participants will maintain significantly increased naming accuracy for treated items, regardless of treatment condition, at testing 1-month post-intervention.

Prediction 2a: Although there will be improvement in naming accuracy for items regardless of treatment condition, participants will show significantly higher naming accuracy for items presented in the gesture condition, than items presented in the no-gesture condition, at testing 1-week post-intervention.

Prediction 2b: As in Prediction 2a, participants will show significantly higher naming accuracy for items presented in the gesture condition, than items presented in the no-gesture condition, at testing 1-month post-intervention.

Following findings from the previous study "Go Fish!" (see Chapter 2), I also predicted a significant interaction between working memory/executive function capacity and gesture effect, in that PWA with poorer working memory scores would show significantly greater naming accuracy for words presented with gestures, than PWA with better working memory scores. I also predicted that words paired with gestures with a higher gesture norm score (see Section 5.3.4 for more details of gesture norming), as better depicting the target word, would result in greater relearning and hence would be named significantly more accurately than words paired with gestures with a lower gesture norm score.

Prediction 3a: Participants with relatively lower working memory/executive function capacity will show significantly greater naming accuracy for items presented in the gesture condition, than participants with relatively higher working memory capacity.

Prediction 3b: For items presented in the gesture condition only, items with a relatively higher gesture norm score will be named significantly more accurately at 1-week and 1-month post-intervention than items with a relatively lower gesture norm score.

Finally, although Marshall et al. (2012) found that naming accuracy for untreated items did not improve significantly following speech and gesture treatment, I predicted that treatment gains may generalise to untreated items

matched for word length and word frequency to the training set in this case due to the high-dose, high-intensity format of the treatment programme (Brady et al., 2022).

Prediction 4: There will be generalisation of improved naming accuracy to items untreated in therapy, at 1-week or 1-month post-intervention.

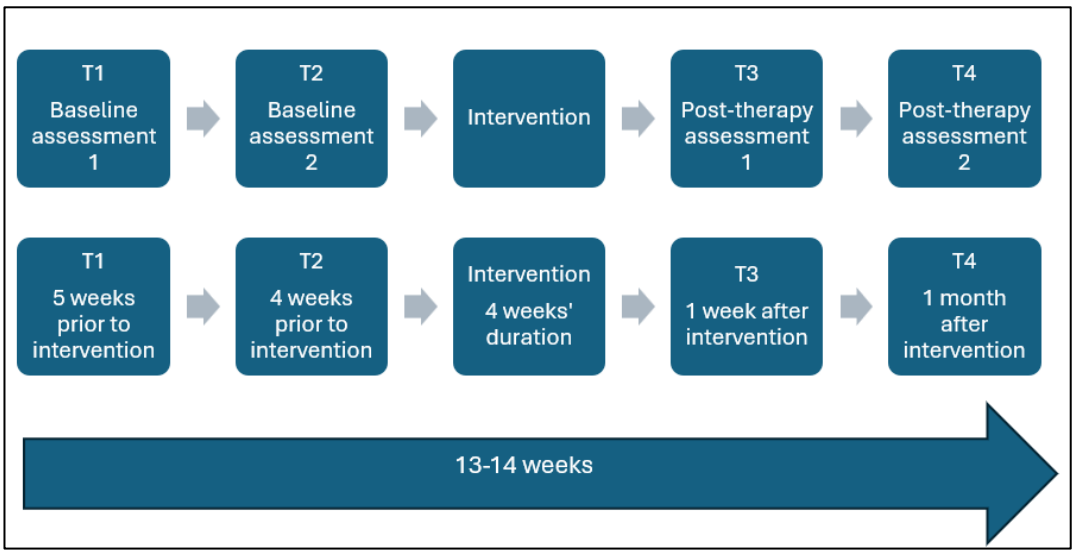
5.3 Methods

5.3.1 The present study

I created a novel, computer-based therapy programme for PWA that aimed to improve word-retrieval and naming accuracy for a discrete set of personalised treatment items. Therapy was self-paced with PWA able to access and complete sessions at a time that suited them without direct input from the researcher. A high-dose treatment design was used with PWA asked to complete (at least) 20 sessions of therapy over four weeks, equivalent to an average of five sessions per week; participants were free to complete up to seven sessions per week if they chose.

The therapy programme involved mass practice of (single word) picture naming, with half of target words presented to PWA in videos clips alongside a meaningful iconic gesture produced by the researcher, and half presented in video clips with the researcher using speech only. A within-subjects repeated measures design was used with all PWA receiving both with-gesture and without-gesture therapy input for separate lexical items; we tested naming accuracy twice one month before therapy input (T1, T2) and twice after therapy sessions were completed (T3, T4). The figure below shows indicative timescales for completion of the intervention. Ethical approval to carry out this study was received from the UCL High Risk Research Ethics Committee (Project 21765/001).

Fig. 1. *Indicative timescales for intervention.*



5.3.2 Participant sample

I recruited 4 PWA with aphasia of at least 12 months' duration caused by left hemisphere stroke (mean time since stroke = 98.75 months; SD = 45 months). PWA were recruited through local stroke support groups in London and via the UCL Research Registry. The inclusion criteria were: (i) chronic post-stroke aphasia (at least 12 months' post-stroke); (ii) aged 18+ years; (iii) English as a first language; (iv) normal or corrected-to-normal vision and hearing; (iv) moderate aphasia as assessed through the WAB-R assessment with a composite Aphasia Quotient score of between 50.1 and 74.9; (v) not currently receiving speech and language therapy intervention from NHS/other services. Exclusion criteria were: (i) diagnosis of developmental language disorders; (ii) major co-existing neurological or psychiatric disorders; (iii) unable to give informed consent. All participants had previously taken part in the studies outlined in Chapters 2 and 3.

All participants presented with predominant deficits in naming and spontaneous speech production and relatively preserved auditory verbal comprehension; three PWA presented with symptoms consistent with Conduction aphasia and one with symptoms consistent with Broca's aphasia. All participants had completed at least 13 years of education (mean = 15.25

years; SD = 2.00). Two were assessed as having mild apraxia of speech based on clinical assessment. All participants had previously received at least one block of speech and language therapy input through the UK NHS. All participants completed a comprehensive background testing battery covering receptive and expressive aphasia deficits (*Western Aphasia Battery-Revised* (Kertesz, 2006)); working memory/executive function (*WAIS-III* Backwards Digit Span test (Wechsler, 1997); confrontation (picture) naming (*Middleton Lab Naming Battery* (Middleton et al., 2016); gesture recognition (*MRRI Apraxia Battery* (Tarhan et al., 2015). Basic participant demographic details may be found in Table 1 below. Full details of background battery assessment scores may be found in Appendix 1. I obtained written informed consent from all participants to take part in this study in line with UCL guidelines.

Table 1. *Participant demographics.*

ID	Aphasia Type	WAB-R AQ	WAB-R Naming	WAB-R Spont. Speech	WAB-R Aud. Comp.	WM/EF	Sex	Educat. (years)	Months post-stroke
P8	Conduction	57.00	3.80	14.00	7.50	1.00	M	16.00	130.00
P14	Conduction	69.00	5.50	16.00	9.20	2.00	M	16.00	48.00
P20	Conduction	59.00	6.10	10.00	8.20	2.00	F	16.00	75.00
P25	Broca's	51.50	6.30	9.00	6.65	5.00	M	13.00	142.00

5.3.3 Materials

Therapy materials consisted of (up to) 120 short video clips of the researcher producing a target object word embedded within a carrier phrase (e.g. "Here is a…."). Clips were recorded in a well-lit sound-proofed room at UCL. All videos depicted the researcher's head, neck and upper body against a plain background. In half of clips the researcher (a native British English speaker) produced an iconic gesture alongside the target word; in the other half, target words were produced without gestures. No target word appeared in both gesture conditions. Video clips were loaded onto Gorilla Experiment Builder (https://gorilla.sc/) for viewing by participants.

5.3.4 Gesture norming

Prior to the start of therapy, to determine how easy to understand the researcher's iconic gestures would be for naive observers, gesture norming ratings were completed for all gestures used with 30 English-speaking neurologically healthy adults (NHPs), aged 18+ years, recruited via Prolific. NHPs were presented with short video clips of the researcher producing each target word with an accompanying iconic gesture embedded within a carrier phrase, e.g. "Now we have a *frisbee*", alongside a "throwing frisbee" gesture. NHPs were then shown the target word in written form and asked to rate (using a Likert scale 1-7) how well the gesture produced by the researcher represented the given target word. Average (mean) gesture norms ratings were then calculated for all iconic gestures produced by the researcher and entered as a predictor in the models for data analysis. Gesture norm scores may be found in Appendix 7.

5.3.5 Online therapy design

A personalised therapy platform was built for each participant using Gorilla Experiment Builder (https://app.gorilla.sc/) with each participant receiving a personalised reusable login link to access their own therapy stimuli. To ensure participant confidentiality while using Gorilla Experiment Builder, names were not used within the therapy platform and each participant was identified only by a number. Participants were only given a link to their own therapy platform and were unable to access the platforms built for any other participants.

Prior to the start of therapy, each participant was provided with written instructions for logging into and navigating through their own therapy platform (see Appendix 8 for an example instruction sheet). Additionally, the researcher held individual sessions with each participant (and communication partner if needed) demonstrating how to access the platform and providing troubleshooting support. Participants were also able to contact

the researcher for additional technical assistance if needed throughout the course of the intervention.

5.3.6 Procedure

5.3.6.1 Pre-therapy naming assessments

Prior to starting therapy sessions, all participants completed two baseline confrontation (picture) naming assessments at least seven days apart, at least four weeks prior to the start of the intervention. The same assessment was administered on both occasions. The assessment chosen was a sub-part of the *Middleton Lab Naming Battery* (Middleton et al., 2016) and consisted of 384 words comprising at least five phonemes. All words chosen exhibited relatively high naming agreement between speakers (mean naming agreement: 0.92). Naming assessments were carried out online using Zoom and assessment sessions were recorded with PWA consent. Participants were presented with pictures of objects and asked to produce the name; the first clear attempt at naming by the participant was taken as the response. Fillers produced by participants immediately after being shown each picture (e.g. *um, ah, uh*) were not classed as naming responses. Participants were permitted to answer either with a single word (e.g. “computer”) or a sentence (e.g. “That looks like it’s a computer to me”). No time limit was imposed and participants were permitted breaks as needed during assessment sessions.

PWA responses were transcribed during the session by the researcher into International Phonetic Alphabet (IPA) form for scoring. Responses were initially scored as 1 (correct) or 0 (incorrect). Participant responses were scored as correct if they produced the target word correctly, produced a British English synonym of the target word (e.g. “spectacles” instead of “glasses”, “cockerel” instead of “rooster”), produced an informal/slang name for the target word (e.g. “brolly” instead of “umbrella”, “fridge” instead of “refrigerator”), or produced a non-British English dialectal equivalent name (e.g. “eggplant” instead of “aubergine”). Responses with incorrect 0 scores were also allocated to an error category: semantic (meaning error),

phonemic (sound error with at least 50% of phonemes in common with the target word), mixed (semantic and phonemic error), omission (no name produced or explicit indication that the participant was unable to produce the name, e.g. “I can’t say that one”, “I don’t know it”, etc.) or unrelated (no clear semantic or phonemic relation to the target word). See Table 2 below for examples of scoring and error categorisation.

Table 2. *Error categorisation.*

Target word	Phonemic error	Semantic error	Mixed error	Unrelated error
lemon	[ˈlɛlən]	orange	melon	jigsaw
tiger	[ˈtaɪdə]	lion		
trumpet		flute	trombone	
volcano	[vɑlˈteɪnəʊ]	mountain		pictures

5.3.6.2 Stimuli selection

All intervention materials were individually selected for participants based on their prior performance in the two baseline naming assessments described in the previous section. Stimuli comprised 112-120 target words designated the treatment set. Of the words selected, 10-15 were “known” control words, which PWA had produced correctly in both baseline naming assessments. These words were included to reduce PWA frustration and to improve motivation by endeavouring to ensure that in each therapy session PWA would be able to correctly produce at least a proportion of the words presented. Naming accuracy for “known” control words was not included in analyses.

The remainder of items were “unknown” words which were targeted during therapy sessions. Given that PWA naming accuracy is known to be variable across time and can be significantly affected by factors such as fatigue (Mirman et al., 2024), only words produced incorrectly on *both* baseline naming assessments were selected for therapy, to ensure that the words targeted represented words that PWA had consistent difficulty producing.

Additionally, a further set of 112-120 words that PWA had produced incorrectly in both assessments was selected and designated the untreated set, which were not targeted in therapy sessions. These words were selected to test for generalisation of naming improvements post-therapy. Words in the untreated set were broadly matched with words in the treated set for length (no. phonemes), frequency, semantic content and (where possible) error type. Frequency measures for both treated and matched untreated words were drawn from the Subtlex-UK database (van Heuven et al., 2014). See Table 3 below for examples of target words and matched untreated words.

Table 3. *Examples of target words and matched untreated words.*

Target word				Matched untreated word			
	Length	Frequency	Semantic category		Length	Frequency	Semantic category
Blueberry ˈblu ˌbɛri	9	3.17	Inanimate Food	Raspberry ˈræz ˌbɛri	9	3.81	Inanimate Food
Dolphin ˈdɑlfən	7	3.77	Animate Animal	Walrus ˈwɔlrəs	7	3.17	Animate Animal
Sailor ˈseɪlər	7	3.75	Animate Human	Pilot ˈpaɪlət	7	4.37	Animate Human

5.3.7 Therapy

The intervention consisted of a single word picture-naming task and included 112-120 words per PWA, selected as described above. Target words were randomised into with-gesture and without-gesture conditions.

PWA were asked to complete (at least) 20 sessions of therapy over four weeks, completing at least one live online session per week with the researcher and four self-paced sessions per week using their personalised therapy platform on Gorilla Experiment Builder. Participants were free to complete additional online sessions if they chose, up to a maximum of seven sessions per week (i.e. one per day). Each session lasted approximately one hour with short breaks built in every 10-12 minutes to reduce participant fatigue; during self-paced sessions PWA were able to skip or lengthen breaks as needed and during live sessions they were also permitted to request additional breaks at any time. The researcher tracked completion of self-paced sessions and automated reminder email messages were sent to

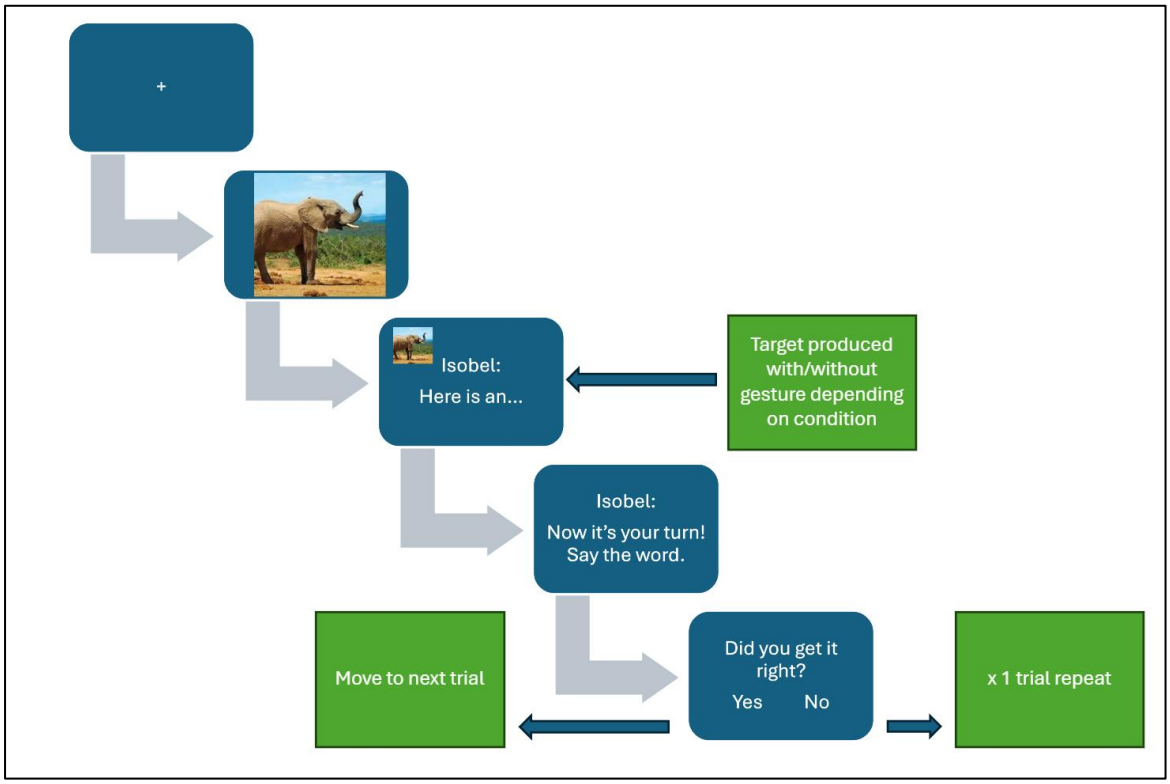
PWA to encourage session completion compliance. Live sessions were recorded by the researcher with PWA consent. Communication partners were free to attend therapy sessions and provide technical support if requested by the PWA, but were asked not to provide assistance if the PWA was unable to produce target words during sessions.

All sessions, whether live or self-paced, followed the same format and used an errorless learning approach. In each trial PWA were initially presented with a fixation cross for 250 milliseconds followed by a picture of an object, which they were instructed not to name. In self-paced sessions, PWA were then presented with a short video clip of the researcher producing the target word with or without an associated iconic gesture; in live sessions, the researcher produced the target word live, again with or without the same iconic gesture presented in the self-paced sessions. In both cases, a smaller picture of the object was present alongside the video of the researcher producing the name. At the end of each video clip the PWA was then asked to repeat the word. In both the live and self-paced sessions, PWA were permitted a single repetition of the target word (and gesture, if applicable) from the researcher if required. Gesture conditions were interleaved and words were presented in a random order in each iteration of the therapy sessions. Iconic gestures paired with target words were kept consistent across sessions and no two target words were produced with the same accompanying iconic gesture to avoid interference.

Target words were produced by the researcher with an accompanying carrier phrase, e.g. “Now we have a [target word]”, to maintain natural prosody. In the self-paced sessions, PWA were then asked to rate whether they thought they had repeated the word correctly; in live sessions, the researcher provided direct feedback. Personalised corrections and production strategies (e.g. the researcher breaking the word into smaller components) were not provided by the researcher during live sessions to ensure consistency of feedback with the self-paced sessions. In both the videos and the live sessions the researcher wore clothes of a contrasting colour to her hands to

ensure they showed up well, wore no make-up, and was seated on a chair against a neutral-coloured background.

Fig. 2. *Intervention design.*



5.3.8 Post-therapy naming assessments

PWA completed two post-therapy picture naming assessments one week and one month after completing the assigned therapy sessions. Assessments comprised both the treated and untreated sets and 10-15 control "known" words which were not scored. To reduce anxiety, participants were specifically informed prior to starting the assessments that pictures would be included that had *not* been targeted in therapy sessions. As in the pre-therapy assessments, participants were presented with a picture of the object and asked to produce the name; pictures were presented in a random order in each assessment. The same scoring protocol for naming was followed as in the pre-therapy picture naming assessments (see previous section for details). No time limit was imposed and participants were permitted as many breaks as required to complete the assessments.

5.4 Data analysis

The same statistical modelling strategy was adopted in this chapter as described in Section 2.5.3.

5.4.1 Power analysis

A priori power analysis to determine the number of items to be presented in each gesture condition was conducted using the *pwr* package (Champely et al., 2022) in RStudio (RStudio Team, 2020). Results indicated that the required sample size to achieve 80% power for detecting a medium effect (0.40) of gesture presence/absence on PWA naming accuracy, at a significance criterion of $\alpha = 0.05$, was 51 items per treatment condition. This resulted in a target set of (at least) 102 items selected per PWA, with (at least) 51 target words to be presented in each gesture condition.

5.4.2 Primary analyses

One PWA (P14) temporarily withdrew from the intervention for medical reasons and was unable to complete the full programme of therapy (completed 17 of 20 planned sessions), but did complete post-therapy naming assessments. Three PWA completed all sessions in line with the outlined protocol and completed post-therapy naming assessments.

Analyses were carried out using the *lme4* package (Bates et al., 2015) in RStudio (RStudio Team, 2020). In all analyses below where working memory/executive function score is included, the score was log transformed to reduce skew, stabilise the models and meet model assumptions for logistic/linear mixed effects regression models. Categorical variables where orthogonally coded to test for effects and continuous variables were scaled and centred to the mean. For all generalised linear mixed effects models the 'logit' link function was used with the *bobyqa* optimizer for model convergence. Results for all PWA are described in full below. Full R code for all models below may be found in Appendix 9.

5.4.2.1 Predictions 1a and 1b: Intervention effect on naming accuracy post-intervention

In order to assess the overall effect of treatment (i.e. regardless of gesture condition) on naming accuracy for treated items as outlined in Predictions 1a and 1b, I used paired *t*-tests with comparison of overall naming accuracy at 1-week and 1-month post-treatment against baseline naming accuracy prior to the intervention. To determine if there was any significant change in naming accuracy between post-intervention testing time points, I then used paired *t*-tests to compare naming accuracy at 1-week and 1-month post-intervention.

5.4.2.2 Predictions 2a and 2b: Gesture effect on naming accuracy post-therapy

In order to assess the effect of gesture (with/without gesture treatment condition) on naming accuracy and to explore whether any observed effects were maintained, I defined global generalised linear mixed effects logistic regression models with (1) naming accuracy at 1-week post-intervention and (2) naming accuracy at 1-month post-intervention as the dependent variables. Gesture condition was entered as the primary independent variable of interest and we also included aphasia severity (WAB-R Aphasia Quotient), working memory/executive function and target word length as control variables, including interactions between these. Finally, I included a random intercept for the grouping variable of participant to account for individual variability in responses. This analysis included treated items only; untreated matched items were excluded (n = 406).

5.4.2.3 Predictions 3a and 3b: Interactions with working memory/executive function and gesture norm scores

In order to investigate whether working memory/executive function (as measured by backwards digit span) predicted gesture benefit, I defined global generalised linear mixed effects logistic regression models with (1)

naming accuracy at 1-week post-intervention and (2) naming accuracy at 1-month post-intervention as the dependent variables. Gesture condition was entered as the primary independent variable of interest and I also included working memory/executive function and target word length as control variables, including interactions between these variables. Finally, I included a random intercept for the grouping variable of participant to account for individual variability in responses.

5.4.2.4 Predictions 4: Generalisation of treatment gains to untreated items

The same paired *t*-test approach as outlined in 5.4.2.1 was used for this analysis. Naming accuracy scores for untreated items (n = 375) at baseline were compared to naming accuracy scores at 1-week and 1-month post-intervention.

5.4.3 Secondary analyses

Further to the primary analyses described above, I also carried out a number of secondary exploratory analyses to identify any other effects of gesture observation during the intervention.

Firstly, on the basis that treatment may increase activation of target words in the lexicon semantically and/or phonemically sufficient to produce an attempt at naming (even if incorrect), I explored whether treated items presented in the gesture present condition for which PWA had previously produced an omission error during baseline testing were subsequently produced correctly or with a semantic, phonemic or mixed error. I tested this at both 1-week and 1-month post-intervention. I included production of a correct target or production of a non-omission error as the dependent variable at each time point, with gesture condition as the primary independent variable of interest. I also included working memory/executive function, target word length and target word frequency as control variables. Finally, I included a random intercept for participant.

Secondly, as I hypothesised that observing iconic gestures may facilitate semantic activation of target items, possibly in the production of either a correct response or a semantic error, I explored whether treated items presented in the gesture condition had a higher cosine similarity (i.e. semantic similarity) score than items produced in the no-gesture condition. I defined a linear mixed effects regression model with cosine similarity at each time point as the dependent variable and gesture condition as the independent variable. I also included aphasia severity (WAB-R Aphasia Quotient), working memory/executive function and target word length as control variables. Finally, I included a random intercept for the grouping variable of participant to account for individual variability in responses.

Finally, as the Lexical Retrieval Hypothesis proposes that gestures facilitate naming at either the stage of phoneme selection or lexical selection, resulting in the production of either a correct response or a phonemic error, I explored whether treated items produced in the gesture condition had more phonemes in common with target words (Levenshtein distance), than items produced in the no-gesture condition. I used the same model as outlined above, with the dependent variable as phonological Levenshtein distance.

5.5 Results

5.5.1 Predictions 1a and 1b: Intervention effect on naming accuracy post-therapy and maintenance of treatment effects

In line with Prediction 1a, I found that all participants showed a significant effect of treatment on naming accuracy for treated items compared to baseline naming accuracy, regardless of treatment condition, at 1-week post-intervention. Further, these gains were maintained at 1-month post-intervention compared to baseline, in line with Prediction 1b. Paired *t*-tests were used to determine change in naming accuracy from baseline at each timepoint for each participant separately.

Notably, although all participants maintained naming accuracy gains at 1-month post-treatment compared to baseline, there were small changes in

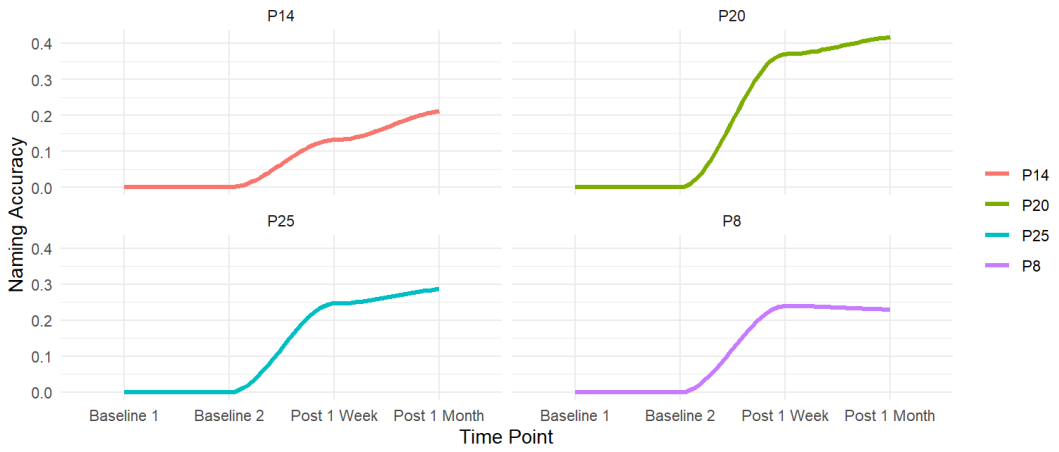
naming accuracy between post-intervention time points. For P14, despite completing fewer intervention sessions overall than other participants, a significant improvement in naming accuracy between time points was observed (p = 0.01). For all other participants, positive or negative change in naming accuracy for treated items between post-intervention time points was non-significant (p > 0.05).

Table 4. *Paired* t*-test results for naming accuracy of treated items at 1-week and 1-month post-intervention.*

Participant	Paired t-test *p* value 1-week post-intervention	Paired t-test *p* value 1-month post-intervention	Paired t-test *p* value between time points
P14	<0.001*	<0.001*	0.01*
P20	<0.001*	<0.001*	> 0.05
P25	<0.001*	<0.001*	> 0.05
P8	<0.001*	<0.001*	> 0.05

Fig. 3. *Change in naming accuracy for treated items by participant at each testing point.*

In this figure each participant's naming accuracy scores are presented separately – see legend. The Y axis presents naming accuracy (0.0-1.0 scale) with time represented on the X axis.



5.5 2 Predictions 2a and 2b: Gesture effect on naming accuracy post-therapy

Contrary to Prediction 2a, I did not find a significant effect of treatment condition on naming accuracy at 1-week post-intervention. Overall, although

naming accuracy for treated items increased for all PWA, items presented in the gesture condition were not named significantly more accurately following treatment then items presented in the no-gesture condition (β=−0.084, SE = 0.236, z=−0.355, *p* = 0.723) Similarly, at 1-month post-intervention I found no effect of treatment condition on naming accuracy (Prediction 2b) (β=−0.137, SE = 0.215, z=−0.640, *p* = 0.522).

However, at both time points post-intervention I found that aphasia severity (WAB-R Aphasia Quotient) significantly predicted treatment response regardless of gesture condition, with PWA with more severe aphasia (i.e. lower WAB-R Aphasia Quotient) benefiting significantly more from therapy than PWA with less severe aphasia (i.e. higher WAB-R Aphasia Quotient). There was also a significant effect of target word length at 1-week post-intervention, with shorter words named more accurately by PWA regardless of the gesture condition in which they were presented (*p* = 0.005). This effect was not observed at 1-month post-intervention.

Table 5. *Naming accuracy at 1-week post-intervention.*

Effect	Estimate	Std.Error	z value	Pr(>|z|)
(Intercept)	-0.663	0.332	-1.997	0.046
Gesture condition (present)	-0.084	0.236	-0.355	0.723
Aphasia Quotient	-0.764	0.381	-2.006	0.045
Working memory	-0.261	0.350	-0.744	0.457
Target word length	-0.553	0.197	-2.808	0.005
Gesture condition (present) * Aphasia Quotient	0.428	0.296	1.446	0.148
Gesture condition (present) * Working memory	0.256	0.243	1.050	0.294
Gesture condition (present) * Target word length	0.363	0.251	1.446	0.148
Aphasia Quotient * Target word length	-0.129	0.258	-0.500	0.617
Working memory * Target word length	-0.270	0.183	-1.473	0.141
Gesture condition (present) * Aphasia Quotient * Target word length	0.424	0.313	1.353	0.176
Gesture condition (present) * Working memory * Target word length	0.489	0.262	1.870	0.061

Table 6. *Naming accuracy at 1-month post-intervention.*

Effect	Estimate	Std.Error	z value	Pr(>|z|)
(Intercept)	-0.393	0.248	-1.584	0.113
Gesture condition (present)	-0.137	0.215	-0.640	0.522
Aphasia Quotient	-0.195	0.273	-0.716	0.474

Working memory	0.206	0.268	0.767	0.443
Target word length	-0.147	0.158	-0.930	0.352
Gesture condition (present) * Aphasia Quotient	0.178	0.241	0.736	0.462
Gesture condition (present) * Working memory	-0.007	0.237	-0.029	0.977
Gesture condition (present) * Target word length	-0.400	0.227	-1.762	0.078
Aphasia Quotient * Target word length	-0.135	0.183	-0.738	0.461
Working memory * Target word length	-0.039	0.164	-0.240	0.811
Gesture condition (present) * Aphasia Quotient * Target word length	0.080	0.261	0.307	0.759
Gesture condition (present) * Working memory * Target word length	0.116	0.250	0.464	0.643

Fig. 4. *Effect plot: Mean naming accuracy for treated items by gesture condition at each testing point.*

In this figure the Y axis represents naming accuracy (0.0-1.0 scale). Bars indicate the range of scores with dots indicating mean naming accuracy. Scores are presented in each gesture condition (i.e. without gestures, with gestures). Red bars indicate scores 1 week post-therapy; blue bars indicate scores 1 month post-therapy.

