Parallel Distributed Processing (PDP) models as a framework for designing cognitive rehabilitation therapy.

A thesis submitted to the University of Manchester for the degree of Doctor of Philosophy in the Faculty of Medical and Human Sciences.

2014

Solomon Eyoka Nte

School of Psychological Sciences

Table of Contents

LIST OF FIGURES **5**

LIST OF TABLES **7**

ABSTRACT **9**

DECLARATION **10**

COPYRIGHT STATEMENT **10**

DEDICATION **12**

ACKNOWLEDGEMENT **12**

ABOUT THE AUTHOR **12**

INTRODUCTION **13**

- THE STRUCTURE OF THE THESIS 13
- THESIS THEME 1: HOW DO PDP MODELS EXTRACT EFFICIENT AND EFFECTIVE REPRESENTATIONS FROM THE LEARNING ENVIRONMENT? 14
- THESIS THEME 2: HOW CAN PDP MODELS BE USED TO UNDERSTAND RECOVERY AND REHABILITATION FOR IMPAIRMENTS IN THE CONTEXT OF PATIENT-RELATED WORK? 23
- ORGANISATION OF THE THESIS 31
- ACKNOWLEDGEMENT OF CONTRIBUTION FROM OTHER AUTHORS 32

CHAPTER 1 - KNOWLEDGE ACQUISITION AND REPRESENTATIONAL ECONOMY: A COMPUTATIONAL INVESTIGATION OF EFFICIENCY IN THE BRAIN'S 'CONVERGENCE ZONES' **33**

- ABSTRACT 33
- INTRODUCTION 34
- AIMS 40
- SIMULATION 1.1: DEVELOPMENTAL LEARNING 41
  - *Method* *41*
    - Training Stimuli 45
    - Training the model 46
    - Calculating representational economy in the semantic convergence zone 51
    - Testing the network and analysing representational economy 52
  - *Results* *53*
- SIMULATION 1.2: ROBUSTNESS TO DAMAGE 56
  - *Method* *56*
  - *Results* *57*
- SIMULATION 1.3: RELEARNING AFTER DAMAGE 59
  - *Method* *59*
  - *Results* *60*
    - Mild Lesioning 60
    - Moderate Lesioning 63
    - Severe Lesioning 66
- DISCUSSION 69
  - *Summary of Results* *69*
  - *Analysis of convergence in development* *73*
  - *Analysis of convergence in relearning after damage* *74*
- CONCLUSION 76

CHAPTER 2 - LEARNING BY DEGREES: EXPLORING THE STRUCTURE AND BENEFITS OF MULTIMODAL LEARNING IN A COMPUTATIONAL MODEL OF SEMANTIC KNOWLEDGE **78**

- ABSTRACT 78
- INTRODUCTION 79
- AIMS 83

SIMULATION 2.1: DEVELOPMENTAL LEARNING 85
Method *85*
Training Stimuli 88
Training the model 90
Calculating representational economy in the semantic convergence zone 96
Testing the network and analysing representational economy 96
Results *97*
SIMULATION 2.2: ROBUSTNESS TO DAMAGE 101
Method *101*
Results *102*
SIMULATION 2.3: RELEARNING AFTER DAMAGE 106
Method *106*
Results *108*
Mild Lesioning 108
Moderate Lesioning 112
Severe Lesioning 116
DISCUSSION 121
Summary of Results *121*
Analysis of multimodal learning in the context of efficiency *122*
CONCLUSION 124
CHAPTER 3 - TECHNICAL REPORT ON THE REPLICATION AND RESCALING OF THE ROGERS ET AL. (2004) MODEL OF SEMANTIC MEMORY, AND SIMULATING SPONTANEOUS RECOVERY TO STABLE BASELINE 126
INTRODUCTION 126
SIMULATION 3.1: DEVELOPMENTAL LEARNING 128
Method *128*
Training the model *130*
Testing the network and analysing representational economy *131*
Results *131*
SIMULATION 3.2: SPONTANEOUS RECOVERY TO STABLE BASELINE 134
Method *134*
Results *138*
Minimal Lesioning 138
Mild Lesioning 141
Moderate Lesioning 144
Severe Lesioning 147
DISCUSSION 148
CONCLUSION 149
CHAPTER 4 - USING COMPUTATIONAL MODELS TO INVESTIGATE COGNITIVE REHABILITATION AFTER BRAIN DAMAGE 150
ABSTRACT 150
INTRODUCTION 151
AIMS 157
SIMULATION 4.1: 158
MANIPULATING THE LEARNING ENVIRONMENT IN REHABILITATION. 158
Method *158*
Training Stimuli 160
Training the model 161
Calculating representational economy in the semantic convergence zone 163
Testing the network and analysing representational economy 164
Results *164*
Rehabilitation After Minimal Lesioning 166
Rehabilitation After Mild Lesioning 170
Rehabilitation After Moderate Lesioning 174
Rehabilitation After Severe Lesioning 178
SIMULATION 4.2: 180

MANIPULATING THE LEARNING ENVIRONMENT AND THERAPY SET SIZE IN REHABILITATION FROM STABLE BASELINE RECOVERY THAT INCORPORATES BACKGROUND SPONTANEOUS RECOVERY. 180
Method *180*
Results *181*
Rehabilitation with a therapy set size of 5 and background spontaneous recovery after a severe lesion 183
Rehabilitation with a therapy set size of 10 and background spontaneous recovery after a severe lesion 187
Rehabilitation with a therapy set size of 20 and background spontaneous recovery after a severe lesion 191
SIMULATION 4.3: 195
MANIPULATING THE LEARNING ENVIRONMENT, THERAPY SET SIZE AND LESION SEVERITY IN REHABILITATION FROM STABLE BASELINE RECOVERY THAT INCORPORATES BACKGROUND SPONTANEOUS RECOVERY. 195
Method *196*
Results *196*
Rehabilitation with a therapy set size of 5 and background spontaneous recovery after a moderate lesion 198
Rehabilitation with a therapy set size of 10 and background spontaneous recovery after a moderate lesion 202
DISCUSSION 206
Summary of Results *206*
Simulation 4.1 206
Simulation 4.2 206
Simulation 4.3 207
CONCLUSION 209
GENERAL DISCUSSION 211
SUMMARISING THE FINDINGS AND THEIR IMPLICATIONS IN TERMS OF THESIS THEME 1 211
SUMMARISING THE FINDINGS AND THEIR IMPLICATIONS IN TERMS OF THESIS THEME 2 213
REVIEW OF THE FINDINGS IN THE CONTEXT OF THESIS THEME 1 215
REVIEW OF THE FINDINGS IN THE CONTEXT OF THESIS THEME 2 220
FUTURE DIRECTIONS 224
CONCLUSION 225
REFERENCES 227

Word count = 50081

List of Figures

Figure 0.1 Rogers et al.'s model of semantic memory 22
Figure 1.1 Architecture of the model adapted from Rogers et al. (2004) 42
Figure 1.2 Architecture of the model redrawn as a generic PDP model with bidirectional mapping between each layer 44
Figure 1.2.1 Representational Economy visualisation diagram......52
Figure 1.3 Mean variation in training performance during developmental learning 55
Figure 1.4 Mean variation in robustness to damage for the fully trained model with each training regime 58
Figure 1.5 Mean variation in training performance during relearning after a mild lesion...62
Figure 1.6 Mean variation in training performance during relearning after a moderate lesion 65
Figure 1.7 Mean variation in training performance during relearning after a severe lesion68
Figure 2.1 Architecture of the model adapted from Rogers et al. (2004) 86
Figure 2.2 Architecture of the model redrawn as a generic PDP model with bidirectional mapping between each layer 87
Figure 2.3 Mean variation in training performance during developmental learning 99
Figure 2.4 Mean variation in robustness to damage for increasingly severe lesions to the fully trained model 104
Figure 2.5 Mean variation in training performance during relearning after a mild lesion.110
Figure 2.6 Mean variation in training performance during relearning after a moderate lesion 114
Figure 2.7 Mean variation in training performance during relearning after a severe lesion 118
Figure 3.1 Architecture of the model adapted from Rogers et al. (2004) 129
Figure 3.2 Mean variation in training performance during developmental learning 133
Figure 3.3 Mean variation in training performance during spontaneous recovery to stable baseline after a minimal lesion......137
Figure 3.4 Mean variation in training performance during spontaneous recovery to stable baseline after a mild lesion......140
Figure 3.5 Mean variation in training performance during spontaneous recovery to stable baseline after a moderate lesion 143
Figure 3.6 Mean variation in training performance during spontaneous recovery to stable baseline after a severe lesion......146
Figure 4.1 Architecture of the model adapted from Rogers et al. (2004) 159
Figure 4.2 Architecture of the model redrawn as a generic PDP model with bidirectional mapping between each layer 160
Figure 4.3 Mean variation in training performance during rehabilitation after a minimal lesion 165
Figure 4.4 Mean variation in training performance during rehabilitation after a mild lesion 169
Figure 4.5 Mean variation in training performance during rehabilitation after a moderate lesion 173
Figure 4.6 Mean variation in training performance during rehabilitation after a severe lesion 177
Figure 4.7 Rehabilitation performance for therapy and control sets for set sizes of 5, 10 and 20 items with background recovery. Performance shown for each of the 10 individual models then the mean performance is given on the far right of each chart 182

Figure 4.8 Mean variation in training performance for a therapy set of 5 with background recovery during rehabilitation after a severe lesion 185
Figure 4.9 Mean variation in training performance for a therapy set of 10 with background recovery during rehabilitation after a severe lesion 189
Figure 4.10 Mean variation in training performance for a therapy set of 20 with background recovery during rehabilitation after a severe lesion 193
Figure 4.11 Mean variation in rehabilitation performance contrasting therapy sets of 5 or 10 items with background recovery after moderate and severe lesions.....................197
Figure 4.12 Mean variation in training performance for a therapy set of 5 with background recovery during rehabilitation after a moderate lesion 200
Figure 4.13 Mean variation in training performance for a therapy set of 10 with background recovery during rehabilitation after a moderate lesion...........................204

List of Tables

Table 1.1 Prototype patterns used to generate binary representations of verbal descriptions and visual features for each category of named object presented to the model for training 46
Table 1.2 Training regime patterns in terms of targeted output layers being manipulated between the multimodal and unimodal conditions....................47
Table 1.3 2x20 ANOVAs indicating significant performance differences between training regimes during developmental learning....................56
Table 1.4 2x11 ANOVAs indicating significant differences in robustness to damage between trained networks....................59
Table 1.5 2x20 ANOVAs indicating significant performance differences between training regimes after a mild lesion....................63
Table 1.6 2x20 ANOVAs indicating significant performance differences between training regimes after a moderate lesion....................66
Table 1.7 2x20 ANOVAs indicating significant performance differences between training regimes after a severe lesion....................69
Table 2.1 Prototypes patterns used to generate binary representations of verbal descriptions and visual features for each category of named object presented to the model for training 90
Table 2.2 Training regime patterns. Arrows indicate order of trial presentation for focused unimodal training. Multimodal training trials are presented in random order....................94
Table 2.3 2x20 ANOVAs indicating significant performance differences between training regimes during developmental learning....................100
Table 2.4 2x11 ANOVAs indicating significant differences in robustness to damage between trained networks....................105
Table 2.5 2x20 ANOVAs indicating significant performance differences between training regimes after a mild lesion....................111
Table 2.6 2x20 ANOVAs indicating significant performance differences between training regimes after a moderate lesion....................115
Table 2.7 2x20 ANOVAs indicating significant performance differences between training regimes after a severe lesion....................119
Table 3.1 2x20 ANOVAs indicating significant performance differences between training regimes during developmental learning....................134
Table 3.2x20 ANOVAs indicating significant performance differences between training regimes after a minimal lesion....................138
Table 3.3 2x20 ANOVAs indicating significant performance differences between training regimes after a mild lesion....................141
Table 3.4 2x20 ANOVAs indicating significant performance differences between training regimes after a moderate lesion....................144
Table 3.5 2x20 ANOVAs indicating significant performance differences between training regimes after a severe lesion....................147
Table 4.1 2x20 ANOVAs indicating significant performance differences between training regimes during rehabilitation after a minimal lesion....................166
Table 4.2 2x20 ANOVAs indicating significant performance differences between training regimes during rehabilitation after a mild lesion. 170
Table 4.3 2x20 ANOVAs indicating significant performance differences between training regimes during rehabilitation after a moderate lesion. Zero comparisons indicated by ……....................174

Table 4.4 2x20 ANOVAs indicating significant performance differences betweem training regimes during rehabilitation after a severe lesion. Zero comparisons indicated by ……..178
Table 4.5 2x20 ANOVAs indicating significant performance differences between training regimes for a therapy set of 5 during rehabilitation after a severe lesion 186
Table 4.6 2x20 ANOVAs indicating significant performance differences between training regimes for a therapy set of 10 during rehabilitation after a severe lesion 190
Table 4.7 2x20 ANOVAs indicating significant performance differences between training regimes for a therapy set of 20 with background recovery during rehabilitation after a severe lesion..194
Table 4.8 2x20 ANOVAs indicating significant performance differences during rehabilitation between training regimes for a therapy set of 5 with background recovery after a moderate lesion 201
Table 4.9 2x20 ANOVAs indicating significant performance differences between training regimes for a therapy set of 10 with background recovery during rehabilitation after a moderate lesion 205

Abstract

THE UNIVERSITY OF MANCHESTER

ABSTRACT OF THESIS submitted by Solomon Eyoka Nte
for the degree of Doctor of Philosophy
and entitled "Parallel Distributed Processing (PDP) models as a framework for designing cognitive rehabilitation therapy"
Month and Year of Submission – August 2014

Parallel Distributed Processing (PDP) modelling has simulated developmental learning across a range of domains such as reading (e.g. Seidenberg & McClelland,1989) or Semantics (e.g. Rogers et al. 2004). However aside from two notable exceptions (Plaut, 1996; Welbourne & Lambon Ralph, 2005b) modelling research has not addressed the simulation of relearning during spontaneous recovery or rehabilitation after brain damage, and no research has considered the effect of the learning environment. This thesis used an established PDP model of semantic memory (Rogers et al., 2004) to simulate the influence of the learning environment. A novel quantitative measure (called representational economy) was developed to monitor efficiency during learning. Developmental learning is considered to be multimodal (e.g. Gogate et al., 2000) whereas rehabilitation is normally carried out through therapy sessions employing unimodal learning tasks (Best & Nickels, 2000). This thesis hoped to discover whether multimodal rehabilitation may be more efficient (as suggested by Howard et al., 1985). Three sets of simulations were conducted: The first set contrasted multimodal and unimodal learning in development and recovery, and tested internal representations for robustness to damage finding multimodal learning to be more efficient in all cases. The second set looked at whether this multimodal advantage could be approximated by reordering unimodal tasks at the item level. Findings indicated that the multimodal advantage is dependent upon simultaneous item presentation across multiple modalities. The third set of simulations contrasted multimodal and unimodal environments during rehabilitation while manipulating background spontaneous recovery, therapy set size and damage severity finding a multimodal advantage for all conditions of rehabilitation. The thesis findings suggest PDP models may be well-suited to predicting the effects of rehabilitation, and that clinical exploration of multimodal learning environments may yield substantial benefits in patient-related work.

Declaration

No portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Copyright Statement

and/or Reproductions.

iv. Further information on the conditions under which disclosure, publication and commercialisation of this thesis, the Copyright and any Intellectual Property and/or Reproductions described in it may take place is available in the University IP Policy (see http://documents.manchester.ac.uk/DocuInfo.aspx?DocID=487), in any relevant Thesis restriction declarations deposited in the University Library, The University Library's regulations (see http://www.manchester.ac.uk/library/aboutus/regulations) and in The University's policy on Presentation of Theses

Dedication

This thesis is dedicated to the memory of my late father Sampson Onuwoyat Nte (1936 – 2014)

Acknowledgement

The research conducted in this PhD was supported by a Cognitive Foresight programme grant (EP/F03430X/1).

About the Author

Solomon Eyoka Nte graduated from Keele University in 1995 with a BSc (Hons) in Mathematics and Philosophy. He subsequently gained an MSc in Digital Musc Technology (Software Development) in 1996, after which he spent a year teaching secondary mathematics. From 1998 to 2010 he worked for Keele University: From 1998 to 2007 he worked full-time in the Psychology Department as a software developer for research and teaching. During this period he also lectured part-time in audio software development for music composition in the Music Department. In 2005 he gained a PgCert in Social and Cultural theory from Staffordshire University having studied part-time while working. From 2007 to 2010 he switched to working on a freelance basis for Keele University in both the Psychology Department and the Research Institute of Life Course Studies, developing software for a variety of research projects. In 2010 he registered for a full-time PhD in Psychology at Manchester University during which time he also worked as a psychology teaching assistant. He is also a keen musician playing free jazz, and other experimental music, in his spare time.

Introduction

The structure of the thesis

This thesis describes a series of computational investigations using a Parallel Distributed Processing (PDP) model of semantic memory to simulate the mechanisms that underpin rehabilitation from brain damage, and to understand how to maximise rehabilitation efficiency. The thesis is presented in alternative format and is structured into an initial introduction, 3 articles (Chapters 1, 2 and 4), a technical report (Chapter 3) on preparing a suitable computational model for the simulations described in Chapter 4. Finally the thesis finishes with a general discussion including some conclusions on the thesis findings and suggestions for future directions implied by those findings. Each of the 3 article chapters (i.e. 1, 2 and 4) is self-contained and provides its own introduction covering the literature relevant to that chapter and its discussion, though technical specifications for the model that have been previously described are only referred to in subsequent chapters, and not repeated.

The goal of this thesis was to use PDP models to understand which factors affect the efficiency with which the brain extracts, and internally represents, information from the learning environment during development, recovery and rehabilitation therapy. The thesis focuses on the role of the learning environment in terms of the sensory modalities targeted. Previous research using PDP models (e.g. Rogers et al., 2004), as well as experimental studies (e.g. Gogate et al., 2006), suggested developmental learning in childhood occurred in a multimodal learning environment. Similarly research into rehabilitation therapy for aphasia suggested that relearning after brain damage in adults may benefit from a multimodal learning environment (e.g. Howard et al, 1985) . This thesis considers whether particular learning environments facilitate learning more efficiently, and if so whether a measure of that efficiency could predict the learning outcome in development, recovery and rehabilitation. Clearly an efficiency measure for learning in computational models, that also possessed some predictive powers, would be useful in understanding and predicting patient rehabilitation behaviour through simulation. The thesis documents the development of a statistic (named representational economy) to measure changes in efficiency alongside learning performance during the time course of learning. This represents a first step towards developing subsequent simulations of rehabilitation that could usefully predict possible outcomes for patient-related interventions. Representational economy yields detailed information regarding the evolution of learning in terms of simulating and assessing the internal representations the brain develops to support learning.

This thesis explored how manipulating the multimodality of the model's learning environment, whilst monitoring representational economy within the model, would help to understand the evolution of efficiency in a variety of situations. Three hypotheses were explored: Firstly, that multimodal learning would be more efficient than unimodal learning in development, recovery and rehabilitation. Secondly, that multimodal learning could be approximated through rearranging the order of unimodal learning tasks so that tasks relating to the same item were always grouped together. Thirdly, that the model could be used to understand efficiency in the context of rehabilitation therapy delivered to patients, and thus predict strategies for improving therapy. The exploration of these hypotheses falls under two broad themes that are discussed in the following sections.

Thesis Theme 1: How do PDP Models extract efficient and effective representations from the learning environment?

This section looks briefly at the history of using PDP models to simulate behaviour. It goes on to consider learning and relearning in PDP models, subsequently considering effective representations in models, distinguishing between input /output representations and the development of internal representations to support learning. Finally the section discusses measuring the efficiency of internal representations and the degree to which this efficiency is dependent upon the learning environment.

Beginning with McClelland, Rumelhart and The PDP Research Group's (1986) seminal work, Parallel Distributed Processing, PDP Models offer testable software simulations of cognitive processes that are capable of independent learning. Implemented in software as neural networks, PDP models are layered collections of simple units, loosely analogous to neurons, possessing weighted connections for intercommunication (Callan, 1999). Traditional models of cognitive processes employed by psychologists have consisted of a "box and arrows" diagrammatic approach. These models are essentially flowcharts where the boxes contain some form of stored representation of the external world, and the arrows indicate the cognitive processes that map between different levels of representation. PDP models offer an advantage over these "box and arrows" models since PDP models genuinely learn (Welbourne & Lambon Ralph, 2005a). PDP models are particularly useful for understanding mechanisms of cognition since they provide testable simulations of how learning can occur in a neural network. Additionally, the architecture of a PDP model makes a concrete proposition about the structure of stored knowledge, and its functional arrangement within such networks. Examining these network architectures makes it immediately obvious which stored representations of the external world are being considered by the model, and the proposed mechanism by which those representations are extracted from the environment, function efficiently, and facilitate a particular cognitive process. For example a PDP model of the process of reading could consist of a network architecture for mapping between orthographic and phonological representations of words (e.g. Seidenberg & McClelland, 1989; Harm & Seidenberg,1999; Zevin & Seidenberg, 2002; Monaghan & Ellis, 2010).

There has been a substantial amount of research using computational models to provide accounts of the underlying mechanisms responsible for developmental learning (for a general introduction see Quinlan, 2003) as well as normal cognitive and language functions (for a general introduction see McLeod, Plunkett, & Rolls, 1998) . Indeed wherever possible PDP research has attempted to produce models whose performance can account for experimental observation, by considering the brain's internal representation processes. For example, learning in PDP models mimics experimental findings in developmental learning such as the vocabulary spurt (McLeod, Plunkett, & Rolls, 1998). Particular progress in producing PDP models of behaviour has been made in the domain of reading and semantics: Seidenberg and McClelland (1989) proposed a triangular general framework for lexical processing which became the basis for their own, and many subsequent, PDP models of linguistic tasks. Often referred to as the triangle model of reading, their framework suggests a bidirectional network architecture such that mapping between orthographic, phonological and semantic representations accounts for most types of linguistic processing. Some examples of such types of linguistic processing are mapping from orthography to phonology corresponding to reading aloud, and mapping from phonology to semantics corresponding to comprehension. Their initial implementation of the model consisted of simulating the orthography- phonology pathway, and thus visual word recognition and pronounciation.. Their orthographic and phonological representations were derived from a system of letter triples whereby MAKE would be orthographically represented as _MA, MAK, AKE, KE_ (where _ indicates beginning and ends of words) and phonologically represented as _ma, mak, ak_ This representation system was developed since merely having a single active unit for each letter would mean words with the same letters (e.g. BAT and TAB) would have the same representation. Thus the model tackles the problem of adequate representation systems for inputs and outputs though does not consider the nature of representations that develop at hidden layers to support mapping between modalities. Seidenberg and McClelland's (1989) initial exploration of the triangle model shows how a PDP model can offer an account of normal cognitive function: Their simulations account for differences between words in terms of processing difficulty, pronunciation of novel items, differences of word recognition skill in different readers, transitions from beginning to skilled reading, and differences in performance on lexical decision and naming tasks. They also suggest how impairments can arise from damage to normal function by considering dyslexia within the model. Plaut, McClelland, Seidenberg, and Patterson's (1996) model of single word reading (PMSP96) made considerable improvements to SM89. The main criticism of SM89 was that it performs poorly when compared to skilled readers at pronouncing nonwords (Besner, Twilley, McCann, & Seergobin, 1990). Plaut et al. (1996) overcame this issue by designing new representations, moving to an onset, vowel, coda structure for each monosyllabic word that more effectively captured the word structure such that the model's non-word reading improved. PMSP96 consequently yielded a PDP model of single word reading that had improved task performance, and as a consequence offered a more substantial account of acquired dyslexias than that suggested by Patterson, Seidenberg, and McClelland's (1989) original proposition. Subsequent work on modelling reading has also discussed the issue of choosing the right representations to capture the task being simulated, or more specifically refining representation structure to improve model performance. For example Monaghan et al. (2004) discuss representations split across input and hidden layers simulating the left and right visual fields in developing an account of hemispheric asymmetries in their split-foveal model of semantic processing. Similarly Chang et al. (2012b) developed letter representations derived from monochrome bitmaps images of each letter in an attempt to achieve greater realism in modelling reading. The increase in realism comes from providing input representations at the level of visual input; Previous orthographic representations (e.g. those used in PMSP96) have been provided at a level that already assumes some type of processing has occurred so orthographic representations are not learned from visual input but instead are learned after a structure has been imposed upon them that makes assumptions about how words are processed visually. This level of representation also shows how the level at which input/output representations are provided can influence, and is indeed directly responsible for, the internal representations that develop as a result of learning.

Many early PDP models of linguistic processing provide accounts of the mechanisms behind tasks involving a single mapping between two modalities, These models show how statistical regularities in the learning environment can be extracted and applied to generalise to new knowledge. However these models do not offer insights into how effective representations are extracted from the learning environment. In terms of considering representations, most models tend to focus upon developing and refining input/output representation structures and do not focus their analyses towards understanding the development of the learnt representations in the hidden layers, or what makes them effective in supporting learning.

These early PDP models are often similar in structure to Seidenberg and McClelland's

(1989) PDP model. Such models include (but are not limited to): Hinton and Shallice's (1991) model of acquired dyslexia;Plaut and Shallice's (1993) models exploring the nature of Deep Dyslexia; Mayall and Humphreys (1996) model of Alexia; Plaut, McClelland, Seidenberg, and Patterson's (1996) model of Word Reading; Plaut's (1997) models of word reading and lexical decision; Cree, McRae and McNorgan's (1999) model simulating semantic priming; Harm and Seidenberg's (1999) model of reading acquisition and dyslexia; Lambon Ralph, McClelland, Patterson, Galton and Hodges' (2001) model of object naming and semantic impairment; Gotts and Plaut's (2002) model of Semantic Impairment; Zevin and Seidenberg's (2002) model of age of acquisition effects in word reading.

Seidenberg and McClelland's (1989) triangular framework for lexical processing contains bidirectional connections between domains. In other words mapping between each of the domains can occur in both directions. Although Seidenberg and McClelland do not implement this bidirectionality, later models have done so with important consequences: Hinton and Shallice's (1991) use of recurrent bidirectional connections between semantics and "clean-up" units allowed the model to capture the semantic similarity of words. Only models with full bidirectional connectivity between all domains can allow attractor states to develop across the whole model that capture similarity relations present in initial representations across different modalities. However there is an important difference between internal representations that support a single mapping between two domains and representations that support mapping between multiple domains and as such capture the similarity structure between multiple connected domains. These similarity relations form the basis of the degree of efficiency within the internal representations that are extracted from the learning environment. When considering bidirectional mappings between several different sensory modalities efficiency can be seen as the degree of convergence to a single robust internal attractor state at the hidden layer which can handle multiple cross-modal mappings. Most bidirectional models do not consider these internal representations however it is through the development of bidirectional models, and thus truly interactive networks, that a greater understanding of efficiency can be considered. It is thus worth identifying that there are a number of bidirectional models that build substantially upon the type of proposed framework offered by Seidenberg and McClelland (1989) and lay the groundwork for the consideration of bidirectionality that supports the modelling work in this thesis: Farah and McClelland's (1991) model of semantic memory impairment; Devlin, Gonnerman, Andersen, and Seidenberg's (1998) model of category specific semantic impairment; Joanisse and Seidenberg's (1999) model of impairments in verb morphology; Lambon Ralph and Howard's (2000) model of anomia and impaired verbal comprehension; McLeod, Shallice and Plaut's (2000) model of attractor dynamics in word recognition; Plaut's (2002) model of optic aphasia; Harm and Seidenberg's (2004) model looking at the division of labour in word reading; Rogers, Lambon Ralph, Garrard, Bozeat, McClelland, Hodges, and Patterson's (2004) model of semantic memory; Dilkina, McClelland, and Plaut's (2008) model of semantic and lexical impairment; Smith, Monaghan and Huettig's (2014) model of an amodal shared resource.

There has been considerable focus on refining the nature of the externally provided input/output representations, for example the development of orthographic representations in reading models as discussed earlier in the context of implementations of the triangle model of reading. Rogers et al. (2004) describe this type of representation process in the context of developing patterns of activation, across a layer of units, that can represent the visual properties of objects in the external world:

"Each unit in the visual layer represents a unique visual property, such as "has limbs" or "is round". Visual representations of objects correspond to patterns of activity across this assembly of visual features, such that objects with similar visual appearances give rise to similar visual representations " (Rogers, Lambon Ralph, Garrard, Bozeat, McClelland, Hodges, & Patterson, 2004: p207).

Very little modelling work however has been done on the internal representations generated at the hidden layers of models as a result of learning. The Rogers et al. (2004) model of semantic memory (see Figure 0.1 for a diagram of its architecture) uses emergent amodal internal re-representations within a model's hidden layer to provide a convincing account of semantics. This model is well suited to the explorations that this thesis wanted to carry out for the following reasons: Firstly, the model is fully bidirectional and has several input/output modalities with semantics as the intermediate hidden layer acting as a convergence zone. Secondly, the analysis Rogers et al. (2004) carried out showed that the internal representations at the hidden layer captured a categorical structure that develops as a result of learning that could possibly account for semantic memory in terms of the learning process itself. Finally, the model simulated developmental learning, and also acquired adult disorders through replicating the patterns of the impaired performance observed in patients on particular semantic tasks. Whilst the effectiveness of the representations the model extracts can be measured in term of performance accuracy, the Rogers et al. (2004) analysis suggested the basis for this thesis to develop the idea that efficiency could be measured as the Euclidean similarity between the internal representations generated by the trained model for the multiple cross-modal mappings belonging to the same item. This thesis fills a gap in the literature, as no modelling work has been done to examine the factors in the learning environment, that affect the efficiency of the internal representations that the model develops, in terms of how well they facilitate mapping between input and output representations.

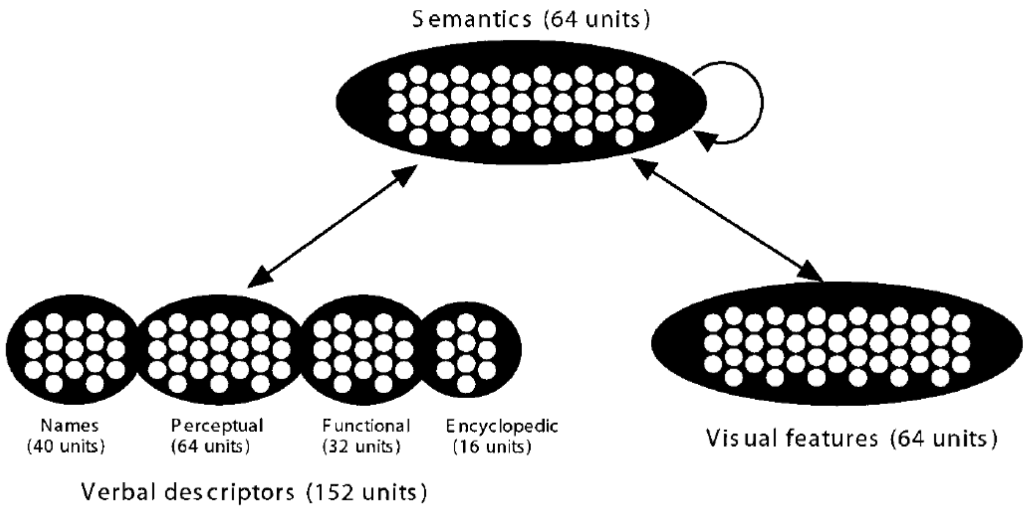

Figure 0.1 Rogers et al.'s (2004) model of semantic memory

Within the thesis the learning environment is considered in terms of the degree of multimodality it possesses as well as the nature of that multimodality. This thesis develops a novel measure of efficiency (representational economy) for the internal representations of PDP models in the context of the structure of increasing differentiation between internal representations of learned items that the model develops as a result of the learning process. This thesis then applies this measure to developmental learning, robustness to damage of the fully trained model and recovery/rehabilitation relearning after damage for a range of learning environments.

The idea that learning requires a multimodal environment arises from a variety of sources. Experimental studies with pre-verbal infants, have shown mothers use temporally synchronous naming when presenting objects or actions for their child to learn (Gogate et al., 2000; Messer, 1978). Similarly experimental work by Sullivan & Horowitz (1983, p210) suggests infant learning of word–object relations is "accomplished in the context of the infant's multimodal interactions with its mother" . The thesis explores this by contrasting multimodal and unimodal learning, thus considering the role of the learning environment on learning efficiency in terms of the modalities targetted during learning episodes and learning environments across the lifespan, development vs recovery/rehabilitation (i.e. learning vs relearning) to see if efficiency is constant across all situations.

Thesis Theme 2: How can PDP models be used to understand recovery and rehabilitation for impairments in the context of patient-related work?

This thesis considers recovery and rehabilitation in relation to patient-related findings in spontaneous recovery and rehabilitation for aphasia. Specifically attempting to simulate issues surrounding therapeutic intervention in anomia. In order to understand how PDP models can simulate aphasic behaviours it is worth briefly considering what constitutes aphasia and aphasia therapy.

Aphasia is generally defined (e.g. Ashcraft, 1989) as a loss of language abilities resulting

from a brain lesion. Such lesions are normally the result of stroke, head injury or other causes such as dementia (Tesak & Code, 2008). Many different types of aphasia have been identified beginning with the two most famous: 1) Broca's aphasia, classified by neuroanatomist Paul Broca (1861) - characterised by minimal speech ability though virtually normal comprehension, and a lesion to a part of the brain in the left inferior frontal gyrus, which has become known as Broca's area. 2) Wernicke's aphasia, classified by neurologist Carl Wernicke (1874) - characterised by jumbled, incoherent speech, impairment to comprehension and repetition ability, and a lesion to an area of the brain surrounding the posterior portion of the superior temporal gyrus, which has become known as Wernicke's area. Three other aphasias types are also relevant to this discussion: 1) Anomia, which is an inability to name objects. 2) Acquired dyslexia, which is a reading impairment resulting from brain lesions. 3) Semantic Dementia, which is a progressive impairment of semantic ability in cognitive tasks such as object naming (Snowden, 1989). It is worth noting that aphasic impairment normally affects ability in multiple, as opposed to single, linguistic tasks, and that although impairments are related to the location of the lesion the exact nature of the impairment is very hard to predict solely on the basis of knowledge of the lesion location. Indeed Geva et al. (2011) show that though identifying lesion-symptom relationships can be done with imaging techniques of Voxel Based Morphometry (VBM) or Voxel-based Lesion Symptom
Mapping (VLSM), the results of each process do not always concur making impairment prediction difficult. Perhaps similarly important for this thesis is the idea that "regardless of their classical diagnostic classification, almost all individuals with aphasia show significant impairment on tasks of single word production, such as picture naming" (Wilshire, 2008: p1020). Accordingly, many theories of aphasia rehabilitation focus on these kind of single word production tasks.

Speech and Language therapists who deliver aphasia therapy use sets of particular tasks to both assess the nature of a patient's aphasia and decide on an appropriate therapy strategy. The stages in this process can be summed up as follows: "Determine the level of breakdown in language processing, identify a therapy that is appropriate for that level of breakdown, predict the pattern of change, and obtain these results" (Best & Nickels, 2000). The level of breakdown is often assessed using test batteries such as PALPA (Psycholinguistic Assessments of Language Processing in Aphasia (Kay, Lesser & Coltheart, 1992)). Following this type of assessment, therapists often employ further diagnostic tools such as psycholinguistic models of language processing (e.g. Ellis & Young, 1988) to select a therapeutic strategy (i.e. an appropriate set of linguistic tasks), according to the processing modalities they wish to target. Therapeutic tasks can include, but are not limited to, picture naming, word-to-picture matching, word repetition, reading aloud, phonemic cueing, listening and writing (Howard & Hatfield, 1987). Predicting the pattern of change, as a result of therapy, is a difficult process (Wilshire, 2008). Much research into the effectiveness of particular therapeutic strategies comes from case studies, yet not all patients respond in the same way to a particular therapeutic task, even if those patients possess similar deficits (Best & Nickels, 2000). Due to these difficulties it is particularly desirable to obtain both measures, and predictive models, of the efficacy of aphasia therapies, especially if they could be understood in terms of the way in which the brain re-establishes internal representations.

Several researchers (e.g. Howard, Patterson, Franklin, Orchard Lisle, & Morton, 1985) have argued the benefits of a "multimodal" approach to therapy. This "multimodal" approach refers to the idea that "a word should be elicited in any one of a variety of ways -

repetition, phonemic cueing, reading, writing or listening - and their use practised in a variety of situational and grammatical contexts" (Howard et al, 1985: p52). In this sense multimodal refers to targeting multiple sensory modalities during the therapeutic process in a manner similar to developmental learning. The rationale behind multimodal therapy is that more durable rehabilitation requires patients to access semantic representations (Howard et al., 1985) of the words, or pictures, they are working with. Thus multimodal therapy measures efficacy as overall performance on a range of single tasks across multiple sensory modalities, and this efficacy is dependent upon rehabilitating access to semantic representations. Despite this many, if not most of the reported cases in the literature use unimodal therapy interventions. Predictive models need to focus on issues in patient-related work that can be easily identified and measured: One such issue is therapy set size which has been explored in some clinical studies but not modeled. Therapy set size is the number of unique items in verbal, pictorial or written form (e.g. dog, car, apple etc.) used for learning during a therapy session. Clinical data on studying the manipulation of therapy set size has been obtained by Snell, Sage and Lambon Ralph (2010) whose meta-analysis of set size within the therapy literature points towards set size being a variable that could be usefully modelled with a view to predicting the variation in therapy efficacy as a function of varying set size. Snell, Sage and Lambon Ralph (2010) also recruited a case-series of aphasic patients with varying severity and gave all of them the same naming therapy which varied in two levels of therapy set size in order to explore the role of set size that their meta-analysis suggested. Since understanding learning during rehabilitation through simulation is one of the goals of this thesis it explores the role of manipulating set size in models of rehabilitation therapy. Therapy set size was chosen as a variable of particular interest since clinical data is available for contrast and comparison with the results of the simulations.

PDP models are well suited to simulating the symptoms of brain damage because they can be selectively and incrementally damaged and their performance under damage shows graceful degradation. Early PDP models (e.g. Rumelhart & McClelland, 1986) focussed on providing accounts of possible mechanisms for a variety of normal cognitive functions but in trying to account for normal functions an understanding of impairments arose (e.g Seidenberg and McClelland, 1989) end up with results that also suggest accounts of developmental and acquired dyslexia). Accounting for neuropsychological symptoms with PDP models thus grows with later modeling work (e.g. Plaut & Shallice, 1993 or Dell, Schwartz, Martin, Saffran & Gagnon, 1997 looked at how such models could be adjusted to simulate specific cognitive impairments). PDP Models of normal function can be "lesioned" (i.e. selectively impaired)by removing units, weighted connections or adjusting other parameters such as weight decay. A model "lesioned" in this manner can display an impaired performance, on its intended mapping, in such a way that the pattern of errors made resembles that of patients with particular types of brain lesion. Indeed several authors have focussed on lesioning models to simulate particular types of cognitive impairments, for example deep dyslexia (Plaut & Shallice, 1993), and their rehabilitation (e.g. Welbourne & Lambon Ralph, 2005b) .