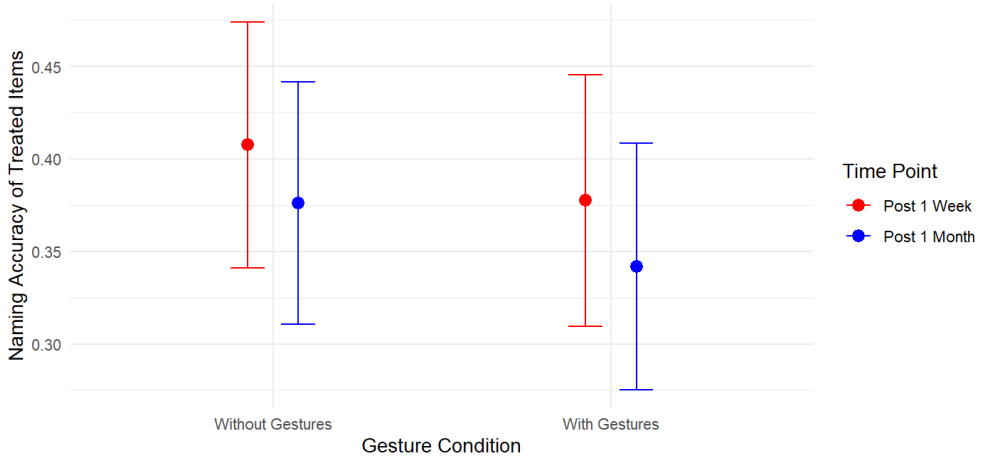
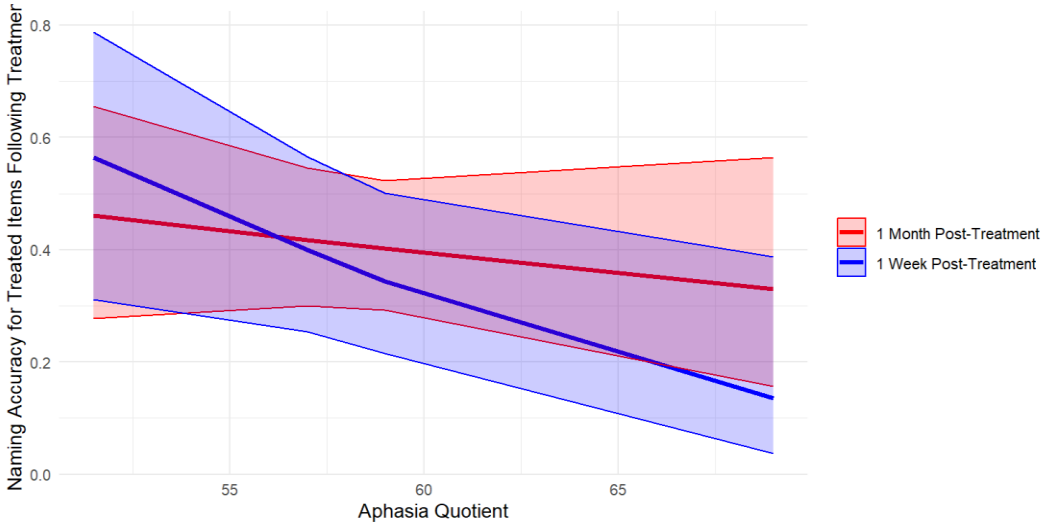


Fig. 5. *Naming accuracy: effect of Aphasia Quotient on naming accuracy for treated items at each testing point.*

In this figure the Y axis represents naming accuracy (0-1.0 scale) and the X axis represents Aphasia Quotient scores. The blue line indicates change in naming accuracy 1 week post-treatment; the red line indicates change in naming accuracy 1 month post-treatment. Shaded areas indicate confidence intervals.



5.5.3 Predictions 3a and 3b: Effects of working memory and gesture norm scores

Contrary to Prediction 3a, I did not find that participants with relatively poorer working memory/executive function capacity showed significantly greater naming accuracy for items presented in the gesture condition, than participants with relatively better memory capacity at either 1-week ($p > 0.05$) or 1-month post-intervention ($p > 0.05$). However, in line with Prediction 3b I found that higher gesture norm score for treated items presented in the gesture condition predicted significantly improved naming accuracy at 1-week post-intervention ($p = 0.026$). At 1-month post-intervention a similar trend was observed which bordered on significance ($p = 0.070$).

Table 7. *Effects of working memory/executive function and gesture norm score on naming accuracy at 1-week post-intervention.*

	Estimate	Std. Error	z value	Pr(>|z|)

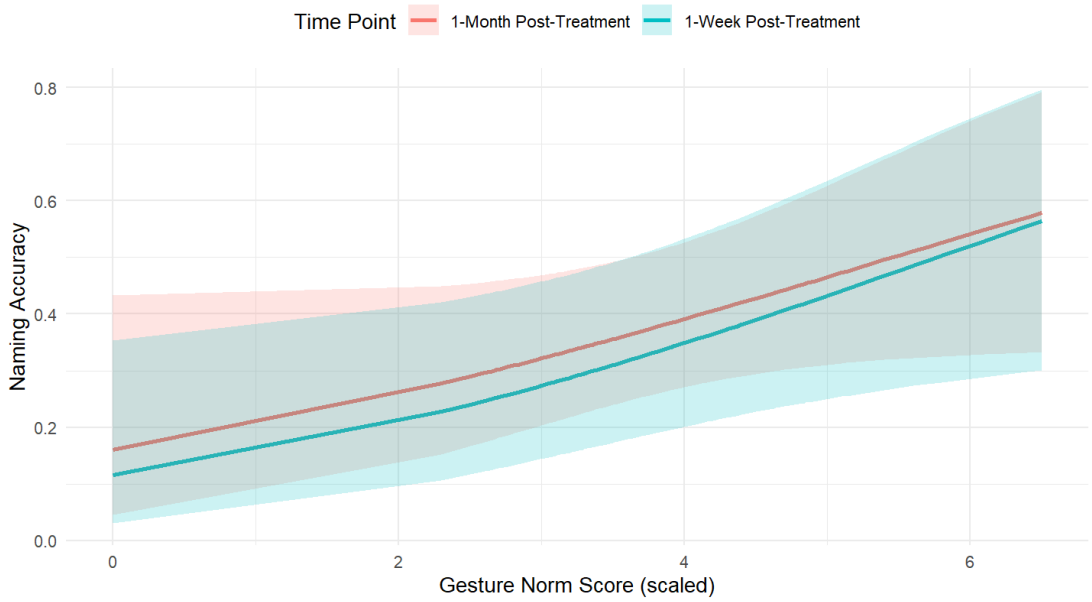
(Intercept)	-0.700	0.385	-1.817	0.069
Gesture norm score	0.451	0.203	2.227	0.026
Working memory	0.014	0.386	0.035	0.972
Target word length	-0.062	0.174	-0.359	0.720
Gesture norm score * Working memory	-0.143	0.209	-0.684	0.494

Table 8. *Effects of working memory/executive function and gesture norm score on naming accuracy at 1-month post-intervention.*

	Estimate	Std.Error	z value	Pr(>|z|)
(Intercept)	-0.507	0.280	-1.813	0.070
Gesture norm score	0.388	0.214	1.813	0.070
Working memory	0.202	0.282	0.716	0.474
Target word length	-0.483	0.184	-2.619	0.009
Gesture norm score * Working memory	-0.442	0.227	-1.953	0.051

Fig. 6. *Naming accuracy: effect of gesture norm score on naming accuracy for treated items at each testing point.*

In this figure the Y axis represents naming accuracy (0.0-1.0 scale) and the X axis represents scaled gesture norm scores. The blue line indicates change in naming accuracy 1 week post-treatment; the red line indicates change in naming accuracy 1 month post-treatment. Shaded areas indicate confidence intervals.



5.5.4 Prediction 4: Generalisation of treatment effects to untreated items

In line with Prediction 4, I found evidence of improvement for untreated items for all PWA at both post-intervention time points. Paired t-tests indicated that there was significant change in naming accuracy of untreated items from baseline naming accuracy at 1-week or 1-month post-intervention ($p > 0.05$ at both time points). For one PWA (P20) there was also significant change in naming accuracy for untreated items between time points. For all other PWA change between time points was non-significant.

Table 9. *Paired* t-*test results for naming accuracy of untreated items at 1-week and 1-month post-intervention.*

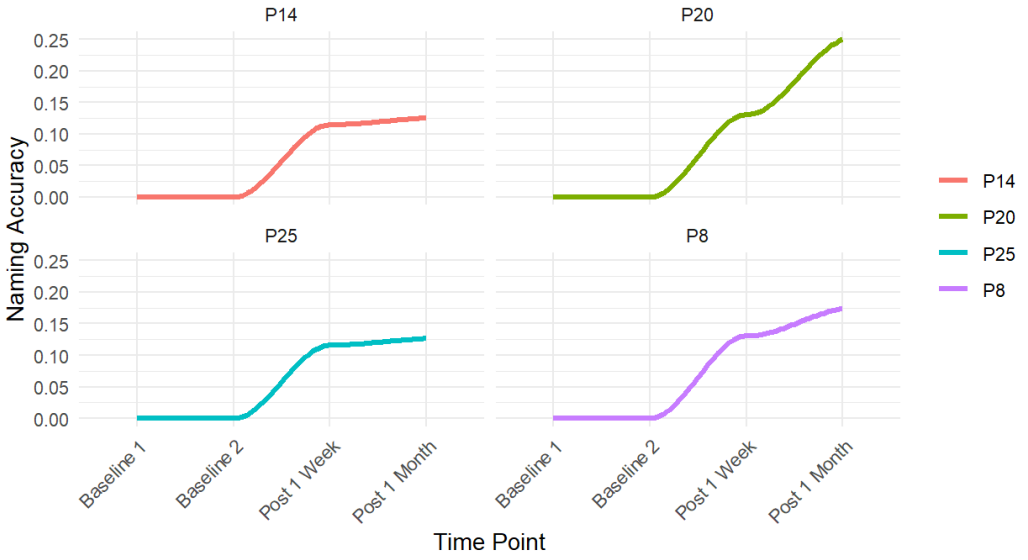
Participant	Paired t-test *p* value 1-week post-intervention	Paired t-test *p* value 1-month post-intervention	Paired t-test *p* value between time points
P14	0.0007*	0.0004*	> 0.05
P20	0.0004*	0.0000*	0.003*
P25	0.0007*	0.0004*	> 0.05
P8	0.0004*	0.0000*	> 0.05

Table 10. *Comparison of number of untreated items named correctly at 1-week and 1-month post-intervention.*

Participant	No. untreated items correct at 1-week post-therapy	No. untreated items correct at 1-month post-therapy
P14	11	12
P20	12	23
P25	11	12
P8	12	16

Fig. 7. *Change in naming accuracy for untreated items by participant at each testing point.*

In this figure each participant's naming accuracy scores are represented separately – see legend. The Y axis presents naming accuracy (0.0-1.0 scale) with time represented on the X axis.



5.5.5 Secondary analyses

I found no significant effect of gesture condition on cosine similarity of PWA responses, or on phonological distance from the target word, at either time point ($p > 0.05$).

Table 11. *Cosine similarity of PWA responses at 1-week post-intervention.*

	Estimate	Std. Error	df	t value	Pr(>|t|)
(Intercept)	0.049	0.283	1.050	0.172	0.890
Gesture condition (present)	-0.095	0.092	401.003	-1.035	0.301
Aphasia Quotient	-0.235	0.303	0.999	-0.775	0.581
Working memory	-0.061	0.304	0.999	-0.200	0.874
Target word length	-0.136	0.046	401.011	-2.943	0.003

Table 12. *Cosine similarity of PWA responses at 1-month post-intervention.*

	Estimate	Std. Error	df	t value	Pr(>|t|)
(Intercept)	0.061	0.214	1.097	0.283	0.821
Gesture condition (present)	-0.121	0.095	401.005	-1.279	0.202
Aphasia Quotient	-0.081	0.226	0.998	-0.356	0.782
Working memory	0.111	0.227	0.998	0.490	0.710

Target word length	-0.154	0.048	401.020	-3.242	0.001

Table 13. *Phonemic similarity of PWA responses at 1-week post-intervention.*

	Estimate	Std. Error	df	t value	Pr(>|t|)
(Intercept)	0.056	0.255	1.063	0.218	0.862
Gesture condition (present)	-0.110	0.093	401.003	-1.186	0.236
Aphasia Quotient	-0.270	0.272	0.999	-0.991	0.503
Working memory	-0.097	0.273	0.999	-0.357	0.782
Target word length	-0.083	0.046	401.013	-1.780	0.076

Table 14. *Phonemic similarity of PWA responses at 1-month post-intervention.*

	Estimate	Std. Error	df	t value	Pr(>|t|)
(Intercept)	0.053	0.217	1.096	0.245	0.844
Gesture condition (present)	-0.106	0.096	401.005	-1.107	0.269
Aphasia Quotient	-0.100	0.230	0.998	-0.436	0.738
Working memory	0.071	0.230	0.998	0.307	0.811
Target word length	-0.102	0.048	401.020	-2.126	0.034

5.6 Discussion

In this study I recruited 4 PWA with moderate aphasia to complete an online computer-based therapy programme targeting word retrieval. Three PWA completed (at least) 20 sessions of therapy, with at least 4 sessions delivered live by the researcher and the remainder completed independently by PWA. The final participant was unable to complete all therapy sessions for medical reasons but nevertheless completed 17 out of a possible 20 sessions. All participants completed two post-intervention confrontation naming assessments following the intervention at 1-week and 1-month post-treatment. Post-intervention naming assessments included both treated and untreated items.

Firstly, I found that PWA naming accuracy for treated items improved significantly, regardless of whether they were presented with or without gestures, at both 1-week and 1-month post-intervention compared to baseline naming accuracy. This suggests that this form of self-paced intervention is a viable intervention method and allows for sufficiently intensive delivery of therapy to achieve significant change. Additionally, I

found a significant change in naming accuracy for matched untreated items. This result should be treated with caution, as although the change was statistically significant the number of untreated items produced correctly at 1-week and 1-month post-intervention was relatively small (mean untreated items named correctly at 1-week post-treatment = 11.5; mean untreated items named correctly at 1-month post-treatment = 16.5) and may partially reflect instability in naming performance. However, it remains promising that evidence for generalisation was found.

Second, contrary to predictions I did not find a significant effect of gesture condition on naming accuracy for treated items at 1-week or 1-month post-intervention. However, there was a significant interaction between gesture norming score and gesture condition on naming accuracy, indicating that items presented with gestures that had been rated as more informative/descriptive were named more accurately by PWA, than items presented with gestures that had been rated as less highly informative. This suggests that while gesture observation may not facilitate naming accuracy overall, iconic gestures which are most closely conceptually aligned to the target word may nevertheless be more facilitatory than iconic gestures which are not.

5.6.1 Theoretical implications

This study supports the proposal that (at least some) iconic gestures can assist PWA to re-learn words. In particular, these results are supportive of the proposal that sub-types of iconic gestures have varying effects on (re-)learning, at least in certain individuals, with some types of iconic gestures having stronger facilitation effects than others. For example, in a study of verb learning in 120 3-year old children, Mumford & Kita (2014) found that children who were taught novel verbs paired with iconic ‘manner’ gestures (gestures depicting how an action is performed) learned them more effectively than verbs paired with iconic ‘end state’ gestures (gestures depicting a change of state). Mumford & Kita (2014) suggest that a possible explanation for this this may be that ‘end state’ gestures are less obviously

imagistically connected to lexical affiliates and therefore require more effort to process. By contrast, ‘manner’ gestures are directly related to their lexical affiliates and are therefore easier to link. Similar learning effects have also been found in other groups including children with language delay (Lüke et al., 2020) and children with Developmental Language Disorder (Vogt & Kauschke, 2017). In support of Mumford & Kita (2014), Lüke & Ritterfield (2014) also suggest that these beneficial learning effects derive from the degree of ‘transparency’ between the iconic gesture and its associated word; the greater the transparency, the lower the cognitive effort required for learning.

In this study, those gestures which scored highest during the gesture norming task and which best facilitated naming accuracy at 1-week post-intervention testing were those that most closely imagistically or directly aligned to the target words, and therefore most easily conceptually mapped to the target words. These results therefore lend support to the idea that not only iconicity of gestures, but *degree* of iconicity, is predictive of a positive re-learning effect in PWA. If so, and following Mumford & Kita (2014), this would also suggest that iconic gestures facilitate re-learning through enhanced semantic activation effects at the level of conceptualisation (in line with the Information Packaging Hypothesis), rather than at the level of phonemic or lexical selection (in line with the Lexical Retrieval Hypothesis).

5.6.2 Clinical implications

Previous studies have consistently shown that aphasia therapy requires a significant dosage to effect change. A meta-analysis of 25 clinical trials of aphasia therapy including 959 individual participants showed that the greatest clinical gains were associated with a total dosage of 20-50 hours of therapy, delivered across a minimum of three days per week, including personalised home-based practice (Brady et al., 2022). Although anomia was not considered in isolation in this meta-analysis, similar results were found by Dignam and colleagues who looked at anomia therapy specifically and found that the greatest gains in single word retrieval were associated with 6 hours

of therapy per week, for up to 8 weeks. In both cases, it was noted that direct delivery of therapy from a speech and language therapist for every session was unrealistic and that computer-based approaches were an appropriate means to deliver intensive therapy (Brady et al., 2022; Dignam et al., 2015). This study has replicated these findings and shows that efficacious anomia therapy can be effectively delivered through an online, self-paced therapy platform. Furthermore, these results indicate that PWA with more severe chronic aphasia (i.e. lower WAB-R Aphasia Quotient) can also show improvements in naming accuracy when therapy is delivered in this way. This is particularly encouraging, as more severe aphasia has been found to correlate with poorer quality of life measures and increased rates of depression (Bullier et al., 2020; Sheppard et al., 2021). These findings suggest that even in cases of relatively severe naming deficits, personalised computer-based therapy can be delivered effectively.

Additionally, these findings suggest that speech and (iconic) gesture observation therapy can facilitate word retrieval in PWA with moderate chronic aphasia with maintenance of improvements for at least one month after the intervention, at least when those gestures depict target words well and are presumably most salient to PWA. Further research is required to determine the characteristics of iconic gestures that render them most likely to facilitate word retrieval in PWA, and indeed the same iconic gestures may not necessarily have the same effects in all PWA and may need to be individually selected alongside target words. It is possible that co-selection of therapy materials with PWA (and communication partners) could result in more salient resources, and AI video/picture generation of stimuli would make this feasible. Indeed, in a recent study of 447 neurologically healthy adults comparing the learning effects of AI-generated videos and human-generated videos, participants exposed to the AI-generated videos achieved equally high learning outcomes with significantly less time required to produce the materials (12-24 hours per video for human-generated; 3-4 hours per video for AI-generated) (Netland et al., 2025). While human- and AI-generated videos are by their nature not interactive, multi-user online aphasia therapy environments such as EVA Park are currently in

development, where PWA can interact with each other informally and therapists can use increasingly naturalistic avatars to deliver programmed therapy (Marshall et al., 2016). EVA Park incorporates multiple virtual 'locations', such as cafes, shops and bars, where PWA can practise communication skills to, e.g. order food or exchange greetings. While EVA Park does not necessarily save therapists' clinical time, because they need to be active online at the same time as PWA to deliver synchronous therapy, it does mean that therapy can be delivered interactively yet remotely and is also sufficiently engaging that high dose practice can be undertaken; reviews have indicated that PWA logged an average of 41 hours of engagement in EVA Park over 5 weeks (Marshall et al., 2016).

Importantly, I saw generalisation of improved naming accuracy to untreated items, at both 1-week and 1-month post-intervention. This result is promising, but should be treated with caution. Although statistically significant, the number of correctly named untreated items was much smaller than the number of correctly named treated items for all PWA. It is possible that although I completed two baseline assessments to attempt to capture PWA variability in naming, this was insufficient. Indeed, in a single case study tracking PWA performance on 9 iterations of the *Boston Naming Test* across 55 months, Stark et al. (2016) found that naming performance varied from 15% to 60% accuracy across time, with great variability in the items named correctly/incorrectly at each time point, despite no treatment between tests. Furthermore, it should be noted that generalisation within the context of this study was controlled: generalisation related to whether PWA showed improvement in naming of a small set of matched items at post-intervention testing, not whether functional change in communication skills occurred. Therefore, in future studies a broader range of metrics could be used to assess change after therapy; Webster et al. (2015) has proposed that post-intervention measures should assess both within-level generalisation (i.e. change at the level being targeted) and across-level generalisation (i.e. change at other levels of output). In addition to impairment-based measures of change post-intervention, change in communicative participation, functioning and quality of life could additionally be investigated.

5.6.3 Limitations

A key limitation of this study is the small number of PWA involved. Although this is not atypical for intervention studies in aphasia (Brady et al., 2014), it remains challenging to draw conclusions about recovery trajectories and the role of individual cognitive differences in treatment response. Further, although PWA in the study were recruited based on pre-selected criteria that were intended to render then as similar as possible behaviourally (i.e. moderate aphasia, relatively unimpaired auditory comprehension in the presence of significant naming deficits, and so on), they nevertheless showed heterogeneity in cognitive profiles, some aspects of which may have impacted gesture benefit effect. Replication of the intervention with a larger sample of PWA would be beneficial. Additionally, neuroimaging data was unavailable and lesion effects were therefore unable to be evaluated.

Notably, I did not replicate the results reported in Chapter 2, where I found a significant improvement in naming accuracy following gesture observation for PWA with poorer working memory/executive function capacity. There are several possible reasons for this. Firstly, compared to the larger PWA cohort described in Chapter 2, there was limited variation in working memory/executive function scores in the PWA who took part in this study and it is possible that the analysis was insufficiently sensitive to detect response. Relatedly, compared to the “Go Fish!” game outlined in Chapter 2 where the maximum time lag between seeing the researcher’s gestures and producing target words was 5 game turns (equating to a maximum time of around 120 seconds), in this study there was a very long time lag between when PWA observed gestures and when they were required to produce target words at post-intervention assessment (i.e. up to 1 month). It is therefore possible that the positive effects of gesture observation on naming accuracy are relatively short-term and gesture priming was insufficient to facilitate consolidation of information into long-term memory for PWA. Finally, due to background distractions at home, PWA may not have consistently closely attended to the researcher’s gestures during online self-paced therapy sessions. In the “Go Fish!” game study outlined in Chapter 2 where positive gesture observation

effects were seen, the researcher was able to ensure that PWA were focused during each trial and were watching when the researcher produced the gestures. While this additional scaffolding could be replicated by the researcher during live therapy sessions, it was not possible to do so during self-paced sessions that PWA completed independently at home. It is therefore possible that PWA did not see the researcher's gestures in each therapy trial and may not have benefited for this reason.

5.7 Conclusion

To sum up, this study demonstrates that a self-paced, computer-based therapy programme targeting word-retrieval is a feasible way of delivering intensive, personalised anomia therapy to PWA, including those with moderate-severe aphasia. Naming accuracy significantly improved at 1-week post-intervention compared to baseline and effects were maintained at 1-month post-intervention; there was also evidence of limited generalisation of naming improvements to a matched untreated set of items. Although gesture observation did not result in improved naming accuracy in every case, gestures which were rated as better depicting target words were found to have a significant beneficial effect on naming accuracy. I suggest that this may be due to transparency effects, with gestures which are more highly imagistically related to target words resulting in enhanced semantic activation of target items and better relearning outcomes for PWA.

Chapter 6: General Discussion

6.1 Summary of goals and findings

Communication is multimodal; in face-to-face exchanges between two (or more) people, speech and gestures always co-occur. Yet, they are often studied in isolation. The primary goals of this thesis were to investigate how gestures, both observed and produced by PWA, integrate with speech to facilitate speech comprehension, speech production and communicative effectiveness in dyadic interactions, and to explore whether gesture effects can be leveraged within therapy to facilitate language rehabilitation.

Across four experimental studies I have demonstrated that individual moderating factors in PWA cognitive profiles, as well as salient features of gestures themselves, play a key role in answering these questions. I have shown that PWA gesture recognition and visual attention capacity moderates whether gesture observation facilitates speech comprehension, while WM/EF moderates gesture observation and production benefit for word-retrieval. I have also shown that while PWA effectively and flexibly produce gestures to convey information to interlocuters that they cannot convey through speech, gesture informativeness is moderated by upper limb apraxia (praxis) severity. Finally, I have demonstrated that while combined gesture observation and naming therapy improves word-retrieval in at least some PWA, the effect is moderated by gesture transparency, with those gestures that are most closely semantically aligned to the target words showing the greatest facilitatory effects for word-retrieval. These results provide new insights and constraints for theories of speech-gesture interaction.

In Chapter 2, I described findings from a novel dyadic LAG, based on the card game “Go Fish!”, designed to investigate whether gestures facilitated speech comprehension and (single word) picture naming in PWA. In this game, players took turns requesting cards from each other to make matching pairs; in half of game turns players were permitted to gesture, while in the other half players were constrained from gesturing. I showed that although

observing gestures did not facilitate speech comprehension across all PWA, a subset benefited; this subset was characterised by better gesture recognition abilities. I also showed dissociating effects of gesture observation and production on PWA word-retrieval: PWA with poorer WM/EF benefited from observing gestures in subsequent game turns, while PWA with better WM/EF did not. Importantly, gesture production did not facilitate word retrieval in the majority of PWA; only a small minority of PWA with significantly better WM/EF benefited from producing gestures.

In Chapter 3, I described findings from a second dyadic LAG, based on the card game “Heads Up!”, developed to explore whether the effects of gesture observation seen in Chapter 2 in a single word naming context extended to a connected speech context. In this game, players took turns describing common objects/actions to each other. As before, in half of turns players were permitted to gesture, while in the other half gesture use was constrained. I also explored whether gesture production showed measurable effects on connected speech output in PWA. I found that although there was no simple significant effect of gesture condition, visual attention capacity moderated gesture observation benefit for naming in (some) PWA; I argued that in this case, gestures in fact targeted comprehension resulting in downstream positive effects on naming accuracy. I also found that PWA gesture production/constraint significantly affected connected speech measures: MLU was significantly shorter when PWA were permitted to gesture, while utterance concreteness was significantly higher. Lexical diversity, as measured by MATTR, did not differ significantly by gesture condition. I have argued that these findings suggest that speech and gestures interact from conceptualisation of the message, rather than from the point of lexical or phonemic selection.

In Chapter 4, I turned to the question of whether gestures produced by PWA contributed to communicative effectiveness from the perspective of the interlocuter. I asked 168 neurologically healthy adults to watch and/or listen to approximately 140 connected speech passages produced by PWA with moderate aphasia. Half of clips were drawn from passages elicited when

PWA produced gestures alongside speech, and half without. I showed that neurologically healthy adults were able to use information encoded in gestures produced by PWA to compensate for impaired speech production in understanding PWA messages, but this effect was moderated by limb apraxia (praxis) severity in PWA; gestures produced by PWA with more severe praxis deficits contributed less to successful communication than gestures produced by PWA with less severe praxis deficits, replicating previous findings by Hogrefe et al. (2012). Thus, PWA may effectively use gestures to communicate information that they cannot convey through speech.

In Chapter 5, I explored the extent to which an intensive computer-based therapy programme combining gesture observation with naming could improve word retrieval in a group of 4 PWA with severe naming deficits but relatively unimpaired speech comprehension. PWA completed a self-paced therapy programme across 20 sessions with half of treatment items delivered alongside gestures and half without. I demonstrated that naming accuracy for treated items (regardless of gesture condition) significantly improved one week after completing the intervention and effects were maintained one month later. I further showed that within the set of treated items presented alongside (iconic) gestures, the greatest facilitation effects came from gestures that were the closest semantically linked to the target words (i.e. the most transparent gestures), replicating learning effects previously found in other clinical populations, including children with language delay (Lüke et al., 2020) and Developmental Language Disorder (Vogt & Kauschke, 2017).

6.2 Theoretical implications

6.2.1 Moderating factors for gesture benefit

6.2.1.1 Comprehension

Representational (iconic and metaphoric) gestures have been argued to support comprehension because they convey semantic information about referent(s) which may help to disambiguate meaning or highlight relevant

properties, e.g. tracing the outline of a ball in the air highlights its shape, while gesturing ‘bouncing’ or ‘throwing’ highlights how a ball may be used (Drijvers & Özyürek, 2017; Holle et al., 2010). The findings presented in this thesis indicate that the semantic information encoded in gestures may be relevant for single-word speech comprehension only in so far as PWA are able to recognise its meaning and integrate this with speech content. Therefore, for PWA with significant impairments in gesture recognition (e.g. as part of ideational apraxia), observing gestures may not contribute to comprehension.

For connected speech, gesture observation benefit for comprehension was moderated by (visual) attention capacity; when gestures were not present, no attention effect was observable, but when gestures were present PWA with poorer attention capacity showed no benefit from gesture observation. This effect may be due to PWA with poorer attention capacity being less able to attend to (and integrate) information from both input streams (i.e. speech and gestures produced by the researcher) simultaneously. Importantly, as the researcher used both representational and non-representational gestures, this finding does not exclude the possibility that non-representational gestures (e.g. beats) also contribute to PWA comprehension, although this would presumably occur via an alternative mechanism than semantic activation (since such gestures are assumed not to carry semantic content), feasibly by directing attention towards key words in an utterance, perhaps particularly when they interact with prosodic peaks (Ferrari & Hagoort, 2025). Furthermore, the fact that PWA attempted to use gestural information when it was present and accessible to them may also constitute tentative evidence for the Integrated-Systems Hypothesis (Kelly et al., 2010), i.e. the proposal that PWA (and neurologically healthy adults) obligatorily attempt to integrate information from gestures with speech for comprehension, although this was not specifically tested.

6.2.1.2 Speech production

Although for most PWA gesture *production* was associated with poorer naming accuracy in a single word/single gesture context, in a minority of PWA with significantly *better* WM/EF gesture production was associated with improved naming. This may be due to better retained ability to manage multiple communicative channels (i.e. speech and gesture) in close temporal proximity, or better retained semantic control processes. Conversely, (iconic) gesture *observation* improved subsequent naming accuracy for PWA with significant *deficits* in WM/EF. This finding aligns with previous evidence from Morsella & Krauss (2004), who found that PWA retained spatial concept words for longer when they were produced with representational gestures, perhaps because the semantic features of the gestures sustained and enhanced activation of the target words in memory. I have suggested that a similar process may be taking place here: observing gestures produced by the researcher may strengthen activation of recently presented words in PWA with more impaired WM/EF, facilitating retrieval. These findings are tentatively supported by observations from the lesion data, although as the number of PWA for whom neuroimaging was available is relatively small these results should be treated with caution. Gesture observation benefit for naming was associated with lesions along the dorsal language processing pathway and the dorsolateral prefrontal cortex, areas associated with (among other things) WM/EF and semantic control; while PWA who showed gesture production benefit for naming presented with relative sparing of the dorsolateral prefrontal cortex (Barbey et al., 2013). For naming in a single word/single gesture context, therefore, there was evidence of dissociation between gesture production and observation benefit, which was moderated by WM/EF.

Gesture transparency, or the extent to which an iconic gesture is semantically associated to a target word, is also relevant for naming facilitation, at least in a therapy context. As I showed in Chapter 5, although pairing observation of an iconic gesture with a word did not result in improved naming outcomes alone, the more transparent the iconic gesture,

the more effective the (re-)learning facilitation. This may be because greater transparency enabled easier mapping between underlying semantic representations, phonological forms, lexical items and gestures, reducing the cognitive effort required for (re-)learning (Lüke & Ritterfeld, 2014). If so, this finding suggests that pairing highly transparent iconic gesture observation with intensive word repetition may be a feasible anomia therapy strategy for PWA with the greatest cognitive deficits.

6.2.1.3 Communicative effectiveness

As shown in Chapter 4, PWA were able to use gestures in a compensatory manner to convey information effectively to interlocuters, even in cases of significant expressive aphasia. I have argued that this suggests that parallel breakdowns in gesture and speech production, perhaps reflecting an underlying loss of semantic representations which has been documented in some people with Alzheimer's (Carlomagno et al., 2005), cannot be assumed post-stroke. Instead, for PWA, limb apraxia (praxis) severity moderated the informativeness of the gestures they were able to produce, from the perspective of neurologically healthy interlocuters.