All PDP models that simulate aphasic behaviour, do so via simulating brain damage by selectively lesioning models of normal function in one of the manners described above. In considering a model's performance under damage, it is usual to look at the pattern of error and compare it to the pattern of error in patients believed to possess analogous brain damage. However this is only possible where clinical data is available for comparison. For example Rogers et al. (2004) compared patient and model performance for confrontation naming, sorting words and pictures, word to picture matching and drawing and delayed copying. Simulating impairments is thus a two way process with the ability to model, and thus account for, patient error patterns as well as looking at recovery of damaged models to account for patient's cognitive recovery and rehabilitation. This thesis considers the process of comparing the results of simulations with clinical data in discussing the results from modeling therapy set size in Chapter 4.

The fact that PDP models learn means they should be useful tools for modelling post stroke recovery. The process of relearning in PDP models can mirror both the process of cognitive recovery and cognitive rehabilitation. At this point it is worth considering the distinction between the two processes. Recovery is generally speaking relearning that occurs as a result of plasticity, that is the ability of the brain to compensate for, and regain, function previously carried out by other tissue if that other tissue has been damaged. An initial discussion of plasticity and the PDP modelling process is presented by Munro (1986): In damaged PDP models recovery is accomplished by allowing the model to relearn, using the same training set as that which resulted in its original (pre-damage) learning but without changing other parameters (Welbourne and Lambon Ralph 2005b). This is analogous to patient recovery where the patient regains cognitive function merely as a result of normal living without specific therapeutic intervention. Welbourne and Lambon Ralph (2005b) suggest that a post-damage learning period observed in PDP models might be equivalent to the period of spontaneous recovery seen in aphasia patients. Indeed other work by these authors (Welbourne & Lambon Ralph, 2005a, 2007) has specifically explored the role of plasticity related recovery in PDP models and how that process relates to patient data on recovery as well as distinguishing between recovery and rehabilitation. Rehabilitation is accomplished within PDP models as intense exposure to a subset of the original items the model learnt by using a higher learning rate (Welbourne and Lambon Ralph, 2005b) than that with which the model was originally trained. This manner of simulating rehabilitation has been followed in this thesis. Hardly any PDP models of rehabilitation exist in the literature and those that do are all in the domain of reading (the main works being Plaut, 1996 and Welbourne & Lambon Ralph, 2005b). Earlier in this review it was stated that efficacy, in aphasia therapy, is considered to be "demonstrable improvement on a specific [linguistic] task" (Best & Nickels, 2000: p232). Thus in order to parallel therapy the few existing PDP models of rehabilitation have focussed on relearning specific tasks (for example Sejnowski & Rosenberg, 1987; Plaut, 1996 or Welbourne & Lambon Ralph, 2005b). Welbourne and Lambon Ralph (2005b: p804) specifcally refer to this in their model stating "rehabilitation is limited to therapy on the same task on which we wish to improve performance". For example picture naming therapy can be modelled by training on representations of visual features of an object, as discussed in Rogers et al. (2004). Whilst they have not modelled rehabilitation some authors (e.g. Dell, Martin & Schwartz, 2007 or Abel, Willmes & Huber, 2007) have considered the possibility of predicting therapeutic outcome, from models of post-damage performance on therapeutic tasks. This work addresses Best and Nickels (2000) concerns regarding the difficulty of predicting therapy outcome yet such models do not explicitly consider the role that variation in the regimes of retraining may play. This thesis prioritises considering the role of the retraining regime in accounting for recovery and rehabilitation. Plaut (1996) looks at relearning and generalization in a small model of the relationship between orthography-semantics that derives from Seidenberg and McClelland's (1989) triangle model of reading. Plaut's methodology really just considers recovery as retraining. It is performed on all words and there is no attempt to emulate clinical selection of items for relearning or the intense exposure to items for relearning identified and implemented by Welbourne and Lambon Ralph (2005b).

In all of the above cases PDP models have measured efficiency as performance accuracy, in the same way as in aphasia therapy, but the models allow for efficiency to be unpacked more subtly in terms of the internal representational structure the model develops as a result of relearning. This idea suggests a new way of conceiving of the therapy process in terms of the type of representational structure therapists are seeking to re-establish in patients who have suffered brain damage. A key question arising from the aphasia literature is that while some authors recommend multimodal therapies most reported studies are unimodal or at least unimodal in terms of the output that is required from the patients. This thesis considers whether therapies that required multimodal outputs might be more efficient and effective.

This thesis intended to explore possible therapeutic interventions by simulating underlying cognitive mechanisms believed to occur during recovery and rehabilitation in patients, and consider factors affecting their efficiency in simulation, with a view to predicting responses to different kinds of therapy intervention. Since it is possible to perform precise manipulations in models that are not possible in patients it is hoped that these models could then be used as part of the design process in creating therapeutic interventions. More generally the aim was to understand how the learning environment can be manipulated to maximise learning efficiency in a range of circumstances. The thesis investigated the effect of the learning environment on recovery and rehabilitation by measuring the efficiency (as representational economy) of the internal representational structure that develops. In addition it looks at issues of specific interest in aphasia therapy such as the number of items used in therapy (therapy set size) again in the context of the effect of the learning environment and at various levels of damage.

Organisation of the thesis

The thesis seeks to answer the following two questions for learning occurring in different life stages: development; spontaneous recovery and rehabilitation.

1) Is multimodal learning more efficient than unimodal learning? 2) Can multimodality be approximated as sequential presentations of the same item in multiple modalities or must it be simultaneous presentations in multiple modalities?

For rehabilitation learning there is an additional question. What effect does varying therapy set size have on learning efficiency in the context of the efficiency of the learning environment and the severity of damage?

As previously described this thesis is submitted in alternative format. This introduction is followed by three self-contained article chapters (1, 2 and 4) documenting a range of simulations that attempt to answer the above questions: Chapter 1 compares multimodal and unimodal learning in terms of measuring efficiency for development, robustness to damage and recovery, and finds the multimodal learning environment to be more efficient. Chapter 2 explores whether the multimodal advantage found in Chapter 2 could be approximated by reordering unimodal learning. Chapter 3 is a technical report detailing the rescaling of the model used in Chapters 1 and 2 ready for use as a suitable starting point in simulating rehabilitation, whilst preserving the findings of Chapter 1. Chapter 4 compares multimodal and unimodal learning in terms of efficiency for rehabilitation. Rehabilitation is also considered in terms of the role of therapy set size, and the degree to which background recovery contributes to rehabilitation performance, for a range of damage severity. The thesis ends with a general discussion of the findings and conclusions reported in Chapters 1, 2 and 4 as well suggestions for future research directions that arise from the work reported in this thesis.

Acknowledgement of contribution from other authors

My supervisors, Dr. Stephen Welbourne and Professor Matthew Lambon Ralph have contributed to some aspects of the design and analysis of the simulations described in this thesis, as well as assisting with the thesis production itself.

Chapter 1 - Knowledge acquisition and representational economy: A computational investigation of efficiency in the brain's 'convergence zones'

Abstract

Various researchers have suggested that knowledge acquisition is dependent upon the convergent interaction of modality-specific perceptual representations (e.g. Damasio, 1989; Tranel et al., 1997; Meyer & Damasio, 2009). It has been suggested that such interaction occurs in the brain's 'convergence zones', where internal representations develop that mediate the association of perceptions arising from different sensory modalities. This study uses an established model of semantic knowledge (Rogers et al., 2004) to investigate temporal and representational efficiency in the convergence process, and to determine how variation in the learning environment of these convergence zones affects the acquisition of novel knowledge, and the re-acquisition of knowledge after neural network damage. The model was trained to associate representations of objects in three domains (names, verbal descriptions and visual features) across two perceptual modalities (verbal and visual): names, verbal descriptions and visual features. During training, a statistical measure of the efficiency of convergence (representational economy) was used to monitor how well the internal 'semantic' representations supported the model's knowledge (i.e. learned associations across the three domains). The findings suggested that the learning environment had a direct effect upon developing convergence: multimodal learning (targeting all possible domains at each presentation of an object) yielded greater efficiency, being faster and developing more convergent semantic representations, than unimodal learning (targeting each domain separately). This was true for both initial learning and relearning after damage. In addition, multimodal learning yielded representations that were more robust to damage. However, when multimodal and

unimodal training was compared on a trial by trial basis (contrasting network performance after each presentation of a single training pattern), as well as in terms of equated error (contrasting network performance after each presentation of all the training patterns) across all training items, multimodal learning was not found to be more efficient than unimodal learning when compared in terms of equated error. This suggests that the greater efficiency of the multimodal learning environment results from the higher error signal obtained for each object presentation during learning. These results suggest a multimodal learning environment should offer greater efficiency in the re-acquisition of knowledge after brain damage (e.g. stroke, head injury), with implications for theories of learning in development and rehabilitation.

Keywords: Convergence Zones; PDP neural network models; Semantic Knowledge; Representational Economy; Multimodal learning

Introduction

Both historical and contemporary neuroscience literatures discuss areas in the brain that merge semantic or conceptual information: Almost a century after Wernicke's (1874) pioneering work on sensory association, Damasio's (1989) proposed 'convergence zones' in the brain offer a contemporary perspective that encapsulates Wernicke's idea as specific regions acting to draw information together. Damasio (1989: p. 130) suggested that "it is not enough for the brain to analyze the world into its components parts: the brain must bind together those parts that make whole entities and events, both for recognition and recall". Various authors, often including Damasio, have engaged in subsequent discussion of convergence zones (e.g. Damasio et al., 1996; Tranel et al., 1997; Moll & Miikkulainen,

1997) and their application in different domains (for an overview see Meyer & Damasio, 2009). Convergence zones can thus possibly account for how the brain binds incoming information from various perceptual modalities (e.g. sensory, motor or linguistic), and how such binding supports both perception and recall of external world experiences. Given that conceptual knowledge is based upon incoming information, and derived from perception and action (e.g. Barsalou, 1999; Gallese & Lakoff, 2005), this information can be categorised and understood in terms of modality-specific perceptual representations (Rogers et al., 2004). Convergence suggests that these modality-specific perceptual representations are then processed, for storage in the brain, into an abstracted intermediate re-representation independent of any modality-specific qualities (Rogers et al., 2004). These intermediate representations can be viewed as storage structures for the learned cross-modal association of modality-specific perceptions that support coherent generalisable concepts (Lambon Ralph et al., 2010). Subsequent perception in a single modality can stimulate the recall of associated information arising from other modalities via this amodal intermediate representation. For example, an intermediate representation could support the conceptual knowledge of a "dog" by facilitating the learned cross-modal associations between perception of a real dog, a picture of a dog, the sound of a dog barking, the phonology of the word "dog", and the orthography of the word "dog". The convergence zone facilitates the binding of these various "dog" perceptions within an intermediate amodal representation. Any one of these modality-specific "dog" perceptions can trigger the others via the learned association encoded within the intermediate representation: If we hear a barking sound we imagine it is being produced by a furry creature with four legs, which we refer to as a dog.

Whilst Damasio suggests the concept of convergence zones he does not describe their function in great detail. Several authors have offered more substantial descriptions for the function of convergence zones (e.g. Rogers et al.. 2004, Lambon Ralph & Patterson, 2008; Lambon Ralph et al., 2010) though little is known regarding how convergence zones operate mechanically at the neuronal level, and what factors within the learning environment affect their efficiency. Furthermore it is only through computational modelling that neuronal behaviour in convergence zones can be explored in a controlled and systematic manner by simulating analogous situations to those cognition mechanisms believed to exist in the brain (e.g. Moll and Miikkulainen, 1997). If the operation of convergence is to be understood, it is essential to know how, and why, it develops during the learning process of knowledge acquisition. This Chapter does not seek to concern itself with the debates regarding the neuroanatomical location and nature of convergence zones (for an overview see Patterson et al, 2007). Instead the intention here is to focus upon the computational role of convergence in the development of knowledge, and thus explore the function of brain areas that merge information and, in doing so, facilitate stored conceptual knowledge. The act of perception in any sensory modality generates activity in the brain's neural networks, which can be modelled in a suitably designed computational 'neural' network. Parallel Distributed Processing (PDP) is one class of computational simulation, that allows the investigation of these computational neural networks. PDP models offer an advantage over other forms of computational simulation since the models actually learn (Welbourne & Lambon Ralph, 2005b). Using PDP models to simulate the performance of convergence zones through initial learning and recovery following damage is a powerful method to learn more about their function (Welbourne & Lambon Ralph, 2005b). Indeed Damasio (1989, p 130) suggested that "modeling studies should illuminate the collective properties of convergence zones and provide us with the intuition we need to sharpen our questions".

This Chapter attempts to do just that by investigating how, and why, convergence zones may develop to support knowledge, and what possible advantages they may offer. It does this by examining their efficiency in direct relation to the learning environment under which they developed. The Chapter adopts a view of knowledge as the merged internal re-representation of multiple modality-specific perceptual representations. This view of knowledge derives from previous modelling literature (Plaut, 2002; Rogers et al., 2004; Dilkina et al, 2008) exploring neural network architectures containing groups of hidden units amongst which convergent representations develop to facilitate learning. The Rogers et al. (2004) model of semantic knowledge was chosen as the most suitable starting point to consider the role of convergence for two reasons: first, because it is an established model of semantic knowledge acquisition, which is a domain that is likely to depend heavily on the convergence of cross-modal information; and secondly, because it already contained some explicit investigation of how convergent representations develop, facilitate learning and act as stored knowledge. The Rogers et al. (2004) model implements a convergence zone that supports the acquisition of conceptual knowledge (semantic memory). Indeed, it provides convincing evidence that convergence zones are required to enable cross-modal associations. Within the convergent zone, amodal semantic representations develop during learning to support the cross-modal association of representations of the visual features and verbal descriptions from a variety of objects. This model also produces an internal category structure, arguing for an emergent view of semantic memory in terms of the representational structure that develops. Additionally when damaged, the model convincingly simulates the behavioural patterns of semantic dementia (Rogers at al., 2004).

The Rogers et al. (2004) model also contains explicit information about its training regime. For PDP models, of this type, the training regime is designed to replicate key features of the external sensory-verbal environment. What then, is the role of learning environment, in terms of convergence? If convergence is a necessary condition for the generation of knowledge representations, then what aspects of sensory experience encourage it? If concepts are being learnt through exposure to sensory impressions, what process binds those incoming impressions, and what is the role of convergence in this binding? To derive an experimental situation where these questions can be answered it was necessary to develop a suitable manipulation of the learning process that would be expected to interact with the convergent architecture to either encourage or discourage the development of convergent representations.

Developmental psychology offers the concept of epigenesis, that infant learning is the result of the interaction between genes and the learning environment (Mareschal et al., 2004). On this view the PDP model architecture could be understood as analogous to the genetic material, and the training regime analogous to the learning environment. Several developmental studies, with pre-verbal infants, have shown that mothers use temporally synchronous naming when presenting new objects or actions that they wish the child to learn (Gogate et al., 2000; Messer, 1978). "During temporally synchronous naming, they [Mothers] speak a word while holding and moving an object rather than moving it out of phase with the spoken word" (Gogate, Bolzani & Betancourt, 2006: p.261). This synchronous presentation to multiple sensory modalities (i.e. multimodal stimulation) follows from the idea that infant learning of word–object relations is "accomplished in the context of the infant's multimodal interactions with its mother" (Sullivan & Horowitz, 1983, p. 210). The evidence from such developmental studies suggests that infant knowledge acquisition is facilitated by a multimodal learning environment. The Rogers et al. (2004: p215) model implements a training regime (learning environment) where “target values were applied across all visual and verbal units in the model, including the units acting as input during the trial”. Clearly this temporally synchronous multimodal presentation, of target values for learning, has much in common with evidence from studies of infant learning. As such it provides an established basis for examining convergence in initial (developmental) learning, as well as illustrating how learning knowledge is extracted from a learning environment from a modelling perspective. It would thus seem reasonable that a reduction in the degree of multimodality within the training regime would affect the degree of convergence that develops. This manipulation of the multimodality of the learning environment could have consequences for a number of areas. Obviously it may affect the speed of knowledge acquisition in development. It may also affect the quality of that knowledge and its robustness to damage. Similarly some of the considerations that apply during development may also be equally applicable in the case of relearning following brain damage. The current investigation undertook to understand the effect upon learning efficiency of manipulating the degree of multimodality in the learning environment, both during development and in relearning after damage.

The case of relearning after damage (i.e. recovery and/or rehabilitation through therapy) offers a contrasting learning environment to that found in the developmental case. Several researchers (e.g. Howard, Patterson, Franklin, Orchard Lisle, & Morton, 1985) have argued the benefits of a "multimodal" approach to therapy when relearning word-object relations. This "multimodal" approach refers to the idea that "a word should be elicited in any one of a variety of ways - repetition, phonemic cueing, reading, writing or listening - and their use practised in a variety of situational and grammatical contexts" (Howard et al, 1985: p52).

In this situation ‘multimodal’ refers to targeting multiple sensory modalities during the therapeutic intervention. Single therapy tasks (e.g. picture naming or reading aloud) concentrate on a single mapping between sensory modalities. In other words they are unimodal since they target a single output modality. The Rogers et al. (2004) model simulates picture naming as well as other therapy tasks in terms of the model’s impaired performance behaviour after damage. Although it does not consider relearning behaviour, exposing the model to further training after damage would easily allow for simulation of spontaneous recovery (Plaut, 1996) and therapy (Welbourne & Lambon Ralph, 2005b). These clinically-related targets were also considered in the current study.

Aims

The current investigation contrasted two learning scenarios using a replication of the Rogers et al. (2004) model: (i) developmental (i.e. infant) learning; (ii) recovery and therapy after brain damage (modeled as relearning post damage) as in the case of stroke or head injury in which therapy seeks to restore knowledge through the presentation of objects for re-learning. The Rogers et al. (2004) multimodal training regime developed convergent internal re-representations and, as discussed above, various authors have argued that human developmental learning is multimodal. The current investigation developed an additional, comparative unimodal training regime (similar to unimodal tasks used in cognitive rehabilitation therapy) to investigate whether manipulation of the uni- vs. multi-modality of the learning environment would affect the development of convergence, and its relationship to learning efficiency. Such a manipulation of multimodality, examined in terms of the effect on learning performance, could provide an account of why convergence may be necessary in terms of the organisational nature of environments in which learning is required. In order to measure convergence a novel statistical measure (representational economy) needed to be developed in order to monitor the degree of convergence among representations in the model's convergence zone at any learning point. Having developed this measure, three simulations were carried out to explore the issues described above. Simulation 1 considered initial (developmental) learning by investigating how convergence is affected by the learning environment, and whether there is a link between convergence and efficiency. Since all neural networks exhibit graceful degradation (McClelland et al., 1986), Simulation 2 explored possible benefits of convergence by examining whether highly convergent representations are also more resistant to damage. In other words do more convergent representations minimise the effects of damage and if so under what conditions does this occur (i.e. does the multimodality of learning affect that learning's robustness. Simulation 3 considered relearning by exploring how convergent representations might best be re-established after damage in terms of targeted modalities.

Simulation 1.1: Developmental learning

Method

The PDP model of semantic knowledge employed for this investigation was a replication of a previously established model of semantic knowledge (Rogers et al., 2004). This model was implemented using the LENS neural network simulator programming environment (Rohde, 2000). Figure 1.1 shows the network architecture of the model:

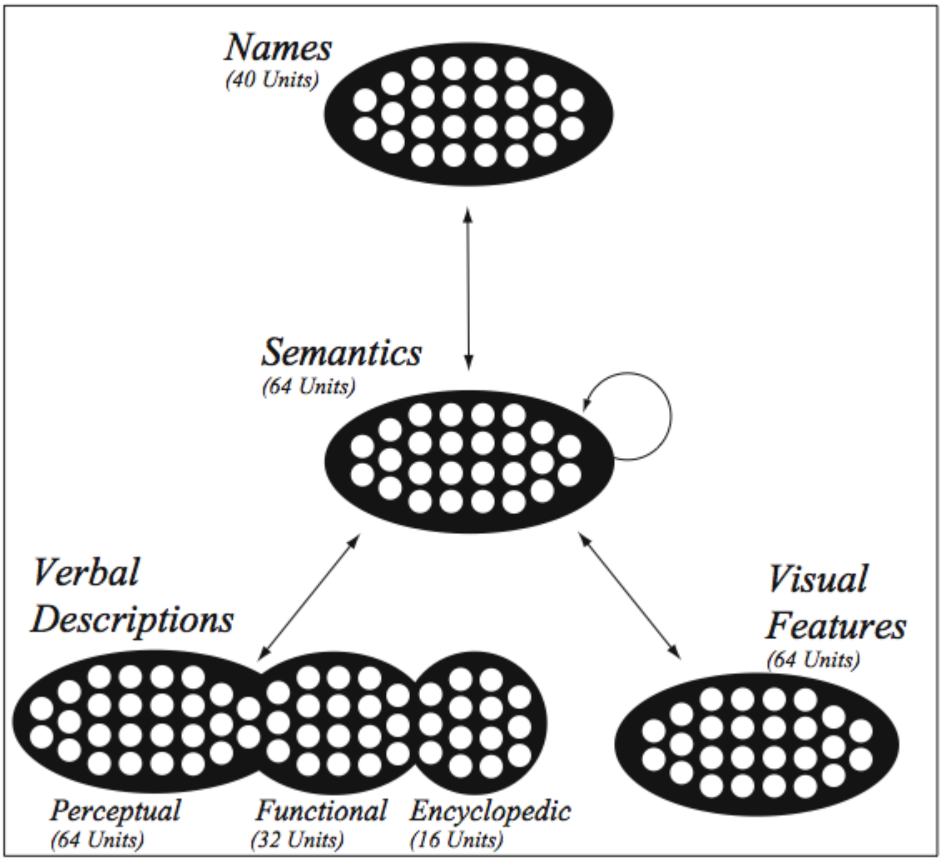

Figure 1.1 Architecture of the model adapted from Rogers et al. (2004)

The model is a fully recurrent network consisting of four layers of units. Three layers of units labelled names, verbal descriptions and visual features are bidirectionally connected via a single layer of hidden units labelled semantics. The names, verbal descriptions and visual features layers are each capable of both input and output, and receive input directly from the environment. The semantics layer is a hidden layer and does not interact directly with the environment, it only receives input from, or outputs to, the names, verbal descriptions and visual features layers. All units in the semantics layer are also recurrently connected to each other. These recurrent connections assist in the development of attractors (Hinton & Shallice, 1991), stabilising representations within the semantics layer. Layers in the model represent specific brain regions in terms of function. The visual features layer represents distinct cortical regions of the brain handling high-level visual information derived from earlier processing. This means that the activation state of each of the units in the visual features layer corresponds directly to visual properties of stimuli received from the environment (e.g. has eyes, has wheels). The verbal descriptions layer is similarly constructed representing specific cortical areas that handle linguistic information obtained from the environment. It is subdivided for convenience into three sections representing particular types of properties that can be expressed verbally: perceptual/structural properties (e.g. has eyes, has wheels), functional properties (e.g can fly, can roll) and encyclopaedic properties (e.g. lives in Africa, found in kitchen). The names layer operates in a similar manner to the verbal descriptions layer except that it only handles the represenions of object names (e.g. animal, bird, chicken). The operation of the semantic layer is such that, as the model learns, the semantic units develop re-representations of the inputs and outputs from the names, verbal descriptions and visual features layers that facilitate mapping between these inputs and outputs. For this reason the units are considered semantic since, as the model learns, they "derive amodal semantic representations that encode the semantic similarity relations among objects regardless of their surface [in this case name, verbal description and visual feature] similarities" (Rogers et al., 2004: p.233). These 'semantic' re-representations are the key subject of interest within the current investigation. Since this is a fully recurrent network as the model learns it develops attractors (Hinton & Shallice, 1991) which yield stable amodal representations across the semantic units. These amodal representations are also convergent as indicated by the Rogers et al.'s (2004) hierarchical cluster analysis of the category structure the model develops as a result of learning. Evidence for this convergence is also provided by the hierarchical cluster analysis that Rogers et al. (2004) perform upon the learned semantic representations with the trained model.

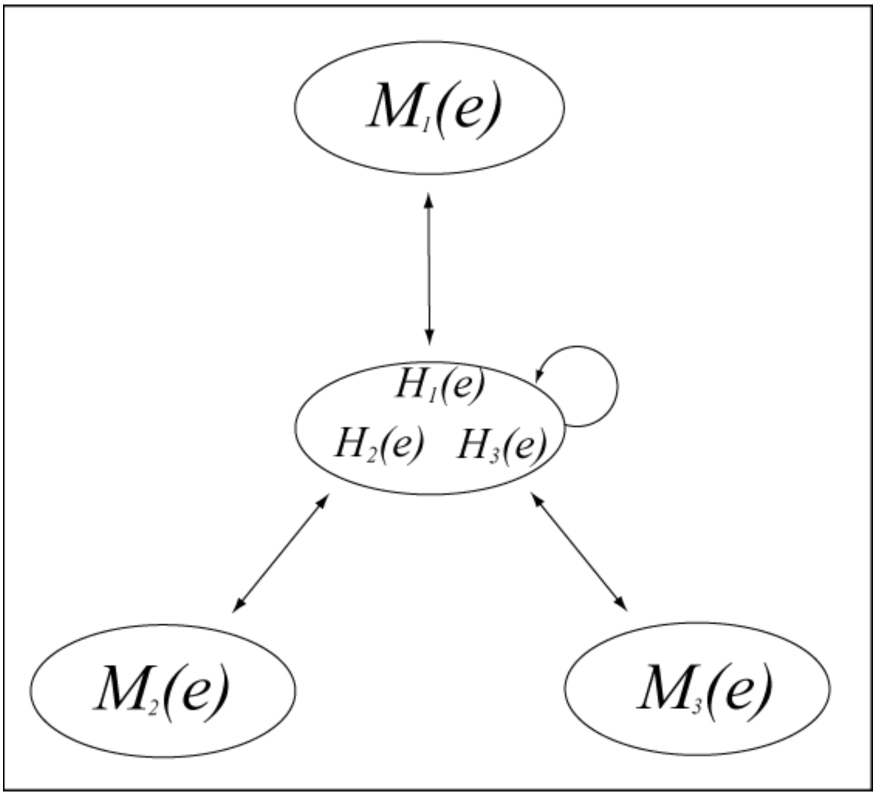

Figure 1.2 Architecture of the model redrawn as a generic PDP model with bidirectional mapping between each layer

Figure 1.2 illustrates the model in generic terms to show how an input representation (example e) of an object in any individual modality domain (layer), generates its own particular re-representation (denoted by H_1,H_2 or H_3 from input in modalities M_1,M_2,M_3 respectively) as a pattern of activity across the hidden semantic units (i.e. the hidden layer H). These re-representations have a tendency to converge towards a single pattern of activation across the units of the hidden layer H as a result of the attractor structure (as illustrated by Rogers et al.'s (2004) cluster analysis of the learned representations) that develops within the model as it learns. So over time the hidden representations (i.e. H_1,H_2, H_3) generated from the same item presented in different modalities begin to move closer together to become increasingly similar for each type of mapping between M modalities (i.e. approaching the idea that for a learned example object e, $H_1(e) = H_2(e) = H_3(e)$ in

Figure 1.2). In this manner the model implements a convergence zone within the semantics H layer.

Training Stimuli

The representations upon which the model was trained were created from prototype patterns in the same manner as in the original model (Rogers et al., 2004). The name representations were localist with 36 of the 40 units directly corresponding to a unique object upon which the model is trained. The four general object names (BIRD, ANIMAL, VEHICLE and TOOL) described in the Rogers et al. model , that would use the remaining 4 localist name units were not used in the current study.

Table 1.1 shows the category prototype patterns used to generate verbal description and visual feature representations for each named object according to the category to which it belongs. The verbal descriptions layer contained 112 units subdivided into 64 perceptual units, 32 functional units and 16 encyclopaedic units. The visual features layer contained 64 units. The prototype patterns were used probabilistically to generate binary representation vectors for each of the input/output layers. Each element of these vectors describes an individual unit's activation state according to the prototype pattern. The symbols in the prototype patterns correspond to the probability that an individual unit is turned on or off:

\+ means units likely to be active with a probability of activation of 0.8

0 means units less likely to be active with a probability of activation of 0.2

\- means units never active (i.e. are always 0) so probability of activation is 0

These prototype patterns were used to generate 6 unique training items for each of the 6 categories.

Table 1.1 Prototype patterns used to generate binary representations of verbal descriptions and visual features for each category of named object presented to the model for training

Visual Features Prototype Patterns	
Birds	++++++++++++++000000000000000000--------------------------------
Mammals	++++++++000000++++++000000000000--------------------------------
Vehicles	--------------------------------++++++00000000000000000000000000
Household Objects	--------------------------------++000000000000000000000000000000
Tools	--------------------------------++0000++++0000000000000000000000
Fruits	--------------------------------++0000++000000000000000000000000

Verbal Descriptions Prototype Patterns	
Perceptual Descriptions	
Birds	++++++++++000000000000000000000---------------------------------
Mammals	+++++00000+++++0000000000000000---------------------------------
Vehicles	-------------------------------+++++0000000000000000000000000000
Household Objects	-------------------------------+00000000000000000000000000000000
Tools	-------------------------------+00000000000000000000000000000000
Fruits	-----+---------0000000000000000+----0000000000000000000000000000
Functional Descriptions	
Birds	++++++0000000000----------------
Mammals	+++000++00000000----------------
Vehicles	----------------+++0000000000000
Household Objects	----------------+000000000000000
Tools	----------------+0+0000000000000
Fruits	------+-00000000--+0000000000000
Encyclopaedic Descriptions	
Birds	+-+-----00000000
Mammals	+--+----00000000
Vehicles	-+--+---00000000
Household Objects	-+---+--00000000
Tools	-+----+-00000000
Fruits	+------+000000++

Training the model

The model was trained on either one of two distinct training regimes developed through a broadly epigenetic consideration of the role of the external environment's influence upon developmental learning. The first regime was simultaneously multimodal (identical to that used in Rogers et al., 2004) since it involved training the model to map from activation of a single input layer to generate the target outputs on all other (multiple) output layers (as illustrated in Table 1.2). The Rogers et al. (2004) model was trained on multimodal outputs. Whilst training on multimodal inputs seems to be closer to the ecological situation of word learning in development as well as therapeutic sessions, this study uses multimodal outputs from a single input. The evidence from the Rogers et al. (2004) model suggests that multimodal output training results in the best possible learning performance for the model, and mimics developmental learning, so this study employs multimodal output training in all simulations. The second regime was unimodal since it involved training input to a single layer to map to a target output in a single layer (as illustrated in Table 1.2).

Table 1.2 Training regime patterns in terms of targeted output layers being manipulated between the multimodal and unimodal conditions

	Possible Cross Modal Mappings		
	No.	**Inputs**	**Target Outputs**
Multimodal (original Rogers et al. 2004)	1	name	name verbal visual
	2	verbal	name verbal visual
	3	visual	name verbal visual
Unimodal	1	name	name
	2	name	verbal
	3	name	visual
	4	verbal	name
	5	verbal	verbal
	6	verbal	visual
	7	visual	name
	8	visual	verbal
	9	visual	visual

The unimodal regime was developed as a result of splitting the multimodal regime into all of its constituent singular unimodal cross-domain mappings. This can be seen in Table 1.2 where the three possible simultaneous multimodal cross-domain mappings, when considered in terms of unimodal cross-domain mappings clearly involve nine singular unimodal mappings, so unimodal training is a reflection of this separation of the multimodal and corresponds to contrasting the two common real-world learning environments, development (multimodal) and post-damage rehabilitation tasks (unimodal). The unimodal regime was considered analogous to the learning environment experienced during adult learning such as that occurring during recovery or rehabilitation after damage (e.g. in the case of stroke etc.). Often therapeutic learning post-damage consists of presentation of learning items in single input and output modalities (e.g. 'spoken word to picture matching' consists of visual input, picture presentation, and verbal output, the spoken word).

Each training trial (i.e. presentation of training patterns for a single item and subsequent weight update) the network was given during learning lasted for seven time steps, each time step lasted for four ticks, meaning each trial lasted for 28 ticks in total, where each tick corresponds to one update of all unit activation values in the network. Each trial consisted of three events. For the first event, a name, verbal description, or visual feature training pattern was given as an input (see Table 1.2 for all possible inputs in the unimodal and multimodal conditions). This input was clamped and the model was then allowed to cycle for 3 time steps (i.e. the first event lasted for 12 ticks). For the second event, all inputs were removed and the model was then allowed to cycle for 2 time steps (i.e. the second event lasted for 8 ticks). Finally, the third event consisted of the application of target values across all the input/output layers (i.e. names, verbal descriptions and visual features) in the multimodal condition, or a single input/output layer in the unimodal condition. For this third event the model was allowed to cycle for 2 time steps (i.e. the third event lasted for 8 ticks). During training all possible training patterns for each of the items in the corpus were presented to the model once in random order. Learning consisted of repeated presentation of the whole training corpus, with the order of presentation re-randomised after each exposure of all items to the network. It should be noted that all aspects of each training trial in the multimodal training regime (as described above) were preserved in the unimodal regimes except that target values were only applied across units in a single output layer (i.e. names, verbal descriptions or visual features) instead of across units in all possible output layers.

Multimodal and unimodal training performance was compared on a trial by trial basis (i.e. at the level of individual items), as well as a comparison in terms of the number of presentations of the entire training corpus. Since presenting the entire corpus in the unimodal condition involved presenting three times as many training trials as in the multimodal condition, comparison in terms of the number of presentations of the entire corpus represented an "equal error comparison". That is to say error was considered across all items to be learned in each condition. The comparison on a trial by trial basis necessarily means that a higher error signal will be present in the multimodal condition since in each trial it provides information on the relation between three times as many cross domain mappings as in the unimodal condition (see Table 1.2). In order to make a detailed comparison between multimodal and unimodal training, results were to be reported for both the item level trial by trial performance comparison as well as the corpus level equated error comparison.

The model was trained with online learning (i.e. using a batch size of one, thus with a weight update after the presentation of each training trial) using recurrent backpropagation through time with a steepest descent algorithm. The learning rate was set to 0.005. A weight decay of 0.0000002 was also used to prevent any weights developing values that were disproportionately high. Similarly, no momentum was used since its process of including a proportion of the previous step in every weight change can cause the effective learning rate to become too high and inhibit learning. Each individual unit within the network was given a fixed, untrainable bias of –2. "This has the effect of deducting 2 from each unit's net input. Thus, in the absence of input, each unit's activation settles to the low end of its activation range." (Rogers et al., 2004: p. 215). Units in all of the input/output layers were clamped to their input values using a soft clamp with a clamp strength of 0.9. A target radius of 0.1 was used during the processing of each batch (in this case each training trial due to a batch size of one) so if an output unit's activation is within 0.1 of the target, no error will be generated. The model was trained until input in a single layer could generate target outputs on all layers to within an accuracy of 0.5.

To simulate development, the model was trained for 648000 weight updates, using either the multimodal and unimodal training regime in each simulation. This equates to 6000 presentations in the multimodal condition (multimodal presentation of the entire corpus takes 108 updates, i.e. 3 training trials for each of the 36 items). In contrast, this amount of training equates to 2000 presentations in the unimodal condition (unimodal presentation of the entire corpus takes 324 updates, i.e. 9 training trials for each of the 36 items). During training, the model was tested on its ability to generate the correct target outputs (to within 0.5) for all patterns, in all output modalities. 648000 weight updates was chosen for the duration of training through initial experimentation, as whenever the model learnt to 100% accuracy (verified through regular testing) it had occurred by this time. The simulation was run ten times in the multimodal condition and ten times in the unimodal condition and the results were averaged. During training, regular testing recorded accuracy (percentage of examples correct upon testing), network error and representational economy (according to the equation derived below).

Calculating representational economy in the semantic convergence zone

As the model learns and structure develops, the semantic representations generated for unrelated objects become increasingly differentiated, whilst the representations of the same item elicited from different domains become more similar. Representational economy (RE) within the semantic layer was formally investigated by developing a statistic that could quantify this process. Representational economy is defined as the average ratio between the similarity of semantic representation for unrelated examples in the same modality and the similarity of semantic representations for the same example from different modalities. The diagram in Figure 1.2.1 provides a visualisation of the representational economy calculation in order to provide a clearer practical understanding of the formulae presented below. This diagram is based upon the diagram of semantic space in Hinton et al. 1993. This is calculated according to the following equation:

$$RE = \frac{1}{n}\sum_{i=1}^{n}\frac{s_{same}(e_i)}{s_{different}(e_i)}$$

where n is the number of item examples *e* (each given by e_i) upon which the model is trained, and *s* is a similarity measure function defined as follows for a model with *m* modalities:

$$s_{different}(e_i) = \overline{d}\left(H_1(e_i), H_2(e_i), \ldots\ldots\ldots, H_m(e_i)\right)$$

where $\overline{d}$ is a function to calculate the average euclidean distance between the representations generated by each of the *m* different input modalities at the hidden layer *H* for the same example (e_i).

$$s_{same}(e_i) = \frac{1}{m}\sum_{j=1}^{m}\overline{d}\left(H_j(e_i), \left\{H_j(e_1), H_j(e_2), \ldots\ldots\ldots, H_j(e_n)\right\}\right)$$

where $\bar{d}$ is a function to calculate the average euclidean distance between the representation generated by each input modality at the hidden layer *H* for an example (e_i) and the representations of all the unrelated examples ($e_1, e_2, \ldots\ldots e_n$) generated at the hidden layer *H* (denoted by $H_j(e_1), H_j(e_2), \ldots H_j(e_n)$ where j indicates the number of the input modality, by that same input modality, averaged across all *m* modalities.