The differential breakdowns in speech and gesture production seen in (at least some) PWA may be due to disrupted connections between semantic representations, phonological forms and speech motor systems, while connections between semantics and motor systems for gesture production may remain (relatively) unaffected despite neural proximity. This may also account for the varying effects of limb apraxia on gesture informativeness, at least for PWA with ideomotor apraxia: PWA with more severe limb apraxia are predicted to be those with more disruptions in connections between semantics and motor systems for gesture production, while those with less severe limb apraxia are predicted to be those with fewer disrupted connections. Indeed, ideomotor apraxia is typically associated with fronto-parietal damage but sparing of the temporal lobe (Buxbaum & Kalénine, 2010), reinforcing the suggestion that for these PWA underlying semantic representations may remain intact, but cannot be accessed.

6.2.2 Models of speech-gesture integration

Findings from this thesis also help to clarify a key issue in research into multimodal communication: in the process of speech production, at what point do speech and gestures interact? Broadly speaking, results from this thesis support predictions from the Information Packaging Hypothesis, rather than the LRH: the evidence suggests that speech and gestures interact from the earliest stages of conceptualisation. Results from this thesis provide new insights into how information is packaged in speech and gesture; the interactions between gesture, WM/EF and visual attention provide new constraints that theories of speech-gesture interaction must seek to explain.

For the majority of PWA included in this thesis, contrary to the predictions of the LRH, gesture production failed to facilitate word-retrieval. Further, production of gestures (at least in a single word naming context) was significantly associated with speech omission errors and unrelated errors (i.e. and not associated with semantic and/or phonemic speech errors), which I have argued suggests that rather than using gestures to enhance activation of lexical items and aid retrieval, PWA used gestures to convey information they were unable to convey through speech, supporting previous findings by Akhavan et al. (2017) and de Beer et al. (2017) and as proposed by the Information Packaging Hypothesis. Indeed, I have argued that this is a primary function of gesture production for PWA. Supporting (albeit indirect) evidence for this suggestion comes from Chapter 3, where I showed that lexical diversity of PWA utterances did not vary depending on PWA gesture use/constraint, contrary to the predictions of the LRH, while MLU was significantly shorter when PWA were permitted to gesture, as predicted by the Information Packaging Hypothesis. I also demonstrated that mean utterance concreteness was higher when PWA used gestures, suggesting that gesture use/constraint *influenced* the semantic content of speech and indicating that speech and gestures interact from the earliest stages of message conceptualisation.

While the focus of this thesis was speech and gesture integration specifically, other channels may also interact at the level of semantics, aiding communication in PWA. For example, certain words exhibit iconic prosody, i.e. a non-arbitrary link between semantic and prosodic content, such as producing the word “long” with an extended vowel sound. Iconic wordforms (e.g. onomatopoeia such as “boom!” or “splash!”) are known to be comprehended and produced more accurately than non-iconic words in PWA, perhaps due to more robust mapping between semantics and phonology (Meteyard et al., 2015); the effect has not been tested for iconic prosodic words specifically, yet they may also convey a processing advantage for PWA if semantics and prosody interact, particularly since prosodic processing is believed to be right hemisphere-lateralised and therefore less vulnerable to the (predominantly) left hemisphere lesions that cause aphasia (Kreitewolf et al., 2014). If so, this suggests that not only gestures, but potentially other communicative channels, may also interact from the earliest stages of message conceptualisation.

6.3 Clinical implications

Against a background of increased interest in personalised language rehabilitation approaches, findings from this thesis underline the importance of comprehensively assessing individual differences in cognitive abilities in PWA in planning appropriate therapy. This is a challenging task. Multiple cognitive and lesion variables have been implicated in the literature in moderating gesture processing abilities in PWA (e.g. Völsch et al., 2021, Carlomagno & Cristilli, 2006; Sekine & Rose, 2013, Kong et al., 2015, Preisig et al., 2018; Dresang et al., 2023). Clinicians have limited time to assess PWA and deciding which cognitive variables to prioritise is vital. As outlined above, findings from this thesis suggest that WM/EF, apraxic deficits and (visual) attention capacity should be included in determining whether PWA may benefit from using and/or observing gestures in interactions.

Relatedly, appropriate goal-setting and outcome measurements for interventions is also critical. Findings from Chapters 2, 3 and 4 broadly

suggest that gestures produced by PWA primarily contribute to communication by conveying information absent from speech. Further, findings from Chapters 2, 3 and 5 suggest that observing (iconic) gestures produced by interlocuters may have beneficial effects for comprehension and naming in (at least some) PWA, provided the gestures are as closely semantically-related to the target words as possible. Therefore, where the goal is to maximise communicative functioning (as opposed to speech production only), multimodal therapy is appropriate. Findings from this thesis therefore broadly support the approach of M-MAT (Rose et al., 2013) over CIAT (Pulvermüller et al., 2001), except where the specific goal of therapy is improved naming. In terms of measuring the success of therapy, findings from this thesis also suggest that functional communication measures which incorporate how effectively PWA communicate in different situations and through different modalities are vital. *The Scenario Test* (Hilari et al., 2018) is an example of such a functional assessment. However, the current version has several limitations: firstly, it was originally designed for PWA with severe aphasia and PWA with mild to moderate aphasia typically score at or near ceiling; secondly, it requires PWA to imagine themselves in a particular situation, which is cognitively challenging (Doedens & Meteyard, 2020). Structured observation frameworks for PWA communicating in real-life situations, including PWA self-assessment and communication partner assessment measures, may therefore be beneficial.

I have also showed that intensive multimodal aphasia therapy can be delivered effectively online, asynchronously. For PWA in harder-to-reach communities (e.g. PWA living in rural areas or who are unable to travel), online delivery may offer an important means of reducing healthcare access inequalities, as well as reducing therapist costs (Pierce et al., 2025). Extending intensive interactive computer-based therapies, such as EVA Park (Marshall et al., 2016), or incorporating enhanced virtual reality environments, are key avenues for further investigation and development.

Since PWA may flexibly use gestures to communicate information effectively, findings from this thesis also highlight the importance of effective training to

support communication partners of PWA to attend to and interpret PWA's use of multimodal cues (not restricted to gestures only). During interaction interlocuters are assumed to automatically engage in *alignment*: a cognitive process by which interlocuters construct (similar) semantic representations about aspects of the world, enabling them to "think like" each other. For communication to be successful, interlocuters' semantic representations must be sufficiently aligned that there is shared understanding about, for example, what words mean. Crucially, alignment may only occur in an interactive context, because it relies on the active participation of interlocuters to form and adjust the shared semantic space for communication (Pickering & Garrod, 2006). One of the main purposes then, of communication partner training, is to facilitate the development of alignment.

As part of the process of alignment, speakers may use what interlocuters say to prime their own speech production. For example, Cleland & Pickering (2003) found that interlocuters copy each other's syntactic structures, even when these are less common (e.g. a speaker producing "Touch the square that is red" instead of "Touch the red square", in response to another speaker producing "Touch the rectangle that is blue"). This also extends to vocabulary choices; for example, in a cooperative maze game, Garrod & Anderson (1987) found that subsequent speakers used the same terminology as the original speaker, such as giving instructions based on map coordinates or referring to "lines of houses" rather than "rows" if the original speaker did so. Alignment therefore occurs across different levels of representation (e.g. syntactic, lexical etc.), and while this has not been investigated in relation to other communicative channels it is possible that gestural alignment (and indeed alignment of other channels) may also occur, e.g. in the form of interlocuters producing progressively more similar representational gestures for target words. As part of training, communication partners should be supported to understand explicitly how and when their own communication may prime the communication of PWA in interaction, and vice versa.

Structured communication partner training programmes (e.g. *SPPARC*, Lock et al., 2001) already exist and could be adapted and extended to include individualised tailored training about PWA use of multimodal cues; again, this could be delivered online to maximise access. Indeed, Beeke et al. (2021) found that communication partners who received communication partner training delivered online ("TeleCPT") rated it as highly effective, with the majority of communication goals either being met or exceeded during the training period. This suggests that alternative methods of delivering communication partner training (e.g. online, in groups etc.) may be a viable option.

6.4 Limitations and future directions

This thesis represents an initial attempt at characterising the extent to which gestures contribute to speech comprehension, production and communicative effectiveness in PWA, and at identifying key cognitive factors and aspects of gesture that mediate these effects. Much more work is needed to investigate how PWA produce and understand multimodal cues, particularly in terms of how they interact together and with speech.

Firstly, findings in this thesis relate largely (though not entirely) to the role of representational (iconic and metaphoric) gestures in PWA communication. The role of other gesture types for comprehension and speech production is under-explored (as indeed is the case in the literature as a whole), particularly gestures which do not carry semantic content, i.e. beats and pragmatic gestures. In interactions between neurologically healthy adults, beat gestures are typically produced in temporal alignment with prosodic peaks and lexical stress, and have been shown to facilitate comprehension in listeners by disambiguating between words with the same phonemes (e.g. "OBject" [object word] vs. "obJECT" [action word]), at least in interactions between first language speakers (Bosker & Peeters, 2021). Interestingly, the opposite effect has been found in non-native speakers, with observation of beat gestures significantly hindering speech comprehension in L2 learners, which the authors argue relates to processing of beats increasing cognitive

load (Rohrer et al., 2020). The effects of representational vs. non-representational gestures on comprehension have not been explored in PWA, although findings from Chapter 3 suggest a combination of representational and non-representational gestures may facilitate comprehension in some PWA.

In terms of speech production, in two studies comparing naming accuracy and naming latency of low frequency words in neurologically healthy adults, Lucero et al. (2014) found that producing beat gestures alongside speech resulted in faster naming latencies compared to baseline, while producing iconic gestures did not. However, naming accuracy did not significantly differ by gesture type. It is unclear why differential effects on naming latency might be, and similar comparisons have not been investigated in PWA. As beat gestures are typically assumed not to carry semantic content, producing beats is not assumed to enhance semantic activation of target words in the lexicon. However, Ferré (2021) found that PWA with non-fluent aphasia used beats at regularly timed intervals in speech production, similarly to a metronome, which may create a rhythmic effect of entrainment that facilitates word-retrieval. Further investigation is needed to determine if these effects occur across PWA and in different communicative contexts, as well as exploration of the cognitive mechanisms involved.

Additionally, communicators use and are exposed to a range of other multimodal cues during interactions, not simply speech and gestures. Facial expressions, vocal tone and eye-gaze are all critical elements of communication but may be altered in clinical populations, both in terms of comprehension and production. For example, Marshall et al. (2019) found that patients with semantic dementia, non-fluent Primary Progressive Aphasia and behavioural-variant fronto-temporal dementia all demonstrated impairments in facial expression identification compared to healthy controls, despite comparable levels of activation of the visual cortex between all groups. Some studies have found similar results for PWA with stroke aphasia, although results are mixed; a review of 92 studies investigating facial expression processing in adults following acquired brain damage found

that PWA with left hemisphere damage presented with significant and persistent deficits in recognising facial expressions in around 10% of studies (Yuvaraj et al., 2013). Impairments in vocal tone production and lexical stress have also been found in PWA in a range of languages, including Arabic, Mandarin and Italian (e.g. Adam, 2014; Chen et al., 2023; Laganaro et al., 2002). Finally, Preisig et al. (2016) showed that PWA have a reduced probability of shifting eye-gaze during turn-taking transitions compared to neurologically healthy controls, which the authors argue may result in interlocuters missing PWA cues to express themselves. Taken together, these findings suggest that comprehension and/or production deficits in a wide range of communicative channels exist and interact (to varying degrees) in PWA, not all of which were accounted for within the scope of this thesis. Currently, research into how these channels interact to affect PWA communication is limited.

Indeed, studying PWA communication in the ecological niche in which it evolved and where most communication today still takes place, face-to-face spontaneous interaction, remains a significant challenge. Maintaining a balance between exerting experimental control while allowing for natural communication between interlocuters is difficult. The LAG-based studies outlined in Chapters 2 and 3 go some way towards representing a middle ground between lab-based controlled studies and analysis of unplanned, spontaneous interactions (as found in conversation analytic studies, for example). But the manipulation of gestural cues and formalised turn-taking nevertheless renders these studies semi-naturalistic, rather than fully naturalistic, which is a step removed from everyday interactions. Additionally, multiple other factors are known to affect how PWA (and indeed, neurologically healthy adults) communicate in everyday life, such as extent of shared experiences/'common ground' and interlocuter familiarity (Doedens et al., 2021). For example, in a study comparing gesture production frequency of 20 PWA (from mild to severe aphasia) interacting with familiar and unfamiliar interlocuters, Kistner (2017) found that PWA used significantly more gestures (of all types) in conversations with unfamiliar listeners, perhaps to compensate because unfamiliar listeners were less able to

comprehend PWA's impaired speech output. Future studies may therefore consider how PWA use and understand gestures and other communicative channels in a range of conversational contexts, with a variety of elicitation tasks and partners, including analysis of entirely spontaneous interactions. While this is computationally and time-intensive to do on a large scale, it may nevertheless yield important insights.

Finally, although neurological lesion information was available for a small number of PWA who participated in the studies described in this thesis, this data was not available for all PWA. As a result, lesion analysis methods such as voxel-based lesion-symptom mapping were unavailable due to lack of sufficient data for statistical power. While admittedly challenging, future studies should attempt to maximise power by using large lesion datasets, such as the PLORAS database (Seghier et al., 2016), to effectively correlate behavioural findings with neuroimaging in order to refine our understanding of the neural correlates of multimodal integration.

6.5 Conclusion

In this thesis, I have investigated how speech and gestures interact to influence comprehension, speech production and communicative effectiveness in PWA. I have shown that PWA produced gestures to convey information effectively in the face of word-finding difficulties, and used the semantic information encoded in gestures produced by others to facilitate speech comprehension. Findings from this thesis indicate that speech and gestures interact from the earliest stages of message conceptualisation, and I have identified key cognitive variables (WM/EF, visual attention, upper limb apraxia) that moderate this interaction. Finally, I have shown that intensive anomia therapy incorporating gesture observation facilitates naming in people with severe word-finding difficulties, at least when the gestures are closely semantically-related to the target word(s). Theories of speech-gesture integration must account for these constraints. Crucially, this thesis highlights the importance of investigating communication multimodally, as a pathway to understanding aphasia, and, ultimately, for developing effective rehabilitation programmes to lessen its impact.

“Research is part of the healing.”

- A participant

References

Adam, H. (2014). Dysprosody in aphasia: An acoustic analysis evidence from Palestinian Arabic. In *Journal of Language and Linguistic Studies* (Vol. 10, Issue 1). www.jlls.org

Akbıyık, S., Karaduman, A., Göksun, T., & Chatterjee, A. (2018). The relationship between co-speech gesture production and macrolinguistic discourse abilities in people with focal brain injury. *Neuropsychologia*, *117*, 440–453. https://doi.org/10.1016/j.neuropsychologia.2018.06.025

Akhavan, N., Göksun, T., & Nozari, N. (2017). Integrity and function of gestures in aphasia. *Aphasiology*, *32*(11), 1310–1335. https://doi.org/10.1080/02687038.2017.1396573

Alyahya, R. S. W. (2018). *Single-Word Processing and Connected Speech Production in Aphasia: Integrating Neuropsychological, Psycholinguistic and Neuroimaging Approaches*. University of Manchester.

Andric, M., & Smal, S. L. (2012). Gesture's neural language. *Frontiers in Psychology*, *3*(APR), 16740. https://doi.org/10.3389/FPSYG.2012.00099/BIBTEX

Arjmandi, M. K., & Behroozmand, R. (2024). On the interplay between speech perception and production: insights from research and theories. *Frontiers in Neuroscience*, *18*, 1347614. https://doi.org/10.3389/FNINS.2024.1347614

Attard, M., Rose, M. L., & Lanyon, L. (2013). The comparative effects of Multi-Modality Aphasia Therapy and Constraint-Induced Aphasia Therapy-Plus for severe chronic Broca's aphasia: An in-depth pilot study. *Aphasiology*, *27*(1), 80–111. https://doi.org/10.1080/02687038.2012.725242

Bak, T. H., & Hodges, J. R. (2003). Kissing and dancing—a test to distinguish the lexical and conceptual contributions to noun/verb and action/object dissociation. Preliminary results in patients with frontotemporal dementia. *Journal of Neurolinguistics*, *16*(2–3), 169–181. https://doi.org/10.1016/S0911-6044(02)00011-8

Bakhtin, M. M. (1986). Towards a Methodology of the Human Sciences. In *Speech Genres and Other Late Essays* (pp. 159–172). University of Texas Press.

Barbey, A. K., Koenigs, M., & Grafman, J. (2013). Dorsolateral prefrontal contributions to human working memory. *Cortex*, *49*(5), 1195–1205. https://doi.org/10.1016/J.CORTEX.2012.05.022

Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/JSS.V067.I01

Baumard, J., Lesourd, M., Remigereau, C., Merck, C., Jarry, C., Etcharry-Bouyx, F., Chauviré, V., Belliard, S., Moreaud, O., Osiurak, F., & le Gall, D. (2019). The – weak – role of memory in tool use: Evidence from neurodegenerative diseases. *Neuropsychologia*, *129*, 117–132. https://doi.org/10.1016/j.neuropsychologia.2019.03.008

Bavelas JB, Chovil N, Lawrie AD, & Wade A. (1992). Interactive gestures. *Discourse Processes*, *15*, 469–489.

Beattie, G., & Coughlan, J. (1999). An experimental investigation of the role of iconic gestures in lexical access using the tip-of-the-tongue phenomenon. *British Journal of Psychology*, *90*(1), 35–56. https://doi.org/10.1348/000712699161251

Beeke, S., Sirman, N., Beckley, F., Maxim, J., Edwards, S., Swinburn, K., & Best, W. (2013). *Better Conversations with Aphasia: an E-Learning Resource | Short courses - UCL – University College London*. https://www.ucl.ac.uk/short-courses/search-courses/better-conversations-aphasia-e-learning-resource

Beeke, S., Volkmer, A., & Farrington-Douglas, C. (2021). TeleCPT: Delivery of a Better Conversations Approach to Communication Partner Training During a Global Pandemic and Beyond. *Perspectives of the ASHA Special Interest Groups*, *6*(6), 1776–1785. https://doi.org/10.1044/2021_PERSP-21-00160

Behrmann, M., & Penn, C. (1984). Non-verbal communication of aphasic patients. *International Journal of Language & Communication Disorders*, *19*(2), 155–168. https://doi.org/10.3109/13682828409007186

Berthier, M. L. (1995). Transcortical sensory aphasia: Dissociation between naming and comprehension. *Aphasiology*, *9*(5), 431–451. https://doi.org/10.1080/02687039508248708

Binder, J. R. (2002). Wernicke Aphasia: A Disorder of Central Language Processing. In M. D'Esposito (Ed.), *Neurological Foundations of Cognitive Neuroscience* (pp. 75–238). The MIT Press. https://philpapers.org/rec/BINWA

Blake, M. L., Duffy, J. R., Myers, P. S., & Tompkins, C. A. (2010). Prevalence and patterns of right hemisphere cognitive/communicative deficits: Retrospective data from an inpatient rehabilitation unit. *Aphasiology*, *16*(4–6), 537–547. https://doi.org/10.1080/02687030244000194

Blonder, L. X., Burns, A. F., Bowers, D., Moore, R. W., & Heilman, K. M. (1995). Spontaneous gestures following right hemisphere infarct. *Neuropsychologia*, *33*(2), 203–213. https://doi.org/10.1016/0028-3932(94)00099-B

Bocanegra, Y., García, A. M., Pineda, D., Buriticá, O., Villegas, A., Lopera, F., Gómez, D., Gómez-Arias, C., Cardona, J. F., Trujillo, N., & Ibáñez, A.

(2015). Syntax, action verbs, action semantics, and object semantics in Parkinson's disease: Dissociability, progression, and executive influences. *Cortex*, *69*, 237–254. https://doi.org/10.1016/J.CORTEX.2015.05.022

Boersma, P. (2002). Praat, a system for doing phonetics by computer. *Glot International*, *5*.

Bonifazi, S., Tomaiuolo, F., Altoè, G., Ceravolo, M., Provinciali, L., & Marangolo, P. (2013). Action observation as a useful approach for enhancing recovery of verb production: new evidence from aphasia of innovative low-cost interventions in language rehabilitation. *Eur J Phys Rehabil Med*, *49*(4), 473–481.

Bonilha, L., Gleichgerrcht, E., Nesland, T., Rorden, C., & Fridriksson, J. (2016). Success of anomia treatment in aphasia is associated with preserved architecture of global and left temporal lobe structural networks. *Neurorehabil Neural Repair*, *30*(3), 266–279. https://doi.org/10.1177/1545968315593808

Boo, M., & Rose, M. L. (2010). The efficacy of repetition, semantic, and gesture treatments for verb retrieval and use in Broca's aphasia. *Aphasiology*, *25*(2), 154–175. https://doi.org/10.1080/02687031003743789

Borod, J. C., Fitzpatrick, P. M., Helm-Estabrooks, N., & Goodglass, H. (1989). The relationship between limb apraxia and the spontaneous use of communicative gesture in aphasia. *Brain and Cognition*, *10*(1), 121–131. https://doi.org/10.1016/0278-2626(89)90079-1

Bosker, H. R., & Peeters, D. (2021). Beat gestures influence which speech sounds you hear. *Proceedings of the Royal Society B: Biological Sciences*, *288*, 20202419. https://doi.org/10.1098/rspb.2020.2419

Bozeat S, Lambon Ralph MA, Patterson K, Garrard P, & Hodges JR. (2000). Non-verbal semantic impairment in semantic dementia. *Neuropsychologia* , *38*, 1207–1215.

Brady, M. C. (2022). Dosage, Intensity, and Frequency of Language Therapy for Aphasia: A Systematic Review-Based, Individual Participant Data Network Meta-Analysis. *Stroke*, Epub ahead of print. https://doi.org/10.1161/STROKEAHA.121.035216

Brady, M. C., Ali, M., Berg, K. Vanden, Williams, L. J., Williams, L. R., Abo, M., Becker, F., Bowen, A., Brandenburg, C., Breitenstein, C., Bruehl, S., Copland, D. A., Cranfill, T. B., Di Pietro-Bachmann, M., Enderby, P., Fillingham, J., Galli, F. L., Gandolfi, M., Glize, B., … Wright, H. H. (2022). Complex speech-language therapy interventions for stroke-related aphasia: the RELEASE study incorporating a systematic review and individual participant data network meta-analysis. *Health and Social Care Delivery Research*, *10*(28). https://doi.org/10.3310/RTLH7522

Brady, M. C., Ali, M., Fyndanis, C., Kambanaros, M., Grohmann, K. K., Laska, A. C., Hernández-Sacristán, C., & Varlokosta, S. (2014). Time for a step change? Improving the efficiency, relevance, reliability, validity and transparency of aphasia rehabilitation research through core outcome measures, a common data set and improved reporting criteria. *Aphasiology*, *28*(11), 1385–1392. https://doi.org/10.1080/02687038.2014.930261

Brady, M. C., Kelly, H., Godwin, J., Enderby, P., & Campbell, P. (2016). Speech and language therapy for aphasia following stroke. *The Cochrane Database of Systematic Reviews*, *2016*(6). https://doi.org/10.1002/14651858.CD000425.PUB4

Breckinridge Church, R., Garber, P., & Rogalski, K. (2007). The role of gesture in memory and social communication. *Gesture*, *7*(2), 137–158. https://doi.org/10.1075/GEST.7.2.02BRE/CITE/REFWORKS

Brown, R. (1973). Development of the first language in the human species. *American Psychologist*, *28*(2), 97–106. https://doi.org/10.1037/H0034209

Brown-Schmidt, S. (2009). Partner-specific interpretation of maintained referential precedents during interactive dialog. *Journal of Memory and Language*, *61*(2), 171–190. https://doi.org/10.1016/J.JML.2009.04.003

Bryant, L., Spencer, E., Ferguson, A., Craig, H., Colyvas, K., & Worrall, L. (2013). Propositional Idea Density in aphasic discourse. *Aphasiology*, *27*(8), 992–1009. https://doi.org/10.1080/02687038.2013.803514

Brysbaert, M., & New, B. (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods*, *41*(4), 977–990. https://doi.org/10.3758/BRM.41.4.977/METRICS

Buck, R., & Duffy, R. J. (1980). Nonverbal Communication of Affect in Brain-Damaged Patients. *Cortex*, *16*(3), 351–362. https://doi.org/10.1016/S0010-9452(80)80037-2

Bullier, B., Cassoudesalle, H., Villain, M., Cogné, M., Mollo, C., De Gabory, I., Dehail, P., Joseph, P. A., Sibon, I., & Glize, B. (2020). New factors that affect quality of life in patients with aphasia. *Annals of Physical and Rehabilitation Medicine*, *63*(1), 33–37. https://doi.org/10.1016/J.REHAB.2019.06.015

Buxbaum, L. J., & Kalénine, S. (2010). Action knowledge, visuomotor activation, and embodiment in the two action systems. *Ann N Y Acad Sci*, *1191*, 201–218. https://doi.org/10.1111/j.1749-6632.2010.05447.x

Buxbaum, L. J., & Kalénine, S. (2020). Cognitive-Motor Interface. In M. Poeppel & D. Mangun R.Gazzaniga (Eds.), *The Cognitive*

Neurosciences (6th ed.). MIT Press. http://direct.mit.edu/books/edited-volume/chapter-pdf/2053724/c040700_9780262356176.pdf

Buxbaum, L. J., & Randerath, J. (2018). Limb apraxia and the left parietal lobe. In *Handbook of Clinical Neurology* (Vol. 151, pp. 349–363). Elsevier B.V. https://doi.org/10.1016/B978-0-444-63622-5.00017-6

Cabiddu, F., Bott, L., Jones, G., & Gambi, C. (2022). *ChiSense-12: An English Sense-Annotated Child-Directed Speech Corpus* (pp. 5198–5205). https://aclanthology.org/2022.lrec-1.557/

Caramelli, P., Mansur, L. L., & Nitrini, R. (1998). Language and Communication Disorders in Dementia of the Alzheimer Type. In *Handbook of Neurolinguistics* (pp. 463–473). Elsevier. https://doi.org/10.1016/B978-012666055-5/50036-8

Carlomagno, S., & Cristilli, C. (2006). Semantic attributes of iconic gestures in fluent and non-fluent aphasic adults. *Brain and Language*, *99*(1–2), 104–105. https://doi.org/10.1016/J.BANDL.2006.06.061

Carlomagno S, Pandolfi M, Marini A, Di Iasi, G., & Cristilli, C. (2005). Coverbal gestures in Alzheimer's type dementia. *Cortex*, *41*(4), 535–546.

Castro, N., Nadeau, S. E., & Kendall, D. L. (2021). The Challenge of Achieving Greater Generalization in Phonological Treatment of Aphasia. *Aphasiology*, *36*(2), 170. https://doi.org/10.1080/02687038.2020.1856327

Cavanaugh, R., Kravetz, C., Jarold, L., Quique, Y., Turner, R., & Evans, W. S. (2021). Is There a Research-Practice Dosage Gap in Aphasia Rehabilitation? *American Journal of Speech-Language Pathology*, *30*(5), 2115–2129. https://doi.org/10.1044/2021_AJSLP-20-00257

Champely, S., Ekstrom, C., Dalgaard, P., Gill, J., Weibelzahl, S., Anandkumar, A., Ford, C., Volcic, R., & De Rosario, H. (2022). *Package 'pwr' Basic Functions for Power Analysis*. https://github.com/heliosdrm/pwr

Chen, W., van de Weijer, J., Qian, Q., Zhu, S., & Wang, M. (2023). Tone and vowel disruptions in Mandarin aphasia and apraxia of speech. *Clinical Linguistics & Phonetics*, *37*(8), 742–765. https://doi.org/10.1080/02699206.2022.2081611

Cicone, M., Wapner, W., Foldi, N., Zurif, E., & Gardner, H. (1979). The relation between gesture and language in aphasic communication. *Brain and Language*, *8*(3), 324–349. https://doi.org/10.1016/0093-934X(79)90060-9

Clark, E. V. (2015). Common Ground. *The Handbook of Language Emergence*, 328–353. https://doi.org/10.1002/9781118346136.CH15

Clarke, A. J., Manser, D., Fleischer, R., Fulham, M., & Ahmed, R. M. (2023). Pearls & Oy-sters: Huntington Disease Presenting as Primary Progressive Aphasia: A Case of Semantics. *Neurology*, *101*(9), 414. https://doi.org/10.1212/WNL.0000000000207428

Cleary, R. A., Poliakoff, E., Galpin, A., Dick, J. P. R., & Holler, J. (2011). An investigation of co-speech gesture production during action description in Parkinson's disease. *Parkinsonism and Related Disorders*, *17*(10), 753–756. https://doi.org/10.1016/J.PARKRELDIS.2011.08.001

Cleland, A. A., & Pickering, M. J. (2003). The use of lexical and syntactic information in language production: Evidence from the priming of noun-phrase structure. *Journal of Memory and Language*, *49*(2), 214–230. https://doi.org/10.1016/S0749-596X(03)00060-3

Clough, S., & Duff, M. C. (2020). The Role of Gesture in Communication and Cognition: Implications for Understanding and Treating Neurogenic Communication Disorders. In *Frontiers in Human Neuroscience* (Vol. 14, p. 323). Frontiers Media S.A. https://doi.org/10.3389/fnhum.2020.00323

Cochet, H., & Vauclair, J. (2014). Deictic gestures and symbolic gestures produced by adults in an experimental context: Hand shapes and hand preferences. *Laterality*, *19*(3), 278–301. https://doi.org/10.1080/1357650X.2013.804079

Cocks, N., Byrne, S., Pritchard, M., Morgan, G., & Dipper, L. (2018). Integration of speech and gesture in aphasia. *International Journal of Language & Communication Disorders*, *53*(3), 584–591. https://doi.org/10.1111/1460-6984.12372

Cocks, N., Dipper, L., Middleton, R., & Morgan, G. (2011). What can iconic gestures tell us about the language system? A case of conduction aphasia. *International Journal of Language and Communication Disorders*, *46*(4), 423–436. https://doi.org/10.3109/13682822.2010.520813

Cocks, N., Hird, K., & Kirsner, K. (2007). The relationship between right hemisphere damage and gesture in spontaneous discourse. *Aphasiology*, *21*(3–4), 299–319. https://doi.org/10.1080/02687030600911393

Cocks, N., Sautin, L., Kita, S., Morgan, G., & Zlotowitz, S. (2009). Gesture and speech integration an exploratory study of a man with aphasia. *International Journal of Language and Communication Disorders*, *44*(5), 795–804. https://doi.org/10.1080/13682820802256965

Cohen, D., Beattie, G., & Shovelton, H. (2011). Tracking the distribution of individual semantic features in gesture across spoken discourse: New perspectives in multi-modal interaction. *Semiotica*, *2011*(185), 147–188. https://doi.org/10.1515/SEMI.2011.037/MACHINEREADABLECITATION/RIS

Cordova Luna, E. P. (2021). *Anomia in people with Relapsing-Remitting Multiple Sclerosis: Investigating the nature and extent of the problem and taking steps toward better assessment and treatment*. University of Manchester.

Côte, H., Payer, M., Giroux, F., & Joanette, Y. (2010). Towards a description of clinical communication impairment profiles following right-hemisphere damage. *Aphasiology*, *21*(6–8), 739–749. https://doi.org/10.1080/02687030701192331

Covington, M. A., & McFall, J. D. (2010). Cutting the gordian knot: The moving-average type-token ratio (MATTR). *Journal of Quantitative Linguistics*, *17*(2), 94–100. https://doi.org/10.1080/09296171003643098

Cravotta, A., Prieto, P., & Busà, M. G. (2021). Exploring the effects of restraining the use of gestures on narrative speech. *Speech Communication*, *135*, 25–36. https://doi.org/10.1016/J.SPECOM.2021.09.005

Cunningham, K. T., & Haley, K. L. (2020). Measuring Lexical Diversity for Discourse Analysis in Aphasia: Moving-Average Type-Token Ratio and Word Information Measure. *Journal of Speech, Language, and Hearing Research : JSLHR*, *63*(3), 710–721. https://doi.org/10.1044/2019_JSLHR-19-00226

Dargue, N., Sweller, N., & Jones, M. P. (2019). When our hands help us understand: A meta-analysis into the effects of gesture on comprehension. *Psychological Bulletin*, *145*(8), 765–784. https://doi.org/10.1037/BUL0000202

Davidson, B., Howe, T., Worrall, L., Hickson, L., & Togher, L. (2008). Social participation for older people with aphasia: the impact of communication disability on friendships. *Topics in Stroke Rehabilitation*, *15*(4), 325–340. https://doi.org/10.1310/TSR1504-325

Davis, G. A., & Wilcox, M. J. (1985). *Adult Aphasia Rehabilitation: Applied Pragmatics*. College-Hill Press.

De Beer, C., Carragher, M., van Nispen, K., Hogrefe, K., de Ruiter, J. P., & Rose, M. L. (2017). How Much Information Do People With Aphasia Convey via Gesture? *American Journal of Speech-Language Pathology*, *26*(2), 483–497. https://doi.org/10.1044/2016_AJSLP-15-0027

de Beer, C., de Ruiter, J. P., Hielscher-Fastabend, M., & Hogrefe, K. (2019). The production of gesture and speech by people with aphasia: Influence of communicative constraints. *Journal of Speech, Language, and Hearing Research*, *62*(12), 4417–4432. https://doi.org/10.1044/2019_JSLHR-L-19-0020

de Beer, C., Hogrefe, K., Hielscher-Fastabend, M., & de Ruiter, J. P. (2020). Evaluating Models of Gesture and Speech Production for People With Aphasia. *Cognitive Science*, *44*(9). https://doi.org/10.1111/cogs.12890

De Renzi, E., & Vignolo, L. A. (1962). THE TOKEN TEST: A SENSITIVE TEST TO DETECT RECEPTIVE DISTURBANCES IN APHASICS. *Brain*, *85*(4), 665–678. https://doi.org/10.1093/BRAIN/85.4.665

de Ruiter, J. (2000). The production of gesture and speech. In D. McNeill (Ed.), *Language and Gesture (Language Culture and Cognition)* (pp. 284–311).

de Ruiter, J., & de Beer, C. (2013). A critical evaluation of models of gesture and speech production for understanding gesture in aphasia. *Aphasiology*, *27*(9), 1015–1030. https://doi.org/10.1080/02687038.2013.797067

Dell, G. S., Chang, F., & Griffin, Z. M. (1999). Connectionist Models of Language Production: Lexical Access and Grammatical Encoding. *COGNITIVE SCIENCE*, *23*(4), 517–542.

Dell, G. S., Martin, N., & Schwartz, M. F. (2007). A case-series test of the interactive two-step model of lexical access: Predicting word repetition from picture naming. *Journal of Memory and Language*, *56*(4), 490–520. https://doi.org/10.1016/J.JML.2006.05.007

Demeyere, N., Riddoch, M. J., Slavkova, E. D., Bickerton, W. L., & Humphreys, G. W. (2015). The Oxford Cognitive Screen (OCS): validation of a stroke-specific short cognitive screening tool. *Psychological Assessment*, *27*(3), 883–894. https://doi.org/10.1037/PAS0000082

Difrancesco, S., Pulvermüller, F., & Mohr, B. (2012). Intensive language-action therapy (ILAT): The methods. *Aphasiology*, *26*(11), 1317–1351. https://doi.org/10.1080/02687038.2012.705815

Dignam, J., Copland, D., McKinnon, E., Burfein, P., O'Brien, K., Farrell, A., & Rodriguez, A. D. (2015). Intensive Versus Distributed Aphasia Therapy: A Nonrandomized, Parallel-Group, Dosage-Controlled Study. *Stroke*, *46*(8), 2206–2211. https://doi.org/10.1161/STROKEAHA.115.009522

Dipper, L., Pritchard, M., Morgan, G., & Cocks, N. (2015). *The language-gesture connection: Evidence from aphasia*. https://doi.org/10.3109/02699206.2015.1036462

Doedens, W., Bose, A., Lambert, L., & Meteyard, L. (2021). Face-to-Face Communication in Aphasia: The Influence of Conversation Partner Familiarity on a Collaborative Communication Task. *Frontiers in Communication*, *0*, 90. https://doi.org/10.3389/FCOMM.2021.574051

Doedens, W. J. &, & Meteyard, L. (2018). The importance of situated language use for aphasia rehabilitation. *PsyArXiv (Preprint)*.

Doedens, W. J., & Meteyard, L. (2020). Measures of functional, real-world communication for aphasia: a critical review. *Aphasiology*, *34*(4), 492–514. https://doi.org/10.1080/02687038.2019.1702848

Donnellan, E., Özder, L. E., Man, H., Grzyb, B., Gu, Y., & Vigliocco, G. (2022). Timing relationships between representational gestures and speech: A corpus based investigation. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *44*(44), 44.

Dresang, H. C., Williamson, R., Kim, H., Hillis, A. E., & Buxbaum, L. J. (n.d.). Gesture profiles distinguish primary progressive aphasia variants. *[Manuscript Submitted for Publication]*.

Dresang, H. C., Wong, A. L., Buxbaum, L. J., & Dresang, H. C. (2023). *Shared and distinct routes in speech and gesture imitation: Evidence from stroke*. https://doi.org/10.1016/j.cortex.2023.01.010

Drijvers, L., & Özyürek, A. (2017). Visual Context Enhanced: The Joint Contribution of Iconic Gestures and Visible Speech to Degraded Speech Comprehension. *Journal of Speech, Language, and Hearing Research*, *60*(1), 212–222. https://doi.org/10.1044/2016_JSLHR-H-16-0101

Eggenberger, N., Preisig, B. C., Schumacher, R., Hopfner, S., Vanbellingen, T., Nyffeler, T., Gutbrod, K., Annoni, J.-M., Bohlhalter, S., Cazzoli, D., & Müri, R. M. (2016). Comprehension of Co-Speech Gestures in Aphasic Patients: An Eye Movement Study. *PLoS ONE*. https://doi.org/10.1371/journal.pone.0146583

Eisner, F., & McQueen, J. M. (2018). Speech Perception. In *Stevens' Handbook of Experimental Psychology and Cognitive Neuroscience* (pp. 1–46). Wiley. https://doi.org/10.1002/9781119170174.EPCN301

Emmorey, K., & Ozyurek, A. (2014). Language in our hands: neural underpinnings of sign language and co-speech gesture. In *The Cognitive Neurosciences* (pp. 656–666). MIT Press.

Evans, A. C., Collins, D. L., Millst, S. R., Brown, E. D., Kelly, R. L., & Peters, T. M. (1993). 3D statistical neuroanatomical models from 305 MRI volumes. *Proc. IEEE-Nuclear Science Symposium and Medical Imaging Conference*, 1813–1817.

Faber-Langendoen, K., Morris, J. C., Knesevich, J. W., LaBarge, E., Miller, J. P., & Berg, L. (1988). Aphasia in senile dementia of the alzheimer type. *Annals of Neurology*, *23*(4), 365–370. https://doi.org/10.1002/ANA.410230409

Fergadiotis, G., Wright, H. H., & West, T. M. (2013). Measuring Lexical Diversity in Narrative Discourse of People With Aphasia. *American Journal of Speech-Language Pathology / American Speech-Language-Hearing Association*, *22*(2), 10.1044/1058-0360(2013/12-0083). https://doi.org/10.1044/1058-0360(2013/12-0083)

Ferrari, A., & Hagoort, P. (2025). Beat gestures and prosodic prominence interactively influence language comprehension. *Cognition*, *256*, 106049. https://doi.org/10.1016/J.COGNITION.2024.106049

Ferré, G. (2021). Gesture, prosody and verbal content in non-fluent aphasic speech. *Multimodal Communication*, *10*(1), 73–91. https://doi.org/10.1515/mc-2020-0016

Ferro, J. M., Santos, M. E., Castro-Caldas, A., & Mariano, M. G. (1980). Gesture recognition in aphasia. *Journal of Clinical and Experimental Neuropsychology*, *2*(4), 277–292. https://doi.org/10.1080/01688638008403800

Feyereisen, P., Barter, D., Goossens, M., & Clerebaut, N. (1988). Gestures and speech in referential communication by aphasic subjects: Channel use and efficiency. *APHASIOLOGY*, *2*(1), 21–32. https://doi.org/10.1080/02687038808248884

Fox, J., & Weisberg, S. (2019). *An R Companion to Applied Regression* (Third edition). Sage.

Franklin, S., Harhen, D., Hayes, M., Demos Mc Manus, S., & Pollock, A. (2018). Top 10 research priorities relating to aphasia following stroke. *Aphasiology*, *32*(11), 1388–1395. https://doi.org/10.1080/02687038.2017.1417539

Frick-Horbury, D., & Guttentag, R. E. (1998). The effects of restricting hand gesture production on lexical retrieval and free recall. *American Journal of Psychology*, *111*(1), 43–62. https://doi.org/10.2307/1423536

Fridriksson, J., Den Ouden, D. B., Hillis, A. E., Hickok, G., Rorden, C., Basilakos, A., Yourganov, G., & Bonilha, L. (2018). Anatomy of aphasia revisited. *Brain : A Journal of Neurology*, *141*(3), 848–862. https://doi.org/10.1093/BRAIN/AWX363

Fromm, D., Forbes, M., Holland, A., Dalton, S. G., Richardson, J., & MacWhinney, B. (2017). Discourse Characteristics in Aphasia Beyond the Western Aphasia Battery Cutoff. *American Journal of Speech-Language Pathology*, *26*(3), 762. https://doi.org/10.1044/2016_AJSLP-16-0071

Gainotti, G., & Lemmo, M. A. (n.d.). *Comprehension of Symbolic Gestures in Aphasia*.

Garrod, S., & Anderson, A. (1987). Saying what you mean in dialogue: a study in conceptual and semantic co-ordination. *Cognition*, *27*(2), 181–218. https://doi.org/10.1016/0010-0277(87)90018-7

Geladó, S., Gómez-Ruiz, I., & Diéguez-Vide, F. (2022). Gestures analysis during a picture description task: Capacity to discriminate between healthy controls, mild cognitive impairment, and Alzheimer's disease.

Journal of Neurolinguistics, *61*. https://doi.org/10.1016/J.JNEUROLING.2021.101038

Geschwind, N. (1970). The organization of language and the brain. *Science*, *170*(3961), 940–944.

Gibson, E., Futrell, R., Piandadosi, S. T., Dautriche, I., Mahowald, K., Bergen, L., & Levy, R. (2019). How Efficiency Shapes Human Language. In *Trends in Cognitive Sciences* (Vol. 23, Issue 5, pp. 389–407). Elsevier Ltd. https://doi.org/10.1016/j.tics.2019.02.003

Gibson, R. N. (2023). *Exploring the evidence-practice gap in post-stroke aphasia in Aotearoa New Zealand* [University of Auckland]. https://hdl.handle.net/2292/66996

Glosser, G., Wiley, M. J., & Barnoski, E. J. (2010). Gestural Communication in Alzheimer's Disease. *Http://Dx.Doi.Org/10.1076/Jcen.20.1.1.1484*, *20*(1), 1–13. https://doi.org/10.1076/JCEN.20.1.1.1484

Goldin-Meadow, S., & Alibali, M. W. (2013). Gesture's role in speaking, learning, and creating language. *Annual Review of Psychology*, *64*, 257–283. https://doi.org/10.1146/ANNUREV-PSYCH-113011-143802

González-Fernández, M., Christian, A. B., Davis, C., & Hillis, A. E. (2012). Role of Aphasia in Discharge Location After Stroke. *Archives of Physical Medicine and Rehabilitation*, *94*(5), 851. https://doi.org/10.1016/J.APMR.2012.11.042

Goodglass, H., & Kaplan, E. (1963). Disturbance of Gesture and Pantomime in Aphasia. *Brain*, *86*, 703–720. https://academic.oup.com/brain/article/86/4/703/321425

Goodglass, H., & Kaplan, E. (1972). *The Assessment of Aphasia and Related Disorders*. Lea & Febiger.

Goodwin, C. (2010). Gesture, aphasia, and interaction. *Language and Gesture*, 84–98. https://doi.org/10.1017/CBO9780511620850.006

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. https://doi.org/10.1212/WNL.0B013E31821103E6

Gullberg, M. (2006). Handling Discourse: Gestures, Reference Tracking, and Communication Strategies in Early L2. *Language Learning*, *56*(1), 155–196.

Hadar, U., Burstein, A., Krauss, R., & Soroker, N. (2010). Ideational Gestures and Speech in Brain-damaged Subjects. *Language and Cognitive Processes*, *13*(1), 59–76. https://doi.org/10.1080/016909698386591

Hadar, U., & Krauss, R. K. (1999). Iconic gestures: the grammatical categories of lexical affiliates. *Journal of Neurolinguistics*, *12*(1), 1–12. https://doi.org/10.1016/S0911-6044(99)00001-9

Halai, A. D., Woollams, A. M., & Lambon Ralph, M. A. (2018). Predicting the pattern and severity of chronic post-stroke language deficits from functionally-partitioned structural lesions. *NeuroImage: Clinical*, *19*, 1–13. https://doi.org/10.1016/J.NICL.2018.03.011

Harciarek, M., Sitek, E. J., & Kertesz, A. (2014). The patterns of progression in primary progressive aphasia—Implications for assessment and management. *Https://Doi.Org/10.1080/02687038.2014.904498*, *28*(8–9), 964–980. https://doi.org/10.1080/02687038.2014.904498

Harmon, T. G., Jacks, A., Haley, K. L., & Bailliard, A. (2019). Dual-Task Effects on Story Retell for Participants With Moderate, Mild, or No Aphasia: Quantitative and Qualitative Findings. *Journal of Speech, Language, and Hearing Research : JSLHR*, *62*(6), 1890–1905. https://doi.org/10.1044/2019_JSLHR-L-18-0399

Henry, M. L., Beeson, P. M., & Rapcsak, S. Z. (2008). Treatment for Anomia in Semantic Dementia. *Seminars in Speech and Language*, *29*(1), 60. https://doi.org/10.1055/S-2008-1061625

Hilari, K., Galante, L., Huck, A., Pritchard, M., Allen, L., & Dipper, L. (2018). Cultural adaptation and psychometric testing of The Scenario Test UK for people with aphasia. *International Journal of Language & Communication Disorders*, *53*(4), 748–760. https://doi.org/10.1111/1460-6984.12379

Hogrefe, K., Rein, R., Skomroch, H., & Lausberg, H. (2016). Co-speech hand movements during narrations: What is the impact of right vs. left hemisphere brain damage? *Neuropsychologia*, *93*, 176–188. https://doi.org/10.1016/J.NEUROPSYCHOLOGIA.2016.10.015

Hogrefe, K., Ziegler, W., Weidinger, N., & Goldenberg, G. (2012). Non-verbal communication in severe aphasia: Influence of aphasia, apraxia, or semantic processing? *Cortex*, *48*(8), 952–962. https://doi.org/10.1016/j.cortex.2011.02.022

Holle, H., Obleser, J., Rueschemeyer, S. A., & Gunter, T. C. (2010). Integration of iconic gestures and speech in left superior temporal areas boosts speech comprehension under adverse listening conditions. *NeuroImage*, *49*(1), 875–884. https://doi.org/10.1016/J.NEUROIMAGE.2009.08.058

Hostetter, A. B., & Alibali, M. W. (2007). Raise your hand if you're spatial: Relations between verbal and spatial skills and gesture production. *Gesture*, *7*(1), 73–95. https://doi.org/10.1075/GEST.7.1.05HOS

Howard, D., & Gatehouse, C. (2006). Distinguishing semantic and lexical word retrieval deficits in people with aphasia. *Aphasiology*, *20*(9–11), 921–950. https://doi.org/10.1080/02687030600782679

Hula, W. D., McNeil, M. R., & Sung, J. E. (2007). Is there an impairment of language-specific attentional processing in aphasia? *Brain and Language*, *103*(1–2), 240–241. https://doi.org/10.1016/J.BANDL.2007.07.023

Humphries, S., Holler, J., Crawford, T. J., Herrera, E., & Poliakoff, E. (2016). A third-person perspective on co-speech action gestures in Parkinson's disease. *Cortex*, *78*, 44–54. https://doi.org/10.1016/J.CORTEX.2016.02.009

Humphries, S., Holler, J., Crawford, T., & Poliakoff, E. (2021). Cospeech Gestures Are a Window Into the Effects of Parkinson's Disease on Action Representations. *Journal of Experimental Psychology: General*, *150*(8), 1581–1597. https://doi.org/10.1037/xge0001002.supp

Hupp, J. M., & Gingras, M. C. (2016). The role of gesture meaningfulness in word learning. *Gesture*, *15*(3), 340–356. https://doi.org/10.1075/GEST.15.3.04HUP

Intercollegiate Stroke Working Party. (2015). *SSNAP Annual Report 2015*. King's College London.

Intercollegiate Stroke Working Party. (2019). *SSNAP Annual Report 2018*. King's Colllege London. https://www.strokeaudit.org/results/Organisational/National-Organisational.aspx

Jefferies, E., & Lambon Ralph, M. A. (2006). Semantic impairment in stroke aphasia versus semantic dementia: a case-series comparison. *Brain*, *129*(8), 2132–2147. https://doi.org/10.1093/BRAIN/AWL153

Kalénine, S., Buxbaum, L. J., & Branch Coslett, H. (2010). Critical brain regions for action recognition: lesion symptom mapping in left hemisphere stroke. *Brain*, *133*, 3269–3280. https://doi.org/10.1093/brain/awq210

Kasselimis, D. S., Simos, P. G., Peppas, C., Evdokimidis, I., & Potagas, C. (2017). The unbridged gap between clinical diagnosis and contemporary research on aphasia: A short discussion on the validity and clinical utility of taxonomic categories. *Brain and Language*, *164*, 63–67. https://doi.org/10.1016/j.bandl.2016.10.005

Kauhanen, M. L., Korpelainen, J. T., Hiltunen, P., Määttä, R., Mononen, H., Brusin, E., Sotaniemi, K. A., & Myllylä, V. V. (2000). Aphasia, depression,

and non-verbal cognitive impairment in ischaemic stroke. *Cerebrovascular Diseases (Basel, Switzerland)*, *10*(6), 455–461. https://doi.org/10.1159/000016107

Kelly, S. D., Barr, D. J., Church, R. B., & Lynch, K. (1999). Offering a Hand to Pragmatic Understanding: The Role of Speech and Gesture in Comprehension and Memory. *Journal of Memory and Language*, *40*(4), 577–592. https://doi.org/10.1006/JMLA.1999.2634

Kelly, S., Ozyürek, A., & Maris, E. (2010). Two sides of the same coin: speech and gesture mutually interact to enhance comprehension. *Psychological Science*, *21*(2), 260–267. https://doi.org/10.1177/0956797609357327

Kendon, A. (1980). Gesticulation and Speech: Two aspects of the process of utterance. In *The Relationship of Verbal and Nonverbal Communication* (pp. 207–227). De Gruyter Mouton. https://doi.org/10.1515/9783110813098.207/MACHINEREADABLECITATION/RIS

Kertesz, A. (2006). *Western Aphasia Battery--Revised*. APA PsycTests.

Kim, H.-Y. (2017). Statistical notes for clinical researchers: Chi-squared test and Fisher's exact test. *Restorative Dentistry & Endodontics*, *42*(2), 152. https://doi.org/10.5395/RDE.2017.42.2.152

Kindell, J., Sage, K., Keady, J., & Wilkinson, R. (2013). Adapting to conversation with semantic dementia: Using enactment as a compensatory strategy in everyday social interaction. *International Journal of Language and Communication Disorders*, *48*(5), 497–507. https://doi.org/10.1111/1460-6984.12023

King, D., Wittenberg, R., Patel, A., Quayyum, Z., Berdunov, V., & Knapp, M. (2020). The future incidence, prevalence and costs of stroke in the UK. *Age and Ageing*, *49*(2), 277. https://doi.org/10.1093/AGEING/AFZ163

Kircher, T., Straube, B., Leube, D., Weis, S., Sachs, O., Willmes, K., Konrad, K., & Green, A. (2009). Neural interaction of speech and gesture: Differential activations of metaphoric co-verbal gestures. *Neuropsychologia*, *47*(1), 169–179. https://doi.org/10.1016/J.NEUROPSYCHOLOGIA.2008.08.009

Kistner, J. (2017). *The use of gestures in the conversations of people with aphasia* [Unpublished PhD thesis]. City, University of London.

Kistner, J., Dipper, L. T., & Marshall, J. (2019). The use and function of gestures in word-finding difficulties in aphasia. *Aphasiology*, *33*(11), 1372–1392. https://doi.org/10.1080/02687038.2018.1541343

Kita, S. (2000). How representational gestures help speaking. In *Language and Gesture (Language Culture and Cognition* (pp. 162–185). https://doi.org/10.1017/CBO9780511620850.011

Kita, S., Alibali, M. W., & Chu, M. (2017). How Do Gestures Influence Thinking and Speaking? The Gesture-for-Conceptualization Hypothesis. *Psychological Review*, *124*(3), 245–266. https://doi.org/10.1037/rev0000059.supp

Kısa, Y. D., Goldin-Meadow, S., & Casasanto, D. (2021). Do Gestures Really Facilitate Speech Production? *Journal of Experimental Psychology: General*, *151*(6), 1252–1271. https://doi.org/10.1037/xge0001135

Klostermann, F., Wyrobnik, M., Boll, M., Ehlen, F., & Tiedt, H. O. (2022). Tracing embodied word production in persons with Parkinson's disease in distinct motor conditions. *Scientific Reports*, *12*, 16669. https://doi.org/10.1038/s41598-022-21106-6

Kok, K. (2016). The grammatical potential of co-speech gesture: A Functional Discourse Grammar perspective. *Functions of Language*, *23*(2), 149–178. https://doi.org/10.1075/FOL.23.2.01KOK

Kong, A. P. H., Law, S. P., & Chak, G. W. C. (2017). A comparison of coverbal gesture use in oral discourse among speakers with fluent and nonfluent aphasia. *Journal of Speech, Language, and Hearing Research*, *60*(7), 2031–2046. https://doi.org/10.1044/2017_JSLHR-L-16-0093

Kong, A. P. H., Law, S. P., & Cheung, C. K. Y. (2019). Use of co-verbal gestures during word-finding difficulty among Cantonese speakers with fluent aphasia and unimpaired controls. *Aphasiology*, *33*(2), 216–233. https://doi.org/10.1080/02687038.2018.1463085

Kong, A. P. H., Law, S. P., Wat, W. K. C., & Lai, C. (2015). Co-verbal gestures among speakers with aphasia: Influence of aphasia severity, linguistic and semantic skills, and hemiplegia on gesture employment in oral discourse. *Journal of Communication Disorders*, *56*, 88–102. https://doi.org/10.1016/j.jcomdis.2015.06.007

Krason, A. M. (2023). *The Role of Visual Cues in Speech Comprehension: Evidence from Brain and Behaviour*. UCL.

Krauss, R. M., Chen, Y., & Gottesman, R. F. (2000). Lexical Gestures and Lexical Access: A Process Model. In D. McNeill (Ed.), *Language and Gesture* (pp. 261–283). Cambridge University Press.

Kreitewolf, J., Friederici, A. D., & von Kriegstein, K. (2014). Hemispheric lateralization of linguistic prosody recognition in comparison to speech and speaker recognition. *NeuroImage*, *102 Pt 2*(P2), 332–344. https://doi.org/10.1016/J.NEUROIMAGE.2014.07.038

Kristensson, J., Longoni, F., Östberg, P., Rödseth Smith, S., Åke, S., & Saldert, C. (2024). Anomia in left hemisphere stroke, multiple sclerosis and Parkinson's disease - a comparative study. *Disability and Rehabilitation*, *46*(11), 2294–2316. https://doi.org/10.1080/09638288.2023.2219902

Kroenke, K. M., Kraft, I., Regenbrecht, F., & Obrig, H. (2013). Lexical learning in mild aphasia: Gesture benefit depends on patholinguistic profile and lesion pattern. *Cortex*, *49*(10), 2637–2649. https://doi.org/10.1016/J.CORTEX.2013.07.012

Kurland, J., Stanek, E. J., Stokes, P., Li, M., & Andrianopoulos, M. (2016). Intensive Language Action Therapy in Chronic Aphasia: A Randomized Clinical Trial Examining Guidance by Constraint. *American Journal of Speech-Language Pathology*, *25*(4S), S798–S812. https://doi.org/10.1044/2016_AJSLP-15-0135

Laganaro, M., Vacheresse, F., & Frauenfelder, U. H. (2002). Selective impairment of lexical stress assignment in an Italian-speaking aphasic patient. *Brain and Language*, *81*(1–3), 601–609. https://doi.org/10.1006/brln.2001.2550

Lam, J. M. C., & Wodchis, W. P. (2010). The relationship of 60 disease diagnoses and 15 conditions to preference-based health-related quality of life in Ontario hospital-based long-term care residents. *Medical Care*, *48*(4), 380–387. https://doi.org/10.1097/MLR.0B013E3181CA2647

Lambon Ralph, M. A., Snell, C., Fillingham, J. K., Conroy, P., & Sage, K. (2010). Predicting the outcome of anomia therapy for people with aphasia post CVA: both language and cognitive status are key predictors. *Neuropsychological Rehabilitation*, *20*(2), 289–305. https://doi.org/10.1080/09602010903237875

Lausberg, H., & Sloetjes, H. (2009). Coding gestural behavior with the NEUROGES-ELAN system. *Behavior Research Methods*, *41*(3), 841–849. https://doi.org/10.3758/BRM.41.3.841

Le Dorze, G., & Brassard, C. (1995). A description of the consequences of aphasia on aphasic persons and their relatives and friends, based on the WHO model of chronic diseases. *Aphasiology*, *9*(3), 239–255. https://doi.org/10.1080/02687039508248198

le May, A., David, R., & Thomas, A. P. (1988). The use of spontaneous gesture by aphasic patients. *Aphasiology*, *2*(2), 137–145. https://doi.org/10.1080/02687038808248899

Lehman Blake, M. (2018). *The Right Hemisphere and Disorders of Cognition and Communication: Theory and Clinical Practice*. Plural Publishing. https://lccn.loc.gov/2017028234

Leonard, T., & Cummins, F. (2011). The temporal relation between beat gestures and speech. *Language and Cognitive Processes*, *26*(10), 1457–1471. https://doi.org/10.1080/01690965.2010.500218

Levelt, W. J. M. (1989). *Speaking: From Intention to Articulation*. The MIT Press. https://doi.org/10.7551/MITPRESS/6393.001.0001

Levinson, S. C., & Holler, J. (2014). The origin of human multi-modal communication. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1651). https://doi.org/10.1098/RSTB.2013.0302

Lock, S., Wilkinson, R., Bryan, K., Maxim, J., Edmundson, A., Bruce, C., & Moir, D. (2001). Supporting Partners of People with Aphasia in Relationships and Cconversation (SPPARC). *International Journal of Language & Communication Disorders*, *36 Suppl*(SUPPL.), 25–30. https://doi.org/10.3109/13682820109177853

Loper, E., & Bird, S. (2002). *NLTK*. 63–70. https://doi.org/10.3115/1118108.1118117

Lucero, C., Zaharchuk, H., & Casasanto, D. (2014). Beat gestures facilitate speech production. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *36*(36), 36. https://escholarship.org/uc/item/8pw1x0sx

Lüke, C., & Ritterfeld, U. (2014). The influence of iconic and arbitrary gestures on novel word learning in children with and without SLI. *Gesture*, *14*(2), 204–225. https://doi.org/10.1075/GEST.14.2.04LUK

Lüke, C., Ritterfeld, U., Grimminger, A., Rohlfing, K. J., & Liszkowski, U. (2020). Integrated Communication System: Gesture and Language Acquisition in Typically Developing Children and Children With LD and DLD. *Frontiers in Psychology*, *11*. https://doi.org/10.3389/FPSYG.2020.00118

Lwi, S. J., Herron, T. J., Curran, B. C., Ivanova, M. V., Schendel, K., Dronkers, N. F., & Baldo, J. V. (2021). Auditory Comprehension Deficits in Post-stroke Aphasia: Neurologic and Demographic Correlates of Outcome and Recovery. *Frontiers in Neurology*, *12*. https://doi.org/10.3389/fneur.2021.680248

Macoir, J., Leroy, M., Routhier, S., Auclair-Ouellet, N., Houde, M., & Laforce, R. (2015). Improving verb anomia in the semantic variant of primary progressive aphasia: the effectiveness of a semantic-phonological cueing treatment. *Neurocase*, *21*(4), 448–456. https://doi.org/10.1080/13554794.2014.917683

Mahmoud, S. S., Pallaud, R. F., Kumar, A., Faisal, S., Wang, Y., & Fang, Q. (2023). A Comparative Investigation of Automatic Speech Recognition Platforms for Aphasia Assessment Batteries. *Sensors 2023, Vol. 23, Page 857*, *23*(2), 857. https://doi.org/10.3390/S23020857

Manning, M., Cuskelly, C., Russ, E., & Franklin, S. (2020). Supporting people with post-stroke aphasia to live well: A cross-sectional survey of Speech & Language Therapists in Ireland. *Health & Social Care in the Community*, *28*(6), 2105–2116. https://doi.org/10.1111/HSC.13021

Marangolo, P., Bonifazi, S., Tomaiuolo, F., Craighero, L., Coccia, M., Altoè, G., Provinciali, L., & Cantagallo, A. (2010). Improving language without words: First evidence from aphasia. *Neuropsychologia*, *48*(13), 3824–3833. https://doi.org/10.1016/J.NEUROPSYCHOLOGIA.2010.09.025

Marangolo, P., Cipollari, S., Fiori, V., Razzano, C., & Caltagirone, C. (2012). Walking but Not Barking Improves Verb Recovery: Implications for Action Observation Treatment in Aphasia Rehabilitation. *PLoS One*, *7*(6), e38610. https://doi.org/10.1371/journal.pone.0038610

Marsh, E. B., & Hillis, A. E. (2006). Recovery from aphasia following brain injury: the role of reorganization. In *Progress in Brain Research* (Vol. 157, pp. 143–156). https://doi.org/10.1016/S0079-6123(06)57009-8

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, J., Booth, T., Devane, N., Galliers, J., Greenwood, H., Hilari, K., Talbot, R., Wilson, S., & Woolf, C. (2016). Evaluating the Benefits of Aphasia Intervention Delivered in Virtual Reality: Results of a Quasi-Randomised Study. *PloS One*, *11*(8). https://doi.org/10.1371/JOURNAL.PONE.0160381

Marshall, J., Roper, A., Galliers, J., Wilson, S., Cocks, N., Muscroft, S., & Pring, T. (2013). Computer delivery of gesture therapy for people with severe aphasia. *Aphasiology*, *27*(9), 1128–1146. https://doi.org/10.1080/02687038.2013.786803

Marslen-Wilson, W. D., & Welsh, A. (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology*, *10*(1), 29–63. https://doi.org/10.1016/0010-0285(78)90018-X

Masterson, J., & Druks, J. (1998). Description of a set of 164 nounsand 102 verbs matched for printed word frequency, familiarityand age-of-acquisition. *Journal of Neurolinguistics*, *11*(4), 331–354. https://doi.org/10.1016/S0911-6044(98)00023-2

Matérne, M., Strandberg, T., & Lundqvist, L. O. (2019). Risk Markers for Not Returning to Work Among Patients with Acquired Brain Injury: A Population-Based Register Study. *Journal of Occupational Rehabilitation*, *29*(4), 728–739. https://doi.org/10.1007/S10926-019-09833-6/TABLES/2

McClelland, J. L., & Elman, J. L. (1986). The TRACE Model of Speech Perception. *Cognitive Psychology*, *18*, 86.