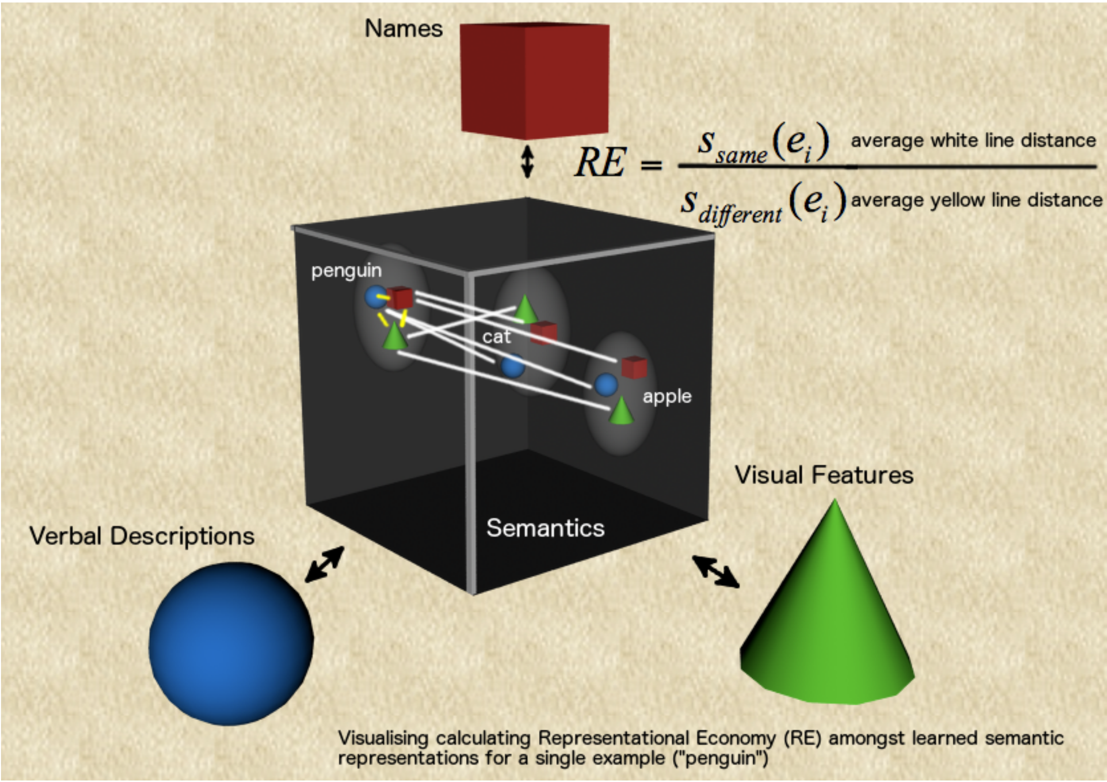

Figure 1.2.1 A visualisation of the representational economy calculation in semantic space derived from Hinton et al.'s (1993) diagram of semantic space.

Testing the network and analysing representational economy

Representational economy, accuracy and error were calculated at regular intervals during the model's training. A 2x20 repeated measures analysis of variance (ANOVA) was conducted on these data to investigate the effect of manipulating the training regime (multimodal or unimodal) at twenty equally spaced time points during training on the dependent variables (accuracy, representational economy and network error). The correlation between representational economy, accuracy, and error was conducted for all simulations carried out with the model to investigate the suggested relationship between Representational Economy, Accuracy and Network Error as the model learns.

Results

The graphs in Figure 1.3 illustrate the model's performance during development, both on a trial by trial basis and for an equated error comparison. It can be clearly seen that multimodal training is more efficient than unimodal training on a trial by trial basis, obtaining greater accuracy and greater convergence (as evidenced by representational economy) as well as greater error reduction. These differences between multimodal and unimodal training are also statistically significant as the results of the 2x20 analysis of variance given in the Table 1.3 clearly show. Figure 1.3 also shows how the multimodal training achieves greater acceleration in the learning process thus showing its greater efficiency at an earlier point. The equal error comparison between multimodal and unimodal training shows no substantial difference and indeed the small differences seen in the graphs in Figure 1.3 are not statistically significant (as indicated in the analysis results in Table 1.3). Table 1.3 shows that multimodal learning does indeed perform better than unimodal learning by indicating a statistically significant difference between multimodal and unimodal training for Accuracy, Representational Economy and Network Error. There is also a statistically significant interaction between training and time for Accuracy and Network Error between Equated Error Multimodal and Unimodal training but not for representational economy. This indicates that training is having an effect on the model, but that there is no statistically significant performance difference between the training regimes independent of time.

Representational Economy was strongly correlated with learning Accuracy for all conditions: Multimodal $r(20)=.820$, $p<.001$; Equated Error Multimodal $r(20)=.878$, $p<.001$; Unimodal $r(20)=.869$, $p<.001$. Representational Economy was strongly correlated with Network Error for all conditions: Multimodal $r(20)=-.789$, $p<.001$; Equated Error Multimodal $r(20)=-.767$, $p<.001$; Unimodal $r(20)=-.760$, $p<.001$.

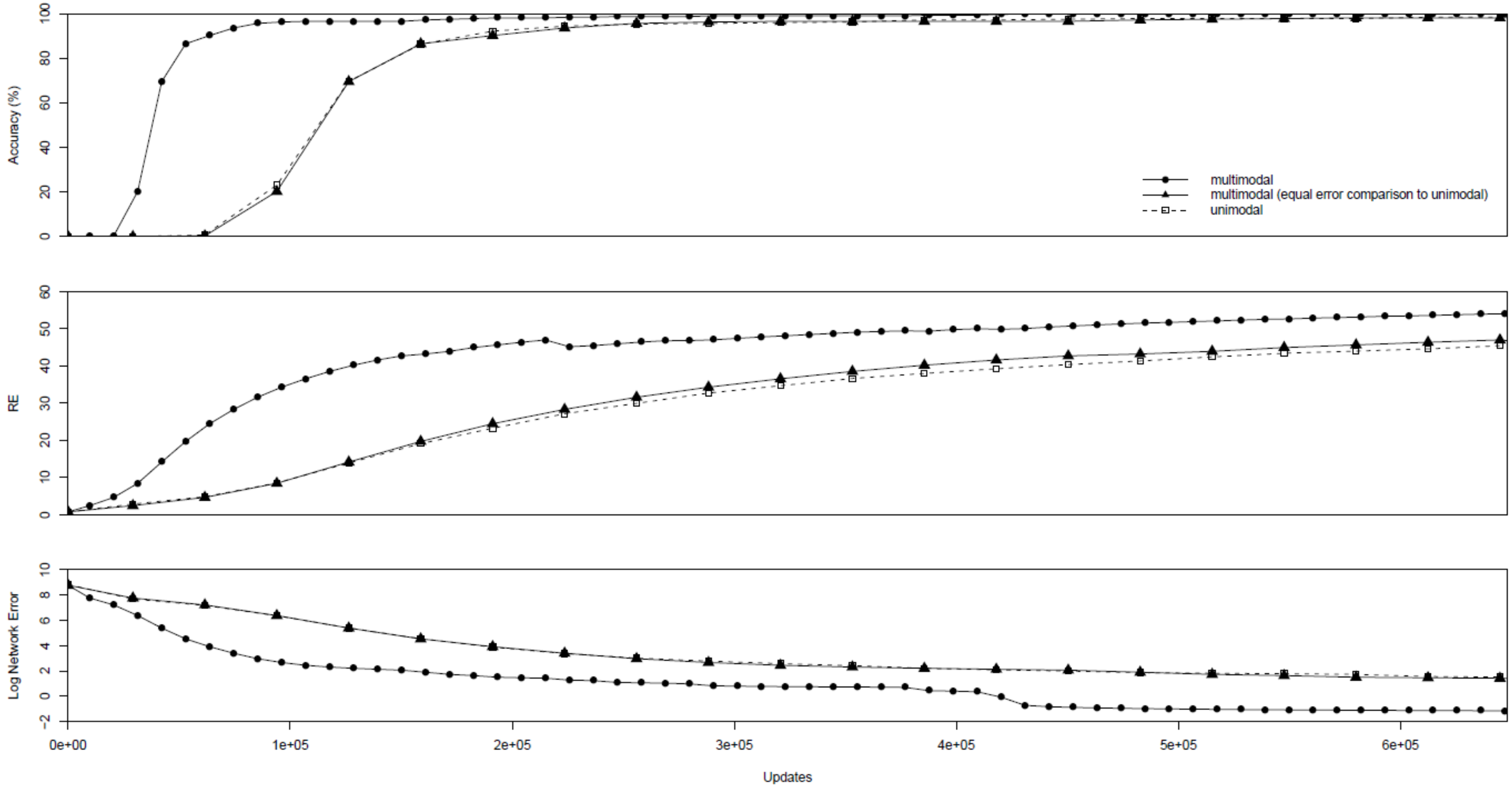

Figure 1.3 Mean variation in training performance during developmental learning

Table 1.3 2x20 ANOVAs indicating significant performance differences between training regimes during developmental learning

Training Regime Comparison	Dependent Variable Measure	Independent Variables	df	F	p	Effect size (η^2)
Multimodal Vs Unimodal	Accuracy	Training	1,9	103.806	<0.001*	0.920
		Time	19,171	959.390	<0.001*	0.991
		Training*Time	19,171	483.020	<0.001*	0.982
	Representational Economy	Training	1,9	114.888	<0.001*	0.927
		Time	19,171	942.068	<0.001*	0.991
		Training*Time	19,171	99.054	<0.001*	0.917
	Network Error	Training	1,9	1195.770	<0.001*	0.993
		Time	19,171	1948.960	<0.001*	0.995
		Training*Time	19,171	1404.116	<0.001*	0,994
Equated Error Multimodal Vs Unimodal	Accuracy	Training	1,9	0.581	0.465	0.061
		Time	19,171	1536.595	<0.001*	0.994
		Training*Time	19,171	1.675	0.045*	0.157
	Representational Economy	Training	1,9	1.105	0.321	0.109
		Time	19,171	1406.016	<0.001*	0.994
		Training*Time	19,171	0.699	0.816	0.072
	Network Error	Training	1,9	0.377	0.554	0.040
		Time	19,171	3089.077	<0.001*	0.997
		Training*Time	19,171	2.878	<0.001*	0.242

Simulation 1.2: Robustness to damage

Method

The model was trained in the manner described in Simulation 1. In order to investigate the model's robustness to damage the 10 multimodal and 10 unimodal trained networks from each of the these simulations were lesioned by removing an increasing proportion of all incoming and outgoing connections across all the units in all layers (i.e. names, verbal descriptions, visual features and semantics) . Each lesion was performed 10 times and scores for accuracy, error and representational economy were recorded. These data were then plotted (see Figure 1.4) and analysed (see Table 1.4) to see if the multimodal training would result in models that were more robust to damage than those generated from unimodal training.

Results

The graphs in Figure 1.4 illustrate the model's robustness to damage as the model was subjected to lesions of increasing severity following the 648000 weight updates of initial training. Across the range of lesions the multimodal regime consistently appeared to be more robust to damage with higher accuracy and lower network error, both for trial by trial, and equal error, comparison to unimodal training. However, for all except the smallest lesions (i.e. less than 2% of connections removed), RE after damage was reduced to the same low level regardless of whether the initial training had been multimodal or unimodal, and regardless of whether the two regimes are compared on a trial by trial or equated error basis. Table 1.4 illustrates that multimodal learning results in a structure that is more robust to damage than unimodal learning, as shown by the statistically significant difference between multimodal and unimodal training for Accuracy, Representational Economy and Network Error. There is also a statistically significant difference in the effect of the lesion severity for Representational Economy, Accuracy and Network Error between Equated Error Multimodal and Unimodal training, but no difference in the effect of training and no significant interaction between the effect of training training and lesion severity. Representational Economy was strongly correlated with learning Accuracy for all conditions: Multimodal $r(11)=.836$, $p=.001$; Equated Error Multimodal $r(11)=.866$, $p=.001$; Unimodal $r(11)=.863$,$p=.001$. Representational Economy was strongly correlated with Network Error for all conditions: Multimodal $r(11)=-.635$,$p=.036$;Equated Error Multimodal $r(11)=-.628$,$p<.039$;Unimodal $r(11)=-.672$,$p<.024$.

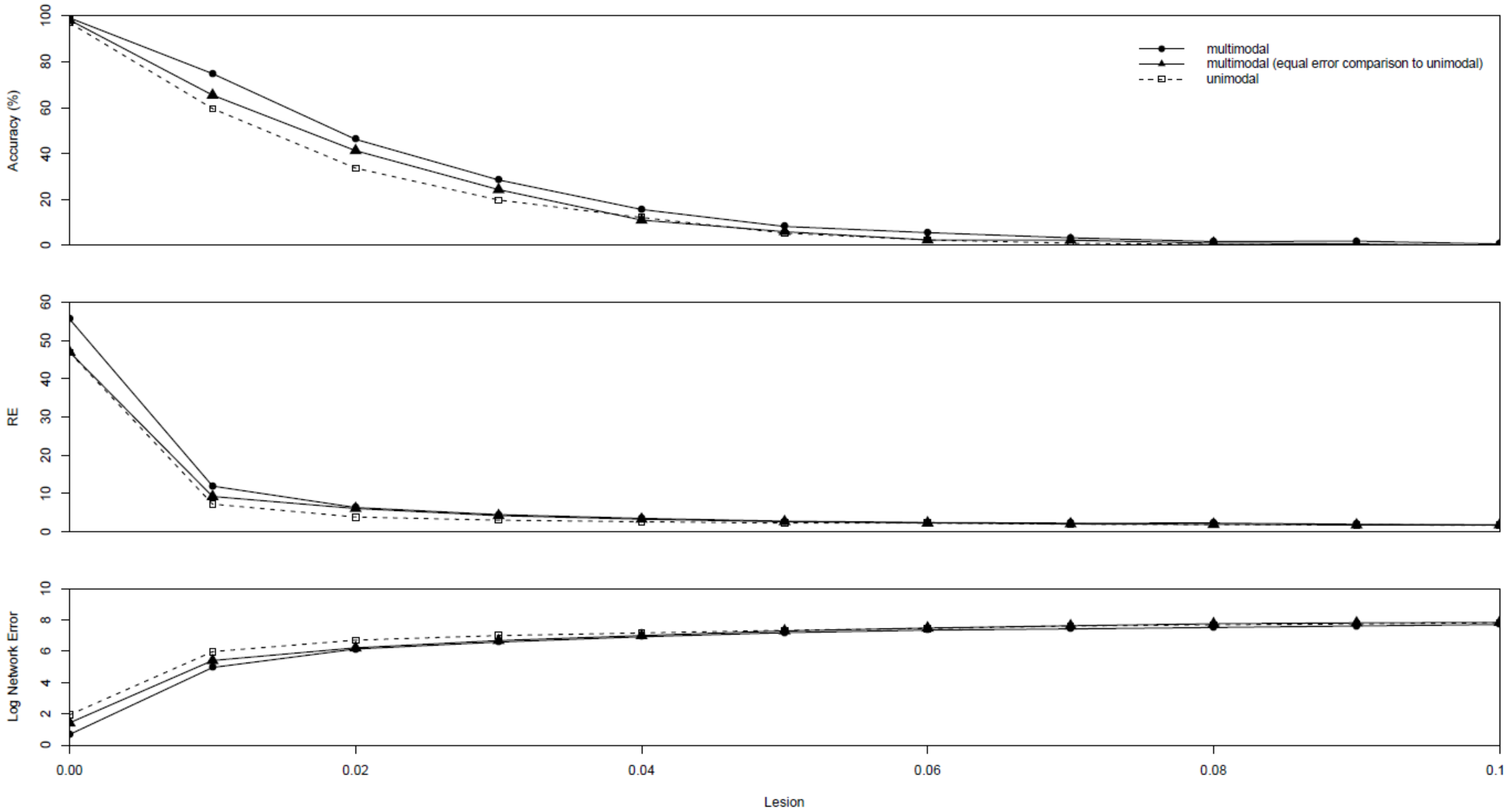

Figure 1.4 Mean variation in robustness to damage for the fully trained model with each training regime

Table 1.4 2x11 ANOVAs indicating significant differences in robustness to damage between trained networks

Training Regime Comparison	Dependent Variable Measure	Independent Variables	df	F	p	Effect size (η^2)
Multimodal Vs Unimodal	Accuracy	Training	1,9	5.536	0.043*	0.381
		Lesion	10,90	1001.493	<0.001*	0.991
		Training*Lesion	10,90	2.338	0.017*	0.206
	Representational Economy	Training	1,9	28.772	<0.001*	0.762
		Lesion	10,90	2256.564	<0.001*	0.996
		Training*Lesion	10,90	6.306	<0.001*	0.412
	Network Error	Training	1,9	8.967	0.015*	0.499
		Lesion	10,90	226.110	<0.001*	0.962
		Training*Lesion	10,90	1.299	0.243	0.126
Equated Error Multimodal Vs Unimodal	Accuracy	Training	1,9	0.296	0.600	0.032
		Lesion	10,90	1296.013	<0.001*	0.993
		Training*Lesion	10,90	0.495	0.889	0.052
	Representational Economy	Training	1,9	1.379	0.270	0.133
		Lesion	10,90	1469.838	<0.001*	0.994
		Training*Lesion	10,90	0.737	0.688	0.076
	Network Error	Training	1,9	0.081	0.782	0.009
		Lesion	10,90	340.135	<0.001*	0.974
		Training*Lesion	10,90	1.909	0.054	0.175

Simulation 1.3: Relearning after damage

Method

The model was trained in the manner described in Simulation 1. However, this time only the multimodal training regime was used to train the model as per the original simulation (see Rogers et al, 2004). In order to investigate the model's re-learning behaviour after damage the 10 multimodally trained networks from the initial training were lesioned by removing a proportion of all incoming and outgoing connections across all the units in all layers (i.e names, verbal descriptions, visual features and semantics). Initial experiments in lesioning the model solely at the connections to semantics to mimic the real situation of localised damage did not yield a range of damage severities appropriate for retraining. In order to consider different levels of damage, lesioning was required such that when the model relearns after damage it would relearn only to a certain level, and no further, essentially giving an accuracy graph that goes to asymptote for several different levels. Three levels of damage were decided upon to examine relearning after increasingly severe damage: Mild where the model relearns to around 90% accuracy, Moderate, where the model relearns to around 60% accuracy, and Severe where the model relearns to around 30% accuracy. In the end this was only achievable in the model by lesioning uniformly across the model rather than a closer simulation to the real situation of localised injury.The trained networks were subjected to three separate lesions of varying degrees of severity: mild (removing 86 % of connections); moderate (removing 88% of connections); and severe (removing 90% of connections). These degrees of severity were intended to encompass the range of damage that typically occurs after a stroke and thus cover a range of behavioural severity outcomes. After lesioning, each network underwent 2592000 updates of retraining with both the multimodal and unimodal training regimes. This corresponds to 24000 multimodal, and 8000 unimodal, presentations of the entire training corpus. This repeated-measures methodology where the same initially trained network was subjected to each lesion and retraining condition allows us to remove any variance associated with individual pre or postmorbid differences. This is important since brain-damaged patients are not a homogenous population, and studying the effects of treatment should be based upon "within-patient comparisons" (Howard & Hatfield, 1987, p.119).

Results

Mild Lesioning

The graphs in Figure 1.5 illustrate the model's performance during relearning after a mild lesion, both on a trial by trial basis and for an equated error comparison. It can be clearly seen that multimodal training is more efficient than unimodal training on a trial by trial basis, obtaining greater accuracy, and greater convergence (as evidenced by representational economy) as well as greater error reduction. These differences between multimodal and unimodal training are also statistically significant as the results of the 2x20 analysis of variance given in the table in Table 1.5 clearly show. Figure 1.5 again shows how the multimodal training achieves greater acceleration in the learning process reflecting greater temporal efficiency. The equal error comparison between multimodal and unimodal training shows no difference, is much noisier than during developmental learning, and not statistically significant (as show in the analysis results in Table 1.5). Examining Table 1.5 shows multimodal learning results in better performance than unimodal learning, as shown by the statistically significant difference between multimodal and unimodal training for Accuracy, Representational Economy and Network Error. There is also a statistically significant difference in the effect of time, on Accuracy, Representational Economy and Network Error, between Equated Error Multimodal and Unimodal training, as well as a significant interaction between training and time. However whilst there is a significant difference in network error there is no difference in the level of accuracy and representational economy achieved between Equated Error Multimodal and Unimodal training. Representational Economy was strongly correlated with learning Accuracy for all conditions: Multimodal $r(20)=.969$, $p<.001$; Equated Error Multimodal $r(20)=.947$, $p<.001$; Unimodal $r(20)=.924$,$p<.001$. Representational Economy was strongly correlated with Network Error for all conditions: Multimodal $r(20)=-.889$,$p<.001$;Equated Error Multimodal $r(20)=-.920$,$p<.001$;Unimodal $r(20)=-.938$,$p<.001$.

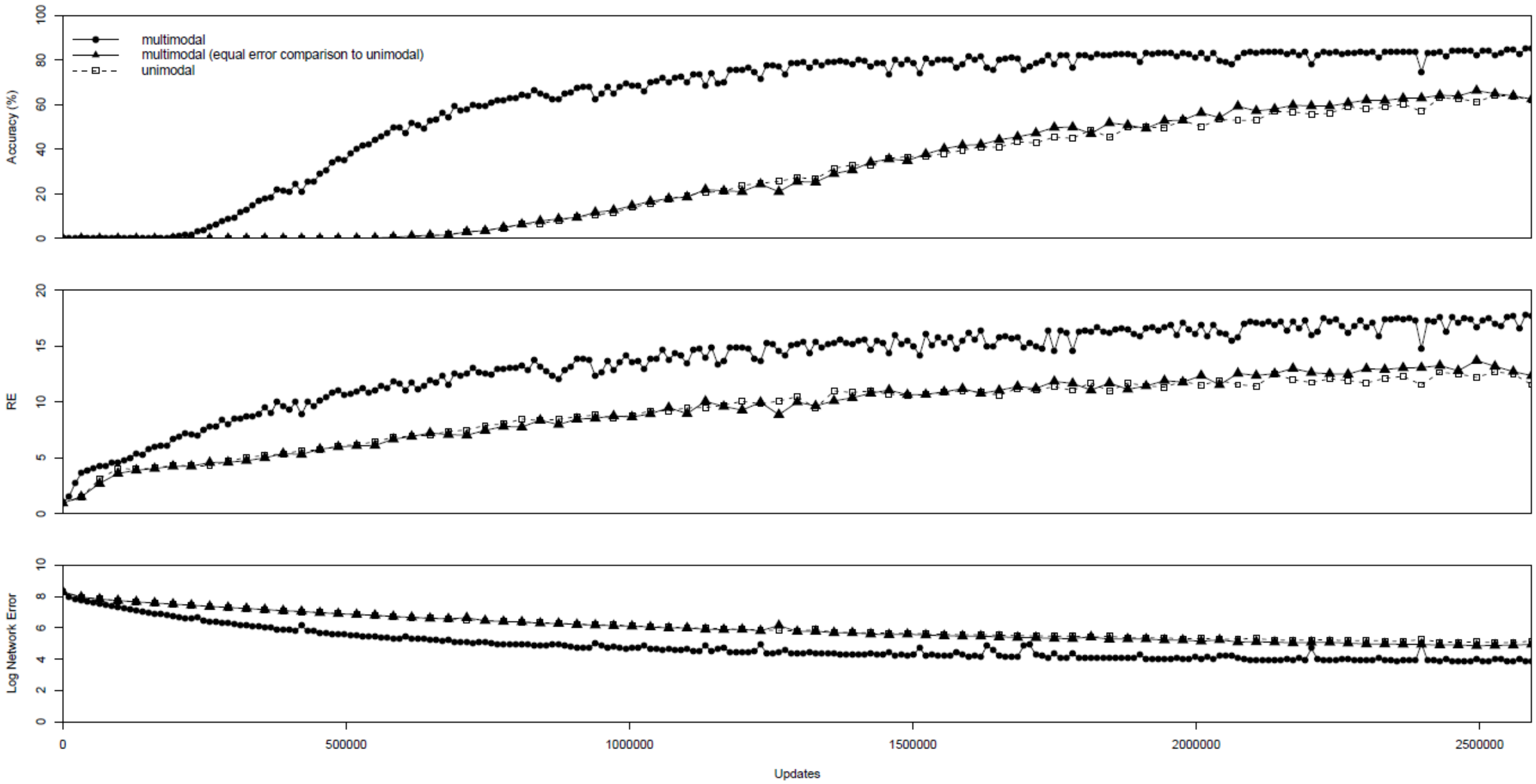

Figure 1.5 Mean variation in training performance during relearning after a mild lesion

Table 1.5 2x20 ANOVAs indicating significant performance differences between training regimes after a mild lesion

ReTraining Regime Comparison	Dependent Variable Measure	Independent Variables	df	F	p	Effect size (η^2)
Multimodal Vs Unimodal	Accuracy	Training	1,9	236.941	<0.001*	0.963
		Time	19,171	361.673	<0.001*	0.976
		Training*Time	19,171	27.997	<0.001*	0.757
	Representational Economy	Training	1,9	451.562	<0.001*	0.980
		Time	19,171	88.263	<0.001*	0.907
		Training*Time	19,171	2.475	<0.001*	0.216
	Network Error	Training	1,9	1778.618	<0.001*	0.995
		Time	19,171	2357.452	<0.001*	0.996
		Training*Time	19,171	322.672	<0.001*	0.973
Equated Error Multimodal Vs Unimodal	Accuracy	Training	1,9	3.357	0.1	0.272
		Time	19,171	348.978	<0.001*	0.975
		Training*Time	19,171	1.738	0.034*	0.162
	Representational Economy	Training	1,9	3.614	0.90	0.287
		Time	19,171	223.372	<0.001*	0.961
		Training*Time	19,171	2.512	0.001*	0.218
	Network Error	Training	1,9	54.635	<0.001*	0.859
		Time	19,171	6478.064	<0.001*	0.999
		Training*Time	19,171	6.549	<0.001*	0.421

Moderate Lesioning

Figure 1.6 shows the model's performance during relearning after a moderate lesion for a trial-by-trial, and an equated error, comparison. It can be clearly seen again that multimodal training was more efficient than unimodal training on a trial-by trial-basis, obtaining greater accuracy and greater convergence (as measured by representational economy) as well as greater error reduction. These differences between multimodal and unimodal training were also statistically significant (see Table 1.6). Figure 1.6 again continues to show the picture of multimodal training achieving greater acceleration in the learning process reflecting greater efficiency in the early phase. The equal error comparison between multimodal and unimodal training again showed no difference, is much noisier than during developmental learning and mild lesioning, and not statistically significant (as show in the analysis results in Table 1.6). The results of the ANOVA in

Table 1.6 gives a statistically significant difference between multimodal and unimodal learning showing that multimodal learning results in better performance for Accuracy, Representational Economy and Network Error. Whilst a significant effect of time can also be seen for Accuracy, Representational Economy and Network Error when comparing Equated Error Multimodal and Unimodal training, there is only a significant effect of training on Network Error and a significant interaction between training and time for Representational Economy and Network Error so there is no difference in the level of Accuracy achieved. Representational Economy was strongly correlated with learning Accuracy for all conditions: Multimodal $r(20)=.932$, $p<.001$; Equated Error Multimodal $r(20)=.807$, $p<.001$; Unimodal $r(20)=.809$, $p<.001$. Representational Economy was strongly correlated with Network Error for all conditions: Multimodal $r(20)=-.948$, $p<.001$; Equated Error Multimodal $r(20)=-.962$, $p<.001$; Unimodal $r(20)=-.955$, $p<.001$.

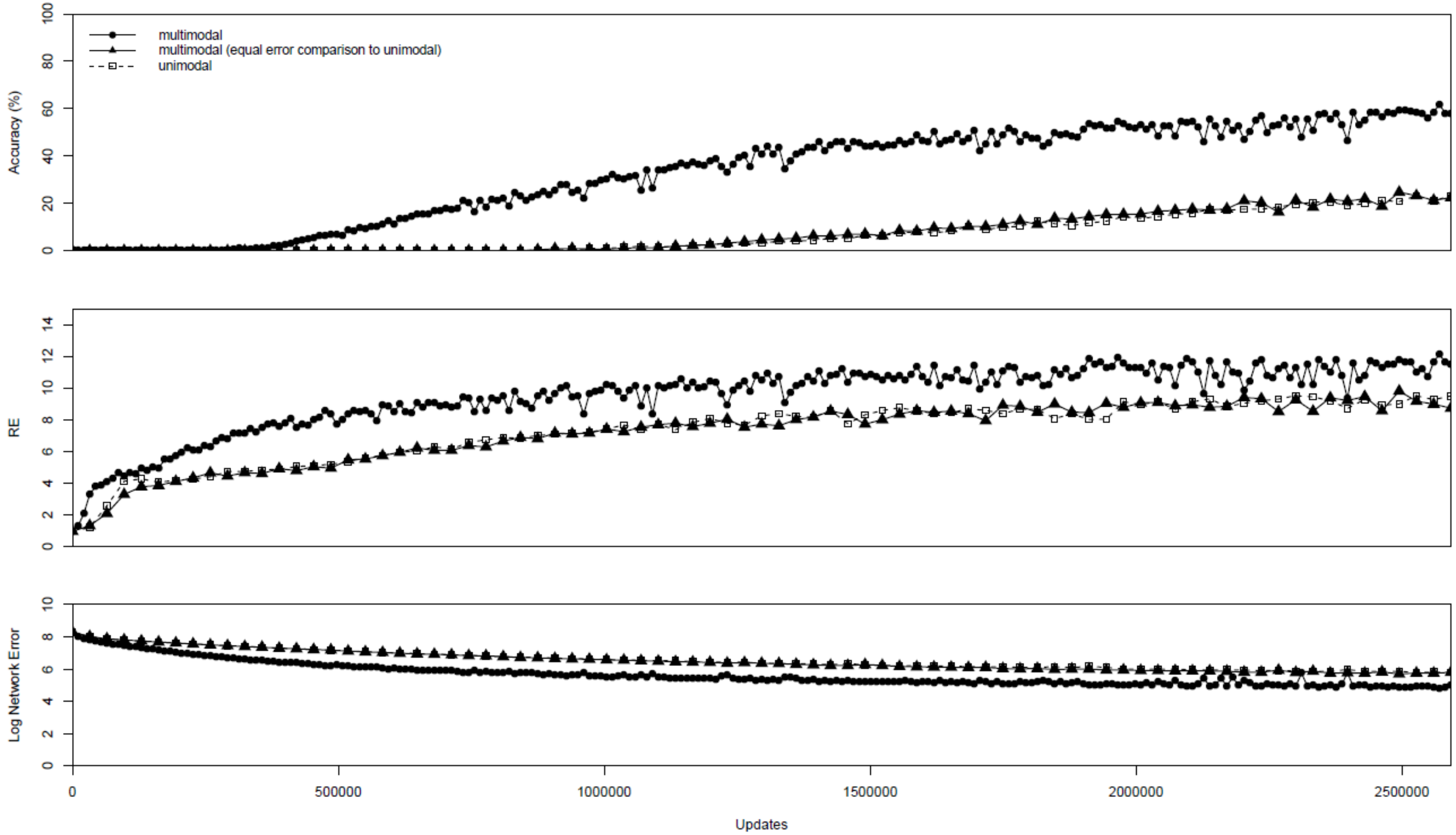

Figure 1.6 Mean variation in training performance during relearning after a moderate lesion

Table 1.6 2x20 ANOVAs indicating significant performance differences between training regimes after a moderate lesion

ReTraining Regime Comparison	Dependent Variable Measure	Independent Variables	df	F	p	Effect size (η^2)
Multimodal Vs Unimodal	Accuracy	Training	1,9	253.503	<0.001*	0.966
		Time	19,171	94.440	<0.001*	0.913
		Training*Time	19,171	22.607	<0.001*	0.715
	Representational Economy	Training	1,9	182.835	<0.001*	0.953
		Time	19,171	51.784	<0.001*	0.852
		Training*Time	19,171	1.624	0.055	0.153
	Network Error	Training	1,9	1892.579	<0.001*	0.995
		Time	19,171	1031.058	<0.001*	0.991
		Training*Time	19,171	89.527	<0.001*	0.909
Equated Error Multimodal Vs Unimodal	Accuracy	Training	1,9	1.104	0.321	0.109
		Time	19,171	114.307	<0.001*	0.927
		Training*Time	19,171	1.220	0.246	0.119
	Representational Economy	Training	1,9	0.466	0.512	0.049
		Time	19,171	135.339	<0.001*	0.938
		Training*Time	19,171	1.840	0.022*	0.170
	Network Error	Training	1,9	40.696	<0.001*	0.819
		Time	19,171	4093.398	<0.001*	0.998
		Training*Time	19,171	2.450	0.001*	0.214

Severe Lesioning

Figure 1.7 shows the model's performance during relearning after a severe lesion. Again data are displayed for a trial-by-trial and an equated error, comparison. Again there is an efficiency advantage of multimodal training over unimodal training on a trial-by-trial basis, obtaining greater accuracy, and greater convergence (as measured by representational economy) as well as greater error reduction. It should also be noted that during experimentation after the severe lesion the model was also left to run for twice the number of weight updates than those shown in Figure 1.7 since it was suspected that the model's performance was not yet at asymptote. However, the model did not show any further gains, thus Figure 1.7 reflects an asymptotic performance which perhaps emphasises the efficiency advantage in multimodal training even more than in the mild and moderate lesion cases. These differences between multimodal and unimodal training were also again statistically significant (see Table 1.7). illustrating greater efficiency in the early phase for multimodal training, and reflecting greater acceleration in the learning process. The equal error comparison between multimodal and unimodal training again shows no difference, is much noisier than during developmental learning, and relearning after mild or moderate lesions, and is not statistically significant (as show in the analysis results in Table 1.7). The results in Table 1.7 give a statistically significant difference between multimodal and unimodal training for Accuracy, Representational Economy and Network Error, showing the superior performance of the multimodal regime. There is also a statistically significant difference in the effect of time on Accuracy, Representational Economy and Network Error between Equated Error Multimodal and Unimodal training. A significant interaction between training and time for network error and representational economy can also be seen, but no difference in the level of accuracy, network error or representational economy achieved between Equated Error Multimodal and Unimodal training. Representational Economy was strongly correlated with learning Accuracy for all conditions: Multimodal $r(20)=.777$, $p<.001$; Equated Error Multimodal $r(20)=.783$, $p<.001$; Unimodal $r(20)=.843$, $p<.001$. Representational Economy was strongly correlated with Network Error for all conditions: Multimodal $r(20)=-.932$, $p<.001$; Equated Error Multimodal $r(20)=-.967$, $p<.001$; Unimodal $r(20)=-.981$, $p<.001$.

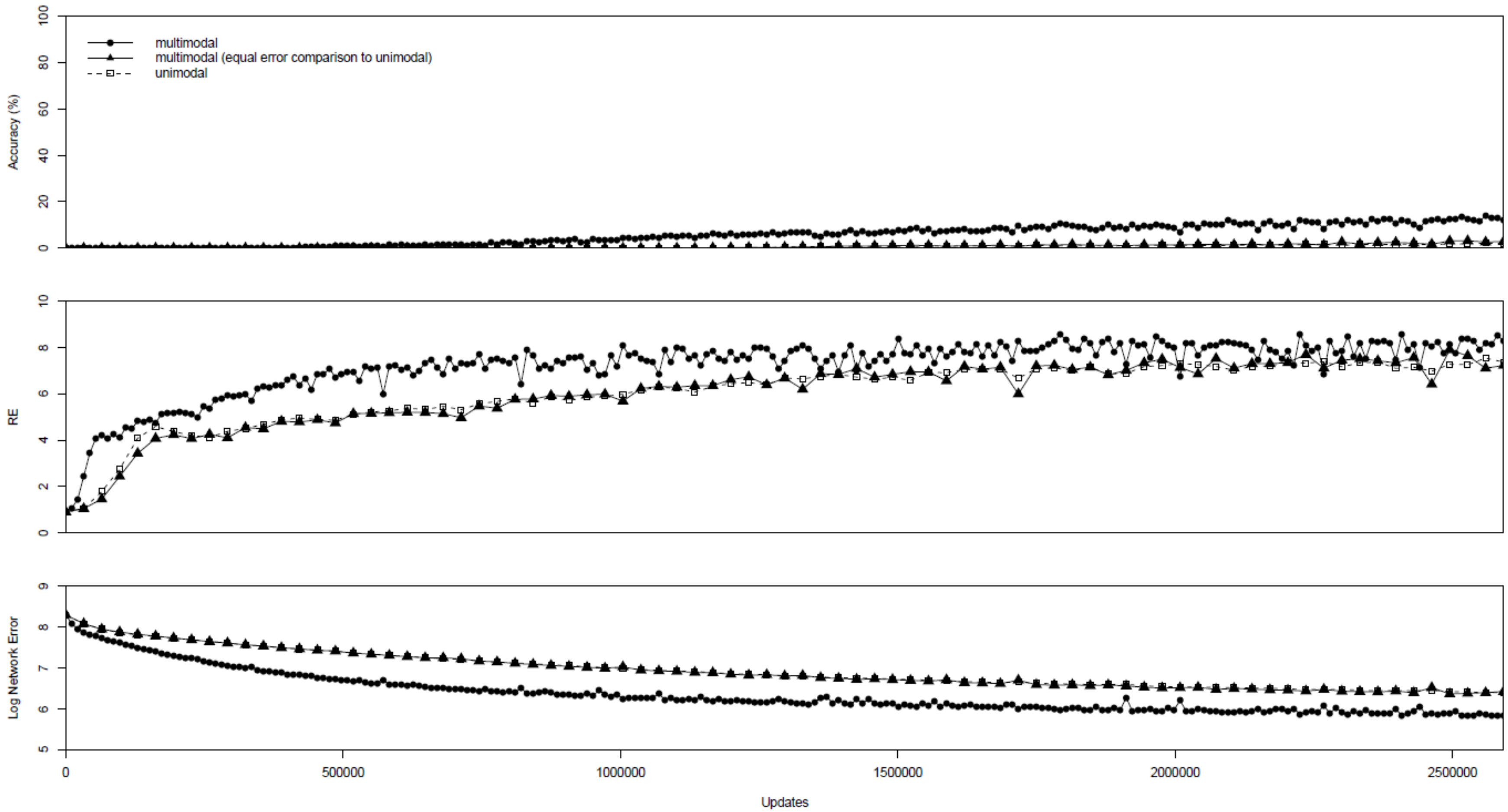

Figure 1.7 Mean variation in training performance during relearning after a severe lesion

Table 1.7 2x20 ANOVAs indicating significant performance differences between training regimes after a severe lesion

ReTraining Regime Comparison	Dependent Variable Measure	Independent Variables	df	F	p	Effect size (η^2)
Multimodal Vs Unimodal	Accuracy	Training	1,9	21.813	<0.001*	0.708
		Time	19,171	17.245	<0.001*	0.657
		Training*Time	19,171	12.539	<0.001*	0.582
	Representational Economy	Training	1,9	45.774	<0.001*	0.836
		Time	19,171	45.823	<0.001*	0.836
		Training*Time	19,171	4.276	<0.001*	0.322
	Network Error	Training	1,9	5304.952	<0.001*	0.998
		Time	19,171	999.062	<0.001*	0.991
		Training*Time	19,171	53.361	<0.001*	0.856
Equated Error Multimodal Vs Unimodal	Accuracy	Training	1,9	0.431	0.528	0.046
		Time	19,171	7.540	<0.001*	0.456
		Training*Time	19,171	0.500	0.960	0.053
	Representational Economy	Training	1,9	1.766	0.217	0.164
		Time	19,171	65.585	<0.001*	0.879
		Training*Time	19,171	2.851	<0.001*	0.241
	Network Error	Training	1,9	0.885	0.372	0.089
		Time	19,171	1779.928	<0.001*	0.995
		Training*Time	19,171	2.559	0.001*	0.221

Discussion

Summary of Results

Simulation 1 illustrated differences in learning efficiency resulting from manipulation of the learning environment (i.e. the multimodal or unimodal conditions). Utilising representational economy as a newly-derived measure of convergence enabled precise monitoring of the evolution of convergence during the learning process. Such monitoring gives a glimpse of the interrelation between the operation of a convergence zone and the quality of learning that it engenders and subsequently supports. Previous modelling studies (e.g. Plaut, 2002; Rogers et al., 2004 or Dilkina et al., 2008) have shown that convergence zones can allow for the development of efficient amodal representations. However, the results of the current study indicate that the efficiency of representation development is also directly affected by the learning environment. Both multimodal and unimodal training regimes yield similar levels of accuracy at the end of the developmental learning period. Yet it is the time course of learning in terms of the relationship between accuracy and the degree of convergence occurring in the convergent zone that is of particular interest in the current investigation. Representational economy gives a measure of the degree of convergence in the computational terms of euclidean similarity between semantic representations supporting a variety of cross-domain mappings, as a function of developing structure. As such, representational economy not only provides information about whether or not convergence is occurring but also about how well that convergence supports learning. The clearly observed efficiency advantages observed in multimodal training as compared to unimodal, in trial-by-trial comparisons of both learning and relearning suggest learning efficiency is directly proportional to the degree of multimodality found in the learning environment. At the item level, the multimodal regime is more efficient in terms of speed of learning (accelerated early development) and the degree of convergence (representational efficiency) it promotes.