McHugh, M. L. (2012). Interrater reliability: The kappa statistic. *Biochemia Medica*, *22*(3), 276–282. https://doi.org/10.11613/BM.2012.031/FULLARTICLE

McKinnon, E. T., Fridriksson, J., Basilakos, A., Hickok, G., Hillis, A. E., Spampinato, M. V., Gleichgerrcht, E., Rorden, C., Jensen, J. H., Helpern, J. A., & Bonilha, L. (2018). Types of naming errors in chronic post-stroke aphasia are dissociated by dual stream axonal loss. *Scientific Reports 2018 8:1*, *8*(1), 1–12. https://doi.org/10.1038/s41598-018-32457-4

McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. University of Chicago Press.

McNeill D, & Pedelty LL. (1995). Right brain and gesture. In K. Emmorey & J. Reilly (Eds.), *Language, gesture, and space* (pp. 63–85). Lawrence Erlbaum Associates, Inc.

Meier, E. L., Sheppard, S. M., Goldberg, E. B., Kelly, C. R., Walker, A., Ubellacker, D. M., Vitti, E., Ruch, K., & Hillis, A. E. (2021). Dysfunctional Tissue Correlates of Unrelated Naming Errors in Acute Left Hemisphere Stroke. *Language, Cognition and Neuroscience*, *37*(3), 330. https://doi.org/10.1080/23273798.2021.1980593

Meinzer, M., Streiftau, S., & Rockstroh, B. (2007). Intensive language training in the rehabilitation of chronic aphasia: Efficient training by laypersons. *Journal of the International Neuropsychological Society*, *13*(5), 846–853. https://doi.org/10.1017/S1355617707071111

Melinger, A., & Levelt, W. J. M. (2005). Gesture and the communicative intention of the speaker. *Gesture*, *4*(2), 119–141. https://doi.org/10.1075/GEST.4.2.02MEL/CITE/REFWORKS

Mesulam, M. M., Thompson, C. K., Weintraub, S., & Rogalski, E. J. (2015). The Wernicke conundrum and the anatomy of language comprehension in primary progressive aphasia. *Brain*, *138*(8), 2423–2437. https://doi.org/10.1093/BRAIN/AWV154

Meteyard, L., Stoppard, E., Snudden, D., Cappa, S. F., & Vigliocco, G. (2015). When semantics aids phonology: A processing advantage for iconic word forms in aphasia. *Neuropsychologia*, *76*, 264–275. https://doi.org/10.1016/J.NEUROPSYCHOLOGIA.2015.01.042

Middleton, E. L., Schwartz, M. F., Rawson, K. A., Traut, H., & Verkuilen, J. (2016). Towards a Theory of Learning for Naming Rehabilitation: Retrieval Practice and Spacing Effects. *Journal of Speech, Language, and Hearing Research : JSLHR*, *59*(5), 1111. https://doi.org/10.1044/2016_JSLHR-L-15-0303

Mirman, D., Britt, A. E., & Chen, Q. (2013). Effects of phonological and semantic deficits on facilitative and inhibitory consequences of item

repetition in spoken word comprehension. *Neuropsychologia*, *51*(10), 1848–1856. https://doi.org/10.1016/J.NEUROPSYCHOLOGIA.2013.06.005

Mirman, D., Krason, A., Thothathiri, M., & Middleton, E. L. (2024). Effect of Fatigue on Word Production in Aphasia. *CogSci ... Annual Conference of the Cognitive Science Society. Cognitive Science Society (U.S.). Conference*, *46*, 2951. https://pmc.ncbi.nlm.nih.gov/articles/PMC11310858/

Mirman, D., Yee, E., Blumstein, S. E., & Magnuson, J. S. (2011). Theories of spoken word recognition deficits in Aphasia: Evidence from eye-tracking and computational modeling. *Brain and Language*, *117*(2), 53–68. https://doi.org/10.1016/J.BANDL.2011.01.004

Moineau, S., Dronkers, N. F., & Bates, E. (2005). Exploring the processing continuum of single-word comprehension in aphasia. *Journal of Speech, Language, and Hearing Research : JSLHR*, *48*(4), 884–896. https://doi.org/10.1044/1092-4388(2005/061)

Mol, L., Krahmer, E., & van de Sandt-Koenderman, M. (2013). Gesturing by Speakers With Aphasia: How Does It Compare? *Journal of Speech, Language, and Hearing Research*, *56*(4), 1224–1236. https://doi.org/10.1044/1092-4388(2012/11-0159)

Monnelly, K., Marshall, J., Dipper, L., & Cruice, M. (2023). Intensive and comprehensive aphasia therapy—a survey of the definitions, practices and views of speech and language therapists in the United Kingdom. *International Journal of Language and Communication Disorders*, *58*(6), 2077–2102. https://doi.org/10.1111/1460-6984.12918

Morsella, E., & Krauss, R. M. (2004). The Role of Gestures in Spatial Working Memory and Speech. *The American Journal of Psychology*, *117*(3), 411–424.

Mueller, K. D., Hermann, B., Mecollari, J., & Turkstra, L. S. (2018). *Connected Speech and Language in Mild Cognitive Impairment and Alzheimer's Disease: A Review of Picture Description Tasks*. https://doi.org/10.1080/13803395.2018.1446513

Mumford, K. H., & Kita, S. (2014). Children Use Gesture to Interpret Novel Verb Meanings. *Child Development*, *85*(3), 1181–1189. https://doi.org/10.1111/CDEV.12188

Munro, P., & Siyambalapitiya, S. (2017). Improved word comprehension in Global aphasia using a modified semantic feature analysis treatment. *Clinical Linguistics & Phonetics*, *31*(2), 119–136. https://doi.org/10.1080/02699206.2016.1198927

Murteira, A., & Nickels, L. (2020). Can gesture observation help people with aphasia name actions? *Cortex*, *123*, 86–112. https://doi.org/10.1016/J.CORTEX.2019.10.005

Murteira, A., Sowman, P. F., & Nickels, L. (2019). Taking action in hand: effects of gesture observation on action verb naming. *Language, Cognition and Neuroscience*, *34*(3), 351–364. https://doi.org/10.1080/23273798.2018.1552978

Nelissen, N., Pazzaglia, M., Vandenbulcke, M., Sunaert, S., Fannes, K., Dupont, P., Aglioti, S. M., & Vandenberghe, R. (2010). Gesture Discrimination in Primary Progressive Aphasia: The Intersection between Gesture and Language Processing Pathways. *The Journal of Neuroscience*, *30*(18), 6334. https://doi.org/10.1523/JNEUROSCI.0321-10.2010

Netland, T., von Dzengelevski, O., Tesch, K., & Kwasnitschka, D. (2025). Comparing human-made and AI-generated teaching videos: An experimental study on learning effects. *Computers & Education*, *224*, 105164. https://doi.org/10.1016/J.COMPEDU.2024.105164

Nicholas, M., Obler, L. K., Albert, M. L., & Helm-Estabrooks, N. (1985). Empty speech in Alzheimer's disease and fluent aphasia. *Journal of Speech and Hearing Research*, *28*(3), 405–410. https://doi.org/10.1044/JSHR.2803.405

Novack, M. A., & Goldin-Meadow, S. (2017). Gesture as representational action: A paper about function. *Psychonomic Bulletin and Review*, *24*(3), 652–665. https://doi.org/10.3758/S13423-016-1145-Z/METRICS

Obermeier, C., Dolk, T., & Gunter, T. C. (2012). The benefit of gestures during communication: Evidence from hearing and hearing-impaired individuals. *Cortex*, *48*(7), 857–870. https://doi.org/10.1016/J.CORTEX.2011.02.007

Oppici, L., Mathias, B., Narciss, S., & Proske, A. (2023). Benefits of Enacting and Observing Gestures on Foreign Language Vocabulary Learning: A Systematic Review and Meta-Analysis. *Behavioral Sciences*, *13*(11), 920. https://doi.org/10.3390/BS13110920

Ozer, D., Tansan, M., & Chatterjee, A. (2017, July). The effects of gesture restriction on spatial language in young and elderly adults. *CogSci*. https://www.researchgate.net/publication/320518128

Paivio, A., Yuille, J. C., & Madigan, S. A. (1968). Concreteness, imagery, and meaningfulness values for 925 nouns. *Journal of Experimental Psychology*, *76*(1 PART 2), 1–25. https://doi.org/10.1037/H0025327

Pak-Hin Kong, A., Law, S.-P., Ching-Yin Kwan, C., Lai, C., Lam, V., P-H Kong, A., & P-H Kong Á S-P Law Á C-Y Kwan Á C Lai Á V Lam, A. C. (2015). A Coding System with Independent Annotations of Gesture

Forms and Functions During Verbal Communication: Development of a Database of Speech and GEsture (DoSaGE). *J Nonverbal Behav*, *39*, 93–111. https://doi.org/10.1007/s10919-014-0200-6

Palmer, R., Witts, H., & Chater, T. (2018). What speech and language therapy do community dwelling stroke survivors with aphasia receive in the UK? *PLOS ONE*, *13*(7), e0200096. https://doi.org/10.1371/JOURNAL.PONE.0200096

Parakh, R., Roy, E., Koo, E., & Black, S. (2004). Pantomime and imitation of limb gestures in relation to the severity of Alzheimer's disease. *Brain and Cognition*, *55*, 272–274. https://doi.org/10.1016/j.bandc.2004.02.049

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury Google, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., Desmaison, A., Xamla, A. K., Yang, E., Devito, Z., Raison Nabla, M., Tejani, A., Chilamkurthy, S., Ai, Q., Steiner, B., … Chintala, S. (2019). PyTorch: An Imperative Style, High-Performance Deep Learning Library. *Advances in Neural Information Processing Systems*, *32*.

Pennington, J., Socher, R., & Manning, C. D. (n.d.). *GloVe: Global Vectors for Word Representation*. Retrieved 13 November 2022, from http://nlp.

Perniss, P., Vinson, D., & Vigliocco, G. (2020). Making Sense of the Hands and Mouth: The Role of "Secondary" Cues to Meaning in British Sign Language and English. *Cognitive Science*, *44*(7), e12868. https://doi.org/10.1111/COGS.12868

Pickering, M. J., & Garrod, S. (2006). Alignment as the basis for successful communication. *Research on Language and Computation*, *4*(2–3), 203–228. https://doi.org/10.1007/s11168-006-9004-0

Pierce, J. E., Hill, A. J., Wong, D., Pitt, R., & Rose, M. L. (2025). Adapting a group-based, multimodal aphasia treatment for telehealth - co-design of M-MAT Tele. *Disability and Rehabilitation Assistive Technology*, *20*(1). https://doi.org/10.1080/17483107.2024.2366423

Pierce, J. E., Menahemi-Falkov, M., O'Halloran, R., Togher, L., & Rose, M. L. (2017). Constraint and multimodal approaches to therapy for chronic aphasia: A systematic review and meta-analysis. *Neuropsychological Rehabilitation*, *29*(7), 1005–1041. https://doi.org/10.1080/09602011.2017.1365730

Pine, K. J., Bird, H., & Kirk, E. (2007). The effects of prohibiting gestures on children's lexical retrieval ability. *Developmental Science*, *10*(6), 747–754. https://doi.org/10.1111/J.1467-7687.2007.00610.X

Plummer, P., Altmann, L., Feld, J., Zukowski, L., Najafi, B., & Giuliani, C. (2020). Attentional prioritization in dual-task walking: effects of stroke, environment, and instructed focus. *Gait & Posture*, *79*, 3. https://doi.org/10.1016/J.GAITPOST.2020.03.013

Poliakoff, E., Marciniak, M., Bettis, A., & Humphries, S. (2021). Speaking with your hands in Parkinson's: Investigating the rate and precision of co-speech gestures. *MDS Virtual Congress 2021*, 1122. https://www.mdsabstracts.org/abstract/speaking-with-your-hands-in-parkinsons-investigating-the-rate-and-precision-of-co-speech-gestures/

Preisig, B. C., Eggenberger, N., Cazzoli, D., Nyffeler, T., Gutbrod, K., Annoni, J.-M., Meichtry, J. R., Nef, T., & Müri, R. M. (2018). Multimodal Communication in Aphasia: Perception and Production of Co-speech Gestures During Face-to-Face Conversation. *Frontiers in Human Neuroscience*, *0*, 200. https://doi.org/10.3389/FNHUM.2018.00200

Pritchard, M., Dipper, L., Morgan, G., & Cocks, N. (2015). Language and iconic gesture use in procedural discourse by speakers with aphasia. *Aphasiology*, *29*(7), 826–844. https://doi.org/10.1080/02687038.2014.993912

Pulvermüller, F. (1990). Analyse aphasischer Kommunikation. *Medizinische Und Therapeutische Kommunikation*, 292–308. https://doi.org/10.1007/978-3-663-14675-9_20

Pulvermüller, F., Neininger, B., Elbert, T., Mohr, B., Rockstroh, B., Koebbel, P., & Taub, E. (2001). Constraint-Induced Therapy of Chronic Aphasia After Stroke. *Stroke*, *32*(7), 1621–1626. https://doi.org/10.1161/01.STR.32.7.1621

Pustina, D., Coslett, H. B., Turkeltaub, P. E., Tustison, N., Schwartz, M. F., & Avants, B. (2016). Automated segmentation of chronic stroke lesions using LINDA: Lesion identification with neighborhood data analysis. *Human Brain Mapping*, *37*(4), 1405–1421. https://doi.org/10.1002/HBM.23110

Ralph, M. A. L., Jefferies, E., Patterson, K., & Rogers, T. T. (2016). The neural and computational bases of semantic cognition. *Nature Reviews Neuroscience*, *18*, 42–55. https://doi.org/10.1038/nrn.2016.150

Raskin, S. A., Sliwinski, M., & Borod, J. C. (1992). Clustering strategies on tasks of verbal fluency in Parkinson's disease. *Neuropsychologia*, *30*(1), 95–99. https://doi.org/10.1016/0028-3932(92)90018-H

Rauscher, F. H., Krauss, R. M., & Chen, Y. (1996). Gesture, Speech, and Lexical Access: The Role of Lexical Movements in Speech Production. *Psychological Science*, *7*(4), 226–231.

Raymer, A. M., Maher, L. M., Foundas, A. L., Rothi, L. J. G., & Heilman, K. M. (2000). Analysis of lexical recovery in an individual with acute anomia. *Aphasiology*, *14*(9), 901–910. https://doi.org/10.1080/02687030050127702

Raymer, A. M., McHose, B., Smith, K. G., Iman, L., Ambrose, A., & Casselton, C. (2012). Contrasting effects of errorless naming treatment

and gestural facilitation for word retrieval in aphasia. *Neuropsychological Rehabilitation*, *22*(2), 235–266. https://doi.org/10.1080/09602011.2011.618306

Raymer, A. M., Singletary, F., Rodriguez, A., Ciampitti, M., Heilman, K. M., & Rothi, I. J. G. (2006). Effects of gesture+verbal treatment for noun and verb retrieval in aphasia. *Journal of the International Neuropsychological Society*, *12*(6), 867–882. https://doi.org/10.1017/S1355617706061042

Records, N. L. (1994). A Measure of the Contribution of a Gesture to the Perception of Speech in Listeners With Aphasia. *Journal of Speech and Hearing Research*, *37*, 1086–1099. https://doi.org/10.1044/jshr.3705.1086

Rice, M. L., Smolik, F., Perpich, D., Thompson, T., Rytting, N., & Blossom, M. (2010). Mean length of utterance levels in 6-month intervals for children 3 to 9 years with and without language impairments. *Journal of Speech, Language, and Hearing Research : JSLHR*, *53*(2), 333–349. https://doi.org/10.1044/1092-4388(2009/08-0183)

Ricker, T. J., & Cowan, N. (2010). Loss of Visual Working Memory within Seconds: The Combined Use of Refreshable and Non-Refreshable Features. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *36*(6), 1355. https://doi.org/10.1037/A0020356

Rodd, J. M., Johnsrude, I. S., & Davis, M. H. (2012). Dissociating frontotemporal contributions to semantic ambiguity resolution in spoken sentences. *Cerebral Cortex (New York, N.Y. : 1991)*, *22*(8), 1761–1773. https://doi.org/10.1093/CERCOR/BHR252

Rodriguez, A. D., Raymer, A. M., & Gonzalez Rothi, L. J. (2006). Effects of gesture+verbal and semantic-phonologic treatments for verb retrieval in aphasia. *Aphasiology*, *20*(2–4), 286–297. https://doi.org/10.1080/02687030500474898

Rohrer, P. (2022). *A temporal and pragmatic analysis of gesture-speech association : a corpus-based approach using the novel MultiModal MultiDimensional (M3D) labeling system* [Nantes Université; Universitat Pompeu Fabra]. https://theses.hal.science/tel-03994053v1

Rohrer, P. L., Delais-Roussarie, E., & Prieto, P. (2020). Beat Gestures for Comprehension and Recall: Differential Effects of Language Learners and Native Listeners. *Frontiers in Psychology*, *11*, 575929. https://doi.org/10.3389/FPSYG.2020.575929/BIBTEX

Romani, C., Thomas, L., Olson, A., & Lander, L. (2018). Playing a team game improves word production in poststroke aphasia. *Aphasiology*, *33*(3), 253–288. https://doi.org/10.1080/02687038.2018.1548205

Roper, A., Marshall, J., & Wilson, S. (2016). Benefits and limitations of computer gesture therapy for the rehabilitation of severe aphasia.

Frontiers in Human Neuroscience, *10*(NOV2016), 217753. https://doi.org/10.3389/FNHUM.2016.00595/BIBTEX

Rorden, C., & Brett, M. (2000). Stereotaxic display of brain lesions. *Behavioural Neurology*, *12*(4), 191–200. https://doi.org/10.1155/2000/421719

Rose, M., & Douglas, J. (2001). The differential facilitatory effects of gesture and visualisation processes on object naming in aphasia. *Aphasiology*, *15*(10–11), 977–990. https://doi.org/10.1080/02687040143000339

Rose, M., & Douglas, J. (2003). Limb apraxia, pantomime, and lexical gesture in aphasic speakers: Preliminary findings. *Aphasiology*, *17*(5), 453–464. https://doi.org/10.1080/02687030344000157

Rose, M., Douglas, J., & Matyas, T. (2002). The comparative effectiveness of gesture and verbal treatments for a specific phonologic naming impairment. *Aphasiology*, *16*(10–11), 1001–1030. https://doi.org/10.1080/02687030143000825

Rose, M., Douglas, J., & Matyas, T. (2010). The comparative effectiveness of gesture and verbal treatments for a specific phonologic naming impairment. *Aphasiology*, *16*(10), 1001–1030. https://doi.org/10.1080/02687030143000825

Rose, M. L. (2013). Releasing the Constraints on Aphasia Therapy: The Positive Impact of Gesture and Multimodality Treatments. *American Journal of Speech-Language Pathology*, *22*(2). https://doi.org/10.1044/1058-0360(2012/12-0091)

Rose, M. L., Attard, M. C., Mok, Z., Lanyon, L. E., & Foster, A. M. (2013). Multi-modality aphasia therapy is as efficacious as a constraint-induced aphasia therapy for chronic aphasia: A phase 1 study. *Aphasiology*, *27*(8), 938–971. https://doi.org/10.1080/02687038.2013.810329

Rose, M. L., Mok, Z., Carragher, M., Katthagen, S., & Attard, M. (2016). Comparing multi-modality and constraint-induced treatment for aphasia: a preliminary investigation of generalisation to discourse. *Aphasiology*, *30*(6), 678–698. https://doi.org/10.1080/02687038.2015.1100706

Rose, M. L., Mok, Z., & Sekine, K. (2017). Communicative effectiveness of pantomime gesture in people with aphasia. *International Journal of Language & Communication Disorders*, *52*(2), 227–237. https://doi.org/10.1111/1460-6984.12268

Rose, M. L., Nickels, L., Copland, D., Togher, L., Godecke, E., Meinzer, M., Rai, T., Cadilhac, D. A., Kim, J., Hurley, M., Foster, A., Carragher, M., Wilcox, C., Pierce, J. E., & Steel, G. (2022). Results of the COMPARE trial of Constraint-induced or Multimodality Aphasia Therapy compared with usual care in chronic post-stroke aphasia. *J Neurol Neurosurg Psychiatry*, *0*, 1–9. https://doi.org/10.1136/jnnp-2021-328422

Rose, M., & Sussmilch, G. (2008). The effects of semantic and gesture treatments on verb retrieval and verb use in aphasia. *Aphasiology*, *22*(7–8), 691–706. https://doi.org/10.1080/02687030701800800

Ross, E. D., & Mesulam, M. M. (1979). Dominant Language Functions of the Right Hemisphere?: Prosody and Emotional Gesturing. *Archives of Neurology*, *36*(3), 144–148. https://doi.org/10.1001/ARCHNEUR.1979.00500390062006

Roth, R., Wilmskoetter, J., & Bonilha, L. (2022). The role of disrupted structural connectivity in aphasia. In *Handbook of Clinical Neurology* (Vol. 185, pp. 121–127). Elsevier B.V. https://doi.org/10.1016/B978-0-12-823384-9.00006-2

Rousseaux, M., Rénier, J., Anicet, L., Pasquier, F., & Mackowiak-Cordoliani, M. A. (2012). Gesture comprehension, knowledge and production in Alzheimer's disease. *European Journal of Neurology*, *19*(7), 1037–1044. https://doi.org/10.1111/J.1468-1331.2012.03674.X

Rousseaux, M., Sève, A., Vallet, M., Pasquier, F., & Mackowiak-Cordoliani, M. A. (2010). An analysis of communication in conversation in patients with dementia. *Neuropsychologia*, *48*, 3884–3890. https://doi.org/10.1016/j.neuropsychologia.2010.09.026

Routhier, S., Bier, N., & Macoir, J. (2015). The contrast between cueing and/or observation in therapy for verb retrieval in post-stroke aphasia. *Journal of Communication Disorders*, *54*, 43–55. https://doi.org/10.1016/J.JCOMDIS.2015.01.003

RStudio Team. (2020). *RStudio: Integrated Development for R*. RStudio, PBC.

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), 7. https://doi.org/10.1007/s11910-021-01097-z/Published

Saatci Yurtsever, S., Öztop Çakmak, Ö., Yapıcı Eser, H., Ertan, S., Demir-Lira, E., & Göksun, T. (2021). Production and comprehension of co-speech gestures in Parkinson's disease. *Neuropsychologia*, *163*, 108061. https://doi.org/10.1016/J.NEUROPSYCHOLOGIA.2021.108061

Saffran, E., Schwartz, M., Linnebarger, M., Martin, N., & Bochetto, P. (1989). Philadelphia Comprehension Battery for Aphasia. In *Unpublished*.

Schmid, M. S., & Jarvis, S. (2014). Lexical access and lexical diversity in first language attrition. *Bilingualism: Language and Cognition*, *17*(4), 729–748. https://doi.org/10.1017/S1366728913000771

Schubotz, L., Holler, J., Drijvers, L., & Özyürek, A. (2021). Aging and working memory modulate the ability to benefit from visible speech and iconic

gestures during speech-in-noise comprehension. *Psychological Research*, *85*(5), 1997–2011. https://doi.org/10.1007/S00426-020-01363-8/FIGURES/3

Schwartz, M. F., & Brecher, A. (2000). A Model-Driven Analysis of Severity, Response Characteristics, and Partial Recovery in Aphasics' Picture Naming. *Brain and Language*, *73*(1), 62–91. https://doi.org/10.1006/BRLN.2000.2310

Schwarz, J., Li, K. K., Sim, J. H., Zhang, Y., Buchanan-Worster, E., Post, B., Gibson, J. L., & McDougall, K. (2022). Semantic Cues Modulate Children's and Adults' Processing of Audio-Visual Face Mask Speech. *Frontiers in Psychology*, *13*, 879156. https://doi.org/10.3389/FPSYG.2022.879156/BIBTEX

Seghier, M. L., Patel, E., Prejawa, S., Ramsden, S., Selmer, A., Lim, L., Browne, R., Rae, J., Haigh, Z., Ezekiel, D., Hope, T. M. H., Leff, A. P., & Price, C. J. (2016). The PLORAS Database: A data repository for Predicting Language Outcome and Recovery After Stroke. *NeuroImage*, *124*(Pt B), 1208–1212. https://doi.org/10.1016/J.NEUROIMAGE.2015.03.083

Seghier, M. L., Ramlackhansingh, A., Crinion, J., Leff, A. P., & Price, C. J. (2008). Lesion identification using unified segmentation-normalisation models and fuzzy clustering. *NeuroImage*, *41*(4), 1253–1266. https://doi.org/10.1016/J.NEUROIMAGE.2008.03.028

Sekine, K., & Rose, M. L. (2013). The relationship of aphasia type and gesture production in people with aphasia. *American Journal of Speech-Language Pathology*, *22*(4), 662–672. https://doi.org/10.1044/1058-0360(2013/12-0030)

Sekine, K., Rose, M. L., Foster, A. M., Attard, M. C., & Lanyon, L. E. (2013). Gesture production patterns in aphasic discourse: In-depth description and preliminary predictions. *Aphasiology*, *27*(9), 1031–1049. https://doi.org/10.1080/02687038.2013.803017

Shrubsole, K., Worrall, L., Power, E., & O'Connor, D. A. (2018). Priorities for Closing the Evidence-Practice Gaps in Poststroke Aphasia Rehabilitation: A Scoping Review. *Archives of Physical Medicine and Rehabilitation*, *99*(7), 1413-1423.e24. https://doi.org/10.1016/J.APMR.2017.08.474

Silkes, J. A. P., Zimmerman, R. M., Greenspan, W., Reinert, L., Kendall, D., & Martin, N. (2021). Identifying verbal short-term memory and working memory impairments in individuals with latent aphasia. *American Journal of Speech-Language Pathology*, *30*(1s), 391–406. https://doi.org/10.1044/2020_AJSLP-19-00105

Simmons-Mackie, N. (2009). Thinking Beyond Language: Intervention for Severe Aphasia. *Perspectives on Neurophysiology and Neurogenic*

Speech and Language Disorders, *19*(1), 15–22. https://doi.org/10.1044/NNSLD19.1.15

Smania, N., Gandolfi, M., Aglioti, S. M., Girardi, P., Fiaschi, A., & Girardi, F. (2010). How long is the recovery of global aphasia? twenty-five years of follow-up in a patient with left hemisphere stroke. *Neurorehabilitation and Neural Repair*, *24*(9), 871–875. https://doi.org/10.1177/1545968310368962

Stasenko, A., Bonn, C., Teghipco, A., Garcea, F. E., Sweet, C., Dombovy, M., McDonough, J., & Mahon, B. Z. (2015). 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

Straube, B., Green, A., Bromberger, B., & Kircher, T. (2011). The differentiation of iconic and metaphoric gestures: Common and unique integration processes. *Human Brain Mapping*, *32*(4), 520–533. https://doi.org/10.1002/HBM.21041

Strong, K. A. (2023). Supporting Mental Health in People Living With Aphasia: A Call to Action. *Perspectives of the ASHA Special Interest Groups*, *8*(6), 1321–1334. https://doi.org/10.1044/2023_PERSP-23-00042

Sueyoshi, A., & Hardison, D. M. (2005). The role of gestures and facial cues in second language listening comprehension. *Language Learning*, *55*(4), 661–699. https://doi.org/10.1111/J.0023-8333.2005.00320.X

Tarhan, L. Y., Watson, C. E., & Buxbaum, L. J. (2015). Shared and Distinct Neuroanatomic Regions Critical for Tool-related Action Production and Recognition: Evidence from 131 Left-hemisphere Stroke Patients. *Journal of Cognitive Neuroscience*, *27*(12), 2491–2511. https://doi.org/10.1162/JOCN_A_00876

Thompson, L. A., & Guzman, F. A. (1999). Some Limits on Encoding Visible Speech and Gestures Using a Dichotic Shadowing Task. *The Journals of Gerontology: Series B*, *54B*(6), P347–P349. https://doi.org/10.1093/GERONB/54B.6.P347

Thye, M., & Mirman, D. (2018). Relative contributions of lesion location and lesion size to predictions of varied language deficits in post-stroke aphasia. *NeuroImage: Clinical*, *20*, 1129–1138. https://doi.org/10.1016/J.NICL.2018.10.017

Tong, Y. (2023). *Embodiment of Concrete and Abstract Concepts: The Role of Gesture* [Vrije Universiteit Amsterdam]. https://research.vu.nl/ws/portalfiles/portal/216318315/yaotongphd2023+-+63f5cd2307557.pdf

Trebuchon, A., Lambert, I., Guisiano, B., McGonigal, A., Perot, C., Bonini, F., Carron, R., Liegeois-Chauvel, C., Chauvel, P., & Bartolomei, F. (2018). The different patterns of seizure-induced aphasia in temporal lobe epilepsies. *Epilepsy & Behavior*, *78*, 256–264. https://doi.org/10.1016/J.YEBEH.2017.08.022

Turken, A. U., & Dronkers, N. F. (2011). The neural architecture of the language comprehension network: Converging evidence from lesion and connectivity analyses. *Frontiers in Systems Neuroscience*, *5*(FEBRUARY 2011), 8794. https://doi.org/10.3389/FNSYS.2011.00001/BIBTEX

Upton, E., Doogan, C., Fleming, V., Leyton, P. Q., Barbera, D., Zeidman, P., Hope, T., Latham, W., Coley-Fisher, H., Price, C., Crinion, J., & Leff, A. (2024). Efficacy of a gamified digital therapy for speech production in people with chronic aphasia (iTalkBetter): behavioural and imaging outcomes of a phase II item-randomised clinical trial. *EClinicalMedicine*, *70*, 102483. https://doi.org/10.1016/J.ECLINM.2024.102483/ATTACHMENT/6EDF8DBF-B22E-4E5A-94A9-F3A1E3DBCA41/MMC2.DOC

van Heuven, W. J. B., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). SUBTLEX-UK: A new and improved word frequency database for British English. *Quarterly Journal of Experimental Psychology*, *67*(6), 1176–1190. https://doi.org/10.1080/17470218.2013.850521/ASSET/IMAGES/LARGE/10.1080_17470218.2013.850521-FIG1.JPEG

van Nispen, K., Mieke, W. M. E., van de Sandt-Koenderman, E., & Krahmer, E. (2018). The comprehensibility of pantomimes produced by people with aphasia. *International Journal of Language and Communication Disorders*, *53*(1), 85–100. https://doi.org/10.1111/1460-6984.12328

Van Nispen, K., Van De Sandt-Koenderman, M., Sekine, K., Krahmer, E., & Rose, M. L. (2017). Part of the message comes in gesture: how people with aphasia convey information in different gesture types as compared with information in their speech. *Aphasiology*, *31*(9), 1078–1103. https://doi.org/10.1080/02687038.2017.1301368

Varkanitsa, M., & Kiran, S. (2022). Understanding, facilitating and predicting aphasia recovery after rehabilitation. *International Journal of Speech-Language Pathology*, *24*(3), 248–259. https://doi.org/10.1080/17549507.2022.2075036

Verbaan, D., Marinus, J., Visser, M., van Rooden, S. M., Stiggelbout, A. M., M Middelkoop, H. A., & van Hilten, J. J. (2007). Cognitive impairment in Parkinson's disease. *J Neurol Neurosurg Psychiatry*, *78*, 1182–1187. https://doi.org/10.1136/jnnp.2006.112367

VideoLan. (2006). *VLC media player*.

Vigliocco, G., Krason, A., Stoll, H., Monti, A., & Buxbaum, L. J. (2020). Multimodal comprehension in left hemisphere stroke patients. *Cortex*, *133*, 309–327. https://doi.org/10.1016/j.cortex.2020.09.025

Vogt, S., & Kauschke, C. (2017). Observing iconic gestures enhances word learning in typically developing children and children with specific language impairment. *Journal of Child Language*, *44*(6), 1458–1484. https://doi.org/10.1017/S0305000916000647

Volkmer, A., Mistry, S., Thompson, D., Warren, J. D., & Beeke, S. (2023). Collaborative turn construction practices of people with Primary Progressive Aphasia and their family conversation partners. *Journal of Interactional Research in Communication Disorders, 14*(3), 456–485.

Völsch, J., Wartenburger, I., & Stadie, N. (2021). Do gestures have a hand in verb retrieval? Investigation of iconic and non-iconic gestures in aphasia. *APHASIOLOGY*, *2022*(7), 781–804. https://doi.org/10.1080/02687038.2021.1907293

Wang, L., & Goodglass, H. (1992). Pantomime, Praxis, and Aphasia. *Brain and Language*, *42*, 402–418.

Wang, X. L., Bernas, R., & Eberhard, P. (2004). Engaging ADHD students in tasks with hand gestures: a pedagogical possibility for teachers. *Educational Studies*, *30*(3), 217–229. https://doi.org/10.1080/0305569042000224189

Webster, J., & Morris, J. (2019). Communicative Informativeness in Aphasia: Investigating the Relationship Between Linguistic and Perceptual Measures. *American Journal of Speech-Language Pathology*, *28*(3), 1115. https://doi.org/10.1044/2019_AJSLP-18-0256

Wechsler, D. (1997). *Wechsler Adult Intelligence Scale-Third Edition*. APA PsycTests.

Winkler, M., Bedford, V., Northcott, S., & Hilari, K. (2014). Aphasia blog talk: How does stroke and aphasia affect the carer and their relationship with the person with aphasia? *Aphasiology*, *28*(11), 1301–1319. https://doi.org/10.1080/02687038.2014.928665

Yap, D., & Casasanto, D. (2018). Beat gestures encode spatial semantics. *40th Annual Conference of the Cognitive Science Society*.

Yliranta, A., & Jehkonen, M. (2020). Limb and face apraxias in frontotemporal dementia: A systematic scoping review. *Cortex*, *129*, 529–547. https://doi.org/10.1016/J.CORTEX.2020.03.023

Yuvaraj, R., Murugappan, M., Norlinah, M. I., Sundaraj, K., & Khairiyah, M. (2013). Review of emotion recognition in stroke patients. In *Dementia and Geriatric Cognitive Disorders* (Vol. 36, Issues 3–4, pp. 179–196). https://doi.org/10.1159/000353440

Zhang, W., & Luck, S. J. (2009). Sudden Death and Gradual Decay in Visual Working Memory. *Psychological Sciences*, *20*(4), 423–428. https://doi.org/10.1111/J.1467-9280.2009.02322.X

Zhang, Y., Ding, R., Frassinelli, D., Tuomainen, J., Klavinskis-Whiting, S., & Vigliocco, G. (2023). The role of multimodal cues in second language comprehension. *Scientific Reports 2023 13:1*, *13*(1), 1–11. https://doi.org/10.1038/s41598-023-47643-2

Zhang, Y., Frassinelli, D., Tuomainen, J., Skipper, J. I., & Vigliocco, G. (2021). More than words: word predictability, prosody, gesture and mouth movemets in natural language comprehension. *Proc. R. Soc. B*, *288*, 20210500.

Appendix 1: PWA Background Battery Assessment Scores

See overleaf – please note that the following pages are in Landscape format.

ID	Aphasia sub-type	WAB-R AQ	CCT	KDT	STT	Naming Battery	STM	WM/EF	TT	ST	Gesture Recog.	Gesture to Sight	AoS
2	Anomic	97.0	59	51	26	317	7	5	34	54	24	NA	Absent
3	Conduction	82.0	61	51	25	303	3	3	27	44	25	28.66	Mild
4	Conduction	81.5	54	40	22	248	5	4	22	51	22	30	Mild
5	Conduction	80.3	60	49	25	290	10	6	28	52	23	27.66	Absent
6	Anomic	84.3	50	49	20	137	8	2	22	51	26	35.66	Mild
8	Conduction	57.0	54	45	15	63	0	1	19	25	24	31.66	Mild
9	Anomic	95.8	52	48	29	341	11	4	32	47	24	30	Absent
10	Anomic	84.2	53	50	28	297	8	3	33	50	25	30	Absent
11	Anomic	62.1	52	46	21	204	4	3	21	41	23	25	Absent
12	Anomic	94.0	54	48	27	333	8	4	33	46	24	31.3	Absent
13	Anomic	78.6	50	46	22	307	6	3	24	53	23	34.66	Mild
14	Conduction	69.0	57	51	22	217	2	2	13	52	26	NA	Absent
17	Anomic	85.7	43	49	20	268	2	3	27	51	21	NA	Absent
18	Anomic	95.2	54	49	28	329	15	8	32	53	25	36.66	Absent
19	Anomic	88.9	51	48	25	252	14	2	29	53	26	34.66	Mild
20	Conduction	59.0	34	41	16	75	5	2	16	36	19	21.84	Absent
22	Anomic	73.1	47	44	16	265	4	2	24	47	23	30.66	Absent
25	Broca’s	51.5	57	44	23	184	6	5	20	40	25	27.651	Mild
27	Anomic	86.3	54	50	26	227	6	2	26	52	23	34.66	Absent
28	Broca’s	68.9	41	45	22	203	2	2	15	47	22	18	Absent
29	Conduction	78.4	61	51	25	249	5	4	23	54	27	28.665	Absent
30	Anomic	91.2	54	49	23	300	6	3	30	48	21	35.66	Absent
32	Conduction	77.5	56	48	29	204	4	3	19	45	23	33.33	Mod
33	Anomic	93.4	42	24	21	315	8	4	29	53	23	34.66	Absent
34	Anomic	80.0	55	41	16	323	5	3	30	52	26	31	Mild
S01	Conduction	79.2	61	51	30	337	1	1	14	48	25	37	Absent
S02	Mixed	87.6	55	50	29	340	8	11	24	43	23	36	Absent

S03	Anomic	82.9	58	51	29	321	10	5	34	52	26	30	Absent
S04	Broca's	51.6	48	44	19	84	0	1	4	48	22	24	Absent
S05	Anomic	94.6	49	47	30	318	9	10	35	53	NA	NA	Absent
S06	Wernicke's	51.8	50	45	23	229	2	0	7	16	19	30	Absent
S07	Anomic	81.3	52	47	26	315	8	2	29	50	26	32	Absent
S08	Anomic	83.0	46	43	25	316	7	1	21	48	20	33	Absent
S09	Conduction	78.1	50	52	30	251	4	3	27	46	26	37	Mod
S10	Wernicke's	77.1	55	50	30	275	2	0	11	51	23	27	Mild
S11	Anomic	92.2	45	48	21	281	8	3	27	51	22	27	Mild
S12	Anomic	88.7	43	46	20	277	7	2	20	50	26	35	Absent

WAB-R AQ = Western Aphasia Battery-Revised Aphasia Quotient; CCT = Camel and Cactus Test; KDT = Kissing and Dancing Test; STT = Synonymy Triplets Test; STM = short-term memory, WAIS III Forward Digit Span; WM/EF = working memory/executive function, WAIS III Backward Digit Span; TT = Token Test; ST = Scenario Test; Gesture Recog. = MRRI Apraxia Battery Gesture Recognition Test; Gesture to Sight = MRRI Apraxia Battery Gesture to Sight (praxis) Test; AoS = Apraxia of Speech severity.

ID	Picture naming	Semantics	Orientation	Visual field	Sentence reading	Number writing	Calculation	Attention	Praxis	Memory (recognition)	Memory (episodic)	Task switching*
2	4	3	4	4	15	3	4	50	12	4	4	-1
3	4	3	4	4	14	3	4	50	12	4	4	3
4	3	3	4	4	12	3	4	50	7	3	4	-1
5	4	3	4	4	13	2	4	45	12	4	4	0
6	4	3	4	4	15	2	4	45	12	4	4	-1
8	1	3	4	4	8	1	4	50	12	3	3	7
9	4	3	4	4	15	3	4	44	8	3	4	-1

10	3	3	4	4	11	3	4	46	11	1	4	-1
11	1	3	3	4	11	1	1	35	7	3	3	0
12	4	3	4	4	13	3	4	47	11	4	4	-1
13	3	3	4	4	14	3	4	40	6	4	3	4
14	1	3	4	4	14	0	2	45	5	1	4	-1
17	2	3	4	4	15	3	4	50	12	3	4	-1
18	4	3	4	4	15	3	4	39	12	3	4	-1
19	4	3	4	4	14	3	4	47	12	3	4	-1
20	1	3	4	4	10	2	2	47	11	3	4	1
22	4	3	4	4	15	1	3	47	12	4	4	3
25	1	3	4	4	2	0	3	20	12	3	4	-1
27	4	3	4	4	15	1	4	50	12	3	4	2
28	2	3	4	4	13	1	4	43	10	0	4	3
29	1	3	4	4	14	1	3	50	12	3	4	-1
30	4	3	4	4	14	3	4	48	12	3	4	4
32	1	3	3	4	12	2	2	48	10	4	4	-1
33	3	3	4	4	12	3	3	49	12	4	4	-1
34	2	3	3	4	10	2	3	46	9	4	4	1

*For the task-switching test, a score of 4 or above is considered evidence of impairment. Negative scores are desirable. MRRI participants did not complete the Oxford Cognitive Screen.

Appendix 2: Chapter 2 "Go Fish!" Gesture Norming Scores

TARGET	MEAN RATING	WORD TYPE
anchor	2.2	object
angel	6.1	object
arm	6.7	object
arrow	5.8	object
axe	3.9	object
baby	6.9	object
bag	4.6	object
ball	3.4	object
balloon	2	object
banana	6.8	object
barking	4.3	action
basket	3	object
bath	2.2	object
bear	5.4	object
beard	6.9	object
bed	5.7	object
bedroom	3.6	object
bee	2.2	object
begging	3.6	action
bell	4.3	object
belt	2.4	object
bending	2.3	action
bird	6.3	object
biting	4.7	action
bleeding	2	action
blowing	2.8	action
bomb	4	object
bone	1.5	object
book	6.6	object
bottle	4.1	object
bouncing	5.9	action
box	6.2	object
boy	1.5	object
brain	4.6	object
bridge	4	object
brush	6	object
bucket	2.4	object

building	3.5	action
bus	4.4	object
butterfly	6.3	object
button	3.3	object
camel	1.7	object
camera	6.8	object
candle	1.8	object
car	5.6	object
carrying	3.2	action
castle	1.7	object
cat	3.7	object
catching	6	action
chain	2.4	object
chair	2.1	object
cheese	2	object
cherries	1.1	object
chest	5	object
church	5.3	object
cigar	5.2	object
cigarette	6.8	object
circle	4.9	object
circus	3.6	object
climbing	6.6	action
clock	2.8	object
clown	4.2	object
collar	5	object
comb	6.3	object
combing	6	action
computer	4	object
conductor	6	object
cooking	4.4	action
cork	5.5	object
cow	4	object
crack	2.7	object
crawling	5.1	action
cross	6.1	object
crossing	2.3	action
crown	5.4	object
crying	6.8	action
cup	5.4	object
curtains	6.2	object

cutting	6.3	action
dancing	6.4	action
desk	3.3	object
devil	3.6	object
digging	6.4	action
diving	5.8	action
doctor	3	object
dog	2.5	object
door	6.4	object
drawing	4.4	action
dreaming	3.5	action
dress	2.8	object
drilling	2.4	action
drinking	6.9	action
dripping	3.4	action
driving	6.8	action
dropping	6.8	action
drum	4.7	object
duck	2.1	object
ear	6.5	object
eating	6.4	action
elephant	6	object
envelope	2.8	object
eye	6.3	object
fan	3.6	object
feather	2	object
fence	2.4	object
finger	6.3	object
fire	3.1	object
fish	6.5	object
fishing	5.8	action
flag	5	object
floating	4.6	action
flower	3.5	object
fly	3.2	object
flying	6.6	action
folding	4.9	action
foot	3.4	object
football	1.1	object
fork	3.9	object
frog	3.7	object

fruit	3	object
garden	2.2	object
gate	4.1	object
ghost	2.8	object
girl	1.5	object
glasses	5.9	object
grapes	1.4	object
guitar	6.9	object
hair	6	object
hammock	3.8	object
hand	5.3	object
hat	6.3	object
heart	6.5	object
hook	2.5	object
horse	3.5	object
hospital	4.8	object
house	5	object
iron	5	object
ironing	4.9	action
jack	1.4	object
jacket	6	object
judge	2	object
juggling	6	action
jumping	4	action
kettle	3.1	object
key	6.3	object
kicking	2.2	action
king	4	object
kissing	5.3	action
kitchen	2.2	object
kneeling	2.4	action
knife	3.6	object
knitting	5.8	action
knocking	7	action
knot	4.7	object
ladder	5.9	object
laughing	1.6	action
leaf	1.8	object
leaning	2.4	action
leg	4.1	object
letter	5.5	object

library	5.4	object
licking	2	action
lighting	3	action
lion	5.1	object
lips	6.4	object
lock	6.7	object
man	4	object
map	3.4	object
marching	5.4	action
match	3.4	object
melting	2.9	action
money	6.6	object
monkey	2.5	object
moon	3.7	object
mouse	2.7	object
mushroom	1.4	object
nail	3.7	object
neck	6.8	object
necklace	5.6	object
needle	6.1	object
nest	1.6	object
net	2.2	object
newspaper	5.5	object
nose	6.8	object
nun	3.5	object
nurse	3.4	object
office	3.1	object
opening	6.7	action
orange	1.8	object
painting	6.5	action
pants	5.3	object
parcel	2.6	object
pear	1.8	object
peeling	5.5	action
pen	5.1	object
pencil	4	object
penny	1.1	object
piano	6.8	object
picnic	2.2	object
picture	4.1	object
pie	2.2	object

pig	2.9	object
pillow	3.9	object
pilot	2.7	object
pinching	6.7	action
pipe	3.2	object
pizza	2.1	object
plant	2.2	object
planting	4.7	action
plate	3	object
playing	2.4	action
plug	4.1	object
pocket	4.2	object
pointing	6.5	action
pond	2.6	object
posting	5.4	action
potato	1.3	object
pouring	6.9	action
pram	2.6	object
praying	6.8	action
pulling	4.3	action
pushing	6.8	action
pyramid	3.9	object
queen	4.6	object
rabbit	5.8	object
radio	3.7	object
rain	6.3	object
raining	6.9	action
rake	2.3	object
reading	6.8	action
record	1.7	object
restaurant	3.4	object
rice	1.1	object
riding	3.3	action
ring	6	object
ringing	4	action
river	3.2	object
road	3	object
roaring	2.5	action
robot	5.8	object
rocket	4.2	object
rocking	2.9	action

roof	5.2	object
roots	2.4	object
rope	2.3	object
running	6.8	action
saddle	1.3	object
sailing	2.5	action
sailor	1.3	object
salad	1.9	object
sandwich	3.2	object
sausages	3.1	object
saw	4.9	object
scissors	7	object
screw	3.8	object
seal	4.3	object
sewing	6.7	action
shadow	2.8	object
shark	2.3	object
shaving	5.6	action
sheep	2.9	object
shirt	5.5	object
shoe	4.1	object
shooting	6.9	action
shower	5.5	object
singing	4	action
sink	3.3	object
sinking	3.2	action
sitting	3.5	action
skating	2.7	action
skiing	3	action
skipping	6.2	action
sleeping	7	action
slide	4.4	object
sliding	3.9	action
smiling	2.5	action
smoke	3.4	object
smoking	6.3	action
snake	4.5	object
sneezing	3.5	action
snowing	4.2	action
soup	4.6	object
spider	7	object

spoon	3.7	object
square	6	object
stamp	2.5	object
star	3.4	object
steak	1.7	object
stirring	6.4	action
stomach	5.5	object
stool	1.8	object
stopping	6.6	action
strawberry	1.7	object
stroking	6.8	action
submarine	1.9	object
sugar	4.8	object
suit	2.7	object
sun	3.8	object
sweeping	4.4	action
swimming	7	action
swinging	3.9	action
sword	3.7	object
table	4.5	object
tank	1.4	object
tear	6.1	object
tennis	5.1	object
tent	2.1	object
thumb	6.8	object
ticket	3.6	object
tickling	1.5	action
tie	5.6	object
tiger	5.6	object
toe	1.9	object
toilet	2.9	object
tongue	3.1	object
tooth	3.6	object
touching	4.8	action
tourist	3.7	object
tractor	3.1	object
train	1.8	object
tray	3.7	object
tree	2	object
triangle	6.6	object
truck	2.8	object

trumpet	5.3	object
trunk	6	object
tunnel	3.3	object
tying	5.2	action
typing	5.3	action
tyre	2.4	object
umbrella	2.9	object
waiter	2.4	object
waitress	2.8	object
walking	6.5	action
washing	6.8	action
watch	6.3	object
watching	4.2	action
watering	3.7	action
wave	3.6	object
waving	6.9	action
weaving	3.4	action
weighing	1.4	action
weight	1.8	object
wheel	4	object
whip	4.4	object
whistle	6.2	object
window	4.8	object
wine	3.6	object
wing	4.4	object
witch	3.2	object
wolf	3.1	object
worm	2.8	object
wrist	5.9	object
writing	6.2	action
yawning	6.8	action

Appendix 3: PWA Gesture and Speech Transcription Coding Manual

Background

This protocol has been put together to assist you in coding the speech and gesture output of aphasic participants in our study. For full details of the study (including details of participant demographics, background test battery and experimental tasks), refer to Isobel's study design document.

Version control

This is Version 1.0 of this document, produced December 2022.

Participants

Participants in this study are all adults (i.e. age 18+) with aphasia, a communication disorder that commonly occurs follow a stroke and which affects the ability to understand and use language. Stroke also often causes hemiplegia (single sided paralysis) and hemiparesis (single sided muscle weakness). Because of this, the speech output and gestures produced by our participants may be atypical when compared to the output produced by neurologically healthy adults. For example, our participants may have long pauses between words due to fluency disruptions, or may produce many gestures with the non-dominant hand. When coding, if you are uncertain about aspects of the participants' output, please talk to Isobel or Gabriella.

Experimental tasks

Participants engaged in two experimental tasks based around interactive card games: a naming to description task (referred to as 'Heads Up') and a picture naming task ('Go Fish'). In each task, the participant engages directly with the researcher. You do not need to code the researcher's gesture output, but we may ask you to code the researcher's speech output. If this is the case, Isobel will let you know.

Pre-coding tasks

Before starting to code the data, please make sure that you:

a. Are familiar with the General Data Protection Regulation (GDPR) (refer to www.ucl.ac.uk/data-protection). Please pay special attention to data protection and individual privacy.

b. Complete GDPR training.

c. Read up on language as a multimodal phenomenon. The following papers may serve as a starting point for your reading.

a. In neurologically healthy children and adults:

i. Vigliocco, G., Perniss, P., & Vinson, D. (2014). Language as a multimodal phenomenon: implications for language learning, processing and evolution. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1651), 20130292.

ii. Levinson, S. C., & Holler, J. (2014). The origin of human multi-modal communication. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1651), 20130302.

b. In adults with aphasia:

i. Vigliocco, G., Krason, A., Stoll, H., Monti, A., & Buxbaum, L. J. (2020). Multimodal comprehension in left hemisphere stroke patients. *Cortex*, *133*, 309-327.

ii. Cocks, N., Byrne, S., Pritchard, M., Morgan, G., & Dipper, L. (2018). Integration of speech and gesture in aphasia. *International Journal of Language & Communication Disorders*, *53*(3), 584-591.

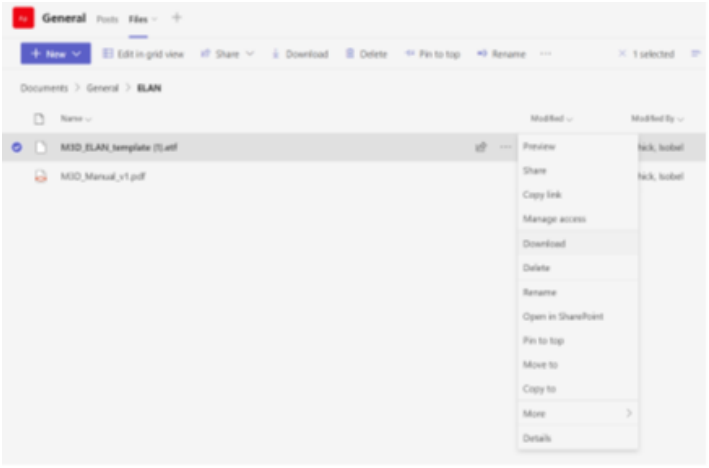
Coding template

Coding for this project utilises the M3D ELAN template. This template can a little overwhelming the first time you use it, because it contains many (many!) tiers. Note that you do not need to use all tiers.

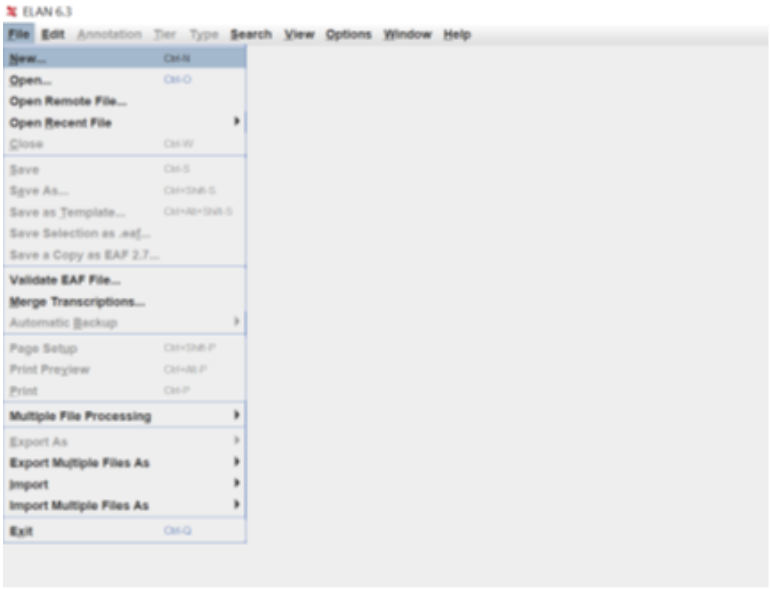
Note that before loading the template, you will need to download the latest version of ELAN. You can download ELAN here: https://archive.mpi.nl/tla/elan/download. Please ensure you select the correct version for your computer (i.e. Windows, Mac, Linux).

To load the M3D template, follow the steps below:

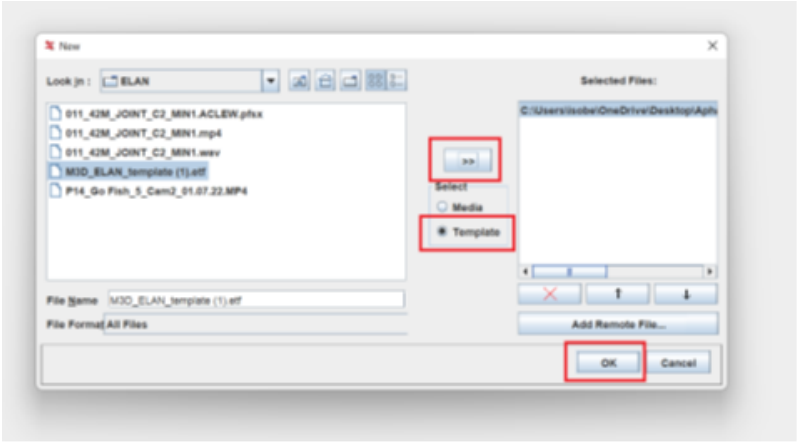
1. Download the template, which is stored on the *Aphasia Project* Teams channel, and save it to your computer.



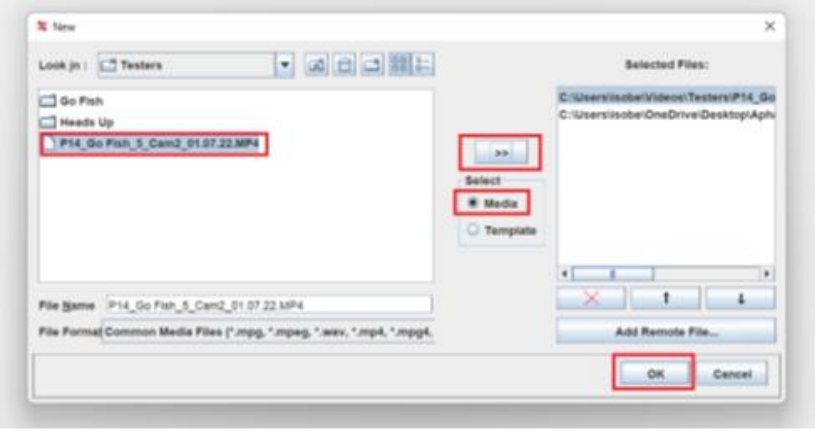
2. Open a new file in ELAN.



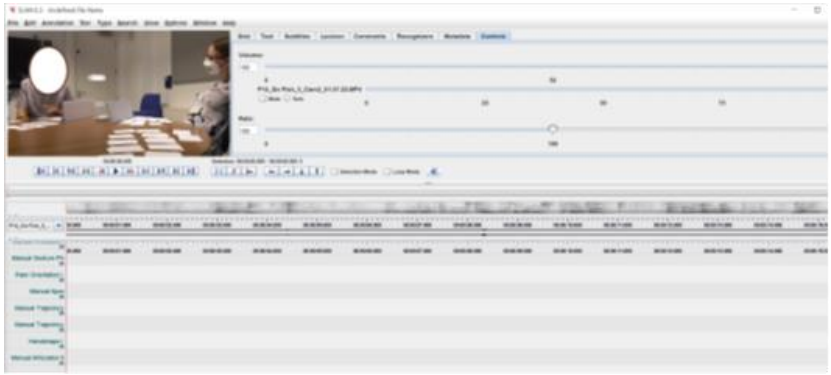
3. Open the template. Note that you need to select the ‘Template’ option for the etf (Template) file type to appear.



4. You also need to select and open the relevant video file.



5. This will produce a screen like this. You can now start coding.



Coding training video

Dr Ed Donnellan (ECOLANG Project Manager) has put together a short training video to introduce new coders to the basics of using ELAN. If you would like access to this video, please ask Isobel.

Data management

UCL Data Safe Haven

All project data (video files, assessment files, audio files) is securely stored on UCL's Data Safe Haven. Note that it is not possible for you to access this data directly.

OneDrive

Any video and audio files that you need to work on will be transferred by Isobel to a protected folder on OneDrive to which you will be given access. You should on no account allow anyone else to access these files. The files contain highly sensitive personal data from vulnerable adults and unauthorized personnel are not permitted to access them.

In addition, you must not:

- Copy data from these files to a different folder on OneDrive.
- Transfer data from these files to a different folder on OneDrive.
- Store any data from these files on your personal computer after you have finished working on them.
- Store any data from these files on any other storage device, e.g. external hard drive, USB drive etc. after you have finished working on them.

Please note that due to slow upload/download speeds on the UCL network, it can take several hours for files to be transferred from the Data Safe Haven to OneDrive (and vice versa). Isobel will let you know when your files have been successfully transferred.

Typographical conventions

Please note that it is important to follow spelling and typographical conventions exactly as they are laid out in this guide to ensure consistency. This is particularly important because we are collecting data from two different sites (UCL and Moss Rehabilitation Research Institute) and we need to ensure consistency across both sites for ease of analysis.

Please avoid adding white space (with spacebar or tab) at the beginning or end of annotations and avoid using accented characters. It is very unlikely that you will need to code speech from a language other than English, but if

you do, transcribe using the closest English approximation (e.g. instead of “frère Jacques”, transcribe “frere Jacques”). Note: if you are working from a keyboard that has shortcuts for accents and diacritics set up, please be careful to not include these in the coding.

File naming conventions

Working files

Files, even if they are work in progress, need to be named following the conventions outlined below to avoid confusion and data loss.

Filenames must start with *P* followed by the participant’s ID number, e.g. *P15*.

The participant identification is followed by the data processing step, e.g. transcription or blurring and the words “in progress” followed by the date on which the file was last edited in the format dd.mm.yy. All parts of the filename should be separated with underscores with no blank spaces.

For example: *P15_blurring_inprogress_15.11.22*

Completed files

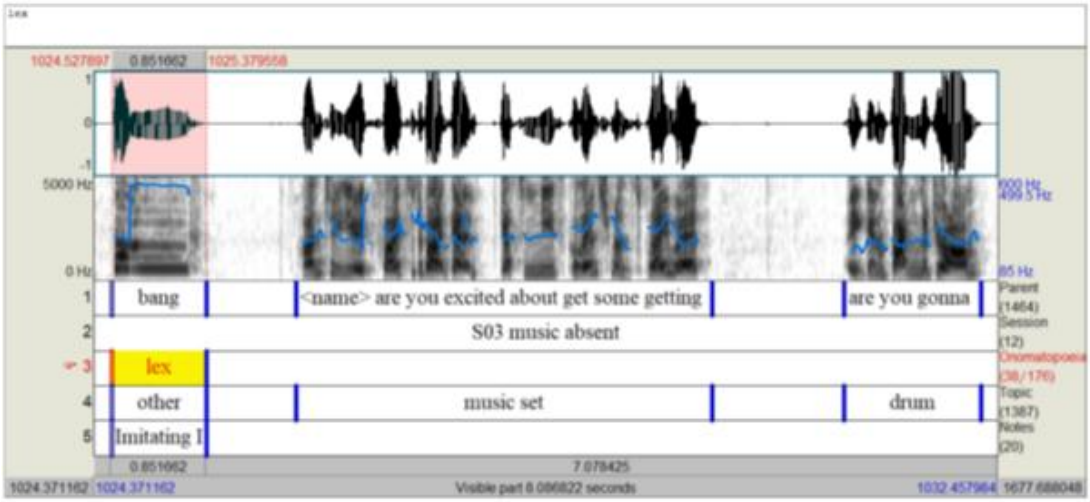
Once the data processing step has been completed, the file should be named as follows:

P15_blurring_complete

Note that the file should only be named in this manner if it is final and ready for further processing. There is no need to specify the date for completed files.

Coding speech output

Speech is transcribed manually in Praat. You can download the most recent version of Praat here: https://www.fon.hum.uva.nl/praat/ Note that we are unable to use automatic speech transcription software, e.g. Google Speech to Text, to transcribe aphasic speech as the results are not accurate enough to analyse. In addition, due to data protection regulations you must not attempt to upload any speech data from our participants to any automatic speech transcription software.