Multimodal and unimodal training was compared on a trial by trial basis (contrasting network performance after each presentation of a single training pattern), as well as in terms of equated error (contrasting network performance after each presentation of all the training patterns) across all training items. The equated error comparison suggests that any comparison of learning environments based on the number of times the network sees the entire set of training patterns, would show no significant difference, regardless of the manner of presentation at the item level. When the network performance on multimodal training is compared to that of unimodal training on a trial by trial basis each multimodal trial gives the network more error information: A single multimodal training trial contains the input and target patterns of all possible crossmodal mappings for an item (see Table 1.2). In contrast a single unimodal training trial contains the input and target patterns of just one crossmodal mapping for an item (see Table 1.2). Clearly the higher error signal available for every multimodal training trial accounts for the observed multimodal advantages in learning performance and the lack of performance difference in the equated error comparison. Nevertheless it is likely that any real world learning environment comparison would involve some degree of inequality during presentation of learning items. Such inequality could possibly arise as a result of attentional or memory issues such that it becomes difficult to achieve multimodal or unimodal conditions with precise control. Simulation 2 suggests other advantages for multimodal learning beyond those of pure efficiency in developmental or representational terms. The results of Simulation 2 suggested that the more efficient convergent representations developed under multimodal learning were also more robust to damage than the representations developed under unimodal learning. It is thus possible to consider whether multimodal learning in development would confer some type of evolutionary advantage such that minor brain damage would still leave an individual with a degree of their previously learned knowledge. Moreover this possible residual degree of knowledge deriving from the more robust representational structure promoted by multimodal learning, would allow for more rapid relearning after damage. Statistical testing confirmed this increased robustness in the multimodal condition only for lesions of extremely minimal severity (removal of just 2% of connections). For more severe lesions multimodal learning appeared to be no more robust to damage than unimodal. Simulation 3 showed that spontaneous recovery can make efficient use of convergence zones by virtue of the manner in which items are perceived during each learning episode resulting in considerably better performance. Detailed analysis of the representational economy within both training regimes demonstrated that there is a very strong relationship between the economy of representation and the performance of the network – those networks that made the most efficient use of their convergent zone also had the lowest error. Overall then the degree of convergence developed within the model's convergence zone, with its attendant efficiency implications, is significantly affected by the manner of item presentation in the learning environment, and the multimodal environment proved more efficient and promoted greater convergence.

The three simulations all showed a clear effect of the benefits of a multimodal as opposed to unimodal training regime in terms of efficiency. As discussed earlier, and illustrated in Figure 1.2, representational economy is a consequence of convergent intermediate 'semantic' re-representations developing as the model learns. In these terms, representational economy is a measure of convergence. It can be seen that representational economy follows the same developmental patterns as accuracy, thus these results establish a clear relationship between convergence and efficiency in the model both in temporal and representational terms. Examining the results then suggests how the semantic units in the model developed to function as a convergence zone, and how convergence can be used as an outcome measure of learning efficiency. The most salient factor at this early stage of investigating the role of the learning environment is that convergence is clearly a characteristic of efficiency.

Analysis of convergence in development

Both the multimodal and unimodal training regimes exhibited well known characteristics of development, such as the ‘vocabulary spurt’ that various authors have noted in modelling terms (Plunkett et al., 1997). Similarly, both the multimodal and unimodal training regimes reached the same levels of accuracy at the end point of the training period. This would perhaps be expected given that both regimes exposed the model to the same material for learning and it is merely the manner of exposure which was subject to variation. The equated error condition illustrated and suggested that any multimodal advantage derived from the manner of exposure on a trial-by-trial basis. More specifically, it is the higher error signal in the multimodal condition, resulting from each training trial containing the input and target patterns of all possible crossmodal mappings for an item (see Table 1.2), at each presentation of an object for learning, that was responsible for the multimodal advantage. The speed of learning was one of the most interesting aspects of initial learning. The multimodal regime resulted in much faster learning which coincided with faster development and greater convergence than the unimodal case (see Figure 1.3). Observations like this allowed an initial consideration of why convergence may occur. As a whole the model’s learning process and properties offer an account of how convergence occurs across a group of units (analogous to a collection of neurons) that handles intermediate re-representation in the facilitation of cross-modal mapping. Rogers et al. (2004) detailed how these convergent intermediate re-representations would form, and how their formation could underpin semantic knowledge. Thus there exists an idea of convergence supporting knowledge as the efficient association of perceptual information arriving via multiple sensory modalities. There is clearly evidence that convergence is a characteristic of efficient learning in the model, and is indeed necessary in some degree for learning, regardless of the training regime. Note that both multimodal and unimodal training generate convergence in the semantics convergence zone. One aim of the current investigation was to consider why convergence might be necessary. The measurement of convergence as an economy of the intermediate re-representations generated during the learning process offers a possible solution: In considering the graphs of initial learning (Figure 1.3), purely in terms of performance, it is obvious that increasing convergence is a characteristic of increasing accuracy. The learning environment of the model, embodied as multimodal or unimodal, in terms of the training patterns employed, primarily affects the speed of learning. In other words, within the model, a multimodal environment promotes the rapid acquisition of knowledge, one characteristic outcome measure of which is convergence. A consequent possibility then arises as a consideration of the degree to which learning can occur without representational economy: Simulation 1.3 offers some insight by considering damaged models where successively increasing the removal of connections between units necessarily reduces the capacity of the model to support convergence.

Analysis of convergence in relearning after damage

The time course of relearning (as illustrated in Figures 1.5, 1.6 and 1.7) presents a picture of how convergence may operate, when the brain has to re-acquire knowledge after damage. Various studies have found evidence of plasticity in adult brains (e.g . Buonomano & Merzenich, 1998). Plasticity offers an account of recovery that suggests that the brain's adaptability after damage allows relearning to be accomplished by the remains of the damaged system. The process of relearning then involves these remaining undamagaed areas acquiring the ability to handle processing previously performed by areas that no longer function. After lesioning the model, the remaining connections facilitate cross-modal mappings previously performed by a substantially larger number of connections. Since damage involves the removal of connections, it necessarily removes a degree of the model's capacity for convergence. This reduced convergent ability makes the damaged model useful for considering the brain's learning behaviour with reduced convergence capacity. In other words, does convergence always accompany efficient learning? During relearning, the substantially lower values of representational economy showed that learning could still occur with lesser convergence. Various implications can be derived from the model's relearning behaviour in terms of convergence. First, it is worth noting that the multimodal training regime yields faster and more accurate learning than a unimodal regime, earlier after damage regardless of damage severity. These benefits of multimodal training became more apparent with the severity of damage. In considering the graphs of relearning after severe damage (see Figure 1.7) it is possible to observe a pronounced increase in the beneficial effect of the multimodal regime. Indeed the picture offered in the case of severe damage clearly shows the multimodal regime yielding substantially greater accuracy and representational economy than the unimodal case. This could be of particular importance where, after severe damage, resources are limited and need to be used as efficiently as possible. Generally speaking, careful observation of both the initial learning and relearning graphs show efficiency manifested in time-of-recovery and representational terms. The multimodal regime can be considered more efficient because it generates more accurate learning in a shorter number of updates. A characteristic of this efficiency is high convergence (as measured by representational economy). Could it not then be suggested that convergence zones exist in order to allow representational economy. What If the brain does indeed develop internal re-representations that integrate incoming perceptions from multiple sensory modalities in this manner? These representations, acting as stored knowledge, would then be an efficient compressed representation that integrated and associated original perceptual information from multiple sensory modalities. The efficiency of such stored knowledge would thus be a direct result of the manner and quality of interaction with the learning environment, as well as the richness with which the learning environment itself could be perceived.

Conclusion

This study set out to determine what factors in the learning environment affect the generation of convergent representations, and thus understand learning efficiency in terms of the number of modalities targeted during learning. The results from this investigation suggest convergence zones allow for the possibility of the formation of economic convergent representations, and that the degree of this economy gives a reliable measure of learning efficiency . This property of convergence zones is clearly suited to maximising performance after damage when neural resources are considerably reduced. Learning environments which can promote greater efficiency within convergence zones (i.e. multimodal learning) result in better performance than environments that generate less efficient convergence (i.e. unimodal learning).

Various practical implications are suggested: If the role of convergent representations as suggested here is an accurate picture of how learning occurs then it would seem that multimodal learning offers particular benefits when normal learning is in some way impaired. The observation that an equated error comparison of multimodal and unimodal training shows no significant difference, suggests the efficiency of multimodal learning is driven by the increased error signal available during each learning episode. Indeed some authors (e.g. Harm, McCandliss, & Seidenberg, 2003; Harm & Seidenberg, 2004) have noted that whilst structuring the training regime has minimal impact on normal performance, it can improve performance in an impaired model. Such impairments are typically observed in aphasia: The results of the current study suggest structuring the training regime multimodally for relearning can improve relearning efficiency in the domain of semantics and have implications for the design of aphasia rehabilitation therapies; suggesting an environment that can maximise convergence results in better relearning performance. Such observations, alongside those made in this paper with regard to convergence, suggest the structure adopted when presenting items for learning affects learning efficiency in a manner which may be a result of variation in convergence zone activity. Varying the learning environment clearly provides a tool for investigating convergence learning efficiency. However such findings would also need to correspond with clinical performance data regarding the manipulation of the learning environment. This investigation is hopefully a first step in developing a fuller methodology for exploring the role of convergence in efficient knowledge acquisition and its implications for rehabilitation.

Chapter 2 - Learning by degrees: Exploring the structure and benefits of multimodal learning in a computational model of semantic knowledge

Abstract

Evidence from a variety of sources (e.g. developmental psychology, computational modelling, and speech and language therapy studies) suggest multimodal learning offers more efficient and robust potential for knowledge acquisition. Traditionally, cognitive rehabilitation does not place specific emphasis upon the number of sensory modalities targeted during episodes of relearning object knowledge. The current study investigates the degree to which multimodality may be important in all learning episodes, and how that multimodality is constructed. Using a computational simulation of semantic knowledge, three types of learning were compared for efficiency: multimodal learning (simultaneous learning of all cross-modal relations for all items), item-focused unimodal learning (targeting single pairwise cross-modal relations, but cycling through all possible pairwise cross-modal relations for each item to be learnt before moving on to the next item), and task-focused unimodal learning (targeting single pairwise cross-modal relations for all items to be learnt before moving on to the next pairwise cross modal relation). The results confirmed previous simulation data from Chapter 1 and results from other researchers (e.g. Rogers et al. 2004) that multimodal learning offers substantial benefits in terms of the efficiency of the representational structure that develops to support knowledge. This implies that simple reorganisation of existing cognitive rehabilitation tasks (i.e. sequential perfomance of all learning tasks for each item to be learnt before moving on to the next item) cannot approximate the hypothesised benefits of multimodal learning

derived from previous computational simulations (see Chapter 1) . These computational results suggest that rehabilitation efficiency could be increased by adopting a multimodal approach during each learning/intervention trial.

Introduction

In our everyday lives we encounter objects and events in the world through our sensory perceptions and verbal experiences (e.g. looking, listening and touching). As a result of cognitive processing we are able to be aware of the unitary nature of such objects and events, that is to say we are aware of whole entities as opposed to separate sensations. For example, if we look at and listen to a barking dog we are aware of a single thing, namely the dog, as opposed to an awareness of the visual image of a dog, and a barking sound as separate entities. Concepts therefore consist of unified multimodal experiences. It thus follows that there must be learning/representational systems for deriving such unified concepts from multimodal experience, and that there may be benefits, in terms of efficiency and storage, from deriving such coherent concepts.

There is considerable evidence that children learn from a multimodal learning environment. For example, studies with pre-verbal infants have shown that mothers use temporally synchronous naming when presenting new objects or actions that they wish the child to learn (e.g. Gogate et al., 2000, Messer, 1978). "During temporally synchronous naming, they [Mothers] speak a word while holding and moving an object rather than moving it out of phase with the spoken word" (Gogate, Bolzani & Betancourt, 2006: p.261). This synchronous presentation to multiple sensory modalities follows from the idea that infant learning of word–object relations is “accomplished in the context of the infant’s multimodal interactions with its mother” (Sullivan & Horowitz, 1983, p. 210). If developmental learning is indeed multimodal or perhaps is more efficiently accomplished multimodally, then multimodality is likely to be beneficial for all learning. Damasio (1989) suggested that convergence zones in the brain provide an account for the mechanism behind the brain's ability to bind incoming information from various perceptual modalities, and how such binding supports both perception and recall of learnt external world experiences. Damasio (1989: p. 130) noted that “it is not enough for the brain to analyze the world into its components parts: the brain must bind together those parts that make whole entities and events, both for recognition and recall”. Maybe our ability to detect relationships between sensory perceptions, which cohere into object knowledge, is dependent upon the brain’s convergence zones working in conjunction with a multimodal learning environment. If the normal system for extracting long-term knowledge representations is based on distilling multimodal information into convergent unitary representations then it is possible that the greater the degree of multimodality in the learning environment, the more efficient the learning will be.

There are two distinct major learning periods that occur across the lifespan. The first is development, which all humans experience, and is based on learning derived from childhood experience of the external world. The second learning period, and one which is only experienced by a small proportion of any population, is that of learning after brain damage (e.g. stroke or head injury). This postdamage learning can be split into spontaneous recovery (i.e. learning derived from exposure to the external world) and cognitive rehabilitation (i.e. discrete learning episodes consisting of intensified exposure to a specific set of concepts). It is possible that, if multimodality is truly central to efficient developmental learning, it is also important for rehabilitation. Indeed some researchers (e.g. Howard et al., 1985) have suggested that multimodality is important for rehabilitation therapy. However, most formal rehabilitation therapeutic tasks are unimodal, that is to say they deal only with re-establishing relationships between paired sensory perceptions. For example, picture naming therapy deals with re-establishing the relationship between the names of objects and their corresponding visual image, in other words between isolated auditory and visual perceptions.

One clear question that arises from the possible benefit of multimodality in learning episodes, is what neural processes would support learning in such a way that interaction with a multimodal learning environment maximises learning capacity. Connectionist models provide simulations of hypothesised neural processes underpinning learning. Parallel distributed processing (PDP) models are one class of connectionist simulation that allows the investigation of these learning processes (in development, recovery and as a part of rehabilitation interventions). PDP models offer an advantage over other forms of computational simulation since the models actually learn (Welbourne & Lambon Ralph, 2005b). Several PDP models of learning (e.g. Plaut, 2002; Rogers et al., 2004; Dilkina et al, 2008) also suggest that multimodality is important for learning. These models implement convergence zones and thus offer computational accounts for the processes proposed by Damasio (1989).

The investigations reported in Chapter 1 used the Rogers et al. (2004) model of semantic memory as a starting point to explore the degree to which multimodal learning contributes to overall learning efficiency. The work attempted to discern, given that convergence is always necessary for learning, whether convergence was actively promoted by multimodal or unimodal learning. The investigation also developed a quantitative measure, termed representational economy, that could track the degree of convergence during learning and relate it to learning accuracy. This measure means that such models may offer a concept of multimodal learning as advantageous due to its potential for generating greater cross-modal interaction within convergence zones, such that a higher error signal (in terms of deriving true relationships between perceptions in different modalities) is available during learning episodes. The findings from Chapter 1 suggested that: multimodal learning was more efficient than unimodal learning; learned representations were more robust and developed more quickly; and that multimodal learning also appeared to be beneficial in simulations of spontaneous recovery after brain damage. What is missing from Chapter 1 is an examination of how multimodal learning works. It has been previously stated that the minimum requirement for multimodality would merely be simultaneous perception in more than two modalities. That is to say any situation where the perceptual load is greater than simple paired sensory perceptions. The Rogers et al. (2004) model of semantic memory implements semantics as a convergence zone. It provides convincing evidence that convergence zones are required to enable the relation of sense perception from different modalities. As they (Rogers et al., 2004) discuss, within the convergent zone amodal semantic representations develop during learning to support the cross-modal association of representations of the visual features, names, and verbal descriptions from a variety of objects. This model also produces a convincing internal category structure, arguing for an emergent view of semantic memory in terms of the representational structure that develops. Most interestingly this model is trained by a particular view of multimodal learning based upon simultaneous perception in multiple modalities. Hence their (Rogers et al., 2004) model was trained to associate input representations in a single modality with simultaneous presentations of target output representations in all possible modalities including the input modality (Rogers et al., 2004). The current investigation takes the view that other possible conceptions of multimodality are possible: By exploring whether simultaneous presentations of target output representations in all possible modalities, is an essential requirement of multimodal learning the current study seeks to establish whether sequential presentations of inputs, and their corresponding target output representations, for all possible cross-modal relations of a given object could approximate the previously observed benefits of simultaneous multimodal learning (see Chapter 1).

Aims

It is worth reiterating that the multimodal learning, in other words multimodal item presentation, explored in Chapter 1 preserves the notion of multimodality suggested by Rogers et al.'s (2004) work since it corresponds well with the observations of multimodality in the developmental literature (i.e. Gogate et al., 2000, Messer, 1978). The aspect of simultaneity in multimodal learning is difficult to reconcile generally with formal learning (e.g. school and other learning situations) and more specifically would run counter to at least some SLT (Speech and Language Therapy) therapeutic interventions based on simple stimulus-response mappings, as it relates to individuals recovering from brain damage. Simply put it is difficult to conceive of how simultaneous perceptions of an object in multiple modalities can be achieved in regular therapeutic settings. The opportunity to gain all possible sensory perceptions of many objects is just not practical. To reproduce all of the modalities in therapy, and to compress the long-term nature of developmental learning into a short, clinic-based rehabilitation seems beyond the scope of therapeutic interventions which are often necessarily short, and last only for a few weeks. For example, objects that are too large (e.g. buildings, lakes etc), or too abstract, to be neatly presented in a home environment (where much post-damage therapy often takes place) can only be presented for learning in their pictorial representation. Yet clearly the opportunity for learning to consist of "hands-on" multisensory perception is certainly achievable for certain small concrete objects. If it were possible that sequential unimodal perceptions of objects could approximate simultaneous multimodal perception then perhaps it would not be necessary to worry about achieving multimodal learning environments. The current study sought to examine what effect particular learning protocols may have in simulation and use this to hypothesise what effects may be found in future patient studies. The current study utilised the same computational simulation as Chapter 1, in which efficiency advantages for multimodal learning were established in terms of both time and the quality of the developed representational structure that supported learnt object knowledge. Evidence from Chapter 1 showed no statistically significant difference between multimodal and unimodal training when compared in a condition where the error signal was equated. In other words no difference if each point of comparison occurs after the same number of presentations of all patterns in the training corpus. So it would seem reasonable to assume that rearranging unimodal tasks (i.e. names to verbal descriptions, names to visual features and verbal descriptions to visual features and the other tasks listed in Table 1.2) so that they occur sequentially for a single item, could approximate simultaneous multimodal presentation for that item and its advantages in terms of convergence. This sequential item focused unimodal training strategy would contrast with existing task focused unimodal therapy strategies (e.g. picture naming, spoken word to picture matching etc).

The current investigation used a replication of the Rogers et al. (2004) model described in Chapter 1 but trains the model with three different training regimes: multimodal, item-focused unimodal, and task-focused unimodal in order to determine which derives the greatest learning efficiency.The effect of these different training regimes were compared with respect to development, robustness to damage, and spontaneous recovery. The investigation sought to answer the following questions: Can sequential item-focused unimodal learning approximate the efficiency of simultaneous multimodal learning? Is there a gain in learning efficiency for sequential item-focused unimodal training compared to the existing task-focused unimodal training strategies that often occur in targeted therapy? Finally, since the benefits of multimodality have already been established, what are the efficiency implications for cognitive rehabilitation based upon simulating spontaneous recovery, in terms of the degree of multimodality that can be achieved within expected patient responses (in terms of output modalities utilised), for any given therapeutic task?

Simulation 2.1: Developmental Learning

Method

For the current study the Rogers et al. (2004) model of semantic knowledge was recreated using LENS neural network simulator programming environment (Rohde, 2000) and the network architecture of the model is shown in Figure 2.1.

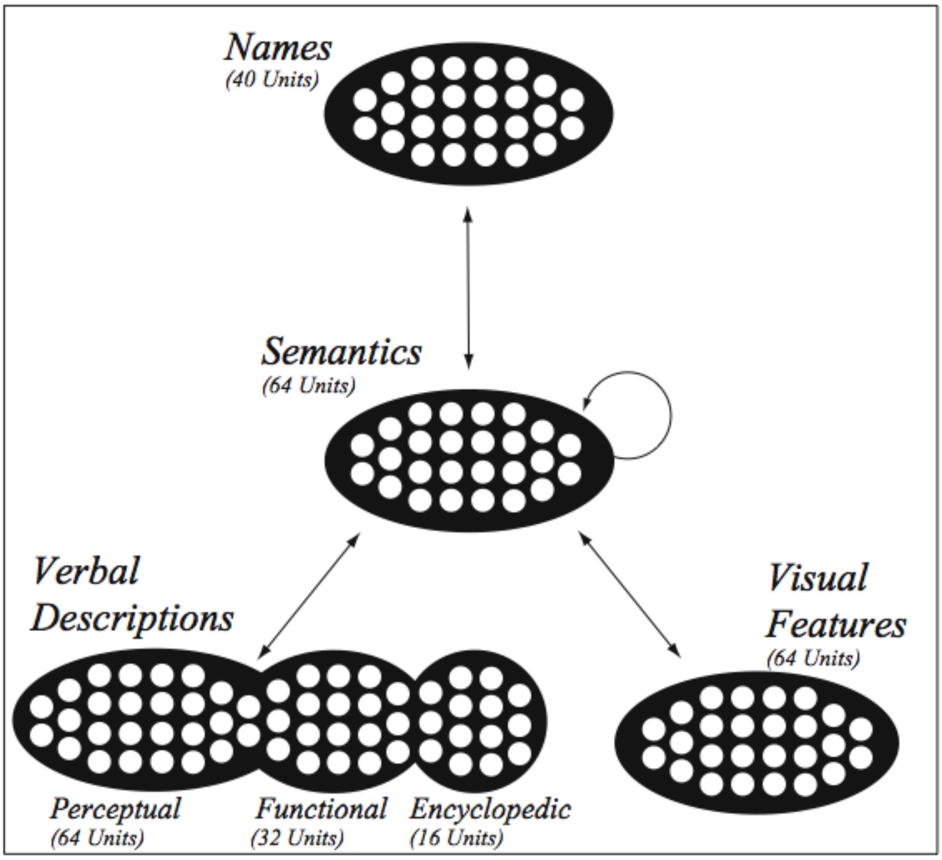

Figure 2.1 Architecture of the model adapted from Rogers et al. (2004)

The Rogers et al. (2004) model contains three layers of units labelled names, verbal descriptions and visual features. Each of these layeers are bidirectionally connected via a single layer labelled semantics and consiting entirely of hidden units. All layers except the semantic hidden layer can receive input and output directly from the environment. The semantics layer receives input from or outputs to, the names, verbal descriptions and visual features layers and is recurrently connected to itself to aid development of attractors (Hinton & Shallice, 1991), as stable semantic representations. The layers represent specific brain regions in terms of function. The visual features layer represents processing high-level visual information so the activation state of the layer corresponds with stimulus properties (e.g. has eyes, has wheels). The verbal descriptions layer handle language information. It contains three sections representing for verbally-expressed properties: perceptual/structural properties (e.g. has eyes, has wheels); functional properties (e.g. can fly, can roll); and encyclopaedic properties (e.g. lives in Africa, found in kitchen). The names layer represents object names (e.g. ‘animal’, ‘bird’, ‘chicken’). As the model learns, the semantic layer’s units “derive amodal semantic representations that encode the semantic similarity relations among objects regardless of their surface [in this case name, verbal description and visual feature] similarities” (Rogers et al., 2004: p.233). These amodal ‘semantic’ re-representations are also convergent as shown by the cluster analysis (Rogers et al, 2004).

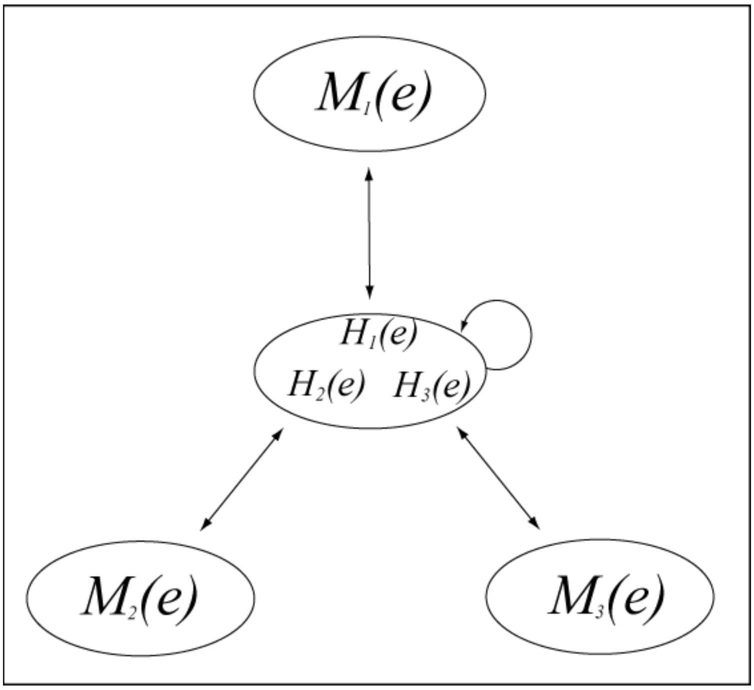

Figure 2.2 Architecture of the model redrawn as a generic PDP model with bidirectional mapping between each layer

Figure 2.2 illustrates the model in generic terms to show how an input representation (example e) of an object in any individual modality domain (layer), generates its own particular re-representation (denoted by H_1,H_2 or H_3 from input in modalities M_1,M_2,M_3 respectively) as a pattern of activity across the hidden semantic units (i.e. the hidden layer H). These re-representations have a tendency to converge towards a single pattern of activation across the units of the hidden layer H as a result of the attractor structure (as illustrated by Rogers et al.'s (2004) cluster analysis of the learned representations) that develops within the model as it learns. So over time the hidden representations (i.e. H_1,H_2, H_3) generated from the same item presented in different modalities begin to move closer together becoming increasingly similar for each type of mapping between M modalities (i.e. approaching the idea that for a learned example object e, $H_1(e) = H_2(e) = H_3(e)$ in Figure 2.2). In this manner the model implements a convergence zone within the semantics H layer.

Training Stimuli

The model was trained using representations made from prototypes. Within the localist name representations 36 of the 40 units represented a single item that the model would learning. The remaining 4 units which were used for general object names in the Rogers et al. (2004) model were not used here. Table 2.1 shows the prototype patterns derived from clinical research into object features (roger et al. 2004) used to generate verbal description and visual feature representations. The verbal descriptions layer contained 112 units subdivided into 64 perceptual units, 32 functional units and 16 encyclopaedic units. The visual features layer contained 64 units. The prototype patterns generate binary representation vectors for each layers.

The symbols in the prototype patterns indicate unit activation as follows:

+ means units likely to be active with a probability of activation of 0.8

0 means units less likely to be active with a probability of activation of 0.2

- means units never active (i.e. are always 0) so probability of activation is 0

These patterns provide each of the 6 unique items for each of the 6 categories of object which the model was going to learn.

Table 2.1 Prototypes patterns used to generate binary representations of verbal descriptions and visual features for each category of named object presented to the model for training

Visual Features Prototype Patterns	
Birds	++++++++++++++000000000000000000--------------------------------
Mammals	++++++++000000++++++000000000000--------------------------------
Vehicles	--------------------------------++++++00000000000000000000000000
Household Objects	--------------------------------++000000000000000000000000000000
Tools	--------------------------------++0000++++0000000000000000000000
Fruits	--------------------------------++0000++000000000000000000000000

Verbal Descriptions Prototype Patterns	
Perceptual Descriptions	
Birds	++++++++++000000000000000000000---------------------------------
Mammals	+++++00000+++++0000000000000000---------------------------------
Vehicles	-------------------------------+++++0000000000000000000000000000
Household Objects	-------------------------------+00000000000000000000000000000000
Tools	-------------------------------+00000000000000000000000000000000
Fruits	-----+---------0000000000000000+----0000000000000000000000000000
Functional Descriptions	
Birds	++++++0000000000----------------
Mammals	+++000++00000000----------------
Vehicles	----------------+++0000000000000
Household Objects	----------------+000000000000000
Tools	----------------+0+0000000000000
Fruits	------+-00000000--+0000000000000
Encyclopaedic Descriptions	
Birds	+-+-----00000000
Mammals	+--+----00000000
Vehicles	-+--+---00000000
Household Objects	-+---+--00000000
Tools	-+----+-00000000
Fruits	+------+000000++

Training the model

Batched learning was used to train the model with three regimes. This was further development of the findings in Chapter 1 which described an advantage for multimodal learning. The first training regime was simultaneously multimodal (identical to that used in Rogers et al., 2004) in which input was trained to map to target outputs in all possible domains (as illustrated in Table 2.2). The second training regime was “item-focused” unimodal since it involved training each possible pairwise combination of single-layer input to a single-layer output (as illustrated in Table 2.2) for a single item before moving on to the next item. All possible pairwise combinations for an item were presented in a single batch. The third regime was “task-focused” unimodal since it involved training one type of single-layer input to single-layer output mapping for all items before moving on to the next pairwise combination of cross-modal mapping. The item-focused and task-focused unimodal regimes were developed by splitting the multimodal regime into all of its constituent singular unimodal cross-domain mappings. This is summarised in Table 2.2.

Each training trial (i.e. presentation of training patterns for a single item and subsequent weight update) lasted for seven time steps. Each time step lasted for four ticks, meaning each trial lasted for 28 ticks in total, where each tick corresponds to one update of all unit activation values in the network. Each trial consisted of three events. For the first event, a name, verbal description, or visual feature pattern was given as an input (see Table 2.2 for all possible inputs in the unimodal and multimodal conditions). This input was clamped and the model was then allowed to cycle for 3 time steps (i.e. the first event lasted for 12 ticks). For the second event, all inputs were removed and the model was allowed to cycle for 2 time steps (i.e. the second event lasted for 8 ticks). Finally, the third event consisted of the application of target values across all the input/output layers (i.e names, verbal descriptions and visual features) in the multimodal condition, or a single input/output layer in the unimodal condition. For this third event the model was allowed to cycle for 2 time steps (i.e. the third event lasted for 8 ticks). During training all possible patterns for each of the items in the training corpus were presented to the model once, in random order. Learning consisted of repeated presentation of the whole training corpus, with the order of presentation re-randomised after each exposure of all items to the network. It should be noted that all aspects of each training trial in the multimodal training regime (as described above) were preserved in the unimodal regimes except that target values were only applied across units in a single output layer (i.e. names, verbal descriptions or visual features) instead of across units in all possible output layers.

Multimodal, item-focused unimodal and task-focused unimodal performance was compared on a trial-by-trial basis (i.e. at the level of individual items), as well as in terms of the number of presentations of the entire training corpus. Since presenting the entire corpus in the item-focused unimodal and task-focused unimodal conditions involved presenting three times as many training trials as in the multimodal condition, comparison in terms of the number of presentations of the entire corpus represented an "equal error comparison". That is to say, error was considered across all items to be learned in each condition. The comparison on a trial-by-trial basis necessarily means that a higher error signal will be present in the multimodal condition since in each trial it provides information on the relation between three times as many cross domain mappings as in the unimodal condition (see Table 2.2). In order to make a detailed comparison between multimodal, item focused unimodal and task focused unimodal training, results were to be reported for both the item level trial by trial performance comparison as well as the corpus level equated error comparison.

The model was trained with batch learning (i.e. using a batch size of 9, representing the total number of possible cross-domain mappings, and allowing for item-focused presentation, with a weight update after the presentation of every batch of 9 training trials) using recurrent backpropagation through time with a steepest descent algorithm. The learning rate was set to 0.005. A weight decay of 0.0000002 was also used to prevent any weights developing values that were disproportionately high. Similarly no momentum was used since it's process of including a proportion of the previous step in every weight change can cause the effective learning rate to become too high and inhibit learning. Each individual unit within the network was given a fixed, untrainable bias of –2. "This has the effect of deducting 2 from each unit's net input. Thus, in the absence of input, each unit's activation settles to the low end of its activation range." (Rogers et al., 2004: p. 215). Units in all of the input/output layers were clamped to their input values using a soft clamp with a clamp strength of 0.9. A target radius of 0.1 was used during the processing of each batch (in this case each training trial due to a batch size of one) so if an output unit's activation is within 0.1 of the target, no error will be generated. The model was trained until input in a single layer could generate target outputs on all layers to within an accuracy of 0.5.

Table 2.2 Training regime patterns. Arrows indicate order of trial presentation for focused unimodal training. Multimodal training trials are presented in random order

Multimodal Training								
Multimodal cross-domain mappings		Items						
		dog	cat	raven	apple	car	boat	hammer
Input	Target Output	each learning batch contains 9 multimodal cross-domain mappings selected at random						
name	verbal descriptors							
	visual features							
	name							
verbal descriptors	name							
	visual features							
	verbal descriptors							
visual features	name							
	verbal descriptors							
	visual features							

Item Focused Unimodal Training								
Unimodal cross-domain mappings		Items						
		dog	cat	raven	apple	car	boat	hammer
Input	Target Output	each learning batch contains all 9 unimodal cross-domain mappings of a particular item selected at random						
name	verbal descriptors	↓	↓	↓	↓	↓	↓	↓
name	visual features	↓	↓	↓	↓	↓	↓	↓
name	name	↓	↓	↓	↓	↓	↓	↓
verbal descriptors	name	↓	↓	↓	↓	↓	↓	↓
verbal descriptors	visual features	↓	↓	↓	↓	↓	↓	↓
verbal descriptors	verbal descriptors	↓	↓	↓	↓	↓	↓	↓
visual features	name	↓	↓	↓	↓	↓	↓	↓
visual features	verbal descriptors	↓	↓	↓	↓	↓	↓	↓
visual features	visual features	↓	↓	↓	↓	↓	↓	↓

Task Focused Unimodal Training								
		Items						
Unimodal cross-domain mappings		dog	cat	raven	apple	car	boat	hammer
Input	Target Output	ordered presentation at task level each learning batch contains the same single cross-domain mapping for 9 items selected at random						
name	verbal descriptors	→	→	→	→	→	→	→
Name	visual features	→	→	→	→	→	→	→
Name	name	→	→	→	→	→	→	→
verbal descriptors	name	→	→	→	→	→	→	→
verbal descriptors	visual features	→	→	→	→	→	→	→
verbal descriptors	verbal descriptors	→	→	→	→	→	→	→
visual features	name	→	→	→	→	→	→	→
visual features	verbal descriptors	→	→	→	→	→	→	→
visual features	visual features	→	→	→	→	→	→	→

To simulate developmental learning, the model was trained for 129600 weight updates, using either the multimodal, item focused unimodal or task focused unimodal training regime in each simulation. Which equates to 10800 presentations in the multimodal condition (multimodal presentation of the entire corpus takes 12 updates i.e. 3 mappings for each of the 36 items, each trial consisting of a batch of 9 mappings). In contrast this equates to 3600 presentations in the item focused and task focused unimodal conditions (unimodal presentation of the entire corpus takes 36 updates i.e. i.e. 9 mappings for each of the 36 items, each trial consisting of a batch of 9 mappings). During training the model was tested on its ability for input in a single domain to generate the correct target outputs (to within 0.5) for all patterns, in all output modalities. 129600 weight updates was chosen for the duration of training through initial experimentation, as whenever the model learnt to 100% accuracy (verified through regular testing) it had occurred by this time. The simulation was run ten times in the multimodal condition and ten times in the unimodal condition and the results were averaged. During training regular testing recorded accuracy (percentage of examples correct upon testing), network error and representational economy (according to the equation derived below).

Calculating representational economy in the semantic convergence zone

Representational economy in the semantic convergence zone was calculated in the exact same manner as described in Chapter 1.

Testing the network and analysing representational economy

Representational economy, accuracy and error, were calculated at regular intervals during the model's training. 2x20 repeated measures analyses of variance (ANOVA) were conducted on this data to compare the effect of manipulating the training regime (multimodal, item focused unimodal or task focused unimodal) at twenty equally spaced time points during training on the dependent variables (accuracy, representational economy and network error). For all simulations the correlation between representational economy, accuracy, and network error was calculated to understand the developing Representational Economy's relationship to Accuracy and Network Error as the model learns.