Tier names

The following tier names should be used:

- Researcher
- PWA
- Onomatopoeia
- Session
- Topic
- Comments

Researcher and PWA

In the researcher and PWA tiers, all speech produced by (1) the researcher and (2) the person with aphasia is transcribed and segmented into utterances. We define an utterance as: a unit that expresses a single situation (activity, event or state) with a predicate that can be implicit or explicit.

The sequence “The spoon has fallen off the table, you should pick it up” consists of two utterances (expressing two separate situations/events). However, “Go and pick it up” expresses a single situation and should be coded as one utterance.

Session

Give the session name in the following format:

Participant ID Game name Gesture condition

For example: P14 Heads Up Gesture permitted

Onomatopoeia

Onomatopoeia should be coded using the codes below. Time boundaries for onomatopoeic annotations should match exactly boundaries for the corresponding utterance. Make a description of the sound effect onomatopoeia in the Comments tier, e.g., imitating the sound of cymbals. Use the following codes

Code	Reference
lex	Lexicalised onomatopoeia - i.e. those that have a conventionalised form in English, where the form represents an acoustic phenomenon e.g. “Chop chop chop”, “bang”, “oink”. Note: Lexicalised onomatopoeia should be coded whether they are

	used as a sound (e.g. “Chop, chop, chop!”) or in their predicate form (e.g. The pig says, “Oink oink!”)
se	Sound effects - noises made with the mouth that represent a sound but do not have a conventional lexical form, e.g. making the noise of eating something or snoring to imitate sleep.
interject	Interjections - noises made with the mouth used to express an emotion or signal discourse breaks.

Topic

Enter the target word for the trial. Refer to the trial spreadsheets if you are unsure of the target word.

Notes

Here you can keep notes of anything you would like to keep track of while coding, e.g. if you are unsure about something.

Other conventions

Interjections

Noises made with the mouth used to express emotion (an emotional response or feeling reaction) or signal discourse breaks. For example: “ouch”, “mmm”, “ah”, “oh”.

Below, we provide a list of common interjections in their conventional English orthography, and the meaning they convey. Please use these to transcribe interjections in the corpus. Transcriptions are taken from the Language Development Project corpus (University of Chicago).

If an utterance contains one of these interjections, the entire utterance should be coded as ‘interject’ in the onomatopoeia tier.

FILLERS: for ‘uh’, ‘er’, and ‘uhm’, we only count this as a true interjection if there was a pause of more than 300 ms as well. If this is not the case, it is not a true interjection but rather just a filler in fluent speech, and we don’t code it as interjections.

EMOTIVE WORDS: words like ‘yeah’ are not coded as interjections at the moment, because they contain emotional valence on top of being an interjection. For interjections such as “god/gosh” where it might have been “oh god/gosh” but there is no “oh” present, you can still code these as being interjections.

ah	“open your mouth and say ah”,
aha	“I have an idea!”
argh	frustration or “I’m a pirate”
aw	“so cute!”
blah	boring
boop	falling down/getting hit noise
brr	“it’s cold in here”
doh	Homer Simpson
duh	“it’s obvious”
eek	“eek! a mouse!”
eh	indifference, “I don’t care”, confirmation, “right?
er	pause in speech, similar to uh

ew	gross!
gee	golly gee whiz!
grr	angry
ha	laughter (not always necessary to transcribe)
hmm	wondering or thinking noise
hmmm	saying no (mouth closed)
hey	attention getting
hooray	“hip hip hooray!”
huh	“what did you say?” or “don’t you agree?”
huhuh	saying no (mouth open)
mmm	yummy noise
mmhm	saying yes (mouth closed)
mmmmmm	prosodic, non-verbal “I don’t know”
nah	saying no
oh	as in “oh my goodness!” or “oh boy!”
ooh	“ooh shiny!”
oops	“something’s wrong” use for woops, woop, oop, etc.
ouch	“it hurts!”

ow	“it hurts!”
psst	“it’s a secret”
shh	quiet noise
tada	“look, I did it!”
uh	filler sound
uhhuh	affirmative (open mouth)
uhoh	“oh no!”
uhuhuh	admonition, “don’t do that”
ugh	disgust, sometimes frustration, grunt of effort
um	filler sound
whee	“we’re on a roller coaster!”
whew	“what a relief”
whoa	“hold your horses” or exclamation
whoo	cheer, “go Bears!”
wow	amazed sound
yeehaw	like a cowboy

Noises

In addition to having a lexical repertoire of onomatopoeia (e.g. woof, meow, ding, crunch), we are also good at imitating real-word sounds. These

imitations are difficult to transcribe; an approximation might be something like rrrr-hrrr-fff as a rendition of a dog barking (i.e. as a more realistic barking sound). When a speaker produces this kind of an onomatopoeic sound effect (that does not correspond to lexicalized onomatopoeia) write as part of the annotation and put an explanation in the note tier, e.g. imitating the sound of a tiger; imitating the sound of knocking at the door etc. Note: only use for this situation. Do not use for noises made by the speaker which don't have a communicative social function, e.g. breathing/ sniffing noise because the participant has a cold. Exceptions are noises which do exhibit a communicative function such as an exaggerated throat clear. Do not use when the recording is unclear; see 'Unclear speech' below for what to do in those cases instead.

Laughs

Code when the participant is laughing in a natural way. If the laugh is instead more conventionalized (e.g. "ha ha ha") transcribe it and code it as an interjection.

Unclear speech

If you cannot understand what the participant is saying, type your best guess and put a note in the Comment tier.

Gesture coding

We have the following tiers for this project.

- Manual Gesture Unit
- Manual Articulator I
- Manual Articulator II
- RepGest

- BeatGest
- Iconicity
- Metaphoricity
- Conventionality
- Pointing
- Deixis
- Pragmatic Function
- Topic/Focus
- Comments
- Other

Manual Gesture Unit

The first step is to code all hand movements which are considered as gestures in the Manual Gesture Unit tier. Note that gestures are communicative hand actions which occur within the context of the interaction. If, for example, a PWA scratches her head between game turns, this would not be considered a gesture. If this same action occurs during a game turn, you will need to consider if the action is communicative (e.g. is the topic the PWA is discussing related to scratching and is the PWA intending to convey this to the researcher?). If so, this would be considered a gesture and should be coded. If not, you do not need to code this movement.

Manual Articulator I and II

We code gestures produced by the right and left hand separately, using the same coding system (described below) for each. Gestures produced by the right hand should be coded as Manual Articulator I. Gestures produced by the left hand should be coded as Manual Articulator II. Where gestures are made with both hands, code on both tiers.

Beat gestures

Code beat gestures in the BeatGest tier. Beat gestures go along with the rhythm of the speech, typically with a repeated up-down movement of the hands (much like a conductor beating time) that stresses what is being talked about. From Kita (2014): “beat gestures are small bi-directional movements that are often performed in the lower periphery of gesture space (e.g., near the lap) as if to beat the rhythm. The form of beat gestures remains more or less the same, regardless of the content of the concurrent speech”.

According to McNeill (1992, p.80), beats have prototypical characteristics. “They are typically biphasic (two movement components), small, low energy, rapid flicks of the fingers or hand; they lack a special gesture space, and are performed indeed wherever the hands happen to find 39 themselves, including rest positions (the lap, next to the cheek, etc.).”

Like pragmatic gestures, beat gestures don’t refer to the semantic content of the speech. In fact, under some accounts, beat gestures are simply a type of pragmatic gesture (e.g., Bavelas et al., 1992), so we need to be clear how to distinguish them from each other.

Here we code beat gestures when:

- The movement is a small gesture (i.e., it does not sweep across the whole space) that rhythmically accompanies a spoken word (or words if there are multiple beats).

- It only has two movement phases, e.g., the hand(s) moves up, then the hand(s) moves down. Beat gestures should simply have two movements, the hand(s) starts in one position (perhaps from resting), moves to another position, and back to the first position.

- Another way to think about this is to consider gesture apexes (e.g., Rohrer et al., 2019). Beat gestures have a single apex (if you move your hand up and down, the apex is the point where the hand changes direction at the top). In comparison a representational gesture may have multiple apexes (the hand moves from rest to the start of the stroke [apex 1] from the end of the stroke to the retraction [apex 2]).

There can be multiple beats in a single beat gesture, as long as they are temporally close together, and the movement is similar and rhythmical (i.e., the movements and apexes are similar for each beat). If there is a clear gap in between beat gestures (e.g., the hand rests for >300ms perhaps while the speaker produces more utterances without rhythmic accompaniment) then code these as separate gestures. Code beats in the BeatGest tier.

Representational gestures

Code representational gestures in the RepGest tier with one sentence describing this gesture (e.g., representing the shape of the avocado). Representational gestures refer to meaningful hand actions performed by a speaker that depict aspects of the object, action or concept that they are talking about. We code three types of representational gestures: iconic, metaphoric and conventionalised/emblems. You can code the type in the corresponding tier,

- Iconicity gestures represent *concrete* referents. For example, if describing ‘stirring’, the speaker may move her hands in a circular motion to imitate the action of stirring ingredients. Code these gestures in the iconicity tier.

- Metaphoricity gestures represent *abstract* referents. For example, if describing ‘love’, the speaker may hold her hands over her heart. Code these gestures in the metaphoricity tier.

- Conventionality gestures are gestures which have a culturally agreed conventional meaning, such as thumbs up or the ‘OK’ sign. Code these gestures in the conventionality tier.

Please note that you can only code types after you code a representational gesture, otherwise there will be no “New Annotation Here” option in these three sub-tiers.

Deixis (points)

The speaker points with an extended index finger (or with other configuration of the hand) to focus the attention on something. Note that PWA commonly experience muscle weakness which affects configuration of the hand position; it is common for PWA to point/indicate using the whole hand rather than with a single finger. To code points, first code all points in the Pointing tier and then determine whether the pointing is concrete or abstract in the Deixis tier.

Note that it is of course possible for gestures to fit within *multiple* categories. For example, if describing ‘thinking’ the speaker may point to her head. Within M3D categories are not mutually exclusive, therefore this gesture would be coded both as deictic and metaphoric. If you are unsure whether a gesture fits into multiple categories, please leave a note in the Comments tier.

Other

Some PWA will use idiosyncratic gesture types which do not fit easily into other gesture categories, such as Spelling gestures, where PWA trace out letters or whole words onto the table or into the air in front of them, then

‘read’ the traced word aloud to cue production of the label. This is an aspect of gestural communication that may be unique to PWA (De Beer et al., 2019). Another example is Homophone gestures where the PWA may, indicate a homophonous word through gesture when they are unable to retrieve the word, e.g. pointing to a wedding ring while saying “I need to [gesture] the doctor”, to be understood as “I need to ring the doctor.” Code all idiosyncratic gestures in the Other tier.

Pragmatic gestures

Pragmatic gestures do not represent the semantic content of the speech - like representational gestures - but instead manage and express the communicative interaction between listener and speaker. Pragmatic gestures have a strong interactional nature: they are usually directed (sometimes with eye gaze) to the interlocutor. In the literature, they are also called interactive gestures (e.g., Bavelas et al., 1992).

We give some examples below, and some examples of speech that may accompany pragmatic gestures (or that is implied by them). From Bavelas et al (1992), there are 4 main types of pragmatic gesture to look out for:

1. Citation: gestures that cite the interlocutor’s contribution

a. e.g., “you were talking about”, “like you said”.

2. Seeking: gestures that seek agreement, understanding or help from the interlocutor

a. e.g., “right?”, “the thing, you know, the thing!”, “See!?”, ”you remember?”, “what’s the word I’m looking for..?”

b. They may also show that the speaker is seeking feedback from the interlocutor; e.g. when making a proposal (“hows about…”), the speaker may move an open palm toward the interlocutor (Abner et al. 2015).

3. Delivery: gestures that deliver new versus shared information, e.g., “well…”, “the thing is”.

4. Regulating turn-taking:

a. these can involve holding the floor, e.g., “if you let me finish”, “don’t interrupt”, “what is more…”, “and another thing”

b. these can also involve ceding the floor, e.g., “do you want to go first?”

Left to Right: Examples of a Citation gesture, a Seeking gesture, a gesture to hold the floor (regulating turn-taking) and a gesture to cede the floor (regulating turn-taking) from Bavelas et al., (1992)

As these gestures can refer to an interlocutor’s role in the conversation, these gestures can often involve the speaker pointing (with finger(s), thumb or whole hand) towards the interlocutor, or otherwise extending their hand in the direction of the interlocutor (for example if seeking feedback, recruiting their help, or deferring to them).

These gestures should be coded with function (1-4 above) specified in the pragmatic function tier. Note that the function of a pragmatic gesture can be subjective; if you are not sure, enter all the functions you feel the gesture may fulfil.

Topic

Within the Topic/Focus tier, enter the target word for the trial. Refer to the trial spreadsheets if you are unsure of the target word.

Enumerations (counting)

If you come across enumerations, please first determine the gesture in the Manual Gesture Unite tier, and leave a comment “enumeration”.

Comments (other coding conventions)

Beyond the coding described above, you can use the Comments tier to code additional aspects of gesture/speech production, including inserting notes if you are unsure about any aspect of coding above.

Blurring faces

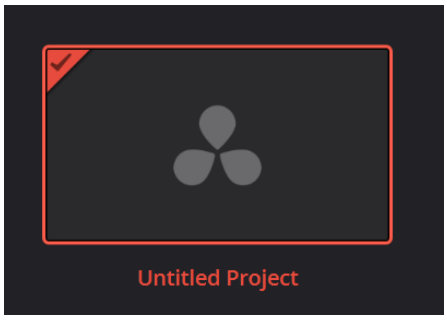
As our participants are vulnerable adults, we need to blur their faces before we can show the videos publicly (e.g. as part of a research presentation), to ensure that they are anonymised and cannot be recognised. Blurring is done using DaVinci Resolve 18. You can download this here: DaVinci Resolve 18 | Blackmagic Design. Note that there are two download options – ensure you select the free download option!

The process is explained fully below, but we also have a short instruction video which summarises how to complete each step of the process. If you would like access to this video, please ask Isobel.

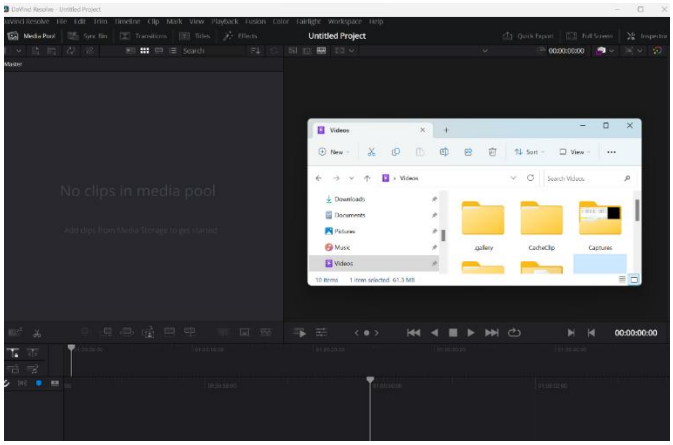
Please be aware that you do not have to blur Isobel’s face in any of the videos.

Setting up a new project

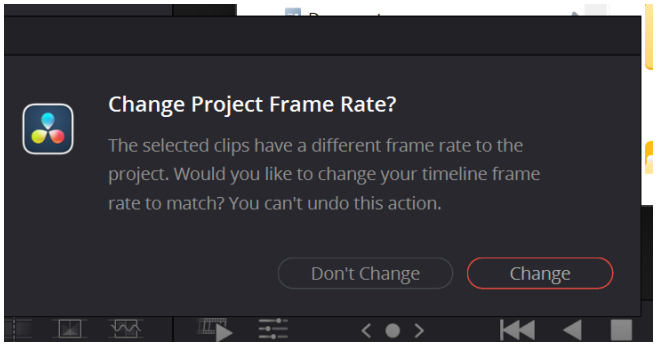
First, create a new project in Da Vinci Resolve and import the video files you want to process. To do this, click Untitled Project.



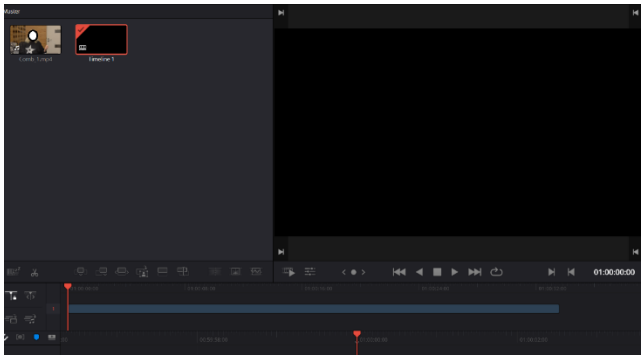
Drag the original video file that you want to blur to the media pool.



A box will appear asking if you want to change the frame rate. Click Change.

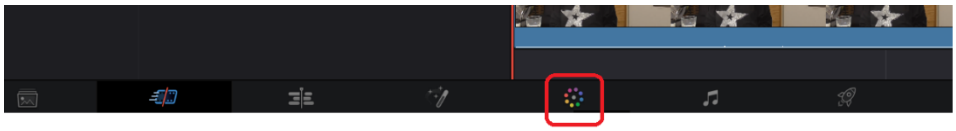


Drag the video clip to the timeline. Now you're ready to start blurring the image.

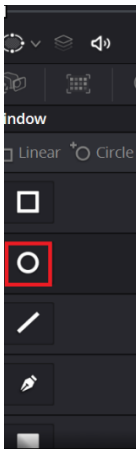


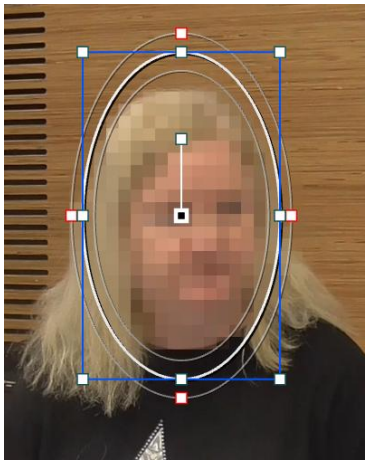
Blurring the face

Click on the colour wheel below the timeline. A new screen will appear showing your video and a range of options for blurring.

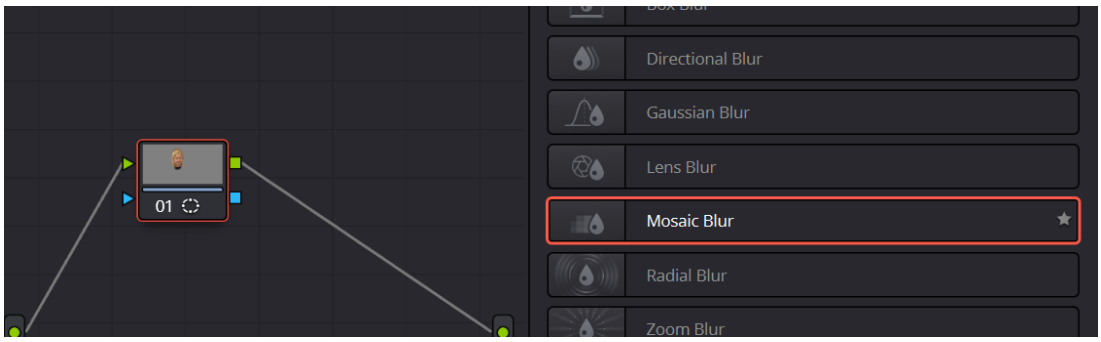


Click on the circle. This will be the shape of the blurring mask. Fit the mask over the face. You can adjust the size of the mask as needed to fit fully over the participant's face. It is OK for hair to be outside the mask, but all facial features must be fully within the mask.



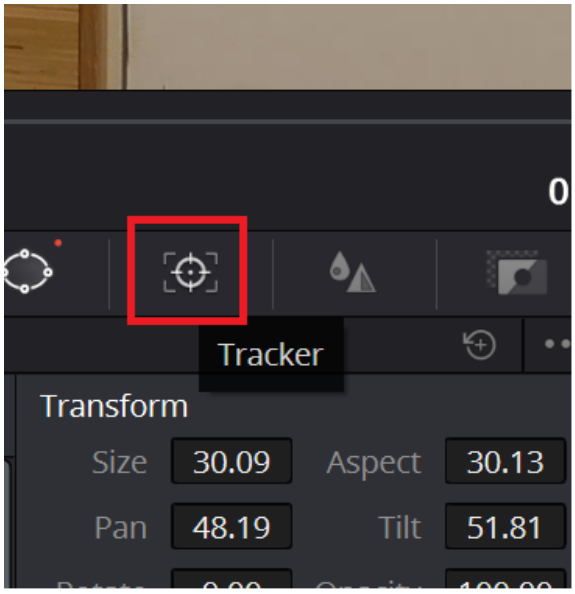


You will notice that in your video, the blur effect has not yet been applied. To apply the blur effect, click on Mosaic Blur and drag it over the effects box. This will add the blur effect to the frame in your video. .

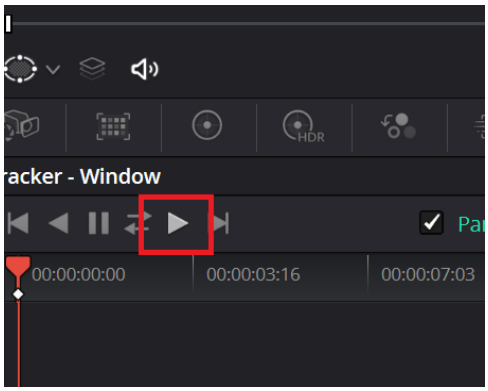


Tracking the blur effect

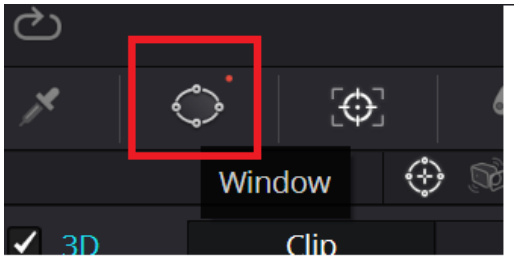
Because our participants do not remain still, we need to ensure that the mask moves as they do so that their facial features are consistently blurred throughout the videos. To do this, we need to use tracking. First, click the Tracker button (just below the video clip).



The Tracking command screen will appear below the video. Click Track Forward to move forward through the video.



You will see the mask move automatically on the screen as the face moves, but the automatic tracking is not perfect and you may have to stop the video and adjust the mask if it moves off the face. To do this, click the Window button. This will take you back to the mask screen and you can adjust the mask as needed.

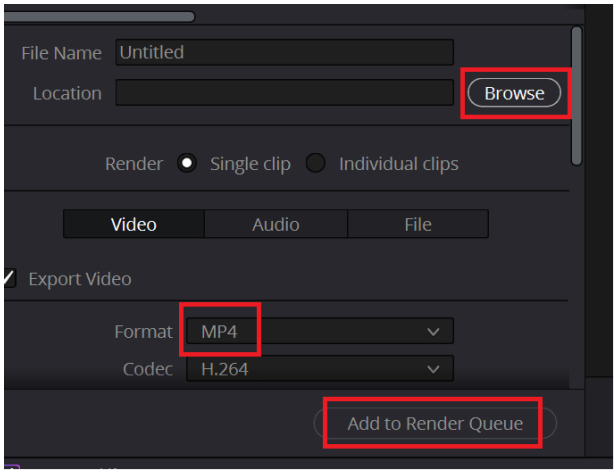


After you have adjusted the mask, click on the Tracker button again to return to the Tracking command screen and continue moving through the video. You can track backwards to move back through the video as well. You may have to stop the video and adjust the mask multiple times. Continue doing this until you have gone through the whole video and are sure that the blurring mask completely covers the face throughout.

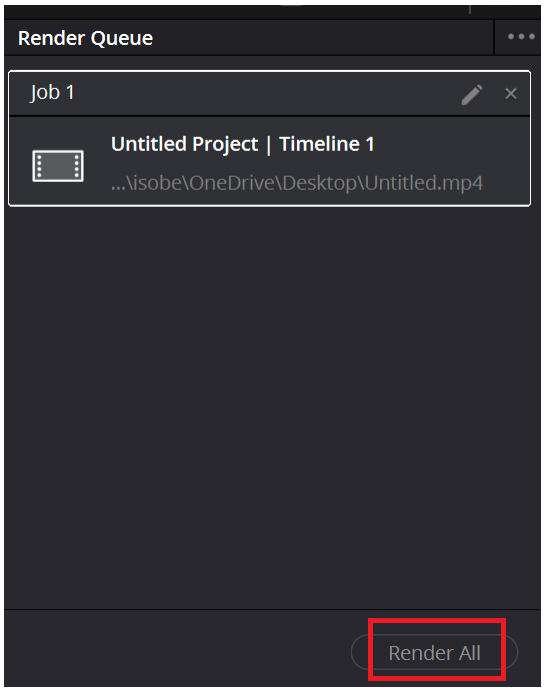
Rendering and saving the blurred file

After you have completed blurring the file, click the Deliver button at the bottom of the screen.

Click Browse and navigate to the location where you want to save the file. The file should be named according to the typographical conventions outlined on page 9. Ensure you change the file type to MP4. Click Add to Render Queue.



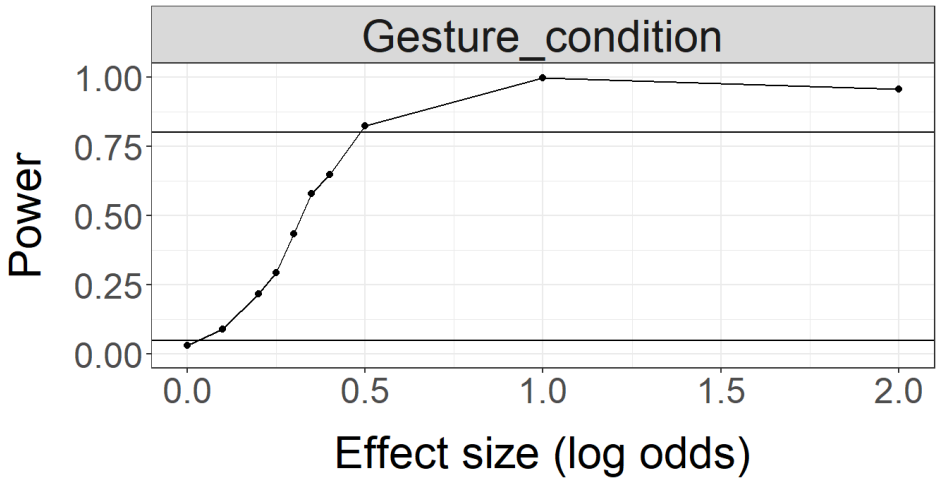
Your file will appear on the right hand side of the screen in the Render Queue. Click Render All to start the rendering and saving process. Note that this process can take several minutes to several hours depending on the length of the video! Ensure you do not close the DaVinci while this process is ongoing or it will not save the file. When the file has finished processing and saved, the white filename will turn green.



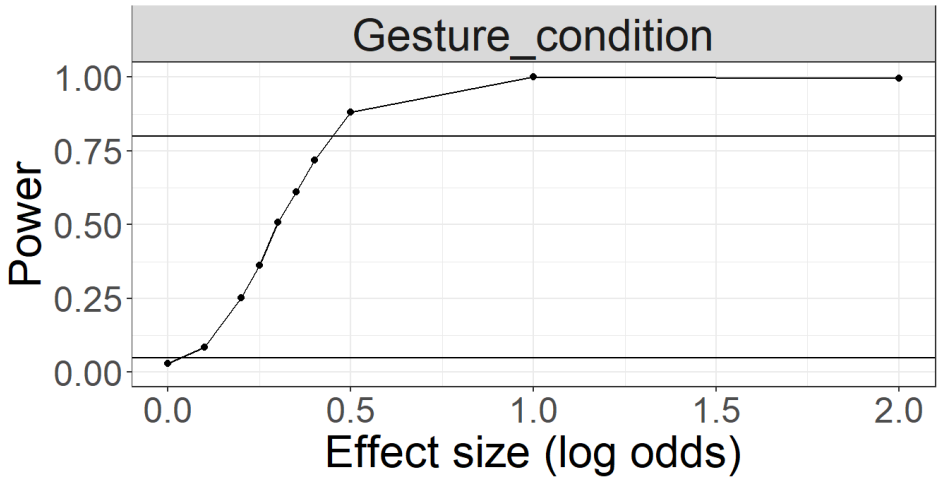
Appendix 4: Chapter 2 "Go Fish!" Power Analyses

Following Cabiddu et al. (2022), for each generalised linear mixed effects model power was simulated (500 model iterations) for the observed data for the following effect sizes: 0.0, 0.1, 0.2, 0.25, 0.3, 0.35, 0.4, 0.5, 1.0, 2.0.