Results

Figure 2.3 illustrates the model's performance during development, both on a trial-by-trial basis and for an equated error comparison. Despite observing a small difference in accuracy performance between item-focused and task-focused unimodal presentation during the "vocabulary spurt" period of developmental learning, overall there were no major benefits for item-focused unimodal presentation. Certainly, neither come close to the performance levels observed in simultaneous multimodal presentation, which is more efficient when considered on a trial-by-trial basis, obtaining greater accuracy, and greater convergence (as evidenced by representational economy) as well as greater error reduction. These differences between multimodal and unimodal training are statistically significant as the 2×20 analyses of variance show. Figure 2.3 also shows how the multimodal training achieves greater acceleration in the learning process thus showing better performance at an earlier point. The equal error comparison between multimodal, item focused and task focused unimodal training shows no substantial difference and indeed the small differences seen in Figure 2.3 are not statistically significant (as shown in the results of the analysis in Table 2.3). Statistical differences between the different learning conditions can be seen in Table 2.3: Multimodal training outperforms both Item-focused and Task-focused Unimodal training in representational economy, accuracy and network error. Comparison of the Equated Error Multimodal and Item-focused Unimodal regimes show significant interaction between training and time for accuracy and network error and a significant effect of time for representational economy, network error and accuracy. However there is no significant performance difference for the effect of training between Equated Error Multimodal and Item-focused Unimodal. Similarly there is no significant performance difference in accuracy, representational economy or network error for the effect of training between Item-focused Unimodal and Task-focused Unimodal training. A significant interaction between training and time for accuracy can however be seen, as well as a significant effect of time on accuracy, representational economy and network error. Representational Economy was strongly correlated with learning Accuracy for all conditions: Multimodal $r(20)=.887$, $p<.001$; Equated Error Multimodal $r(20)=.888$, $p<.001$; Item-focused Unimodal $r(20)=.906$, $p<.001$; Task-focused Unimodal $r(20)=.888$, $p<.001$. Representational Economy was strongly correlated with Network Error for all conditions: Multimodal $r(20)=-.865$, $p<.001$; Equated Error Multimodal $r(20)=-.702$, $p=.001$; Item-focused Unimodal $r(20)=-.702$, $p=.001$; Task-focused Unimodal $r(20)=-.680$, $p=.001$.

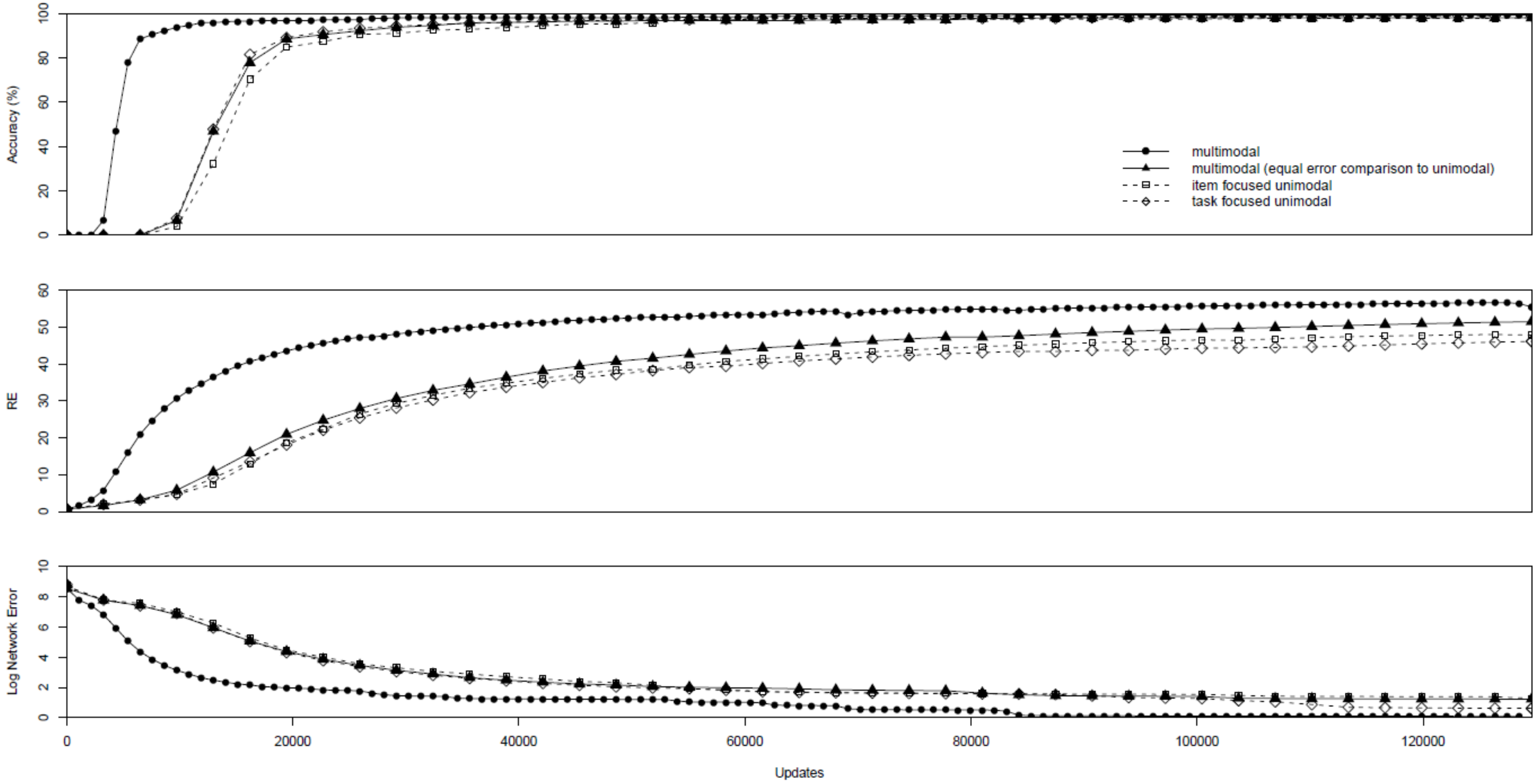

Figure 2.3 Mean variation in training performance during developmental learning

Table 2.3 2x20 ANOVAs indicating significant performance differences between training regimes during developmental learning

Training Regime Comparison	Dependent Variable Measure	Independent Variables	df	F	p	Effect size (η^2)
Multimodal Vs Item Focused Unimodal	Accuracy	Training	1,9	204.194	<0.001*	0.958
		Time	19,171	1233.681	<0.001*	0.993
		Training*Time	19,171	407.577	<0.001*	0.978
	Representational Economy	Training	1,9	81.712	<0.001*	0.901
		Time	19,171	963.772	<0.001*	0.991
		Training*Time	19,171	80.772	<0.001*	0.900
	Network Error	Training	1,9	506.474	<0.001*	0.983
		Time	19,171	447.579	<0.001*	0.980
		Training*Time	19,171	334.652	<0.001*	0.974
Equated Error Multimodal Vs Item Focused Unimodal	Accuracy	Training	1,9	1.381	0.270	0.133
		Time	19,171	2650.994	<0.001*	0.997
		Training*Time	19,171	4.977	<0.001*	0.356
	Representational Economy	Training	1,9	3.522	0.093	0.281
		Time	19,171	1447.881	<0.001*	0.994
		Training*Time	19,171	1.400	0.132	0.135
	Network Error	Training	1,9	5.296	0.047	0.370
		Time	19,171	1255.340	<0.001*	0.993
		Training*Time	19,171	3.306	<0.001*	0.269
Multimodal Vs Task Focused Unimodal	Accuracy	Training	1,9	316.016	<0.001*	0.972
		Time	19,171	871.370	<0.001*	0.990
		Training*Time	19,171	448.648	<0.001*	0.980
	Representational Economy	Training	1,9	182.117	<0.001*	0.953
		Time	19,171	139.297	<0.001*	0.994
		Training*Time	19,171	169.592	<0.001*	0.950
	Network Error	Training	1,9	1583.328	<0.001*	0.994
		Time	19,171	917.542	<0.001*	0.990
		Training*Time	19,171	741.087	<0.001*	0.988

Training Regime Comparison [continued]	Dependent Variable Measure	Independent Variables	df	F	p	Effect size (η^2)
Item Focused Unimodal Vs Task Focused Unimodal	Accuracy	Training	1,9	3.705	0.086	0.292
		Time	19,171	2726.844	<0.001*	0.997
		Training*Time	19,171	7.444	<0.001*	0.453
	Representational Economy	Training	1,9	0.327	0.581	0.035
		Time	19,171	2071.204	<0.001*	0.996
		Training*Time	19,171	1.263	0.214	0.123
	Network Error	Training	1,9	0.397	0.544	0.042
		Time	19,171	2119.819	<0.001*	0.996
		Training*Time	19,171	1.009	0.454	0.101

Simulation 2.2: Robustness to Damage

Method

The model was trained in the manner described in Simulation 2.1. In order to investigate the model's robustness to damage the 10 multimodal, 10 item-focused unimodal and 10 task-focused unimodal trained networks were lesioned by removing an increasing proportion of all incoming and outgoing connections across all the units in all layers (i.e. names, verbal descriptions, visual features and semantics). Each lesion was performed 10 times and scores for accuracy, error and representational economy were recorded. The data were then plotted (see Figure 2.4) and analysed (results in Table 2.4) to test if the multimodal-trained models were more robust to damage than those generated from item-focused and task-focused unimodal training. A 2x11 repeated-measures analysis of variance (ANOVA) was conducted on these data, using the dependent measures of representational economy, accuracy and error at each lesion point, to investigate the effect of manipulating the training regime (multimodal, item-focused or task-focused unimodal).

Results

Figure 2.4 illustrates the model's robustness to damage as the model was subjected to lesions of increasing severity. Across the range of lesions, the multimodal regime consistently appeared to be more robust to damage, with higher accuracy and lower network error, both for trial-by-trial, and equal error, when compared to item-focused or task-focused unimodal training. There was a statistically significant effect of training type on robustness to damage (see Table 2.4) However, for all except the smallest lesions (i.e. less than 2% of connections removed) RE after damage was reduced to the same low level regardless of training type. As found in Chapter 1, a clear advantage in terms of robustness to damage was present for the simultaneous multimodal training. Item-focused unimodal training generated no statistically significant greater robustness to damage than task-focused unimodal training. Thus, the differences in order and grouping between these two forms of unimodal training did not alter the degree of robustness to damage that unimodal training was capable of achieving. From Table 2.4 it can be seen that multimodal training yields a structure that is more robust to damage than those structures generated by Item-focused or Task-focused Unimodal training. There is a statistically significant difference between multimodal and item-focused unimodal for accuracy and representational economy, as well as a significant interaction between training and time for accuracy, representational economy and network error. There is also a statistically significant difference in the effect of training between multimodal and task-focused unimodal for accuracy and representational economy. There is only a significant effect of training on representational economy when comparing equated error multimodal and item-focused unimodal training as well as an effect of time and a significant interaction between training and time on accuracy, representational economy and network error. In comparing Item-focused Unimodal and Task-focused Unimodal there is a significant effect of time on accuracy, representational economy and network error as well as a significant interaction between training and time for network error only. Representational Economy was strongly correlated with learning Accuracy for all conditions: Multimodal $r(11)=.870$, $p<.001$; Equated Error Multimodal $r(11)=.877$, $p<.001$; Item-focused Unimodal $r(11)=.873$,$p<.001$; Task-focused Unimodal $r(11)=.893$,$p<.001$. Representational Economy was correlated with Network Error for all conditions: Multimodal $r(11)=-.591$,$p<.006$;Equated Error Multimodal $r(11)=-.595$,$p=.006$;Item-focused Unimodal $r(11)=-.617$,$p=.004$; Task-focused Unimodal $r(11)=-.569$,$p=.009$.

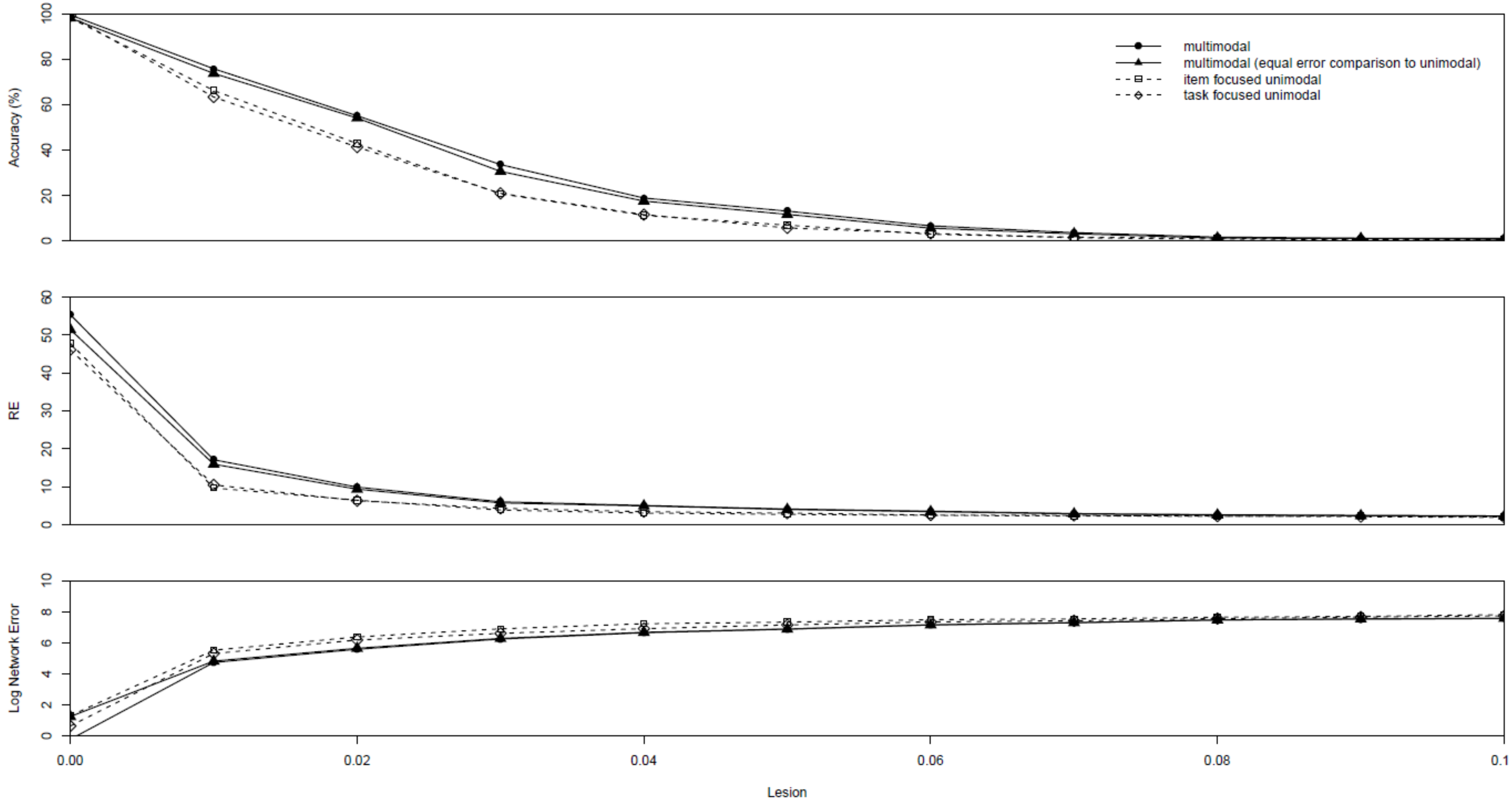

Figure 2.4 Mean variation in robustness to damage for increasingly severe lesions to the fully trained model

Table 2.4 2x11 ANOVAs indicating significant differences in robustness to damage between trained networks

Training Regime Comparison	Dependent Variable Measure	Independent Variables	df	F	p	Effect size (η^2)
Multimodal Vs Item Focused Unimodal	Accuracy	Training	1,9	6.747	0.029	0.428
		Time	10,90	693.544	<0.001*	0.987
		Training*Time	10,90	3.356	<0.001*	0.272
	Representational Economy	Training	1,9	9.680	0.012	0.518
		Time	10,90	761.322	<0.001*	0.988
		Training*Time	10,90	7.597	<0.001*	0.458
	Network Error	Training	1,9	1.819	0.210	0.168
		Time	10,90	365.321	<0.001*	0.976
		Training*Time	10,90	1.901	0.017	0.174
Equated Error Multimodal Vs Item Focused Unimodal	Accuracy	Training	1,9	4.137	0.072	0.315
		Time	10,90	664.804	<0.001*	0.987
		Training*Time	10,90	2.326	0.002	0.205
	Representational Economy	Training	1,9	5.913	0.038	0.397
		Time	10,90	955.219	<0.001*	0.991
		Training*Time	10,90	4.520	<0.001*	0.334
	Network Error	Training	1,9	1.607	0.237	0.151
		Time	10,90	375.897	<0.001*	0.977
		Training*Time	10,90	1.816	0.024	0.168
Multimodal Vs Task Focused Unimodal	Accuracy	Training	1,9	11.286	0.008	0.556
		Time	10,90	830.903	<0.001*	0.989
		Training*Time	10,90	3.733	<0.001*	0.293
	Representational Economy	Training	1,9	19.213	0.002	0.681
		Time	10,90	875.057	<0.001*	0.990
		Training*Time	10,90	10.382	<0.001*	0.536
	Network Error	Training	1,9	4.079	0.074	0.312
		Time	10,90	350.308	<0.001*	0.975
		Training*Time	10,90	0.703	0.813	0.072

Training Regime Comparison [continued]	Dependent Variable Measure	Independent Variables	df	F	p	Effect size (η^2)
Item Focused Unimodal Vs Task Focused Unimodal	Accuracy	Training	1,9	0.042	0.842	0.005
		Time	10,90	870.573	<0.001*	0.990
		Training*Time	10,90	0.156	1.000	0.017
	Representational Economy	Training	1,9	0.044	0.839	0.005
		Time	10,90	1487.438	<0.001*	0.994
		Training*Time	10,90	0.499	0.961	0.052
	Network Error	Training	1,9	0.025	0.879	0.003
		Time	10,90	326.095	<0.001*	0.973
		Training*Time	10,90	3.387	<0.001*	0.273

Simulation 2.3: Relearning after damage

Method

The model was trained in the manner described in Simulation 2.1. However, this time only the multimodal training regime was used to train the model as per the original simulation (see Chapter 1; and Rogers et al, 2004). In order to investigate the model's re-learning behaviour after damage, the 10 multimodally-trained networks from the initial training were lesioned by removing a proportion of all incoming and outgoing connections across all the units in all layers (i.e. names, verbal descriptions, visual features and semantics). Based upon the findings from Chapter 1 lesioning was again required such that relearning would only occur to a certain level after damage to enable the exploration of relearning for a range of damage. The same three levels of damage from Chapter 1 were used: Mild where the model relearns to around 90% accuracy, Moderate, where the model relearns to around 60% accuracy, and Severe where the model relearns to around 30% accuracy. It is again acknowledged that this is not identical to the real situation of localised injury. As described in Chapter 1 this lesioning was uniform across the model since it was the only form of lesioning that yielded the practical situation in the model that the study wanted to explore. The trained networks were subjected to three separate lesions of varying degrees of severity; mild (removing 86 % of connections), moderate (removing 88% of connections) and severe (removing 90% of connections). These degrees of severity were intended to encompass a wide range of damage. After lesioning, each network underwent 2592000 updates of retraining with either multimodal and unimodal training regimes. This corresponds to 21600 multimodal and 7200 unimodal presentations of the entire training corpus. This repeated-measures methodology, where the same initially trained network was subjected to each lesion and retraining conditions, allowed us to remove any variance associated with individual pre- or postmorbid differences. This is important since brain-damaged patients are not a homogenous population and studying the effects of treatment should be based upon "within-patient comparisons" (Howard & Hatfield, 1987, p.119).

Representational economy, accuracy and error were calculated at regular intervals during the model's training. A 2x20 repeated measures analysis of variance (ANOVA) was conducted on these data to investigate the effect of manipulating the training regime (multimodal, item-focused unimodal or task-focused unimodal) at twenty equally-spaced time points during training on the dependent variables (accuracy, representational economy and network error) for each lesion condition.

Results

In all graphs of the results relearning is shown on a trial by trial basis (where each trial consists of the presentation of 9 cross-modal mappings to the model in batch learning) for each of the training regimes (i.e. multimodal, item-focused unimodal and task-focused unimodal). The multimodal regime is also plotted for an equated error comparison to the unimodal regimes. A table of results also shows the results of 2x20 ANOVAs carried out on each of the dependent measures (i.e. accuracy, representational economy and network error) for the 4 pairwise comparisons of the training regime that are necessary to confirm the findings displayed in the graphs.

Mild Lesioning

The graphs in Figure 2.5 illustrate the model's performance during relearning after a mild lesion, which removed 86% of the model's connections. It can be clearly seen that multimodal training is more efficient than either of the unimodal training regimes on a trial by trial basis, obtaining greater accuracy, and greater convergence (as shown by higher values of representational economy) as well as greater error reduction. These differences between multimodal and unimodal training are also statistically significant as the results of the 2x20 ANOVA given in Table 2.5 clearly show. The steeper slope of the line showing accuracy during multimodal training in Figure 2.5, also shows how the multimodal training achieves greater acceleration during relearning. The equal error comparison between multimodal and the two unimodal training regimes shows no difference, this lack of difference is confirmed by the results of the ANOVAs shown in Table 2.5. In examining Table 2.5 the advantage of multimodal training over Item-focused and Task-Focused Unimodal training can be seen in the statistically significant difference between multimodal and item-focused or task focused unimodal learning in accuracy, representational economy and network error. Comparing Equated Error Multmodal and Item-focused unimodal shows a significant effect of time on accuracy, representational economy and network error, as well as an effect of training on representational economy. Contrasting Item-focused and Task-focused Unimodal training gives a significant effect of time as well as an effect of training on network error. Representational Economy was strongly correlated with learning Accuracy for all conditions: Multimodal $r(20)=.970$, $p<.001$; Equated Error Multimodal $r(20)=.899$, $p<.001$; Item-focused Unimodal $r(20)=.938, p<.001$; Task-focused Unimodal $r(20)=.924, p<.001$. Representational Economy was strongly correlated with Network Error for all conditions: Multimodal $r(20)=-.849, p<.001$;Equated Error Multimodal $r(20)=-.932, p<.001$;Item-focused Unimodal $r(20)=-.901, p<.001$; Task-focused Unimodal $r(20)=-.919, p<.001$.

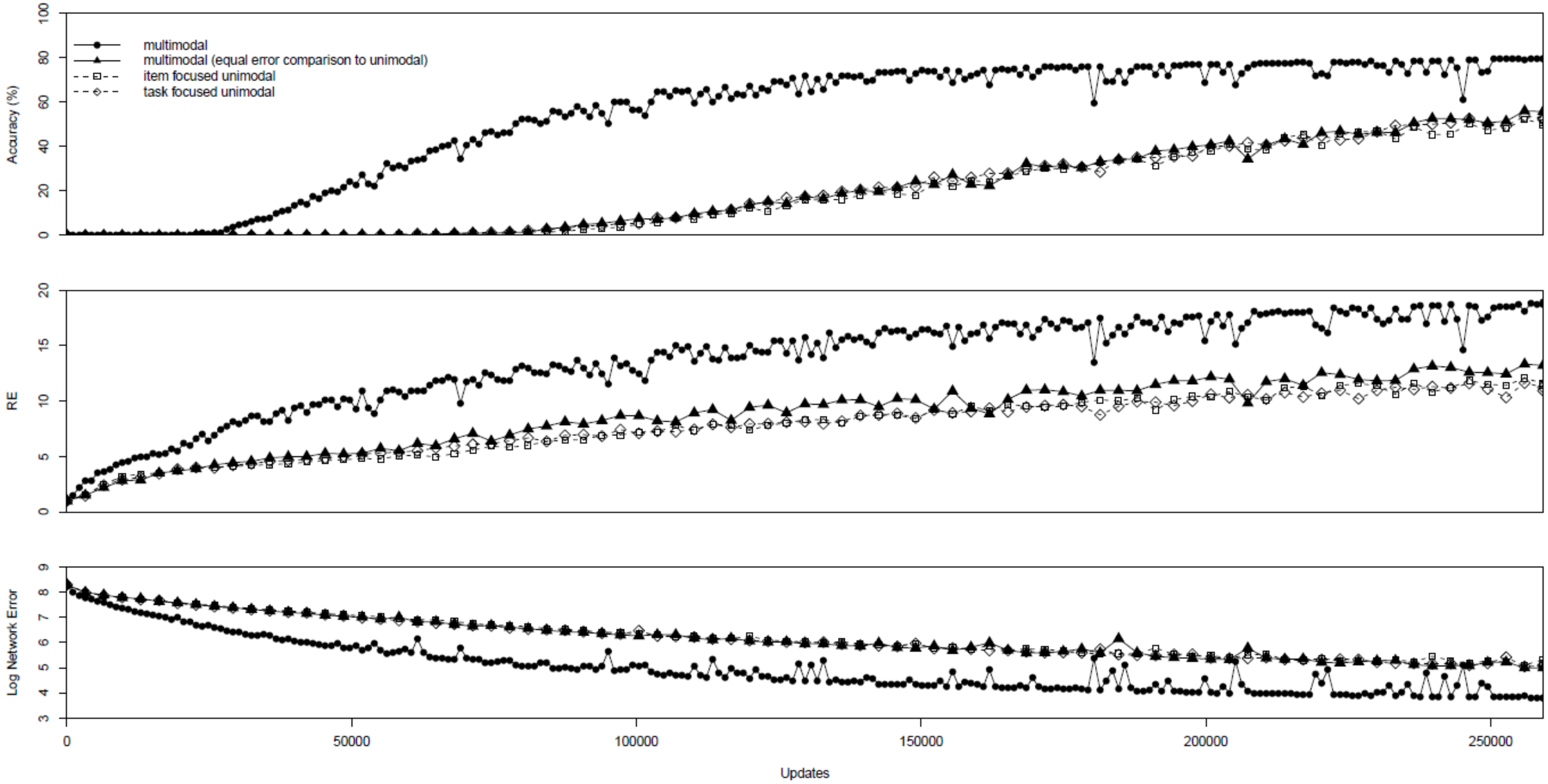

Figure 2.5 Mean variation in training performance during relearning after a mild lesion

Table 2.5 2x20 ANOVAs indicating significant performance differences between training regimes after a mild lesion

Training Regime Comparison	Dependent Variable Measure	Independent Variables	df	F	p	Effect size (η^2)
Multimodal Vs Item Focused Unimodal	Accuracy	Training	1,9	777.117	<0.001*	0.989
		Time	19,171	149.019	<0.001*	0.943
		Training*Time	19,171	23.496	<0.001*	0.723
	Representational Economy	Training	1,9	160.503	<0.001*	0.947
		Time	19,171	63.577	<0.001*	0.876
		Training*Time	19,171	3.790	<0.001*	0.296
	Network Error	Training	1,9	1056.707	<0.001*	0.992
		Time	19,171	601.826	<0.001*	0.985
		Training*Time	19,171	73.147	<0.001*	0.890
Equated Error Multimodal Vs Item Focused Unimodal	Accuracy	Training	1,9	2.543	0.145	0.220
		Time	19,171	118.922	<0.001*	0.930
		Training*Time	19,171	0.874	0.615	0.089
	Representational Economy	Training	1,9	33.195	<0.001*	0.787
		Time	19,171	91.086	<0.001*	0.910
		Training*Time	19,171	1.625	0.055	0.153
	Network Error	Training	1,9	3.345	0.101	0.271
		Time	19,171	861.787	<0.001*	0.990
		Training*Time	19,171	0.887	0.600	0.090
Multimodal Vs Task Focused Unimodal	Accuracy	Training	1,9	525.672	<0.001*	0.983
		Time	19,171	165.614	<0.001*	0.948
		Training*Time	19,171	23.393	<0.001*	0.722
	Representational Economy	Training	1,9	256.310	<0.001*	0.966
		Time	19,171	64.953	<0.001*	0.878
		Training*Time	19,171	4.543	<0.001*	0.335
	Network Error	Training	1,9	4114.769	<0.001*	0.998
		Time	19,171	648.042	<0.001*	0.986

Training Regime Comparison [continued]	Dependent Variable Measure	Independent Variables	df	F	p	Effect size (η^2)
Item Focused Unimodal Vs Task Focused Unimodal	Accuracy	Training	1,9	1.884	0.203	0.173
		Time	19,171	129.217	<0.001*	0.935
		Training*Time	19,171	0.707	0.808	0.073
	Representational Economy	Training	1,9	0.002	0.970	<0.001
		Time	19,171	111.850	<0.001*	0.926
		Training*Time	19,171	0.911	0.570	0.092
	Network Error	Training	1,9	5.293	0.047	0.370
		Time	19,171	1406.008	<0.001*	0.994
		Training*Time	19,171	1.035	0.424	0.103

Moderate Lesioning

Here the picture is very similar to the results shown for milder lesioning. Though it is worth noting that at end point of relearning the difference between the multimodal training regime and all other training regimes is more pronounced, The graphs in Figure 2.6 show the model's relearning performance after a moderate lesion, which removed 88% of the model's connections. As in the case of mild lesioning multimodal training is more efficient than either of the unimodal training regimes on a trial by trial basis, obtaining greater accuracy, convergence (as shown by higher values of representational economy) as well as greater error reduction. These differences between multimodal and unimodal training are also statistically significant as the results of the 2x20 ANOVA given in Table 2.6 clearly show. The line showing accuracy during multimodal training in Figure 2.6, also shows how the multimodal training achieves faster learning. The equal error comparison between multimodal and the two unimodal training regimes does shows a slightly better performance which is confirmed by the results of the ANOVAs shown in Table 2.6. In Table 2.6 the multimodal advantage can be observed in the statistically significant difference between multimodal and item-focused or task focused unimodal learning in accuracy, representational economy and network error. There is also a statistically significant difference between equated error multimodal and item-focused unimodal learning in accuracy, representational economy and network error. When comparing Item-focused and Task-focused Unimodal there is an effect of time on accuracy, representational economy and network error and a significant interaction between training and time on network error. Representational Economy was strongly correlated with learning Accuracy for all conditions: Multimodal $r(20)=.961$, $p<.001$; Equated Error Multimodal $r(20)=.827$, $p<.001$; Item-focused Unimodal $r(20)=.792$, $p<.001$; Task-focused Unimodal $r(20)=.793$, $p<.001$. Representational Economy was strongly correlated with Network Error for all conditions: Multimodal $r(20)=-.900$, $p<.001$; Equated Error Multimodal $r(20)=-.949$, $p<.001$; Item-focused Unimodal $r(20)=-.964$, $p<.001$; Task-focused Unimodal $r(20)=-.972$, $p<.001$.

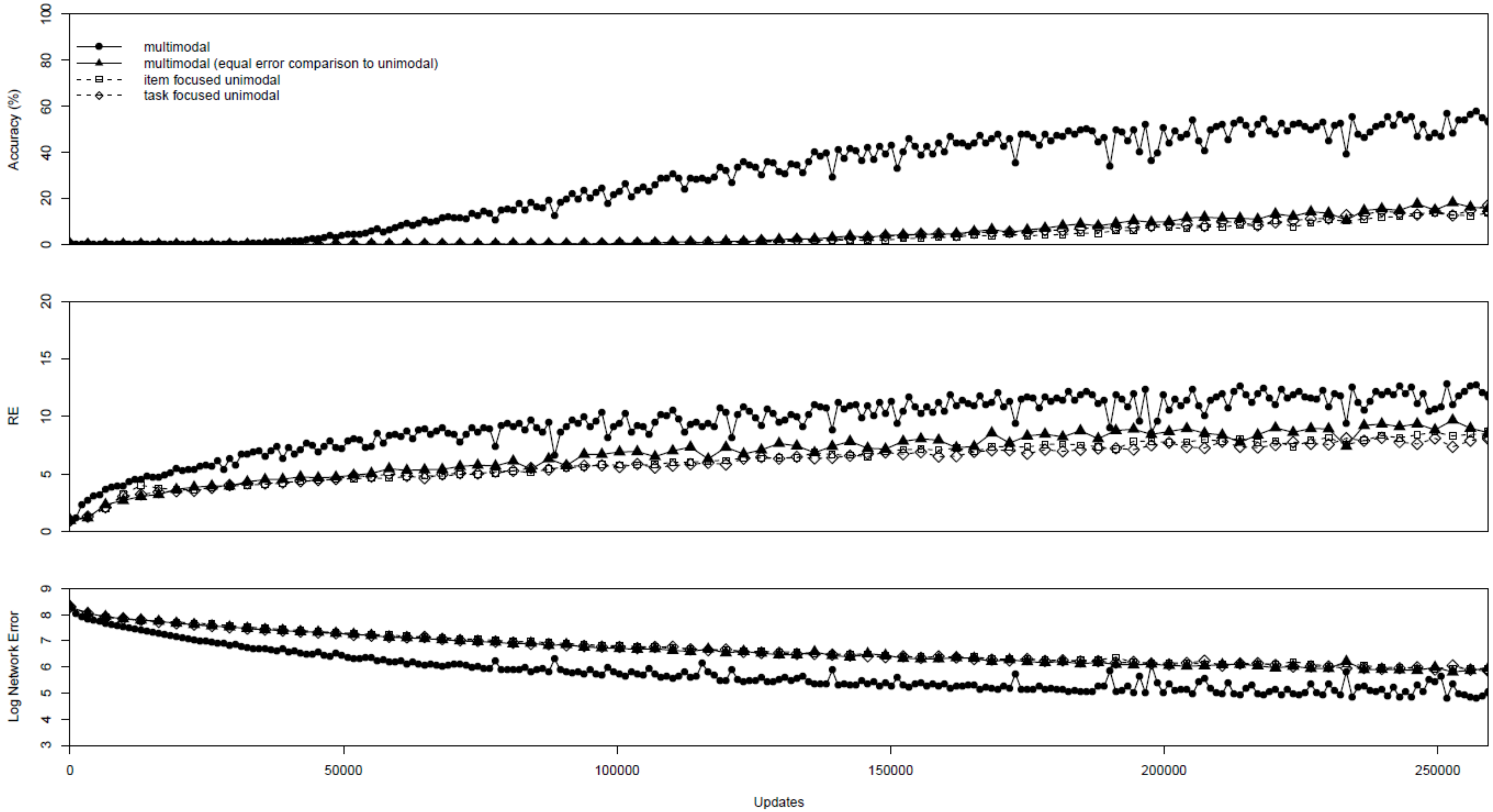

Figure 2.6 Mean variation in training performance during relearning after a moderate lesion

Table 2.6 2x20 ANOVAs indicating significant performance differences between training regimes after a moderate lesion

Training Regime Comparison	Dependent Variable Measure	Independent Variables	df	F	p	Effect size (η^2)
Multimodal Vs Item Focused Unimodal	Accuracy	Training	1,9	330.307	<0.001*	0.973
		Time	19,171	171.052	<0.001*	0.950
		Training*Time	19,171	58.747	<0.001*	0.867
	Representational Economy	Training	1,9	529.309	<0.001*	0.983
		Time	19,171	107.246	<0.001*	0.923
		Training*Time	19,171	4.089	<0.001*	0.312
	Network Error	Training	1,9	10752.065	<0.001*	0.999
		Time	19,171	869.472	<0.001*	0.990
		Training*Time	19,171	59.646	<0.001*	0.869
Equated Error Multimodal Vs Item Focused Unimodal	Accuracy	Training	1,9	14.634	0.004	0.619
		Time	19,171	58.217	<0.001*	0.866
		Training*Time	19,171	4.175	<0.001*	0.317
	Representational Economy	Training	1,9	25.473	<0.001*	0.739
		Time	19,171	105.642	<0.001*	0.921
		Training*Time	19,171	2.152	0.005	0.193
	Network Error	Training	1,9	31.856	<0.001*	0.780
		Time	19,171	1276.168	<0.001*	0.993
		Training*Time	19,171	2.063	0.008	0.186
Multimodal Vs Task Focused Unimodal	Accuracy	Training	1,9	419.091	<0.001*	0.979
		Time	19,171	159.264	<0.001*	0.947
		Training*Time	19,171	63.518	<0.001*	0.876
	Representational Economy	Training	1,9	267.988	<0.001*	0.968
		Time	19,171	83.998	<0.001*	0.903
		Training*Time	19,171	7.184	<0.001*	0.444
	Network Error	Training	1,9	1631.492	<0.001*	0.995
		Time	19,171	708.504	<0.001*	0.987
		Training*Time	19,171	46.706	<0.001*	0.838

Training Regime Comparison [continued]	Dependent Variable Measure	Independent Variables	df	F	p	Effect size (η^2)
Item Focused Unimodal Vs Task Focused Unimodal	Accuracy	Training	1,9	1.259	0.291	0.123
		Time	19,171	62.128	<0.001*	0.873
		Training*Time	19,171	1.041	0.417	0.104
	Representational Economy	Training	1,9	1.619	0.235	0.152
		Time	19,171	127.135	<0.001*	0.934
		Training*Time	19,171	1.292	0.194	0.126
	Network Error	Training	1,9	3.968	0.078	0.306
		Time	19,171	1068.487	<0.001*	0.992
		Training*Time	19,171	1.656	0.048	0.155

Severe Lesioning

Severe lesioning, which removed 90% of the model's connections, echoes the findings of a multimodal advantage from the mild and moderate lesioning conditions. At end point of relearning the difference between the multimodal training regime and all other training regimes is very small, thought still statistically significant (see Table 2.7). The graphs in Figure 2.7 show the model's relearning performance after a severe lesion. As in the case of mild, and moderate, lesioning multimodal training is more efficient than either of the unimodal training regimes on a trial by trial basis. These differences between multimodal and unimodal training are also statistically significant as the results of the 2x20 ANOVA given in Table 2.7 clearly show. Multimodal training obtains greater accuracy, convergence (as shown by higher values of representational economy despite the obvious noise which is reflected in the small effect sizes for the multimodal comparisons in Table 2.7), and greater error reduction. The equal error comparison between multimodal and the two unimodal training regimes does shows a slightly better performance in representational economy and network error, but not in accuracy which is confirmed by the results of the ANOVAs shown in Table 2.7. Table 2.7 highlights the statistically significant difference between multimodal and item-focused or task focused unimodal learning in accuracy, representational economy and network error that signifies the advantage of multimodal training over the other regimes. Comparing Equated Error Multmodal and Item-focused unimodal shows a significant difference in representational economy and network error but only an effect of time on accuracy. Similarly the comparison of item-focused and task-focused unimodal shows a significant effect of time on accuracy, representational economy and network error and an effect of training on network error. Representational Economy was strongly correlated with learning Accuracy for all conditions: Multimodal $r(20)=.770$, $p<.001$; Equated Error Multimodal $r(20)=.702$, $p<.001$; Item-focused Unimodal $r(20)=.819, p<.001$; Task-focused Unimodal $r(20)=.662, p<.001$. Representational Economy was strongly correlated with Network Error for all conditions: Multimodal $r(20)=-.937, p<.001$;Equated Error Multimodal $r(20)=-.982, p<.001$;Item-focused Unimodal $r(20)=-.970, p<.001$; Task-focused Unimodal $r(20)=-.980, p<.001$.