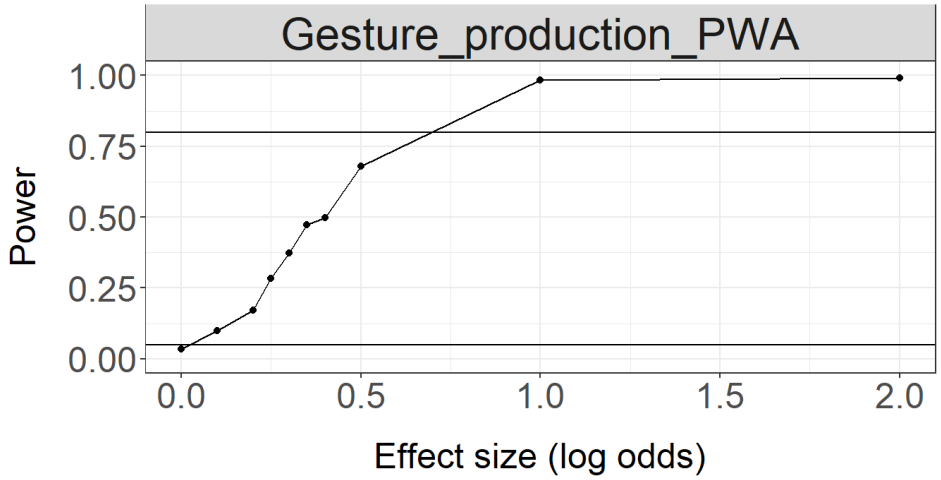
Speech comprehension



Naming accuracy (gesture observation)



Naming accuracy (gesture production)



Sample code:

```
# Display the model summary and Odds Ratios
summary(mod); tab_model(mod)

# Check for singularity
isSingular(mod)

# Check for multicollinearity
car::vif(mod)

# Simulate residuals for model assumptions
simulationOutput <- simulateResiduals(fittedModel = mod, plot = TRUE)

# Number of iterations for power simulations
M <- 500

# Effect sizes to explore
effect_sizes <- c(0, 0.1, 0.2, 0.25, 0.3, 0.35, 0.4, 0.5, 1, 2)

# Extract model design components
beta <- fixef(mod)
sigma_id_intercept <- sqrt(unlist(VarCorr(mod))[1])
sigma_error <- attr(VarCorr(mod), "sc")

# Prepare participant random effects
pt_effects_intercept <- tibble(intercept = ranef(mod)$Participant_ID[, 1],
                               name_effect = row.names(ranef(mod)$Participant_ID))
%>%
  column_to_rownames(var = "name_effect")

# Calculate non-target fixed effects
X <- model.matrix(mod)
```

```
fixed_effects <- X[, -2] %*% beta[-2]

# Power simulation function
do_power_sim <- function(effect_size_sim,
                         data_for_power_sim,
                         pt_effects_intercept_sim,
                         sigma_id_intercept_sim,
                         sigma_error_sim,
                         fixed_effects_sim,
                         replication, conditionFilter, conditionLabel) {
  options(warn = 2)
  on.exit(options(warn = 1))

  replicate(replication, {
    repeat {
      sim_data <- data_for_power_sim

      # Simulate participant-specific effects
      pt_effects_intercept_sim[, 1] <- rnorm(pt_effects_intercept_sim[, 1], 0,
sigma_id_intercept_sim)

      # Combine simulated intercepts
      random_effects_sim <-
pt_effects_intercept_sim[as.character(data_for_power_sim$Participant_ID),
1]

      # Sample model errors
      error_sim <- rnorm(nrow(X), 0, sigma_error_sim)

      # Simulate response variable
      sim_response <- (random_effects_sim + fixed_effects_sim + error_sim) +
        X[, 2] * effect_size_sim

      # Convert to probabilities and simulate new responses
      probs <- plogis(sim_response)
      sim_data$new_response <- rbinom(length(probs), 1, probs)

      # Refit model to simulated data
      sim_x <- try({
        lme4::glmer(new_response ~
                      Gesture_condition * scale(Working_memory) +
                      Gesture_condition * scale(Semantic_access) +
                      Gesture_condition * scale(Gesture_recognition) +
                      Gesture_condition * scale(Target_word_frequency) +
                      Gesture_condition * scale(Target_word_length) +
                      Gesture_condition * scale(Aphasia_quotient) +
                      Researcher_ID +
                      (1 | Participant_ID),
                    data = sim_data,
                    family = binomial(link = 'logit'),
```

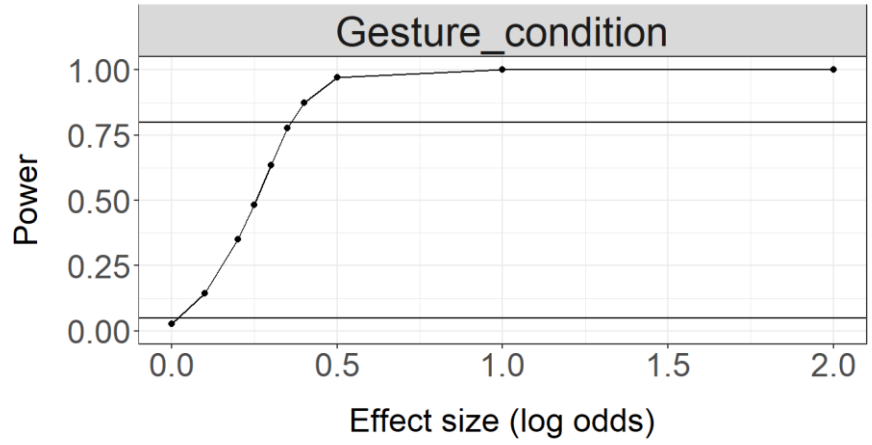
```
                control = glmerControl(optimizer = "bobyqa", optCtrl =
list(maxfun = 100000)))
    })

    # Extract results for the effect of interest
    if (inherits(sim_x, "try-error") == FALSE) {
      sim_temp <- try({
        broom.mixed::tidy(sim_x, conf.int = TRUE, exponentiate = TRUE,
effects = "fixed") %>%
          filter(term == conditionFilter) %>%
          unlist
      })
    } else {
      sim_temp <- c(effect = "fixed", term = conditionFilter, estimate = NA,
std.error = NA,
                    statistic = NA, p.value = NA, conf.low = NA, conf.high = NA)
    }

    break
  }

  sim_temp
}) -> sim1

sim1 <- as.data.frame(t(sim1))
sim1$sim <- 1:replication
sim1$effect_size_sim <- effect_size_sim
sim1$comparison <- conditionLabel
return(sim1)
}

# Initialise parallel processing
plan(multisession)

# Start timing
tic()

# Initialise empty list for simulation results
sim_results_final <- list()

# Perform the simulations in parallel
set.seed(123)
sim_results_final[["gesture_effect"]] <-
  effect_sizes %>%
  future_map_dfr(~ do_power_sim(., data_for_power,
                                pt_effects_intercept,
                                sigma_id_intercept,
                                sigma_error,
                                fixed_effects,
                                replication = M,
```

```
                        conditionFilter = "Gesture_condition",
                        conditionLabel = "Gesture_condition"),
            .progress = TRUE,
            .options = furrr_options(seed = TRUE)) %>%
  mutate(condition = "Gesture_condition")

# Combine results
sim_results_final[["gesture_effect"]] %<>% bind_rows
sim_results_final %<>% bind_rows
sim_results_final %<>% rename(effect_size = effect_size_sim)

# Return simulation time
toc()

# Plot simulation results
sim_results_final %>%
  mutate(p.value = as.numeric(p.value),
         effect_size = as.numeric(effect_size),
         estimate = as.numeric(estimate) %>% log) %>%
  mutate(significance = case_when(
    p.value < .05 & estimate > 0 ~ TRUE,
    TRUE ~ FALSE
  )) %>%
  group_by(condition, effect_size) %>%
  summarise(power = sum(significance) / n()) %>%
  ungroup %>%
  ggplot(aes(x = effect_size, y = power)) +
  geom_point() +
  geom_line() +
  geom_hline(yintercept = 0.8) +
  geom_hline(yintercept = 0.05) +
  xlab("Effect size (log odds)") +
  ylab("Power") +
  scale_y_continuous(limits = c(0, 1)) +
  facet_grid(~condition) +
  theme_bw() +
  theme(axis.text.x = element_text(angle = 0, size = 20),
        axis.text.y = element_text(angle = 0, size = 20),
        strip.text = element_text(size = 25),
        plot.margin = unit(c(1, 1, 1, 1), "lines"),
        panel.spacing = unit(2, "lines"),
        axis.title.y = element_text(margin = unit(c(0, 5, 0, 0), "mm"), size = 25),
        axis.title.x = element_text(margin = unit(c(5, 0, 0, 0), "mm"), size = 25),
        legend.title = element_text(size = 25),
        legend.text = element_text(size = 20)) +
  scale_color_viridis(discrete = TRUE, option = "turbo")
```

Appendix 5: Chapter 3 “Heads Up!” Power Analysis

Following Cabiddu et al. (2022), for the generalised linear mixed effects model power was simulated (500 model iterations) for the observed data for the following effect sizes: 0.0, 0.1, 0.2, 0.25, 0.3, 0.35, 0.4, 0.5, 1.0, 2.0.

Naming accuracy (gesture observation)



Sample code:

```
# Display the model summary and Odds Ratios
summary(mod); tab_model(mod)

# Check for singularity
isSingular(mod)

# Check for multicollinearity
car::vif(mod)

# Simulate residuals for model assumptions
simulationOutput <- simulateResiduals(fittedModel = mod, plot = TRUE)

# Number of iterations for power simulations
M <- 500

# Effect sizes to explore
effect_sizes <- c(0, 0.1, 0.2, 0.25, 0.3, 0.35, 0.4, 0.5, 1, 2)

# Extract model design components
beta <- fixef(mod)
sigma_participant_intercept <- sqrt(unlist(VarCorr(mod)$Participant_ID)[1])
sigma_participant_slope <- sqrt(unlist(VarCorr(mod)$Participant_ID)[2])
```

```
sigma_error <- attr(VarCorr(mod), "sc")

# Prepare participant random effects
participant_effects <- tibble(
  intercept = ranef(mod)$Participant_ID[, 1],
  slope = ranef(mod)$Participant_ID[, 2],
  name_effect = row.names(ranef(mod)$Participant_ID)
) %>%
  column_to_rownames(var = "name_effect")

# Calculate non-target fixed effects
X <- model.matrix(mod)
fixed_effects <- X[, -2] %*% beta[-2]

# Power simulation function
do_power_sim <- function(effect_size_sim,
                         data_for_power_sim,
                         participant_effects_sim,
                         sigma_participant_intercept_sim,
                         sigma_participant_slope_sim,
                         sigma_error_sim,
                         fixed_effects_sim,
                         replication, conditionFilter, conditionLabel) {
  options(warn = 2)
  on.exit(options(warn = 1))

  replicate(replication, {
    repeat {
      sim_data <- data_for_power_sim

      # Simulate participant-specific random effects
      participant_effects_sim[, 1] <- rnorm(participant_effects_sim[, 1], 0,
sigma_participant_intercept_sim)
      participant_effects_sim[, 2] <- rnorm(participant_effects_sim[, 2], 0,
sigma_participant_slope_sim)

      # Combine simulated intercepts and slopes
      random_effects_sim <-
participant_effects_sim[as.character(data_for_power_sim$Participant_ID), 1]
+
        X[, 2] *
participant_effects_sim[as.character(data_for_power_sim$Participant_ID), 2]

      # Sample model errors
      error_sim <- rnorm(nrow(X), 0, sigma_error_sim)

      # Simulate response variable
      sim_response <- (random_effects_sim + fixed_effects_sim + error_sim) +
        X[, 2] * effect_size_sim
```

```
    # Convert to probabilities and simulate new responses
    probs <- plogis(sim_response)
    sim_data$new_response <- rbinom(length(probs), 1, probs)

    # Refit model to simulated data
    sim_x <- try({
      lme4::glmer(new_response ~
                    Gesture_condition * scale(Working_memory) +
                    Gesture_condition * scale(STM) +
                    Gesture_condition * scale(Semantic_access) +
                    Gesture_condition * scale(Gesture_recognition) +
                    Gesture_condition * scale(WABR_SponSpeech) +
                    scale(Trial_order) +
                    Researcher_ID +
                    First_gesture_condition +
                    Gesture_condition * Condition_object1_action0 +
                    Gesture_condition * scale(Target_word_length) +
                    Gesture_condition * scale(Target_word_frequency) +
                    (Gesture_condition | Participant_ID),
                  data = sim_data,
                  family = binomial(link = 'logit'),
                  control = glmerControl(optimizer = "bobyqa", optCtrl =
list(maxfun = 100000)))
    })

    # Extract results for the effect of interest
    if (inherits(sim_x, "try-error") == FALSE) {
      sim_temp <- try({
        broom.mixed::tidy(sim_x, conf.int = TRUE, exponentiate = TRUE,
effects = "fixed") %>%
          filter(term == conditionFilter) %>%
          unlist
      })
    } else {
      sim_temp <- c(effect = "fixed", term = conditionFilter, estimate = NA,
std.error = NA,
                    statistic = NA, p.value = NA, conf.low = NA, conf.high = NA)
    }

    break
  }

  sim_temp
}) -> sim1

sim1 <- as.data.frame(t(sim1))
sim1$sim <- 1:replication
sim1$effect_size_sim <- effect_size_sim
sim1$comparison <- conditionLabel
return(sim1)
```

```
}

# Initialise parallel processing
plan(multisession)

# Start timing
tic()

# Initialise empty list for simulation results
sim_results_final <- list()

# Perform the simulations in parallel
set.seed(123)
sim_results_final[["gesture_condition_effect"]] <-
  effect_sizes %>%
  future_map_dfr(~ do_power_sim(., data_for_power_imputed,
                                participant_effects,
                                sigma_participant_intercept,
                                sigma_participant_slope,
                                sigma_error,
                                fixed_effects,
                                replication = M,
                                conditionFilter = "Gesture_condition",
                                conditionLabel = "Gesture_condition"),
                 .progress = TRUE,
                 .options = furrr_options(seed = TRUE)) %>%
  mutate(condition = "Gesture_condition")

# Combine results
sim_results_final %<>% bind_rows
sim_results_final %<>% rename(effect_size = effect_size_sim)

# Return simulation time
toc()

# Plot simulation results
sim_results_final %>%
  mutate(p.value = as.numeric(p.value),
         effect_size = as.numeric(effect_size),
         estimate = as.numeric(estimate) %>% log) %>%
  mutate(significance = case_when(
    p.value < .05 & estimate > 0 ~ TRUE,
    TRUE ~ FALSE
  )) %>%
  group_by(condition, effect_size) %>%
  summarise(power = sum(significance) / n()) %>%
  ungroup %>%
  ggplot(aes(x = effect_size, y = power)) +
  geom_point() +
  geom_line() +
```

```
geom_hline(yintercept = 0.8) +
geom_hline(yintercept = 0.05) +
xlab("Effect size (log odds)") +
ylab("Power") +
scale_y_continuous(limits = c(0, 1)) +
facet_grid(~condition) +
theme_bw() +
theme(axis.text.x = element_text(angle = 0, size = 20),
      axis.text.y = element_text(angle = 0, size = 20),
      strip.text = element_text(size = 25),
      plot.margin = unit(c(1, 1, 1, 1), "lines"),
      panel.spacing = unit(2, "lines"),
      axis.title.y = element_text(margin = margin(r = 15), size = 20),
      axis.title.x = element_text(margin = margin(t = 15), size = 20),
      legend.title = element_text(size = 25),
      legend.text = element_text(size = 20))
```

Appendix 6: Chapter 3 MATTR Python Code

```
def calculate_mattr(tokens, window_size=50):
    # Optional: check if list of tokens is smaller than planned window size. Edit window size if so.
   # if len(tokens) < window_size:
        # raise ValueError("Token list is smaller than the window size.")

    # Build list to store TTR values for each window
    ttr_values = []

    # Slide window across tokens
    for i in range(len(tokens) - window_size + 1):
        window = tokens[i: i + window_size]
        unique_types = len(set(window))  # Count unique words in the window
        ttr = unique_types / window_size  # Calculate TTR for the window
        ttr_values.append(ttr)

    # Calculate the mean of TTR values across all windows
    mattr = sum(ttr_values) / len(ttr_values)
    return mattr


# Calculate MATTR with a specified window size
window_size = 50  # Adjust if needed
mattr_value = calculate_mattr(lemmatized_tokens,
window_size=window_size)
print("Moving-Average Type-Token Ratio (MATTR):", mattr_value)
```

Appendix 7: Chapter 5 Gesture Norming Scores (Therapy)

Label	Score
angel	4.0
banjo	4.5
binoculars	6.2
bubble	3.1
calculator	4.1
can opener	3.9
captain	5.1
caterpillar	4.9
cheerleaders	4.3
clarinet	4.7
computer	5.7
cymbals	4.3
dustpan	4.3
earmuffs	4.8
fly swatter	3.7
frisbee	5.5
hammock	2.3
ladder	5.6
ladle	3.7
lawnmower	2.6
lightning	4.7
locker	3.0
magnifying glass	3.9
marker	4.8
mascara	5.1
measuring cup	3.0
medal	4.4
menu	3.8
microphone	4.7
mobile/cell phone	6.5
mustache	5.6
nail clippers	6.1

Label	Score
necklace	5.4
needle	5.1
notebook	5.4
picture	4.4
pocket	5.8
rainbow	4.1
raincoat	4.1
roller coaster	4.5
scarf	6.1
shaving cream	4.4
shoelace	4.8
slingshot	6.2
slinky	2.7
spider	5.3
staples	2.5
sugar	3.8
table	3.0
tap	4.7
teabag	4.3
telescope	6.4
toaster	3.5
toothpaste	3.8
trunk	4.8
waitress	2.7
walker	3.4
watermelon	3.2
windmill	3.2
window	3.7
xylophone	3.3

Appendix 8: Chapter 5 PWA Therapy Instruction Sheet

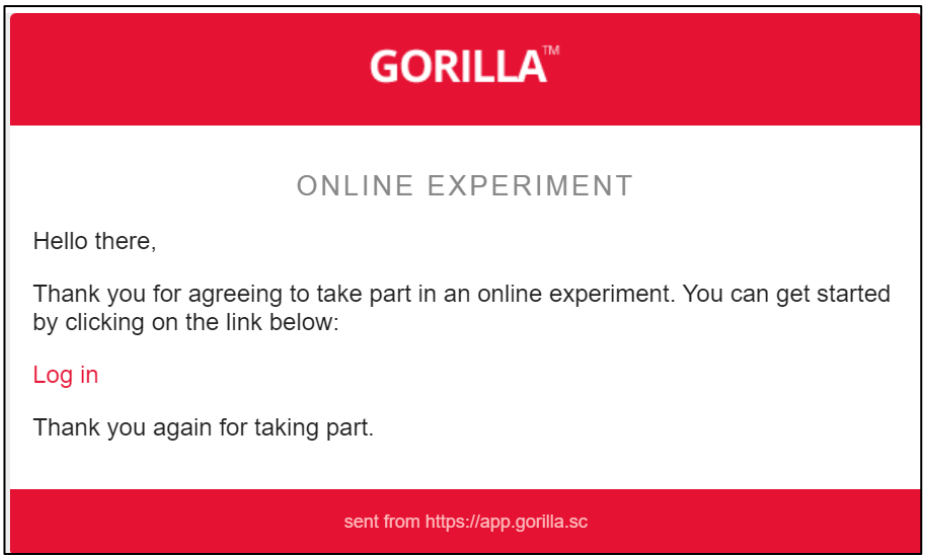
Welcome to your therapy programme!

Your first session

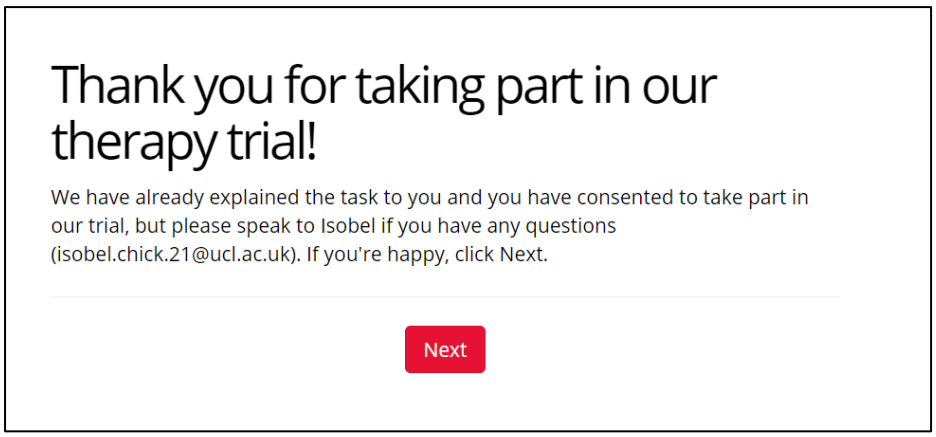
- Your online sessions will run on a website called Gorilla. You will need to access this on a computer, you cannot use a mobile phone.

- Before you start your sessions, you will receive an email like this from Gorilla. Isobel will tell you when to expect this email. Please contact Isobel immediately if you do not receive this email.

- Click the red link to log in for the first time.

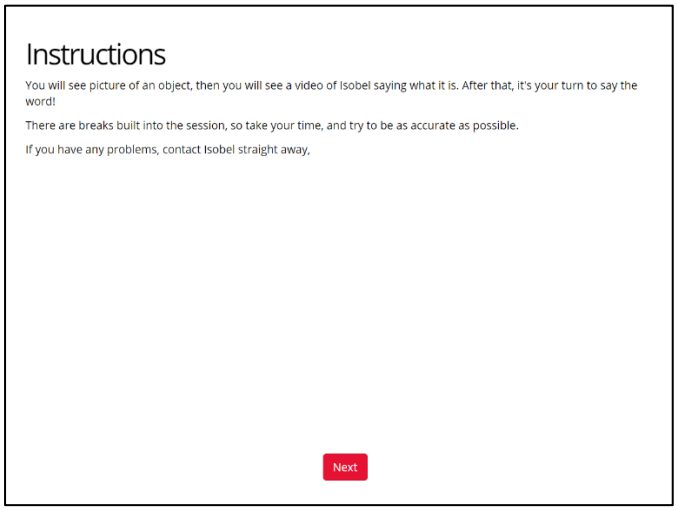


- You will see this screen. Click Next.

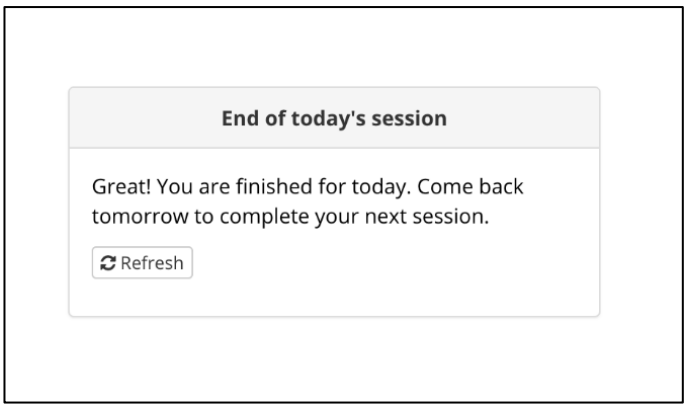


- The first time you log in, you will be asked to consent to take part. Tick the box to consent.

- You will see an instruction screen like this. Click Next to start the therapy session.



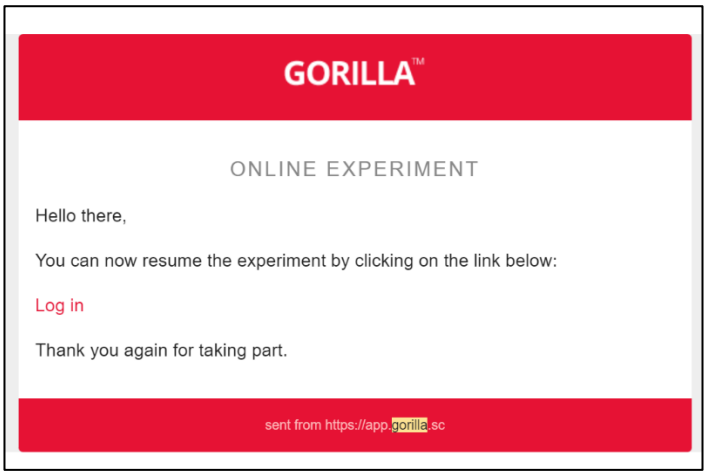
- At the end of the session, you will see this screen. This means you can close your browser and end the session.



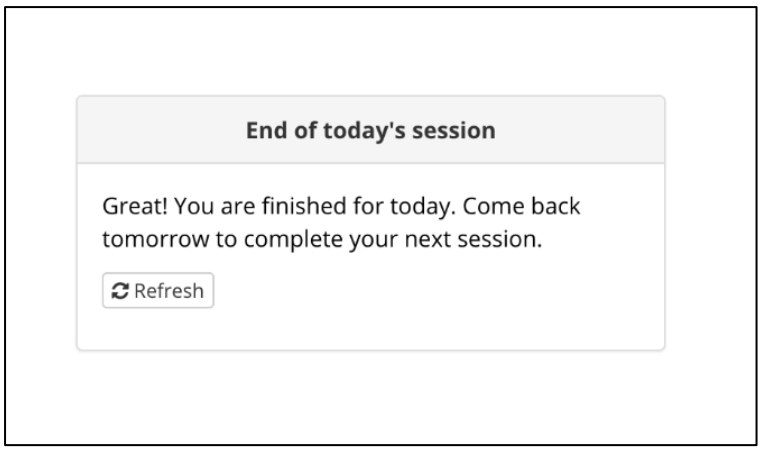
Your next sessions

- The next day, you will receive a reminder email telling you to complete your next session. You will not receive a reminder for the next day unless you complete the session for that day. For example, if you do not complete your session on Wednesday, you will not receive another reminder on Thursday.

- To complete your next session, click on the red link in the reminder email.



- You can then complete your session like before. You will not need to consent again. At the end of all your sessions, you will see this screen. This means you can close the browser and end the session.



- Please complete *at least* four online sessions per week, for at least four weeks. You can complete more if you choose. This is because the evidence shows people with aphasia need to complete at least 20 hours of speech and language therapy, with sessions at least three times a week, to see the greatest change in their communication.

- You will also have Zoom sessions with Isobel approximately once per week. Isobel will email you separately to arrange these sessions. The Zoom sessions will work the same way as the online sessions.

- Isobel will keep track of the sessions you have completed and will email you if you need to complete extra sessions.

Note that this is a pilot project and sometimes things do go wrong! If you have any technical problems at all or problems accessing the Gorilla website please let Isobel know *immediately*.

If you have technical problems Isobel will re-set your sessions and send you a new access link.

Appendix 9: Data Analysis Models

Chapter 2

Model name	Model code
Effects of gesture observation on comprehension accuracy	mod <-glmer(Comprehension ~ Gesture_condition * scale(WM_EF) + Gesture_condition * scale(Semantic_access) + Gesture_condition * scale(Gesture_recognition) + Gesture_condition * scale(Target_word_frequency) + Gesture_condition * scale(Target_word_length) + Gesture_condition * scale(Aphasia_quotient) + Site + (1 | Participant_ID), data = GoFish, na.action = "na.omit", family = binomial(link = 'logit'), control = glmerControl(optimizer = "bobyqa", optCtrl = list(maxfun = 100000)))
Effects of gesture observation on naming accuracy	mod <-glmer(Naming_accuracy ~ Gesture_condition * scale(WM_EF) + Gesture_condition * scale(STM) + Gesture_condition * scale(Semantic_access) + Gesture_condition * scale(Gesture_recognition) + Gesture_condition * scale(WABR_SponSpeech) + scale(Back_count) + Gesture_condition * scale(Target_word_frequency) + Gesture_condition * scale(Target_word_length) + First_gesture_condition + Site + (1 | Participant_ID), data = GoFishMay, na.action = "na.omit", family = binomial(link = 'logit'), control = glmerControl(optimizer = "bobyqa", optCtrl = list(maxfun = 100000)))
Semantic and phonemic similarity of naming errors (gesture observation)	mod <- lmer(Similarity ~ Gesture_condition + scale(WM_EF) + scale(Gesture_to_sight) + First_gesture_condition + scale(Target_word_length) + scale(Target_word_frequency) + (1 | Participant_ID), data = GoFish_subset)
Effects of gesture production on naming accuracy	mod<-glmer(Naming_accuracy ~ Gesture_production_PWA * scale(WM_EF) + Gesture_production_PWA * scale(STM) + Gesture_production_PWA * scale(Semantic_access) + Gesture_production_PWA * scale(Gesture_to_sight) + Gesture_production_PWA * scale(WABR_SponSpeech) + First_gesture_condition + Gesture_production_PWA * scale(Target_word_length) + Gesture_production_PWA * scale(Target_word_frequency) + (1 | Researcher_ID) + (1 | Participant_ID), data = GoFish, na.action = "na.omit", family = binomial(link = 'logit'), control = glmerControl(optimizer = "Nelder_Mead", optCtrl = list(maxfun = 100000)))
Semantic and phonemic similarity of naming errors (gesture production)	mod <-lmer(Similarity ~ Gesture_production_PWA * scale(WM_EF) + Gesture_production_PWA * scale(STM) + Gesture_production_PWA * scale(Semantic_access) + Gesture_production_PWA * scale(Gesture_to_sight) + Gesture_production_PWA * scale(WABR_SponSpeech) + scale(Trial_order) + Gesture_production_PWA * scale(Target_word_frequency) + Gesture_production_PWA *

	scale(Target_word_length) + First_gesture_condition + Researcher_ID + (1 | Participant_ID), data = GoFish_subset, na.action = "na.omit")

Chapter 3

Model name	Model code
Effects of gesture observation on naming accuracy	mod <-glmer(Naming_accuracy ~ Gesture_condition * scale(WM_EF) + Gesture_condition * scale(STM) + Gesture_condition * scale(Semantic_access) + Gesture_condition * scale(Gesture_recognition) + Gesture_condition * scale(WABR_SponSpeech) + scale(Trial_order) + Researcher_ID + First_gesture_condition + Gesture_condition * Condition_object1_action0 + Gesture_condition * scale(Target_word_length) + Gesture_condition * scale(Target_word_frequency) + (Gesture_condition | Participant_ID), data = file, na.action = "na.omit", family = binomial(link = 'logit'), control = glmerControl(optimizer = "bobyqa", optCtrl = list(maxfun = 100000)))
Semantic and phonemic similarity of naming errors	mod <- lmer(Similarity ~ Gesture_condition + scale(WM_EF) + scale(Aphasia_Quotient) + scale(Semantic_access) + First_gesture_condition + scale(Target_word_length) + scale(Target_word_frequency) + (1 | Participant_ID), data = file_subset)
MLU variance by gesture condition	mod<-lmer(scale(Utterance_length) ~ Gesture_condition + scale(AQ) + scale(Working_memory) + scale(STT) + scale(CCT) + scale(KDT) + (1 | PWA_ID) + (1 | Researcher_ID), data = file)
Utterance concreteness by gesture condition	mod<-lmer(scale(Concreteness) ~ Gesture_condition + scale(AQ) + scale(Working_memory) + scale(STT) + scale(KDT) + scale(CCT) + Researcher_ID + (1 | PWA_ID), data = file)

Chapter 4

Model name	Model code
Cosine similarity	mod<- lmer(scale(Cosine_similarity) ~ Gesture_condition * Median_GTS + Object_action + scale(Target_word _frequency) + scale(Target_word_length) + scale(AQ) + (1 | PWA), data = CosineSimilarity)
Communicative effectiveness ratings	mod<- lmer(scale(Ratings) ~ Gesture_condition * Median_GTS + Object_action + scale(Target_word _frequency) + scale(Target_word_length) + scale(AQ) + (1 | PWA), data = Ratings)
Error analysis	mod<- lmer(scale(Cosine_similarity) ~ Gesture_condition * Median_GTS + Object_action + scale(Target_word _frequency) + scale(Target_word_length) + scale(AQ) + (1 | PWA), data = CosineSimilarity_subset)

Reaction time analysis	anova <- aov(scale(`Reaction Time`) ~ Gesture_condition * Study_Version, data = data)

Chapter 5

Model name	Model code
Gesture effect on naming accuracy post-therapy	mod <-glmer(Naming_accuracy ~ Gesture_condition + Gesture_condition * scale(AQ) + Gesture_condition * scale(Working_memory_transformed) + Gesture_condition * scale(Target_word_length) + scale(AQ) * scale(Target_word_length) + Gesture_condition * scale(AQ) * scale(Target_word_length) + Gesture_condition * scale(Working_memory_transformed) * scale(Target_word_length) + (1 | PWA), na.action = "na.omit", data = treated_data, family = binomial(link = 'logit'), control = glmerControl(optimizer = "bobyqa", optCtrl = list(maxfun = 200000)))
Interactions with working memory/executive function and gesture norm scores	mod <-glmer(Naming_accuracy ~ scale(Gesture_norm_score) * scale(Working_memory_transformed) + scale(Target_word_length) + (1 | PWA), na.action = "na.omit", data = treated_data, family = binomial(link = 'logit'), control = glmerControl(optimizer = "bobyqa", optCtrl = list(maxfun = 100000)))
Sub-analysis: correct responses and non-omission errors	mod<-glmer(Correct_non_omission ~ Gesture_condition + Gesture_condition * scale(Working_memory_transformed) + scale(Target_word_length) + scale(Target_word_frequency) + (1 | PWA), na.action = "na.omit", data = omission_subset, family = binomial(link = 'logit'), control = glmerControl(optimizer = "bobyqa", optCtrl = list(maxfun = 200000)))
Error analysis (cosine similarity)	mod<-lmer(scale(Cosine_similarity) ~ Gesture_condition + scale(AQ) + scale(Working_memory_transformed) + scale(Target_word_length) + (1 | PWA), na.action = "na.omit", data = treated_data)
Error analysis (phonemic Levenshtein distance)	mod<-lmer(scale(Phonemic_distance) ~ Gesture_condition + scale(AQ) + scale(Working_memory_transformed) + scale(Target_word_length) + (1 | PWA), na.action = "na.omit", data = treated_data)