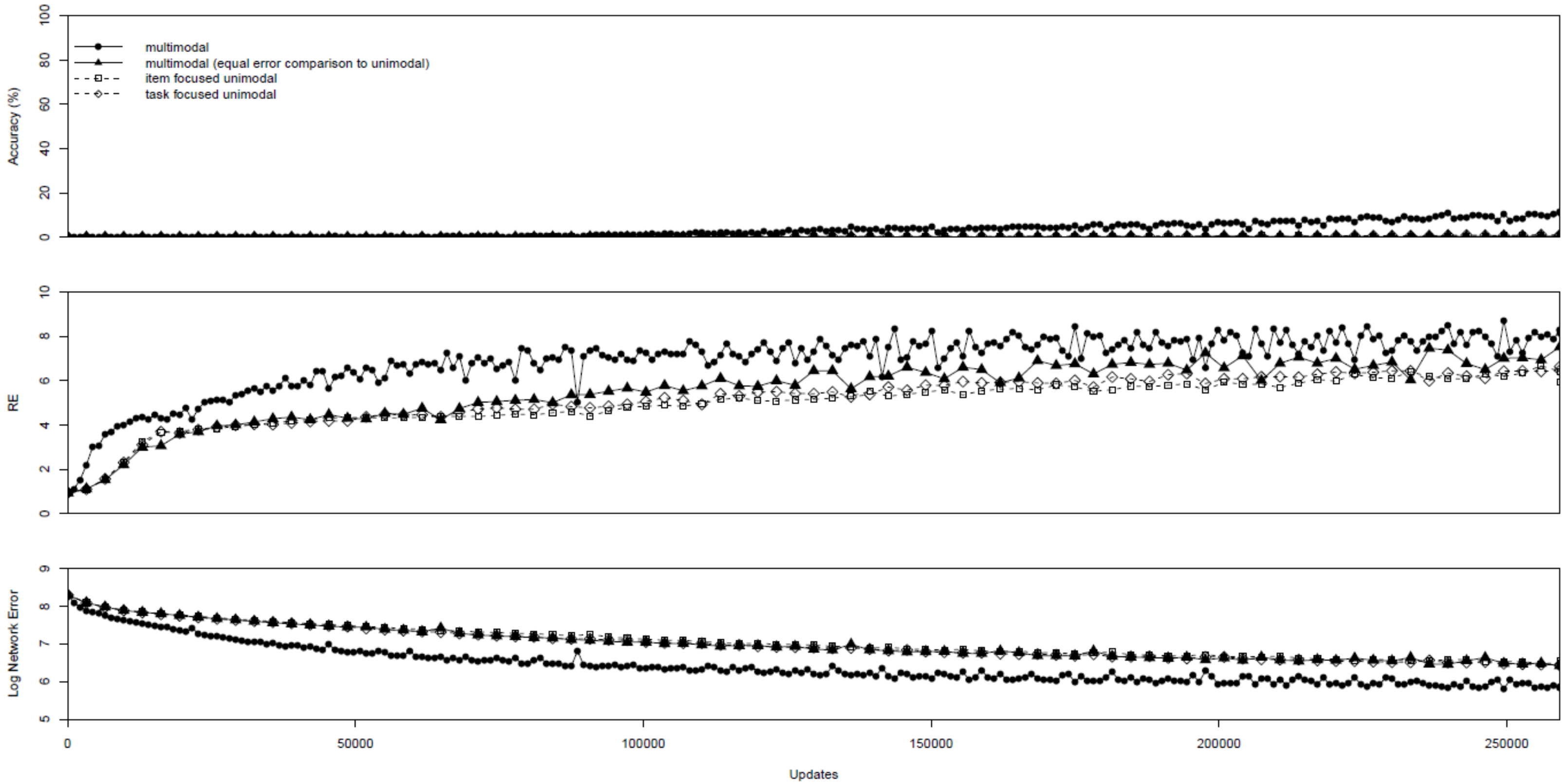

Figure 2.7 Mean variation in training performance during relearning after a severe lesion

Table 2.7 2x20 ANOVAs indicating significant performance differences between training regimes after a severe lesion

Training Regime Comparison	Dependent Variable Measure	Independent Variables	df	F	p	Effect size (η^2)
Multimodal Vs Item Focused Unimodal	Accuracy	Training	1,9	24.745	<0.001*	0.733
		Time	19,171	17.103	<0.001*	0.655
		Training*Time	19,171	13.251	<0.001*	0.596
	Representational Economy	Training	1,9	194.458	<0.001*	0.956
		Time	19,171	50.814	<0.001*	0.850
		Training*Time	19,171	1.979	0.012	0.180
	Network Error	Training	1,9	1726.137	<0.001*	0.995
		Time	19,171	648.350	<0.001*	0.987
		Training*Time	19,171	31.488	<0.001*	0.778
Equated Error Multimodal Vs Item Focused Unimodal	Accuracy	Training	1,9	2.481	0.150	0.216
		Time	19,171	2.735	<0.001*	0.233
		Training*Time	19,171	0.950	0.522	0.096
	Representational Economy	Training	1,9	57.100	<0.001*	0.864
		Time	19,171	94.201	<0.001*	0.913
		Training*Time	19,171	3.307	<0.001*	0.269
	Network Error	Training	1,9	19.639	0.002	0.686
		Time	19,171	738.045	<0.001*	0.988
		Training*Time	19,171	1.536	0.079	0.146
Multimodal Vs Task Focused Unimodal	Accuracy	Training	1,9	25.720	0.001	0.741
		Time	19,171	17.061	<0.001*	0.655
		Training*Time	19,171	13.216	<0.001*	0.595
	Representational Economy	Training	1,9	213.703	<0.001*	0.960
		Time	19,171	38.977	<0.001*	0.812
		Training*Time	19,171	3.070	<0.001*	0.254
	Network Error	Training	1,9	1811.254	<0.001*	0.995
		Time	19,171	723.156	<0.001*	0.988
		Training*Time	19,171	42.187	<0.001*	0.824

Training Regime Comparison [continued]	Dependent Variable Measure	Independent Variables	df	F	p	Effect size (η^2)
Item Focused Unimodal Vs Task Focused Unimodal	Accuracy	Training	1,9	0.854	0.380	0.087
		Time	19,171	3.369	<0.001*	0.272
		Training*Time	19,171	1.091	0.364	0.108
	Representational Economy	Training	1,9	1.540	0.246	0.146
		Time	19,171	109.582	<0.001*	0.924
		Training*Time	19,171	0.736	0.778	0.076
	Network Error	Training	1,9	16.311	0.003	0.644
		Time	19,171	1380.713	<0.001*	0.994
		Training*Time	19,171	1.299	0.190	0.126

Similar to the picture observed in development, item-focused unimodal training has no real advantage over task-focused unimodal training. Yet again, we observed a statistically significant multimodal advantage. There is a significant effect of training although it is worth noting that the differences between the two unimodal regimes observed during the "vocabulary spurt" period of development learning do not manifest themselves in recovery learning. Most importantly it can be seen that the levels of efficiency achievable through simultaneous multimodal item presentation cannot be achieved through rearrangement of unimodal item presentation nor through existing forms of unimodal presentation (i.e. task focused) commonly observed during therapies (e.g. picture naming etc). The error equated comparisons echo the finding from Chapter 1, that the observed multimodal benefits are a function of the higher error signal available for simultaneous multimodal presentation of learning items.

Discussion

Summary of Results

Reorganising unimodal learning sequentially at the item level, to make it item-focused, does not approximate the benefits of multimodal learning, nor does it prove more efficient than task-focused unimodal learning. In development (see Figure 2.3), during the ‘vocabulary spurt’ period (between 10000 and 20000 updates), item-focused unimodal learning performed worse than all other forms of learning, in terms of accuracy. Whilst convergence (as measured with representational economy) appeared to benefit slightly (see Figure 2.3) from item-focused unimodal learning, such benefits were not statistically significant. When considering robustness to damage (see Figure 2.4) multimodal learning yielded a more robust representational structure. There was no significant difference between the level of robustness derived from item-focused and task-focused unimodal learning. The picture for spontaneous recovery after damage (see Figures 2.5,2.6 and 2.7) remained consistent for all levels of lesioning in terms of the performance of the different training regimes. It is also worth noting that in the recovery picture there was no separation, at any point, between the accuracy achieved by item-focused or task-focused unimodal learning. Whilst minor differences in representational economy were observed between item-focused and task-focused unimodal learning, such differences were not statistically significant (see Tables 2.5,2.6 and 2.7). The picture emerging from the current investigation supports those observations in Chapter 1 regarding multimodal learning. Most interestingly, the current investigation cleared up the question of whether reorganised unimodal learning could approximate multimodal learning, suggesting that it is the simultaneity of multimodal item presentation that is important, as opposed to merely presentation order. The results suggest that simultaneous multimodal learning achieves a consistent, statistically significant, advantage in efficiency compared to unimodal learning regardless of different forms of item organisation within the unimodal learning strategy. The lack of any statistically significant difference between the equated error comparison to the item-focused and task-focused unimodal conditions shows these advantages be a function of the higher error signal available during simultaneous multimodal presentation of learning items.

Analysis of multimodal learning in the context of efficiency

The results of the current study, and indeed of the investigations in Chapter 1, are clear – knowledge about an object can be learnt more efficiently, if more information about that object is provided at the time at which it is learnt. The current investigation attempted to uncover the computational basis for why this should be the case. At this point, it is worth recalling that the model used in this study is one of semantic memory. Semantic memory, as part of declarative memory (Squire, 1992), has been extensively investigated not only in terms of how a semantic system can arise, but also in terms of what kinds of cognitive processing could account for such a system. Various cognitive researchers (e.g. Elio and Reutener,1970; Tulving, 1962) suggest that there is an advantage for memory in terms of how items are organised at the point when they are encoded. Hunt and McDaniel (1993) offer the idea that two different types of processing are at work when items are learnt: Firstly, the detection of similarity through 'relational processing'. Secondly, the detection of item distinctiveness through ‘item-specific processing’. Previous research by Mandler (1979) described item-specific processes as ‘integration’ whereby increasingly fluent detail was learnt about a specific item through its continual processing (e.g. maintenance rehearsal). Mandler (1979) also described relational processes as ‘elaboration’, in other words processing that establishes relationships between item representations in memory and elaborates knowledge of those items (i.e. performs semantic processing). In the context of discussing processing the Rogers et al. (2004) model of semantic memory incorporates both item-specific and relational processing. The category structure that develops amongst the model’s learnt internal semantic representations, as shown by Rogers et al. (2004) using cluster analysis, is evidence of relational processing. However relational processing relies upon preceding item specific processing. In other words all cross modal perceptual relations for a particular item need to be processed in order for that item to be correctly related to another item during further processing. As the model learns there is then a transition from item specific processing to relational processing. Simultaneous multimodal learning means that all possible cross modal relations for a given item are presented simultaneously for learning. It is likely then, that this simultaneous presentation reduces some of the requirement for item specific processing. For example in unimodal learning the model must combine separate cross modal relations for a given item to be aware of all possible cross modal relations for that item. This means in unimodal learning the model must multimodally combine incoming unimodal information before it can go on to develop the relations between items that are responsible for the emergent category structure amongst its semantic representations. If the model is considered in this way then the benefits of simultaneous multimodal learning would appear to be a reduction in the need for item specific processing allowing faster progress to relational processing. In real world terms, this means an idea of the multimodal learning environment reducing the level of item integration that will be required from processing in the brain, and so allowing item elaboration to begin sooner than would occur in a unimodal learning environment. This would explain why no amount of reorganisation of unimodal learning tasks would approximate multimodal learning, since there would always be the need for the same level of integration before elaboration could begin.

Conclusion

One of the aims of the current investigation was to think about learning during spontaneous recovery and whether the nature of such learning would offer insights into cognitive rehabilitation: To consider possible requirements for clinical practice in terms of maximising learning efficiency after brain damage. If the benefits of multimodal learning are to be translated into clinical practice it seems that it will be necessary to provide tasks that allow simultaneous feedback across several modalities. Despite suggesting possible outcomes for rehabilitation as a result of observations made during the current study these simulations only consider spontaneous recovery not rehabilitation. It is thus possible that in focusing on spontaneous recovery it is only really possible to offer accounts of relearning as a result of re-exposure to previously learnt/known concepts in the external world. Models of the rehabilitation process as distinct from recovery are necessary to explore whether learning strategies in terms of targeted modalities have identical relevance in development, recovery and rehabilitation or whether each learning situation involves processes behaving in a distinct situation-specific manner. Simply put, is rehabilitation a separate case that due to its special circumstances as a learning situation is dependent upon different factors in the learning environment for its efficiency from those factors observed be relevant in development and recovery. Whilst the current study offers interesting insights into possible mechanical accounts for the construction and benefits of multimodal learning, further research is required in order to discover how to construct multimodal learning environments in clinical practice that possess the necessary level of simultaneous multimodality. Computational simulation has allowed an understanding of the ideal conditions in which learning efficiency can be maximised by taking advantage of multimodality. It remains for clinical practice, in conjunction with further theoretical investigation, to adopt the challenge of achieving such beneficial levels of multimodality within therapeutic tasks.

Chapter 3 - Technical report on the replication and rescaling of the Rogers et al. (2004) model of semantic memory, and simulating spontaneous recovery to stable baseline

Introduction

The purpose of this Chapter is to provide a technical report detailing the replication and rescaling of the Rogers et al. (2004) model of semantic memory that was used in Chapter 1, to provide a model which could be used to explore rehabilitation in Chapter 4. Chapter 1 yielded the finding that multimodal learning was more efficient than unimodal learning, both in development and in recovery following damage. This Chapter sets out to replicate the model in the same manner as Chapter 1, except for making potentially non-significant adjustments to the model in order for it to be capable of learning twice the number of items. To begin simulating rehabilitation, it is necessary to have a model with enough learnt items so that after a period of recovery to stable baseline, the model will still have a proportion of items that it consistently gets wrong. The number of these consistently incorrect items needs to be large enough to be split into a rehabilitation training set and a control training set, if simulated rehabilitation is to mimic the clinical processes for which it is trying to provide an account.

Whilst there have been many models of developmental learning, to the extent that certain characteristics of children's learning behaviour (e.g. the vocabulary spurt - Plunkett et al., 1997) are easily observed in any Parallel Distributed Processing (PDP) model, there have been very few simulations of recovery and rehabilitation after neural network damage.

Relearning in PDP models can mirror both the process of cognitive rehabilitation (e.g. Plaut , 1996) and cognitive rehabilitation after recovery (e.g. Welbourne and Lambon Ralph, 2005b). In damaged PDP models, recovery is accomplished by allowing the model to relearn using the same training set as that which resulted in its original (pre-damage) learning. Central to this usage of PDP models is the idea that long-term recovery would employ the same learning mechansims as those used in development. This means that through relearning the model finds a set of connection weights that maximises performance given a balance between available computational resources and the training environment and these weights stabilise into a new 'equilibrium' in the same way weights stabilise in models of developmental learning. This post-damage relearning is analogous to spontaneous recovery where the patient regains cognitive function merely as a result of normal living without specific therapeutic intervention.

This chapter details development and recovery simulations that parallel the simulations of development and recovery in Chapter 1. The purpose of the simulations within this chapter is to replicate the findings from Chapter 1 in a larger version (i.e. learning 72 items instead of 36) of the Rogers et al. (2004) model, and then to use the end point of the recovery simulations as the starting point for rehabilitation simulations in Chapter 4. The simulations in this chapter were essential to verify that doubling the number of items the model could learn did not in any way affect the previously observed performance comparison of the multimodal and unimodal learning environments. The simulations in Chapter 4 will attempt to simulate clinical observations so it is important to simulate this in a model capable of exhibiting a range of recovery behaviour such that different levels of damage severity will result in different numbers of items that the model consistently gets wrong on naming (notes on the importance of therapy items being consistently wrong before therapy on those items commences can be found in Conroy et al., 2006). Thus rehabilitation learning performance can be investigated for a range of damage severity, and with different numbers of therapy items, all in the context of the original comparison of multimodal and unimodal learning environments established in Chapter 1.

Simulation 3.1: Developmental learning

Method

For this simulation the PDP model of semantic knowledge employed in Chapter 1, a replication of a previously established model of semantic knowledge (Rogers et al., 2004), was scaled up so it was capable of learning twice as many items. This model was implemented using the LENS neural network simulator programming environment (Rohde, 2000). Figure 3.1 shows the network architecture of the model:

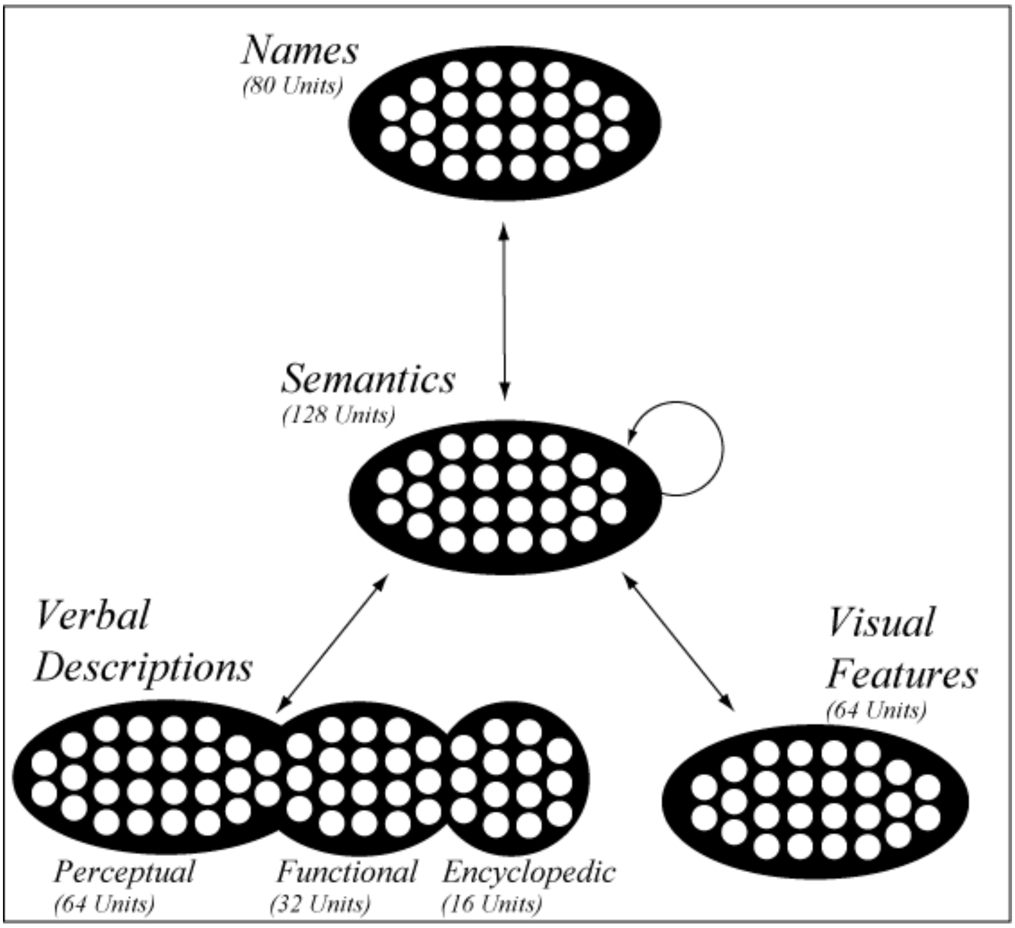

Figure 3.1 Architecture of the model adapted from Rogers et al. (2004)

Firstly it is important to note the changes to the model's architecture. The original model learns 36 unique items (as detailed in chapter 3). In order to scale the model so it was capable of learning twice this number of items, the quantity of localist name units had to be doubled (to 80 from 40). In addition to this, the original use of 64 semantic units was not enough to allow learning of twice as many items so the number of semantic units was also doubled (to 128 from 64), after initial experimentation with a smaller number of units (100) proved unsucessful. All other aspects of the model were the same as the original simulations. The model was a fully recurrent network consisting of four layers of units. Three layers of units labelled names, verbal descriptions and visual features are bidirectionally connected via a single layer of hidden units labelled semantics. The names, verbal descriptions and visual features layers are each capable of both input and output, and receive input directly from the environment. The semantics layer is a hidden layer and does not interact directly with the environment, it only receives input from, or outputs to, the names, verbal descriptions and visual features layers. All units in the semantics layer are also recurrently connected to each other. These recurrent connections assist in the development of attractors (Hinton & Shallice, 1991), stabilising representations within the semantics layer. Representational economy is measured in exactly the same way as described in Chapter 1.

Training the model

Identically to Chapter 1 the model was trained on either one of two distinct training regimes:simultaneously multimodal (identical to that used in Rogers et al., 2004) or unimodal The model was trained with online learning (i.e. using a batch size of one, thus with a weight update after the presentation of each training trial) using recurrent backpropagation through time with a steepest descent algorithm. The learning rate was set to 0.005. A weight decay of 0.0000002 was also used to prevent any weights developing values that were disproportionately high.

To simulate development, the model was trained for 1036800 weight updates, using either the multimodal and unimodal training regime in each simulation. This equates to 4800 presentations in the multimodal condition (multimodal presentation of the entire corpus takes 216 updates, i.e. 3 training trials for each of the 72 items). In contrast, this amount of training equates to 1600 presentations in the unimodal condition (unimodal presentation of the entire corpus takes 648 updates, i.e. 9 training trials for each of the 72 items). During training, the model was tested on its ability to generate the correct target outputs (to within 0.5) for all patterns, in all output modalities. 1036800 weight updates was chosen for the duration of training through initial experimentation, as whenever the model learnt to 100% accuracy (verified through regular testing) it had occurred by this time.

Testing the network and analysing representational economy

Representational economy (calculated as described in Chapter 1), accuracy and error were calculated at regular intervals during the model's training. A 2x20 repeated measures analysis of variance (ANOVA) was conducted on these data to investigate the effect of manipulating the training regime (multimodal or unimodal) at twenty equally spaced time points during training on the dependent variables (accuracy, representational economy and network error). The resulting data from the simulations was also used to calculate the correlation between representational economy, accuracy, and network error as the model learns in order to examine Representational Economy's relationship to Accuracy and Network Error.

Results

The graphs in Figure 3.2 show performance of the model during developmental learning, both on a trial by trial basis and for an equated error comparison. It can be clearly seen that (as found during simulations in Chapter 1, see Chapter 1 - Figure 1.3) multimodal training is more efficient than unimodal training on a trial by trial basis, obtaining greater accuracy and greater convergence (as shown by representational economy) as well as greater error reduction. Again this mirrors the findings of Chapter 1. These differences between multimodal and unimodal training are also statistically significant as the results of the 2x20 analysis of variance confirm (see Table 3.1). In addition Figure 3.2 shows how multimodal training achieves greater acceleration in the learning process thus showing its greater efficiency at an earlier point in learning. This also mirrors the findings of Chapter 1. The equal error comparison between multimodal and unimodal training shows no substantial overall difference and indeed the small differences seen in the graphs in Figure 3.2 are not statistically significant (as indicated in the analysis results in Table 3.1). However it is worth noting that during the vocabulary spurt period, and only here, the equal error comparison performs worse than unimodal training. Table 3.1 shows multimodal training outperforming unimodal training with statistically significant differences in accuracy, representational economy and network error. When comparing equated error multimodal and unimodal there is an effect of time and significant interaction between training and time on accuracy, representational economy and network error. Representational Economy was strongly correlated with learning Accuracy for all conditions: Multimodal $r(20)=.964$, $p<.001$; Equated Error Multimodal $r(20)=.926$, $p<.001$; Unimodal $r(20)=.926, p<.001$. Representational Economy was strongly correlated with Network Error for all conditions: Multimodal $r(20)=-.895, p<.001$;Equated Error Multimodal $r(20)=-.816, p<.001$;Unimodal $r(20)=-.795, p<.001$.

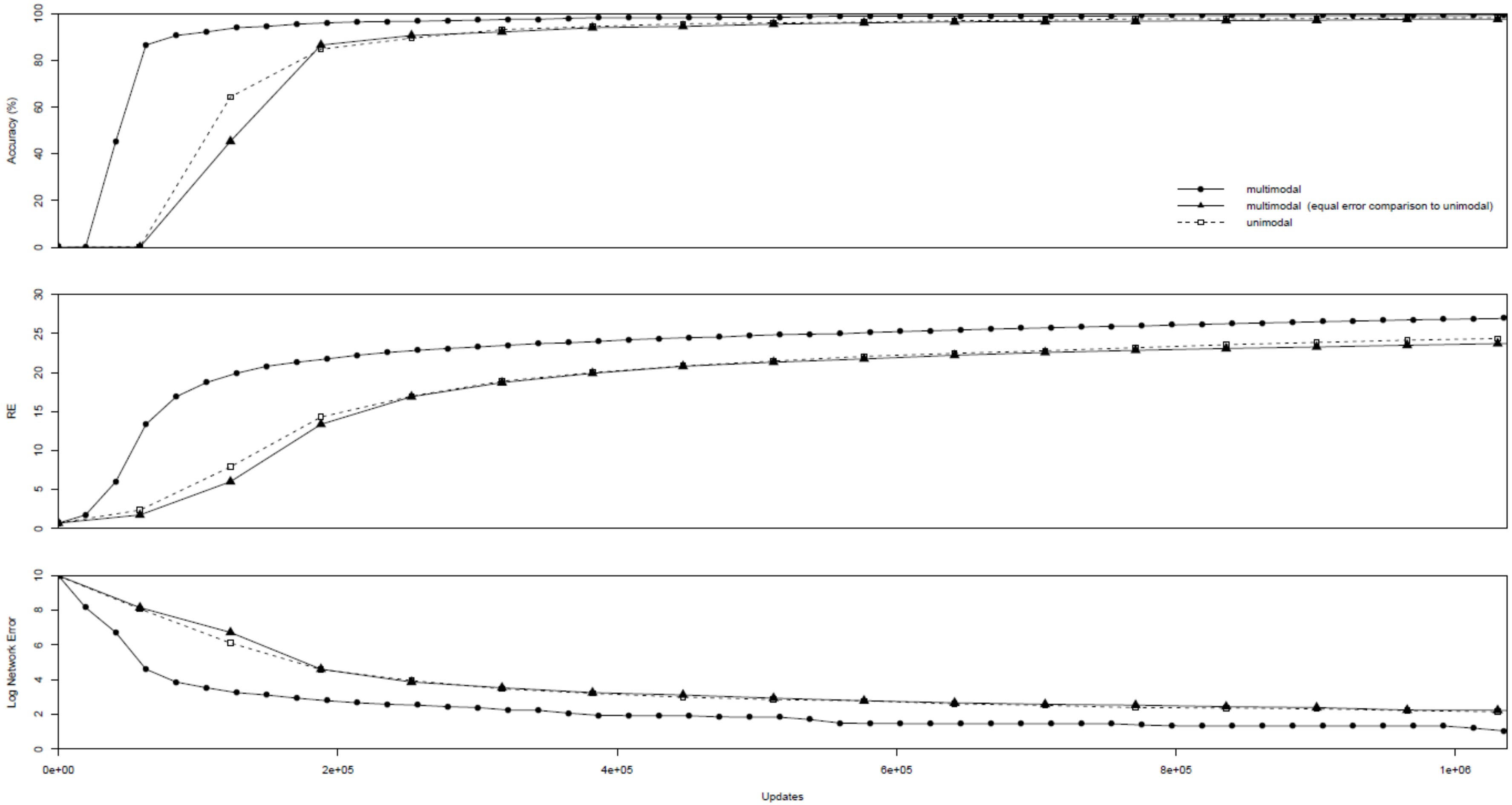

Figure 3.2 Mean variation in training performance during developmental learning

Table 3.1 2x20 ANOVAs indicating significant performance differences between training regimes during developmental learning

Training Regime Comparison	Dependent Variable Measure	Independent Variables	df	F	p	Effect size (η^2)
Multimodal Vs Unimodal	Accuracy	Training	1,9	114.536	<0.001*	0.927
		Time	19,171	479.967	<0.001*	0.982
		Training*Time	19,171	213.605	<0.001*	0.960
	Representational Economy	Training	1,9	192.012	<0.001*	0.955
		Time	19,171	1039.641	<0.001*	0.991
		Training*Time	19,171	117.331	<0.001*	0.929
	Network Error	Training	1,9	350.789	<0.001*	0.975
		Time	19,171	994.417	<0.001*	0.991
		Training*Time	19,171	529.937	<0.001*	0.983
Equated Error Multimodal Vs Unimodal	Accuracy	Training	1,9	0.853	0.380	0.087
		Time	19,171	2040.479	<0.001*	0.996
		Training*Time	19,171	1.753	0.032*	0.163
	Representational Economy	Training	1,9	2.171	0.175	0.194
		Time	19,171	1125.866	<0.001*	0.992
		Training*Time	19,171	1.703	0.040*	0.159
	Network Error	Training	1,9	4.671	0.059	0.342
		Time	19,171	2336.111	<0.001*	0.996
		Training*Time	19,171	5.251	<0.001*	0.368

Simulation 3.2: Spontaneous recovery to stable baseline

Method

The model was trained in the manner described in Simulation 3.1. However, this time only the multimodal training regime was used to train the model as per the original simulation (see Rogers et al, 2004). In order to investigate the model's re-learning behaviour (spontaneous recovery) after damage, the 10 multimodally trained networks from the initial training were lesioned by removing a proportion of all incoming and outgoing connections across all the units in all layers (i.e., names, verbal descriptions, visual features and semantics). Based upon the findings from Chapter 1 lesioning was again required but this time four levels of damage were needed as the simulations in Chapters 1 & 2 suggest that four levels of damage would capture the more detailed picture of relearning that would be required for the rehabilitation simulations in the following Chapter. The same four levels of damage chosen were: Minimal - where the model relearns to 90% accuracy or higher. Mild - where the model relearns to around 75% accuracy, Moderate - where the model relearns to around 55% accuracy, and Severe where the model relearns to around 30% accuracy or less. Again it should be pointed out this is not identical to the real situation of localised injury since this lesioning was uniform across the model to generate the levels of relearning this study wanted to explore. The trained networks were subjected to four separate lesions of varying degrees of severity: minimal (removing 88 % of connections); mild (removing 90 % of connections); moderate (removing 92% of connections); and severe (removing 93% of connections). These degrees of severity were intended to cover a range of behavioural severity outcomes. After lesioning, each network underwent 2073600 updates of retraining with both the multimodal and unimodal training regimes. This corresponds to 9600 multimodal, and 3200 unimodal, presentations of the entire training corpus. 2073600 updates was chosen through experimentation in order to allow the network to recovery to a stable baseline level of accuracy that would provide a starting point for simulating rehabilitation. Rehabilitation learning focuses upon consistently failed items to start from a zero naming baseline for each participant which means effects of different rehabilitation regimes can be contrasted (Conroy et al., 2006). Since these simulations will form the structural basis of the model for use in a comparison of multimodal and unimodal rehabilitation in Chapter 4, it was essential to have achieved stable baseline performances in recovery. This repeated-measures methodology where the same initially trained network was subjected to each lesion and retraining condition also allows us to remove any variance associated with individual pre or postmorbid differences. This is important since these simulations are an attempt to capture the relearning behaviour of brain-damaged patients, however these patients are not a homogenous population, and studying the effects of treatment should ideally be based upon “within-patient comparisons” (Howard & Hatfield, 1987, p.119).

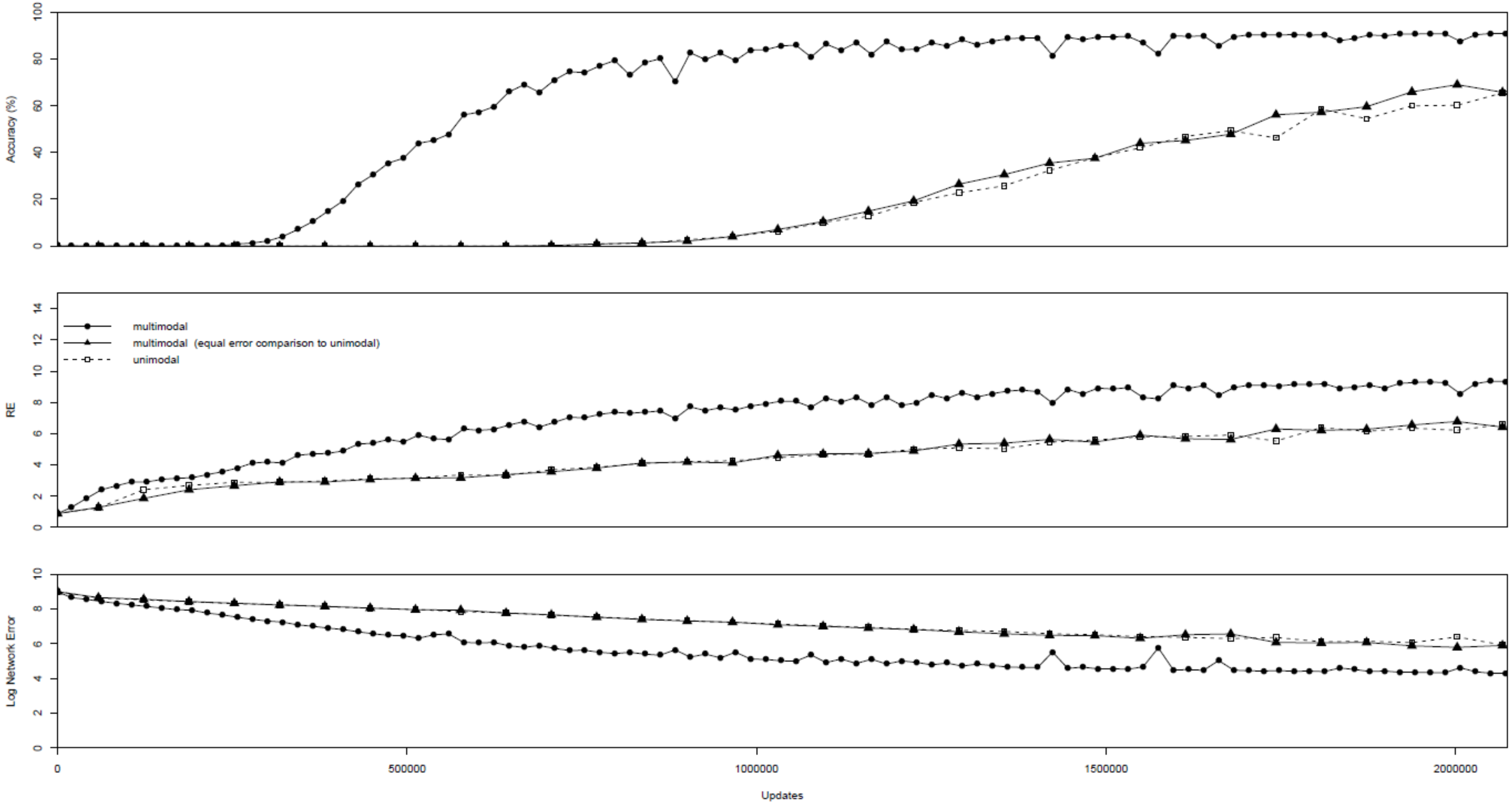

Figure 3.3 Mean variation in training performance during spontaneous recovery to stable baseline after a minimal lesion

Table 3.2 2x20 ANOVAs indicating significant performance differences between training regimes after a minimal lesion

Training Regime Comparison	Dependent Variable Measure	Independent Variables	df	F	p	Effect size (η^2)
Multimodal Vs Unimodal	Accuracy	Training	1,9	2953.537	<0.001*	0.997
		Time	19,171	472.046	<0.001*	0.981
		Training*Time	19,171	103.449	<0.001*	0.920
	Representational Economy	Training	1,9	1033.339	<0.001*	0.991
		Time	19,171	155.000	<0.001*	0.945
		Training*Time	19,171	11.196	<0.001*	0.554
	Network Error	Training	1,9	12493.587	<0.001*	0.999
		Time	19,171	2720.140	<0.001*	0.997
		Training*Time	19,171	235.713	<0.001*	0.963
Equated Error Multimodal Vs Unimodal	Accuracy	Training	1,9	2.345	0.160	0.207
		Time	19,171	487.855	<0.001*	0.982
		Training*Time	19,171	0.366	0.993	0.039
	Representational Economy	Training	1,9	1.831	0.209	0.169
		Time	19,171	347.650	<0.001*	0.975
		Training*Time	19,171	0.734	0.780	0.075
	Network Error	Training	1,9	0.957	0.353	0.096
		Time	19,171	2855.395	<0.001*	0.997
		Training*Time	19,171	0.951	0.522	0.096

Results

Minimal Lesioning

The graphs in Figure 3.3 show the model's recovery performance to stable baseline after a minimal lesion, both on a trial by trial basis and for an equated error comparison. As observed during development it can be clearly seen that multimodal training is more efficient than unimodal training on a trial by trial basis, obtaining greater accuracy, and greater convergence (as shown by representational economy) as well as greater error reduction. These differences between multimodal and unimodal training are also statistically significant as confirmed by the results of the 2x20 analysis of variance (see

Table 3.2). Figure 3.3 again shows how the multimodal training achieved much greater acceleration in the learning process reflecting greater temporal efficiency. The equal error comparison between multimodal and unimodal training showed no overall difference, and is not statistically significant (as shown in the analysis results in Table 3.2). The superior performance of multimodal training over unimodal is supported by the results in Table 3.2 which show a statistically significant difference between the two regimes in accuracy, representational economy and network error. The comparison of equated-error multimodal and unimodal also shows a significant effect of time on accuracy, representational economy and network error. Representational Economy was strongly correlated with learning Accuracy for all conditions: Multimodal $r(20)=.977$, $p<.001$; Equated Error Multimodal $r(20)=.908$, $p<.001$; Unimodal $r(20)=.907$, $p<.001$. Representational Economy was strongly correlated with Network Error for all conditions: Multimodal $r(20)=-.879$, $p<.001$; Equated Error Multimodal $r(20)=-.950$, $p<.001$; Unimodal $r(20)=-.947$, $p<.001$.

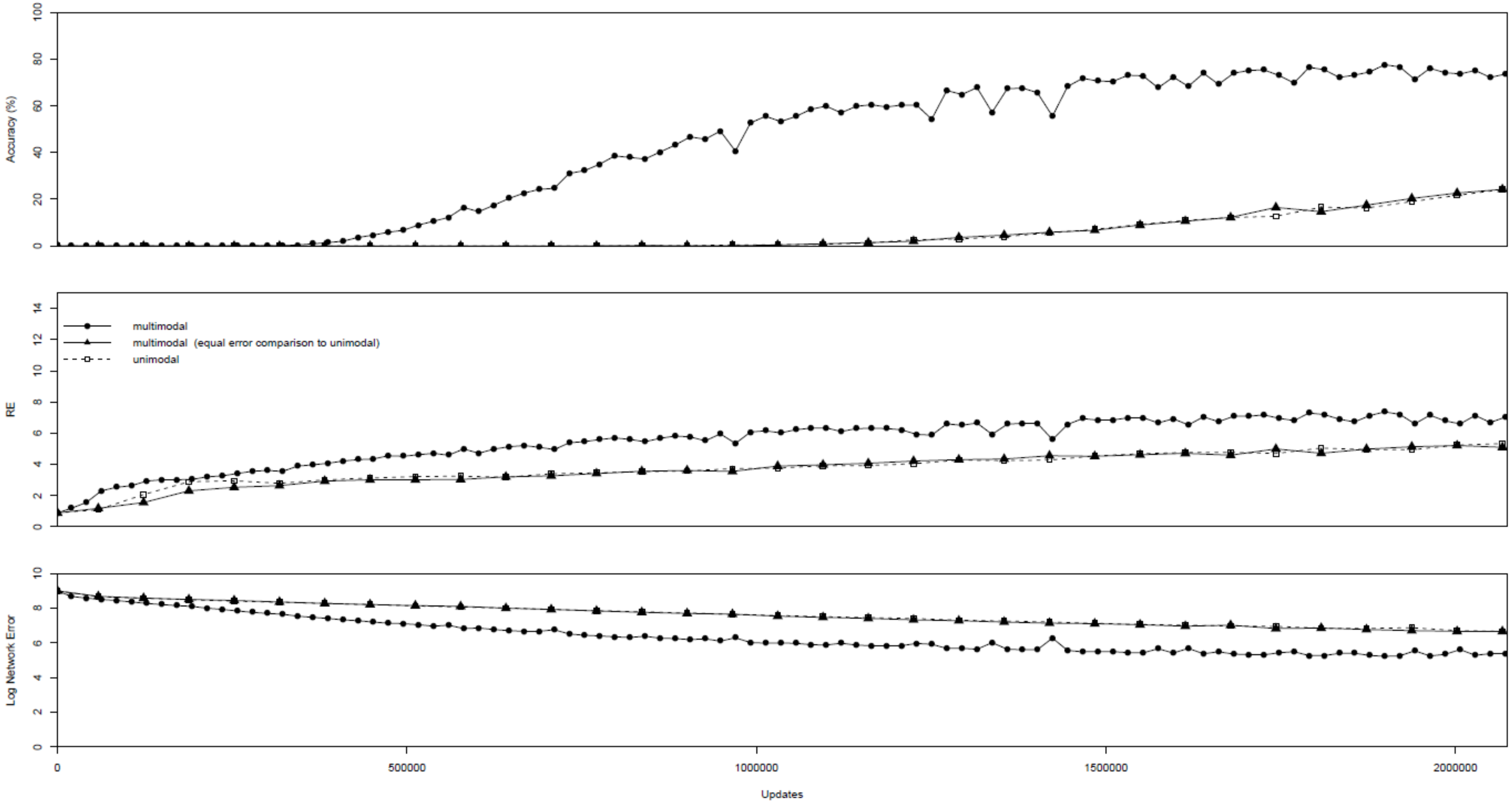

Figure 3.4 Mean variation in training performance during spontaneous recovery to stable baseline after a mild lesion

Table 3.3 2x20 ANOVAs indicating significant performance differences between training regimes after a mild lesion

Training Regime Comparison	Dependent Variable Measure	Independent Variables	df	F	p	Effect size (η^2)
Multimodal Vs Unimodal	Accuracy	Training	1,9	938.547	<0.001*	0.991
		Time	19,171	150.985	<0.001*	0.944
		Training*Time	19,171	69.382	<0.001*	0.885
	Representational Economy	Training	1,9	2065.153	<0.001*	0.996
		Time	19,171	110.923	<0.001*	0.925
		Training*Time	19,171	7.577	<0.001*	0.457
	Network Error	Training	1,9	8042.993	<0.001*	0.999
		Time	19,171	777.629	<0.001*	0.989
		Training*Time	19,171	51.758	<0.001*	0.852
Equated Error Multimodal Vs Unimodal	Accuracy	Training	1,9	0.809	0.392	0.082
		Time	19,171	150.190	<0.001*	0.943
		Training*Time	19,171	1.843	0.022*	0.170
	Representational Economy	Training	1,9	7.969	0.020*	0.470
		Time	19,171	256.162	<0.001*	0.966
		Training*Time	19,171	2.412	<0.001*	0.211
	Network Error	Training	1,9	0.842	0.383	0.086
		Time	19,171	2851.404	<0.001*	0.997
		Training*Time	19,171	1.125	0.330	0.111

Mild Lesioning

The graphs in Figure 3.4 show the model's recovery performance to stable baseline after a mild lesion, both on a trial by trial basis and for an equated error comparison. Multimodal training was again observed to be more efficient than unimodal training on a trial by trial basis, obtaining greater accuracy, and greater convergence (as shown by representational economy) as well as greater error reduction. These differences are statistically significant as the results of the 2x20 analysis of variance, given in Table 3.3, confirmed. Figure 3.4 again shows how the multimodal training achieves much greater acceleration in the learning process reflecting greater temporal efficiency. The equal error comparison between multimodal and unimodal training shows no overall difference in accuracy or error, and is not statistically significant (as show in the analysis results in Table 3.3) for these measures. However a statistically significant difference was found in representational economy (see Table 3.3) for the equal error comparison. Table 3.3 gives the results from statistically comparing multimodal and unimodal training showing a multimodal advantage with the significant difference between the two in accuracy, representational economy and network error. The comparison of the Equated Error Multimodal and unimodal training showed an effect of time on accuracy, representational economy and network error and a significant interaction between training and time for accuracy and representational economy. Representational Economy was strongly correlated with learning Accuracy for all conditions: Multimodal $r(20)=.972$, $p<.001$; Equated Error Multimodal $r(20)=.769$, $p<.001$; Unimodal $r(20)=.791$, $p<.001$. Representational Economy was strongly correlated with Network Error for all conditions: Multimodal $r(20)=-.900$, $p<.001$; Equated Error Multimodal $r(20)=-.977$, $p<.001$; Unimodal $r(20)=-.958$, $p<.001$.

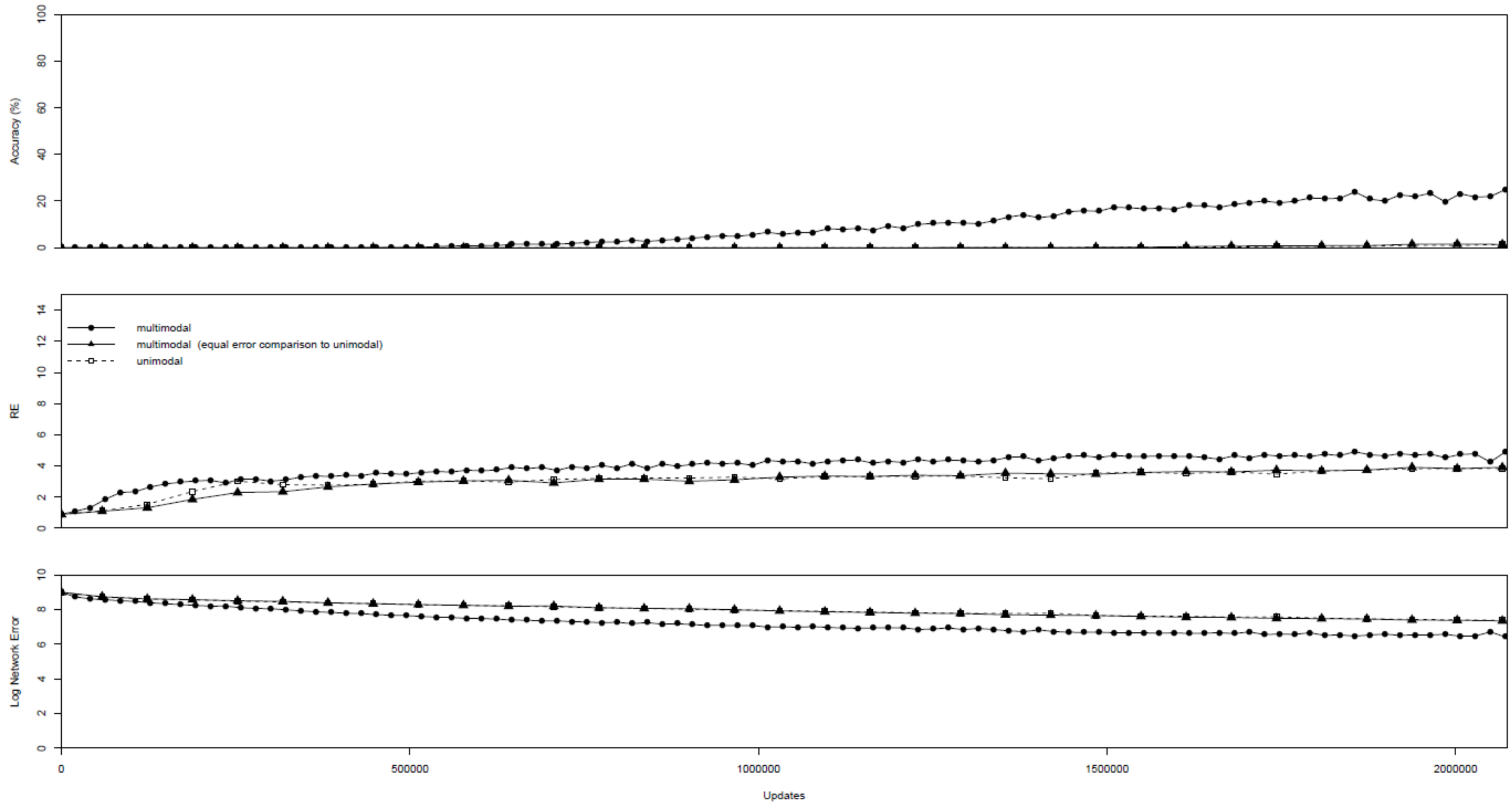

Figure 3.5 Mean variation in training performance during spontaneous recovery to stable baseline after a moderate lesion

Table 3.4 2x20 ANOVAs indicating significant performance differences between training regimes after a moderate lesion

Training Regime Comparison	Dependent Variable Measure	Independent Variables	df	F	p	Effect size (η^2)
Multimodal Vs Unimodal	Accuracy	Training	1,9	229.071	<0.001*	0.962
		Time	19,171	117.160	<0.001*	0.929
		Training*Time	19,171	101.985	<0.001*	0.919
	Representational Economy	Training	1,9	364.888	<0.001*	0.976
		Time	19,171	169.446	<0.001*	0.950
		Training*Time	19,171	6.020	<0.001*	0.401
	Network Error	Training	1,9	18439.087	<0.001*	1.000
		Time	19,171	5666.687	<0.001*	0.998
		Training*Time	19,171	127.694	<0.001*	0.934
Equated Error Multimodal Vs Unimodal	Accuracy	Training	1,9	0.433	0.527	0.046
		Time	19,171	28.595	<0.001*	0.761
		Training*Time	19,171	0.795	0.710	0.081
	Representational Economy	Training	1,9	11.557	0.008*	0.562
		Time	19,171	181.216	<0.001*	0.953
		Training*Time	19,171	3.931	<0.001*	0.304
	Network Error	Training	1,9	8.509	0.017*	0.486
		Time	19,171	5152.525	<0.001*	0.998
		Training*Time	19,171	1.970	0.012*	0.180

Moderate Lesioning

Figure 3.5 show the model's recovery performance to stable baseline after a moderate lesion, both on a trial by trial basis and for an equated error comparison. Multimodal training is again observed to be more efficient than unimodal training on a trial by trial basis, obtaining greater accuracy, and greater convergence (as shown by representational economy) as well as greater error reduction. These differences are statistically significant as the results of the 2x20 analysis of variance, given in Table 3.4, confirm. Figure 3.5 again shows how the multimodal training achieves much greater acceleration in the learning process reflecting greater temporal efficiency. However with the greater lesion severity the unimodal accuracy performance is almost zero. The equal error comparison between multimodal and unimodal training shows no overall difference in accuracy, and is not statistically significant (as show in the analysis results in Table 3.4) for these measures. However, a statistically significant difference was found in error and representational economy (see Table 3.4) for the equal error comparison. Table 3.4 again indicates a multimodal advantage with a statistically significant difference between multimodal and unimodal training in accuracy, representational economy and network error. There is also a statistically significant difference between equated error multimodal and unimodal in representational economy and network error but not In accuracy where there is only a significant effect of time. Representational Economy was strongly correlated with learning Accuracy for all conditions: Multimodal $r(20)=.848$, $p<.001$; Equated Error Multimodal $r(20)=.600$, $p=.005$; Unimodal $r(20)=.518$, $p=.019$. Representational Economy was strongly correlated with Network Error for all conditions: Multimodal $r(20)=-.964$, $p<.001$; Equated Error Multimodal $r(20)=-.961$, $p<.001$; Unimodal $r(20)=-.926$, $p<.001$.

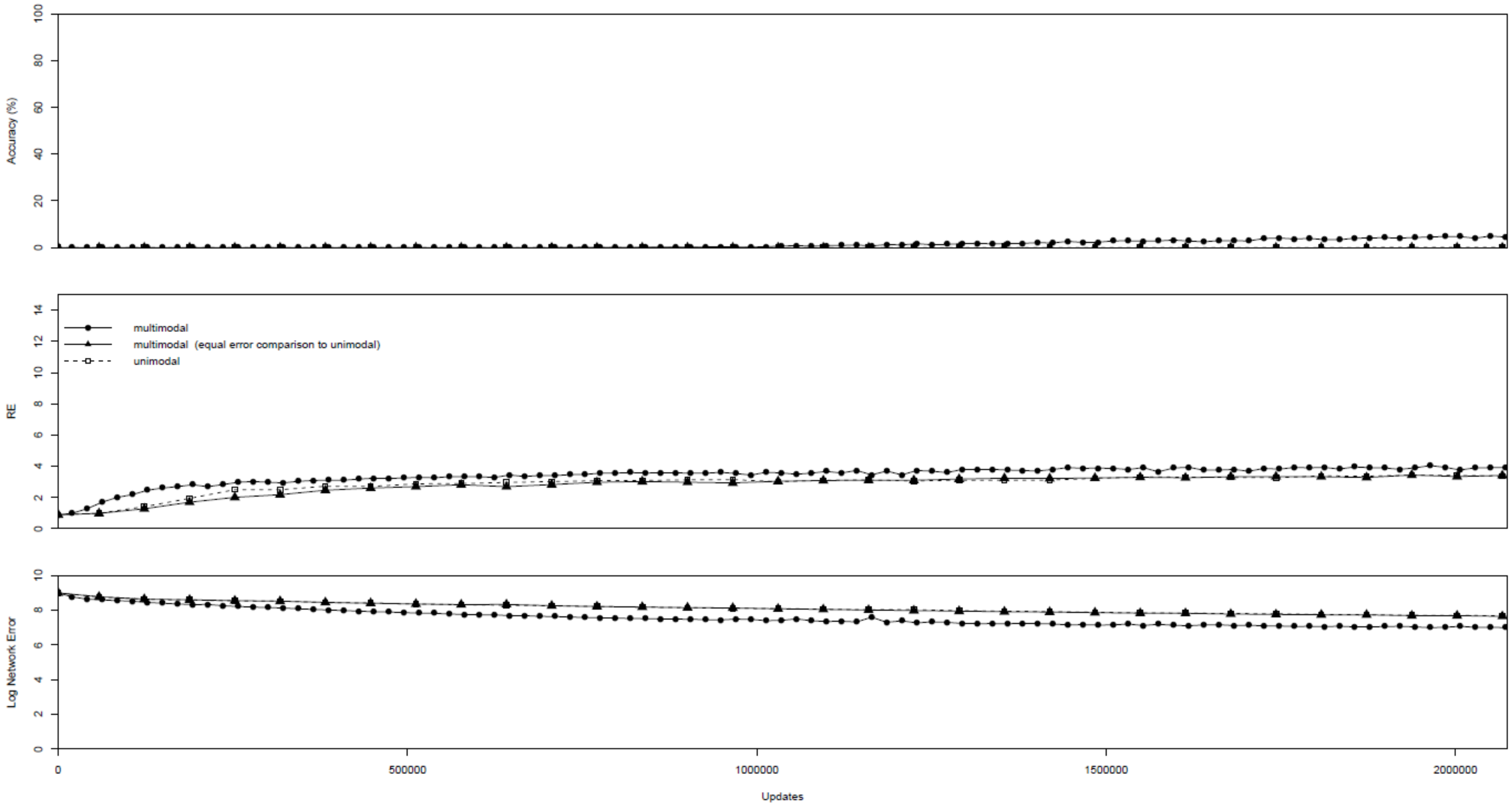

Figure 3.6 Mean variation in training performance during spontaneous recovery to stable baseline after a severe lesion

Table 3.5 2x20 ANOVAs indicating significant performance differences between training regimes after a severe lesion

Training Regime Comparison	Dependent Variable Measure	Independent Variables	df	F	p	Effect size (η^2)
Multimodal Vs Unimodal	Accuracy	Training	1,9	53.992	<0.001*	0.857
		Time	19,171	37.520	<0.001*	0.807
		Training*Time	19,171	35.552	<0.001*	0.798
	Representational Economy	Training	1,9	423.361	<0.001*	0.979
		Time	19,171	184.113	<0.001*	0.953
		Training*Time	19,171	6.619	<0.001*	0.424
	Network Error	Training	1,9	9391.928	<0.001*	0.999
		Time	19,171	4227.733	<0.001*	0.998
		Training*Time	19,171	61.354	<0.001*	0.872
Equated Error Multimodal Vs Unimodal	Accuracy	Training	1,9	1.976	0.193	0.180
		Time	19,171	1.784	0.028*	0.165
		Training*Time	19,171	1.784	0.028*	0.165
	Representational Economy	Training	1,9	26.567	0.001*	0.747
		Time	19,171	158.702	<0.001*	0.946
		Training*Time	19,171	3.416	<0.001*	0.275
	Network Error	Training	1,9	0.007	0.935	0.001
		Time	19,171	2352.847	<0.001*	0.996
		Training*Time	19,171	1.169	0.289	0.115

Severe Lesioning

Figure 3.6 show the model's recovery performance to stable baseline after a severe lesion, both on a trial by trial basis and for an equated error comparison. Recovery performance is very low. Multimodal training is again observed to be more efficient than unimodal training on a trial by trial basis, obtaining greater accuracy, and greater convergence (as shown by representational economy) as well as greater error reduction. Though the difference between multimodal and unimodal performance is much less for higher severity lesion. These differences are statistically significant as the results of the 2x20 analysis of variance, given in Table 3.5, confirm. The equal error comparison between multimodal and unimodal training shows no overall difference in accuracy or error, and is not statistically significant (as show in the analysis results in Table 3.5) for these measures. However a statistically significant difference was found in representational economy (see Table 3.5) for the equal error comparison. The superior multimodal performance is reflected in the statistics in Table 3.5 where there is a statistically significant difference between multimodal and unimodal training in accuracy, representational economy and network error. In addition there is a significant difference between equated error multimodal and unimodal training in representational economy as well as an effect of time on network error and accuracy and a significant interaction between training and time for accuracy. Representational Economy was correlated with learning Accuracy for all Multimodal $r(20)=.739$, $p<.001$; but not Equated Error Multimodal (zero values so no calculation) or Unimodal $r(20)=.265, p=.258$. Representational Economy was strongly correlated with Network Error for all conditions: Multimodal $r(20)=-.980, p<.001$;Equated Error Multimodal $r(20)=-.950, p<.001$;Unimodal $r(20)=-.913, p<.001$.

Discussion

Simulation 3.1 showed multimodal training to be more efficient than unimodal training for developmental learning in an identical manner to previous observations in Chapter 1. This scaled up replication of the Rogers et al. (2004) model of semantic memory behaves almost identically to the original model. It should be noted that the equated error comparison shows slightly poorer performance than unimodal during the vocabulary spurt period but is otherwise virtually identical to unimodal performance suggesting as previously observed in Chapter 1, the higher error signal available during multimodal training is responsible for its efficiency advantages in a model where larger numbers of items need to be learned as well as for smaller numbers of items.

Simulation 3.2 showed multimodal training to be more efficient than unimodal training for recovery to stable baseline, exhibited similar recovery behaviour to the performance of the smaller model in Chapter 1. The equated error comparison showed no difference over the course of recovery learning, however for certain lesions there was a difference in representational economy. It is very likely that this difference in convergence arises from the model being required to learn a larger number of items, and having semantic representations consisting of a larger number of units. Nevertheless, the overall recovery behaviour of the larger model seen here mirrors the findings of Chapter 1.

Conclusion

The intention of the simulations, reported in this Chapter, was to provide a replication of the Rogers et al. (2004) model used in Chapter 1, with appropriate adaptations to allow simulating rehabilitation after a period of spontaneous recovery to stable baseline. The goal of simulating rehabilitation was to consider rehabilitation learning behaviour under multimodal or unimodal training for different levels of lesion severity, and consequently for different numbers of items consistently wrong after recovery. The simulations reported here provide a suitable starting point for simulating rehabilitation given that they preserved the features and findings from Chapter 1 regarding the advantage of multimodal training.

Chapter 4 - Using computational models to investigate cognitive rehabilitation after brain damage

Abstract

Despite the large number of connectionist models that have attempted to account for developmental learning, very few models (notably Plaut,1996 and Welbourne & Lambon Ralph, 2005b) have looked at recovery and rehabilitative relearning after brain damage. This study firstly seeks to replicate and extend previous approaches to modelling rehabilitation, but in a fully recurrent convergence zone model, which has not previously been attempted. Secondly, the study seeks to extend previous studies by incorporating both background spontaneous recovery and manipulations of therapy set size during relearning, after different levels of lesion damage. An enlarged version of the Rogers et al (2004) model of semantic memory was trained,damaged with varying severity (four levels), and subsequently allowed to recover to a stable baseline performance accuracy. At stable baseline, the model was tested on object naming and the error scores for each item were rank ordered. For each simulated rehabilitation, a therapy set and a control set (counterbalanced across simulations), each containing an identical number of items was created. These items were selected as those with the highest error scores from the list of rank-ordered items. Three simulations were conducted: In the first simulation the model was trained on either a multimodal or unimodal rehabilitation regime, where intervention was simulated by raising the learning rate for rehabilitation items (Welbourne and Lambon Ralph, 2005b) but for the target set only. In the second and third simulations, as well as training on the target items the model was also re-exposed to ‘continued’ spontaneous

recovery at the original learning rate for the remainder of items. In the second simulation, the effect of therapy set size (three levels) was explored at the severe level of lesioning. In the third simulation, the effect of therapy set size (two levels) was crossed with two degrees of lesioning to explore possible interactions between these variables.

Simulation 1, a replication of previous rehabilitation learning methodology (Plaut, 1996) found that multimodal rehabilitation learning was more efficient than unimodal learning in all conditions. Simulation 2, an investigation of the effect of the manipulation of therapy set size on rehabilitation learning showed greater performance at the end of therapy for smaller therapy set sizes. In addition multimodal learning outperformed unimodal learning in all conditions. Simulation 3, considered the role of the level of damage in conjunction with manipulation of therapy set size and targeted modalities during rehabilitation learning: an interaction between therapy set size and the level of damage meant that a greater performance at the end point of therapy was seen for smaller set sizes with lower levels of damage. Again multimodal learning proved more effective than unimodal learning in all conditions.

These results suggest that some consideration of the level of severity of damage observed in patients would be necessary in the context of selecting the number of items to be used in therapy. In addition, based on the model performance, multimodal rehabilitation strategies would be predicted to be more efficient than unimodal strategies (currently the predominant practice in aphasia therapy).

Introduction

Across the lifespan it is perhaps possible to classify learning into two distinct situations. First, there is developmental learning. Such learning occurs during childhood and underpins knowledge and knowledge acquisition, for the remainder of the lifespan. Secondly, there is the type of learning that occurs after brain damage. This type of learning will necessarily only be undergone by a relatively small proportion of the broader population, nevertheless it marks a distinct situation in which learning consists of recovering or rehabilitating pre-existing knowledge. Following the development of computational parallel distributed processing (PDP) models in the early 1980s (Rumelhardt & McClelland, 1986), it was possible to simulate learning using computers. In the 20 or so years following the development of PDP modelling, most modelling activity has focused on situations from developmental learning, that is to say using models to account for learning processes observed during childhood (e.g., vocabulary spurt, learning the past tense, etc). However, beginning with the triangle model (Seidenberg and McClelland, 1989) a body of research began to develop that used models to understand larger processes such as reading. One consequence of this was considering developmental processes and disorders, for example the range of issues described in Philip Quinlan's (2003) book "Connectionist models of development: Developmental process in real and artificial neural networks". Similarly PDP models have been applied to investigations of neuropsychological situations such as the acquired dyslexias (see Welbourne and Lambon Ralph, 2005b) that can follow brain damage. In other words, PDP models could be used to investigate the breakdown of learning that occurs following brain damage, and the recovery and rehabilitation learning that is possible in such situations. Since PDP models actually learn, they are able to consider developmental learning, spontaneous recovery after brain damage, and rehabilitation all within the same model. However, despite this capacity, very few researchers have pursued investigations into recovery and rehabilitation learning using PDP models. Despite an early attempt by Hinton and Sejnowski (1986), the only substantial considerations of recovery and rehabilitation learning are those by Plaut (1996), who focussed on recovery learning as a first step towards full simulation of rehabilitation, and Welbourne and Lambon Ralph (2005b) who developed the most realistic rehabilitation simulation to date. The history of Aphasia therapy has seen the development of models that assist in understanding the mechanisms behind language learning in development, as well as the type of rehabilitation learning that occurs during therapy, for language impairments following brain damage. Early models are “boxes and arrows” (e.g. Ellis & Young, 1988 or Kay et al., 1992). These models derive from the idea that opposing abilities on two different language tasks implies that a minimum of two distinct mechanisms are involved in those tasks (Wilshire, 2008). Over the last three decades these early theoretical models have given way to computational models that can simulate language learning and the mechanisms that support it. The earliest truly successful computational model that has influenced aphasia understanding is Seidenberg and McClelland's (1989) triangle model (SM89) which shows how a PDP model can offer an account of normal cognitive function and suggest how impairments can arise from damage by considering dyslexia within the model. Plaut, McClelland, Seidenberg, and Patterson's (1996) model of single word reading (PMSP96) made considerable improvements to SM89 by designing new representations such that the model’s non-word reading improved. Various early models early PDP models are often similar in structure to Seidenberg and McClelland's (1989) PDP model. Such models include (but are not limited to): Hinton and Shallice's (1991) model of acquired dyslexia;Plaut and Shallice's (1993) models exploring the nature of Deep Dyslexia; Mayall and Humphreys (1996) model of Alexia; Plaut, McClelland, Seidenberg, and Patterson's (1996) model of Word Reading; Plaut's (1997) models of word reading and lexical decision; Cree, McRae and McNorgan's

(1999) model simulating semantic priming; Harm and Seidenberg's (1999) model of reading acquisition and dyslexia; Lambon Ralph, McClelland, Patterson, Galton and Hodges' (2001) model of object naming and semantic impairment; Gotts and Plaut's (2002) model of Semantic Impairment; Zevin and Seidenberg's (2002) model of age of acquisition effects in word reading. Similarly Farah and McClelland's (1991) model of semantic memory impairment; Devlin, Gonnerman, Andersen, and Seidenberg's (1998) model of category specific semantic impairment; Joanisse and Seidenberg's (1999) model of impairments in verb morphology; Lambon Ralph and Howard's (2000) model of anomia and impaired verbal comprehension; McLeod, Shallice and Plaut's (2000) model of attractor dynamics in word recognition; Plaut's (2002) model of optic aphasia; Harm and Seidenberg's (2004) model looking at the division of labour in word reading; Rogers, Lambon Ralph, Garrard, Bozeat, McClelland, Hodges, and Patterson's (2004) model of semantic memory; Dilkina, McClelland, and Plaut's (2008) model of semantic and lexical impairment.

Clearly various type of computational model exist that inform our understanding of aphasic impairments, however it becomes clear that it is PDP models that can provide the most ecologically valid simulations since as Welbourne and Lambon Ralph (2005b) point out, PDP models actually learn. Abel et al. (2007) used Dell et al.'s (1997) model of word retrieval simulating naming difficulty in aphasia: the weight-decay model (WD model; Dell et al., 1997) and the semantic-phonological model (SP model; Foygel & Dell, 2000) to predict therapy performance in patients. However these studies used models that do not learn. The current study uses PDP models that learn to predict therapy outcome since it is an area that has not previously been explored but seems the most logical next step in understanding aphasia through modeling and simulation. Indeed the only step so far towards using PDP models that learn to inform aphasia therapy comes from Kiran and Thompson (2003) using Plaut's (1996) conclusions from simulating relearning in PDP models. Plaut noticed that during retraining atypical learning items resulted in greater generalization. Kiran and Thompson (2003) employed this discovery in creating a semantically based naming treatment which prioritised semantic features of typical items over those of atypical items in four participants who had fluent aphasia, All participants improved, three of whom showed some evidence of generalization however Plaut's (1996) model is technically more one of recovery rather than rehabilitation since he re-exposes the network to further training after damage but does nothing to technically distinguish that further training from the initial training. Welbourne and Lambon Ralph (2005b) used selected items (those which the model consistently fails on), and an increased learning rate in their model to simulate rehabilitation as intense exposure to a subset of the items the model had learnt before it was damaged. This situation is much closer to the ecology of the clinical setting. Whilst it can be seen that several researchers have explored using computational models of language that do not learn to design aphasia therapy (e.g Abel, Willmes & Huber, 2007) or the results of the models of others (i.e. Kiran & Thompson, 2003), the current study explores using PDP models of learning that go through development, spontaneous recovery and rehabilitation in order to simulate anomia therapy with greater ecological validity. This is accomplished by attempting to treat the model in exactly the same manner as a patient in terms of initial assessment, level of damage (but not type of damage since the model used in this study is not one of neuroanatomy), establishing a baseline performance before therapy, incorporating background spontaneous recovery and selecting therapy items from those items the patient gets wrong on naming. In addition this study simulates the effects of lesioning in a PDP model as oppose to trying to approximate the neuroanatomy of real lesions. In the version of the Rogers et al. (2004)

model used in the current study lesioning is uniform across the model rather than a closer simulation to the real situation where lesions would be localised. The reason for this is that the Roger et al. (2004) model simulates learning mechanisms and the effects of damages but in order to achieve different levels of damage such that performance after a period of recovery or rehabilitation will be at a range of points between full relearning and almost no relearning lesioning requires removal of random connections across the whole model. Experimentation during the model's development showed that localized lesions (e.g. just the incoming connection to semantics) did not produce a range of relearning performance instead giving an 'all or nothing situation' where the model either relearnt to 100% accuracy or failed to learn anything. In this study the model's training attempts ecological validity in terms of using online learning, in other words presenting each item to be learnt separately. Similarly the model's knowledge is tested for all possible cross-domain mappings on each individual item in the same manner as a patient would be presented with items individually for naming or any other task (e.g. those from the PALPA (Kay et al. 1992)). The current study chose to try an explore a particular therapy issue as a way of demonstrating how PDP models can be used to predict therapy outcome by comparison to pre-existing clinical findings. Therapy set size, the number of items presented to a patient for learning during a therapy session, was chosen as a suitable therapy issue to be explored through modelling since there were already clinical findings and a meta-analysis (Snell, Sage and Lambon Ralph, 2010) that showed what would be expected to happen in patients when therapy set size was varied. Thus the model could first simulate therapy and predict outcome performance before comparing results with clinical data to see how close the model comes to clinical findings so far. The current study uses a replication of a fully recurrent model of semantic memory (Rogers et al., 2004), which implements semantics as a convergence zone, as a starting point to consider recovery and rehabilitation learning.

This particular model was chosen because it proposed an account of semantics in terms of the emergence of amodal intermediate representational structures that facilitate the storage and transformation of information arising in different modalities. Furthermore, this representational structure was shown to be an emergent property of learning, such that the representations possessed an emergent category structure even though category was not coded per se in the training patterns. This model can also be used to account for the behaviour of brain damaged patients in rehabilitation tasks such as picture naming, because it provides an ideal basis for considering recovery and rehabilitation learning.

Aims

The current study set out to explore two factors with regard to rehabilitation. The first aim was to simulate previous observations regarding recovery and rehabilitation and to consider such findings with regard to the role of the learning environment and, specifically, the number of sensory modalities targeted during relearning. Most current therapy strategies involve unimodal learning tasks (Best and Nickels, 2000), yet many accounts of developmental learning prioritise the role of multimodality (e.g., discussions of temporally synchronous naming such as Gogate, Bolzani & Betancourt, 2006). Indeed, the model used for this study also adopts the view that developmental learning is multimodal,.

The second aim of the simulations was to explore the degree to which the number of items used in rehabilitation affects the outcome and how this interacts with the level of damage. Whilst therapy set size often varies across therapy studies, it has only been explored formally in one case-series study (for a meta-analysis of 21 therapy studies, involving 109 patients and a targetted case-series investigation, see Snell, Sage and Lambon Ralph, 2010) and as a result there are no clear conclusions as to best practice or how therapy set size should be varied according to the severity of the patient. More generally, if it were possible to predict the level of benefit achievable for a given level of damage severity in relation to manipulating the learning environment and therapy set size, therapists would have a particularly useful tool at their disposal. This investigation seeks to provide that tool and comment on its consequences for guiding clinical practice.

Simulation 4.1:
Manipulating the learning environment in rehabilitation.

This simulation consists of examining manipulating the learning environment in rehabilitation. The multimodal versus unimodal comparison that is first explored for developmental and recovery learning in Chapter 1, and verified in a larger model in Chapter 3, is carried out for rehabilitation learning in this simulation.

Method

The Rogers et al. (2004) model of semantic memory was replicated and scaled up to support twice as many items for this study. Details of the implementation and initial testing were reported in Chapter 3, as such this method section only details model parameters that have changed since Chapter 3

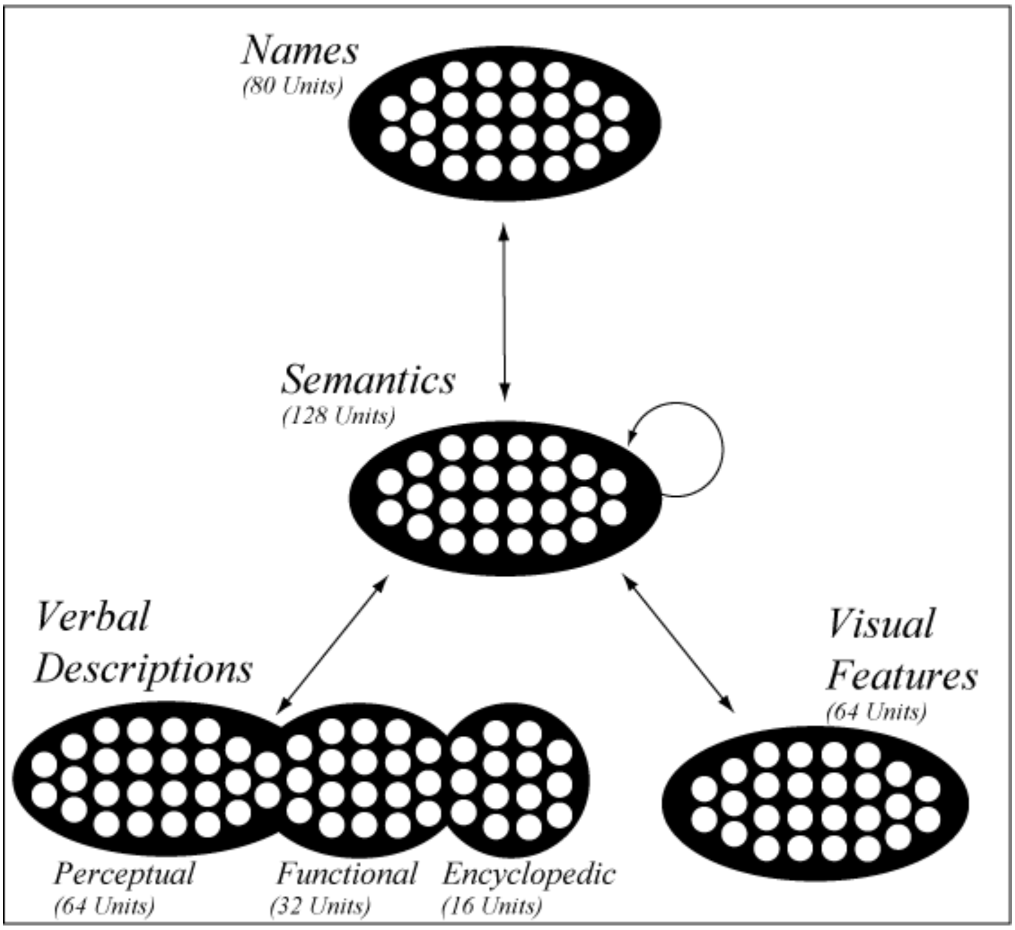

Figure 4.1 Architecture of the model adapted from Rogers et al. (2004)

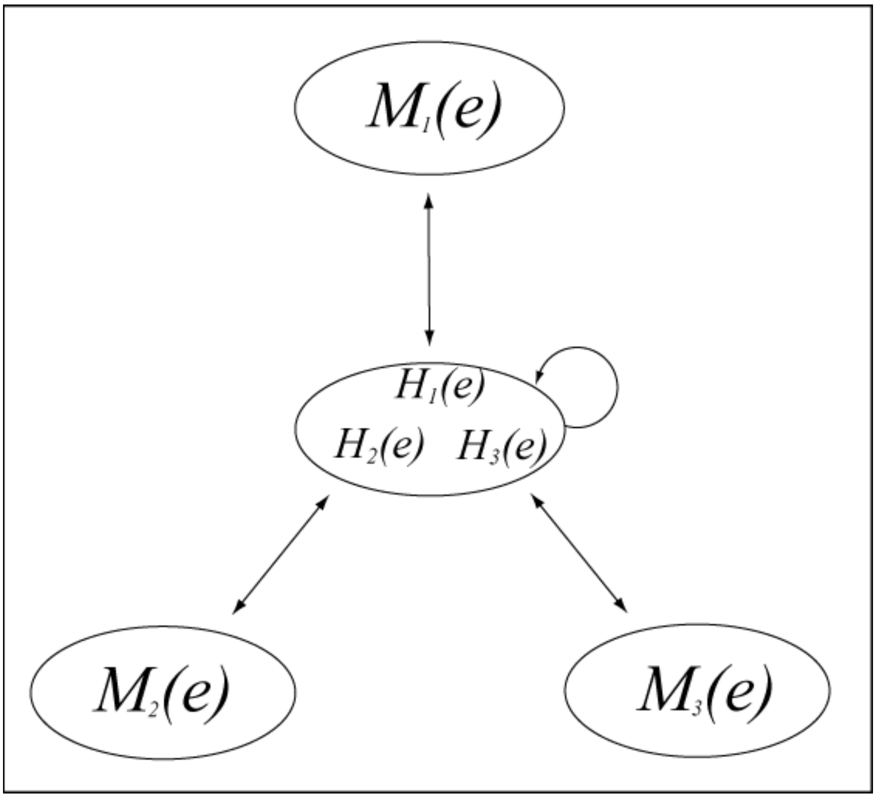

Figure 4.2 Architecture of the model redrawn as a generic PDP model with bidirectional mapping between each layer

Figure 4.2 illustrates the model in generic terms to show how an input representation (example e) of an object in any individual modality domain (layer), generates its own particular re-representation (denoted by H_1,H_2 or H_3 from input in modalities M_1,M_2,M_3 respectively) as a pattern of activity across the hidden semantic units (i.e. the hidden layer H).

Training Stimuli

The representations upon which the model was trained were created from prototype patterns in the same manner as in the original model (Rogers et al., 2004) described in Chapter 1. As previously stated the model in the current study is a scaled up version of the original capable of learning twice as many items (see Chapter 3 for technical details regarding the scaling process). The name representations were localist with 72 of the 80 units directly corresponding to a unique object upon which the model is trained. The four general object names (BIRD, ANIMAL, VEHICLE and TOOL) described in the Rogers et al. model , that would use the remaining name units were not used in the current study.

Training the model

As detailed in Chapter 3 the model was trained with the multimodal training regime until it had learnt all items accurately. Following the findings of the technical report detailed in Chapter 3 the identical four levels of lesion severity were again used for the simulations in the current study: Minimal - where the model relearns to 90% accuracy or higher. Mild - where the model relearns to around 75% accuracy, Moderate - where the model relearns to around 55% accuracy, and Severe where the model relearns to around 30% accuracy or less. As previously discussed in Chapters 1, 2 and 3 this is not identical to the real situation of localised injury but uniform lesioning across the model was the only practical solution to generate an appropriate range of relearning abilities necessary for the current set of simulations. The model was then lesioned at four different degrees of damage severity: Minimal (88% of the model's connections removed), Mild (90% of the model's connections removed), Moderate (92% of the model's connections removed) and Severe (93% of the model's connection removed). It should be noted that small variations in the percentage of connections removed have a substantial impact upon the model's learning behaviour. The model was then allowed to recover to a stable baseline (a sustained level of performance accuracy as shown in Chapter 3), by re-exposure to its original multimodal training regime with identical parameters. Therapy studies in patients test that patients are at a stable baseline performance, before therapy commences, for the items that will be learnt during therapy. At the end point of recovery to stable baseline the model is tested on its ability to name the items it had originally learnt (i.e. for name input to generate correct target outputs in all other modalities). All the items were then rank ordered by error (largest error first), and the 20 items with the highet error values were selected for use in rehabilitation training. These 20 items were then alternately assigned to either the therapy set or the control set, 10 items in each set. Alternately assigning the items ensured that the magnitude of the overall error for each set was as equal as possible to avoid one set containing an excess of items with a high error value. As far as possible modelling has thus attempted to simulate clinical practice.

As previously in Chapter 3, the model was trained on either one of two distinct training regimes multimodal or unimodal.
It is worth reiterating that the unimodal regime was developed as a result of splitting the multimodal regime into all of its constituent singular unimodal cross-domain mappings, and reflects the type of tasks used in rehabiltation therapy. Unimodal training is a reflection of this separation of the multimodal and corresponds to contrasting the two common real-world learning environments, development (multimodal) and post-damage rehabilitation tasks (unimodal). The unimodal regime was considered analogous to the learning environment experienced during adult learning such as that occurring during recovery or rehabilitation after damage (e.g. in the case of stroke etc.). Often therapeutic learning post-damage consists of presentation of learning items in single input and output modalities (e.g. ‘spoken word to picture matching’ consists of visual input, picture presentation, and verbal output, the spoken word).

For rehabilitation the learning rate was set to 0.01 (double the learning rate of 0.005 used in recovery, see Welbourne and Lambon Ralph, 2005b for the origins of this technique and a discussion of its validity) To simulate rehabilitation, the model was trained for 90000 weight updates (therapy often only occurs for an extremely short number of hours when contrasted with developmental learning), using either the multimodal and unimodal training regime, on the therapy set, for the four different levels of damage. This equates to 3000 presentations in the multimodal condition (multimodal presentation of the entire therapy, or control, set takes 30 updates, i.e. 3 training trials for each of the 10 items). In contrast, this amount of training equates to 1000 presentations in the unimodal condition (unimodal presentation of the entire therapy, or control, set takes 90 updates, i.e. 9 training trials for each of the 10 items). During training, the model was tested on its ability to generate the correct target outputs (to within 0.5) for items in the therapy set, as well as items in the control set, in all output modalities. Since the model received no training on the control set, the control items effectively correspond to untreated items in therapy. 90000 weight updates was chosen for the duration of training through initial experimentation, as whenever the model's learning had stabilised it had occurred by this time. The simulation was run ten times in the multimodal condition and ten times in the unimodal condition, for the therapy and control sets, and the results were averaged. During training, regular testing on each set recorded accuracy (percentage of examples correct upon testing), network error and representational economy (according to the equation derived below).

Calculating representational economy in the semantic convergence zone

As discussed previously in Chapter 1 as the model learns and structure develops, the semantic representations generated for unrelated objects become increasingly differentiated, whilst the representations of the same item elicited from different domains become more similar. Representational economy (RE) within the semantic layer was formally investigated by developing a statistic that could quantify this process. RE was measured in the same manner during this rehabilitation investigation.

Testing the network and analysing representational economy

Representational economy, accuracy and error were calculated at regular intervals during the model's training for items in the therapy and control sets. A 2x20 repeated measures analysis of variance (ANOVA) was conducted on these data to investigate the effect of manipulating the training regime (multimodal or unimodal) for each level of damage severity (minimal, mild, moderate or severe) at twenty equally spaced time points during training on the dependent variables (accuracy, representational economy and network error). For each of the simulations the correlation between representational economy, accuracy, and network error was calculated as the model learns to explore Representational Economy's relationship to Accuracy and Network Error for a range of relearning scenarios.

Results

This simulation took a damaged model that recovers to baseline (analogous to therapy beginning after recovery to stable baseline) and exposed it to both multimodal and unimodal learning environments for training on a targetted set of items. It is worth remembering that the purpose is to compare performance of multimodal and unimodal learning environments, the specific interest of this thesis is to see whether the efficiency advantages of the multimodal environment observed in developmental and recovery learning (see Chapters, 1,2, and 3) will also occur in rehabilitation learning.

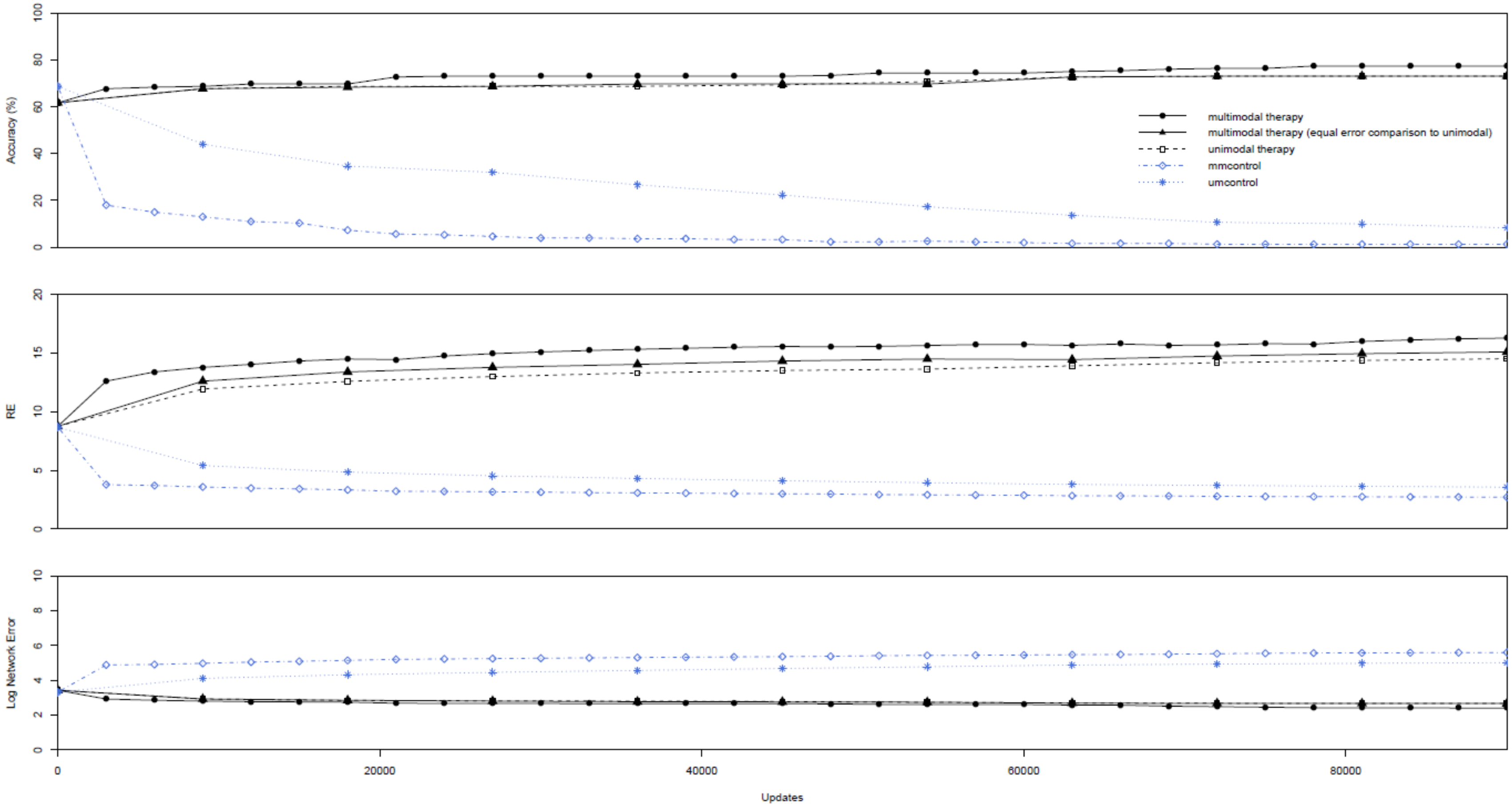

Figure 4.3 Mean variation in training performance during rehabilitation after a minimal lesion

Table 4.1 2x20 ANOVAs indicating significant performance differences between training regimes during rehabilitation after a minimal lesion

Training Regime Comparison	Dependent Variable Measure	Independent Variables	df	F	p	Effect size (η^2)
Multimodal Therapy Vs Unimodal Therapy	Accuracy	Training	1,9	18.297	0.002*	0.670
		Time	19,171	13.374	<0.001*	0.598
		Training*Time	19,171	1.390	0.137	0.134
	Representational Economy	Training	1,9	87.027	<0.001*	0.906
		Time	19,171	199.837	<0.001*	0.957
		Training*Time	19,171	2.547	0.001*	0.221
	Network Error	Training	1,9	33.181	<0.001*	0.787
		Time	19,171	26.690	<0.001*	0.748
		Training*Time	19,171	1.249	0.224	0.122
Equated Error Multimodal Therapy Vs Unimodal Therapy	Accuracy	Training	1,9	0.007	0.934	0.001
		Time	19,171	4.638	<0.001*	0.340
		Training*Time	19,171	1.178	0.282	0.116
	Representational Economy	Training	1,9	14.397	0.004*	0.615
		Time	19,171	259.579	<0.001*	0.966
		Training*Time	19,171	2.367	0.002*	0.208
	Network Error	Training	1,9	2.328	0.161	0.206
		Time	19,171	20.106	<0.001*	0.691
		Training*Time	19,171	0.995	0.469	0.100
Multimodal Control Vs Unimodal Control	Accuracy	Training	1,9	46.185	<0.001*	0.837
		Time	19,171	40.353	<0.001*	0.818
		Training*Time	19,171	15.878	<0.001*	0.638
	Representational Economy	Training	1,9	143.890	<0.001*	0.941
		Time	19,171	51.448	<0.001*	0.851
		Training*Time	19,171	20.059	<0.001*	0.690
	Network Error	Training	1,9	286.559	<0.001*	0.970
		Time	19,171	82.477	<0.001*	0.902
		Training*Time	19,171	4.283	<0.001*	0.322

Rehabilitation After Minimal Lesioning

After minimal lesioning, performance on the therapy set improved only slightly from its initial performance in terms of accuracy and network error, however there was a large improvement in convergence amongst semantic representations measured as representational economy (as illustrated in Figure 4.3) . Multimodal training achieved significantly higher accuracy, representational economy and lower network error than unimodal training on the therapy set (a summary of the ANOVAs is provided in Table 4.1). The control set demonstrated no improvement in performance, despite starting with higher accuracy, representational economy, and lower network error. Indeed, there was a deterioration of performance over the rehabilitation period which is due to catastrophic interference (McCloskey and Cohen, 1989) within the network. The equated error comparison only shows a statistically significant difference in representational economy (see Table 4.1) between multimodal and unimodal training on the therapy set with no difference for accuracy or network error.

Table 4.1 shows a statistically significant advantage for multimodal therapy over unimodal therapy with a significant effect of training and time on accuracy, representational economy and network error, and a significant interaction between training and time for representational economy. There is a statistically significant difference between equated error multimodal therapy and unimodal therapy in representational economy and a significant effect of time on accuracy, representational economy and network error. It is also clear to see that the multimodal control outperformed the unimodal control with a significant difference in accuracy, representational economy and network error. Representational Economy was strongly correlated with learning Accuracy for all conditions: Multimodal Therapy $r(20)=.860$, $p<.001$; Equated Error Multimodal Therapy $r(20)=.884$, $p<.001$; Unimodal Therapy $r(20)=.900, p<.001$;Multimodal Control $r(20)=.958, p<.001$;Unimodal Control $r(20)=.987, p<.001$. Representational Economy was strongly correlated with Network

Error for all conditions: Multimodal Therapy $r(20)=-.895$, $p<.001$; Equated Error Multimodal Therapy $r(20)=-.990$, $p<.001$; Unimodal Therapy $r(20)=-.988$, $p<.001$; Multimodal Control $r(20)=-.988$, $p<.001$; Unimodal Control $r(20)=-.962$, $p<.001$.

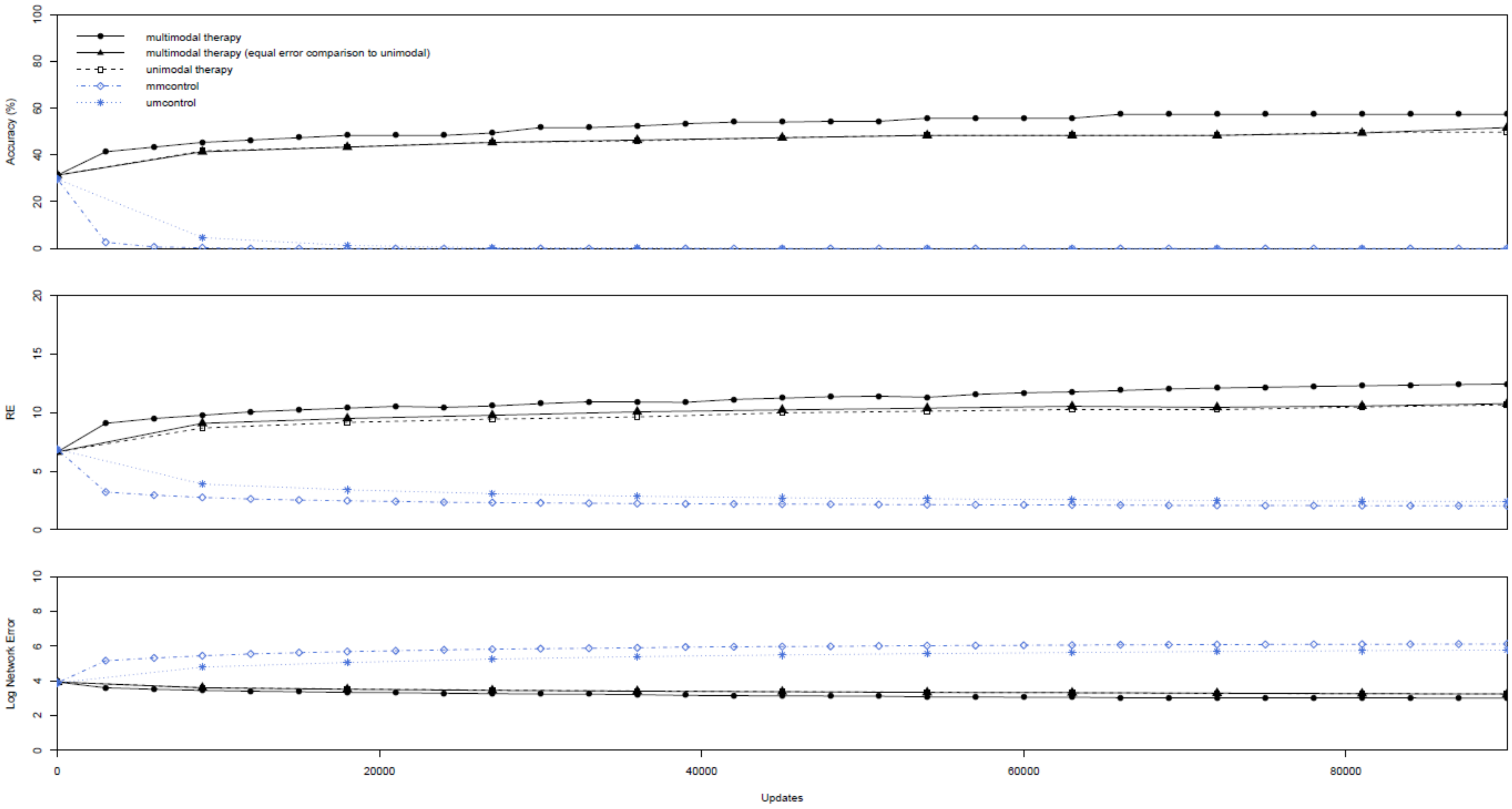

Figure 4.4 Mean variation in training performance during rehabilitation after a mild lesion

Table 4.2 2x20 ANOVAs indicating significant performance differences between training regimes during rehabilitation after a mild lesion.

Training Regime Comparison	Dependent Variable Measure	Independent Variables	df	F	p	Effect size (η^2)
Multimodal Therapy Vs Unimodal Therapy	Accuracy	Training	1,9	27.621	0.001*	0.754
		Time	19,171	21.378	<0.001*	0.704
		Training*Time	19,171	0.735	0.779	0.075
	Representational Economy	Training	1,9	183.168	<0.001*	0.953
		Time	19,171	68.072	<0.001*	0.883
		Training*Time	19,171	1.726	0.036*	0.161
	Network Error	Training	1,9	98.664	<0.001*	0.916
		Time	19,171	86.139	<0.001*	0.905
		Training*Time	19,171	0.663	0.851	0.069
Equated Error Multimodal Therapy Vs Unimodal Therapy	Accuracy	Training	1,9	2.912	0.122	0.244
		Time	19,171	14.167	<0.001*	0.612
		Training*Time	19,171	1.189	0.272	0.117
	Representational Economy	Training	1,9	35.868	<0.001*	0.799
		Time	19,171	35.381	<0.001*	0.797
		Training*Time	19,171	1.317	0.178	0.128
	Network Error	Training	1,9	2.680	0.136	0.229
		Time	19,171	77.639	<0.001*	0.896
		Training*Time	19,171	1.546	0.076	0.147
Multimodal Control Vs Unimodal Control	Accuracy	Training	1,9	26.638	0.001*	0.747
		Time	19,171	20.892	<0.001*	0.699
		Training*Time	19,171	10.532	<0.001*	0.539
	Representational Economy	Training	1,9	316.024	<0.001*	0.972
		Time	19,171	214.212	<0.001*	0.960
		Training*Time	19,171	49.520	<0.001*	0.846
	Network Error	Training	1,9	578.869	<0.001*	0.985
		Time	19,171	526.339	<0.001*	0.983
		Training*Time	19,171	4.195	<0.001*	0.318

Rehabilitation After Mild Lesioning

After mild lesioning, performance on the therapy set improved from its initial performance in terms of accuracy and network error, and in convergence amongst semantic representations measured as representational economy (as illustrated in

Figure 4.4) . This improvement is greater than after minimal lesioning. Multimodal training achieves higher accuracy, representational economy and lower network error than unimodal training on the therapy set and this difference is confirmed as statistically significant by the results of the ANOVA in Table 4.2. The control set exhibited no improvement in performance, despite starting with higher accuracy, representational economy and lower network error. Again, the deterioration of performance is again due to catastrophic interference within the network. The equated error comparison only shows a statistically significant difference in representational economy (see Table 4.2) between multimodal and unimodal training on the therapy set with no difference for accuracy or network error.

In Table 4.2 the results give a statistically significant effect of training and time on accuracy, representational economy and network error for multimodal therapy over unimodal therapy, in addition there is a significant interaction between training and time for representational economy. When comparing equated error multimodal therapy and unimodal therapy there is a significant effect of time on on accuracy, representational economy and network error for as well as a significant effect of training for representational economy. A multimodal advantage over the unimodal control is also shown by the statistically significant difference in accuracy, representational economy and network error. Representational Economy was strongly correlated with learning Accuracy for all conditions: Multimodal Therapy $r(20)=.981$, $p<.001$; Equated Error Multimodal Therapy $r(20)=.968$, $p<.001$; Unimodal Therapy $r(20)=.984$,$p<.001$; Multimodal Control $r(20)=.788$,$p<.001$;Unimodal Control $r(20)=.922$,$p<.001$. Representational Economy was strongly correlated with Network Error for all conditions: Multimodal Therapy $r(20)=-.968$, $p<.001$; Equated Error

Multimodal Therapy $r(20)=-.995$, $p<.001$; Unimodal Therapy $r(20)=-.996$,$p<.001$; Multimodal Control $r(20)=-.976$,$p<.001$;Unimodal Control $r(20)=-.965$,$p<.001$.

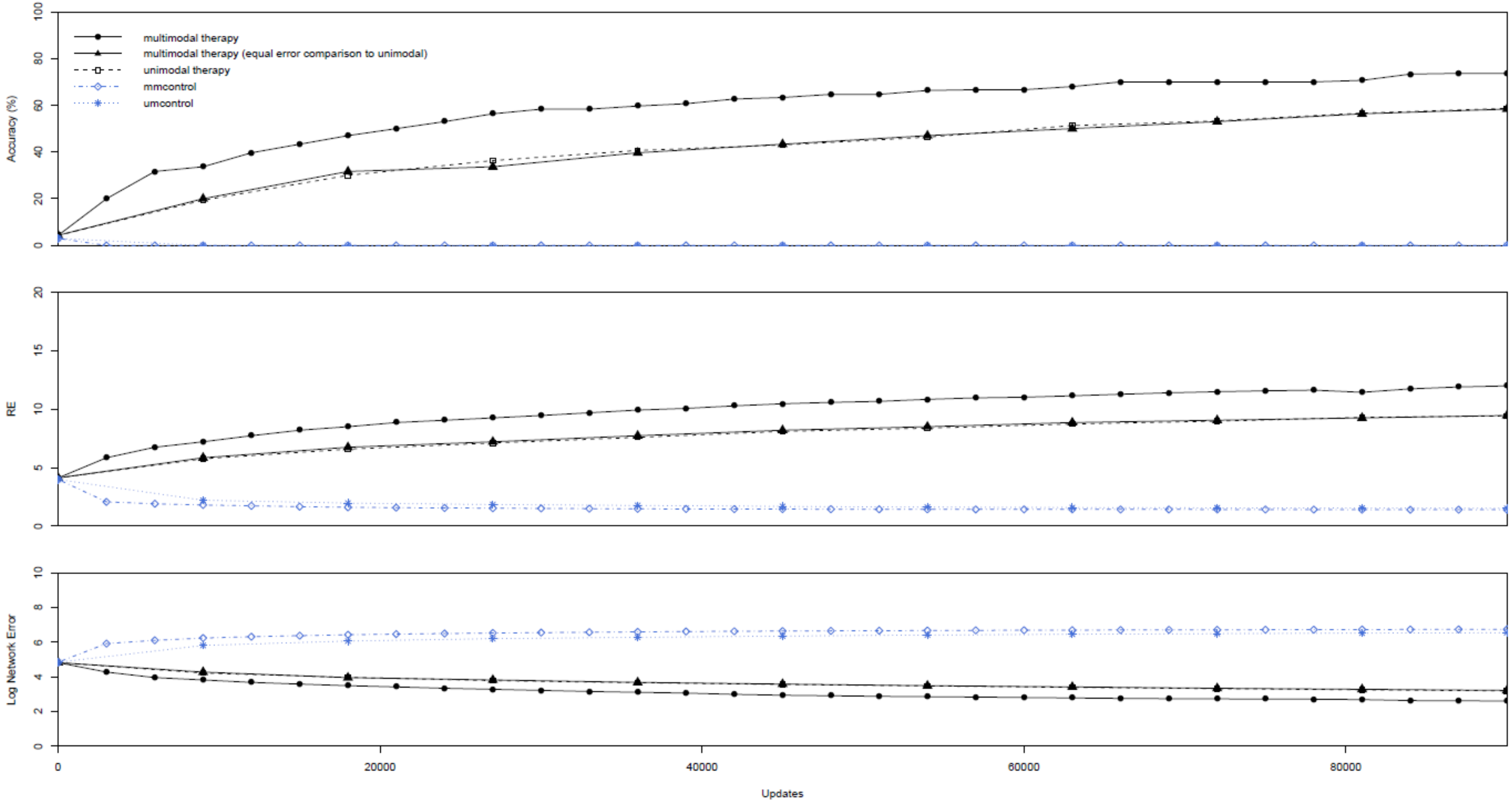

Figure 4.5 Mean variation in training performance during rehabilitation after a moderate lesion

Table 4.3 2x20 ANOVAs indicating significant performance differences between training regimes during rehabilitation after a moderate lesion. Zero comparisons indicated by ……

Training Regime Comparison	Dependent Variable Measure	Independent Variables	df	F	p	Effect size (η^2)
Multimodal Therapy Vs Unimodal Therapy	Accuracy	Training	1,9	137.170	<0.001*	0.938
		Time	19,171	85.301	<0.001*	0.905
		Training*Time	19,171	0.958	0.513	0.096
	Representational Economy	Training	1,9	316.866	<0.001*	0.972
		Time	19,171	218.662	<0.001*	0.960
		Training*Time	19,171	5.306	<0.001*	0.371
	Network Error	Training	1,9	239.710	<0.001*	0.964
		Time	19,171	298.847	<0.001*	0.971
		Training*Time	19,171	15.355	<0.001*	0.630
Equated Error Multimodal Therapy Vs Unimodal Therapy	Accuracy	Training	1,9	0.266	0.619	0.029
		Time	19,171	38.854	<0.001*	0.812
		Training*Time	19,171	1.620	0.056	0.153
	Representational Economy	Training	1,9	9.479	0.013*	0.513
		Time	19,171	203.963	<0.001*	0.958
		Training*Time	19,171	2.191	0.004*	0.196
	Network Error	Training	1,9	0.049	0.829	0.005
		Time	19,171	187.180	<0.001*	0.954
		Training*Time	19,171	2.093	0.007*	0.189
Multimodal Control Vs Unimodal Control	Accuracy	Training	1,9	……	……	……
		Time	19,171	……	……	……
		Training*Time	19,171	……	……	……
	Representational Economy	Training	1,9	151.614	<0.001*	0.944
		Time	19,171	123.475	<0.001*	0.932
		Training*Time	19,171	30.052	<0.001*	0.770
	Network Error	Training	1,9	376.800	<0.001*	0.977
		Time	19,171	379.682	<0.001*	0.977
		Training*Time	19,171	4.989	<0.001*	0.357

Rehabilitation After Moderate Lesioning

After moderate lesioning, performance on the therapy set improved considerably from its initial performance in terms of accuracy and network error, and in convergence amongst semantic representations measured as representational economy (as illustrated in Figure 4.5) . This improvement wass much greater than after minimal and mild lesioning. Multimodal training achieved higher accuracy, representational economy and lower network error than unimodal training on the therapy set and this difference is confirmed as statistically significant by the results of the ANOVA in Table 4.3. There was no improvement in performance on the control set. The equated error comparison again only shows a statistically significant difference in representational economy (see Table 4.3) between multimodal and unimodal training on the therapy set, with no difference for accuracy or network error.

Table 4.3 presents results showing that the better performance of multimodal therapy is a statistically significant difference from the unimodal in accuracy, representational economy and network error except for the non-significant interaction between training and time for accuracy. The table also shows a significant effect of time and an interaction between training and time for the comparison of equated error multimodal and unimodal therapy as well as a significant effect of training on representational economy. The multimodal control performs better than the unimodal control with significant differences in representational economy and network error though neither training yielded any accuracy hence no statistical comparison was possible for accuracy. Representational Economy was strongly correlated with learning Accuracy for all conditions: Multimodal Therapy $r(20)=.986$, $p<.001$; Equated Error Multimodal Therapy $r(20)=.996$, $p<.001$; Unimodal Therapy $r(20)=.998$,$p<.001$; Multimodal Control (zero values so no calculation);Unimodal Control (zero values so no calculation). Representational Economy was strongly correlated with Network Error for all conditions: Multimodal Therapy $r(20)=-.968$, $p<.001$; Equated Error

Multimodal Therapy $r(20)=-.970$, $p<.001$; Unimodal Therapy $r(20)=-.972$, $p<.001$; Multimodal Control $r(20)=-.979$, $p<.001$; Unimodal Control $r(20)=-.980$, $p<.001$.

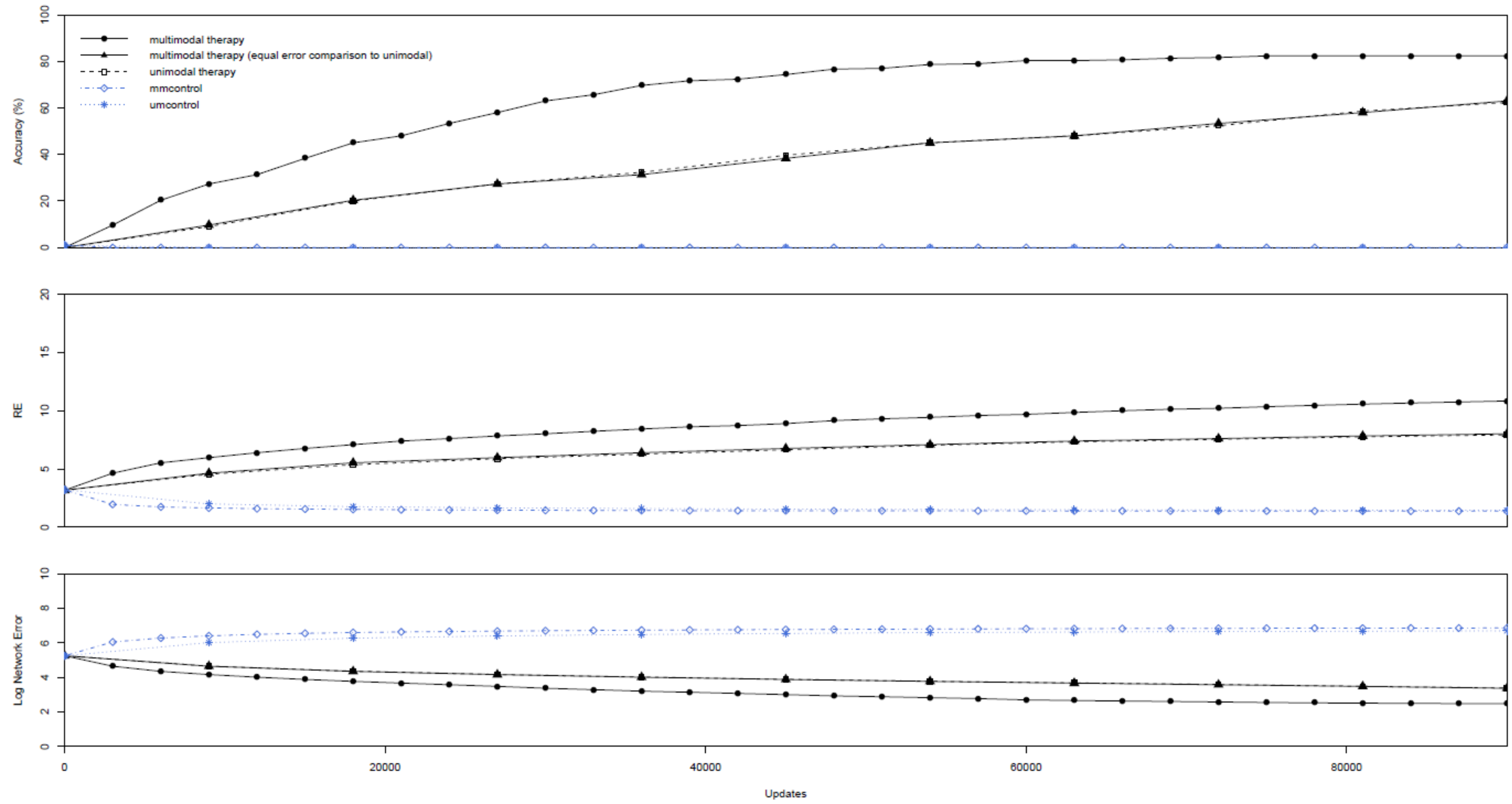

Figure 4.6 Mean variation in training performance during rehabilitation after a severe lesion

Table 4.4 2x20 ANOVAs indicating significant performance differences betweem training regimes during rehabilitation after a severe lesion. Zero comparisons indicated by ……

Training Regime Comparison	Dependent Variable Measure	Independent Variables	df	F	p	Effect size (η^2)
Multimodal Therapy Vs Unimodal Therapy	Accuracy	Training	1,9	370.592	<0.001*	0.976
		Time	19,171	188.864	<0.001*	0.955
		Training*Time	19,171	7.381	<0.001*	0.451
	Representational Economy	Training	1,9	852.454	<0.001*	0.990
		Time	19,171	568.231	<0.001*	0.984
		Training*Time	19,171	17.788	<0.001*	0.664
	Network Error	Training	1,9	410.685	<0.001*	0.979
		Time	19,171	334.504	<0.001*	0.974
		Training*Time	19,171	25.626	<0.001*	0.740
Equated Error Multimodal Therapy Vs Unimodal Therapy	Accuracy	Training	1,9	0.017	0.900	0.002
		Time	19,171	80.227	<0.001*	0.899
		Training*Time	19,171	0.431	0.982	0.046
	Representational Economy	Training	1,9	2.499	0.148	0.217
		Time	19,171	242.239	<0.001*	0.964
		Training*Time	19,171	1.475	0.100	0.141
	Network Error	Training	1,9	0.041	0.844	0.005
		Time	19,171	246.587	<0.001*	0.965
		Training*Time	19,171	1.870	0.019*	0.172
Multimodal Control Vs Unimodal Control	Accuracy	Training	1,9	……	……	……
		Time	19,171	……	……	……
		Training*Time	19,171	……	……	……
	Representational Economy	Training	1,9	229.850	<0.001*	0.962
		Time	19,171	152.488	<0.001*	0.944
		Training*Time	19,171	16.004	<0.001*	0.640
	Network Error	Training	1,9	486.625	<0.001*	0.982
		Time	19,171	748.662	<0.001*	0.988
		Training*Time	19,171	13.009	<0.001*	0.591

Rehabilitation After Severe Lesioning

After severe lesioning, there was a large improvement in performance on the therapy set in comparison to its initial performance in terms of accuracy and network error, and in convergence amongst semantic representations measured as representational economy (as illustrated in Figure 4.6). This improvement is much greater than after minimal,mild and moderate lesioning. Multimodal training achieves higher accuracy, representational economy and lower network error than unimodal training on the therapy set and this difference is confirmed as statistically significant by the results of the ANOVA in Table 4.4. There was no change in performance for the control set. The equated error comparison shows no statistically significant difference in accuracy, network error or representational economy (see Table 4.4) between multimodal and unimodal training on the therapy set.

Table 4.4 again shows the multimodal advantage as a statistically significant difference between multimodal and unimodal therapy in representational economy, accuracy and network error. Comparing equated error multimodal and unimodal showed an effect of time in accuracy, representational economy and network error as well as a significant interaction between training and time for network error. The comparison of the multimodal and unimodal control shows a multimodal advantage with significant differences in representational economy and network error though accuracy remained zero for both forms of training hence no statistical comparison was possible for accuracy. Representational Economy was strongly correlated with learning Accuracy for all conditions: Multimodal Therapy $r(20)=.967$, $p<.001$; Equated Error Multimodal Therapy $r(20)=.990$, $p<.001$; Unimodal Therapy $r(20)=.991$,$p<.001$: Multimodal Control (zero values so no calculation):Unimodal Control (zero values so no calculation). Representational Economy was strongly correlated with Network Error for all conditions: Multimodal Therapy $r(20)=-.945$, $p<.001$; Equated Error Multimodal Therapy $r(20)=-.971$, $p<.001$; Unimodal Therapy $r(20)=-.972$,$p<.001$: Multimodal Control $r(20)=-.985$,$p<.001$:Unimodal Control $r(20)=-.973$,$p<.001$.

Simulation 4.2:

Manipulating the learning environment and therapy set size in rehabilitation from stable baseline recovery that incorporates background spontaneous recovery.

This simulation extends simulation 4.1 by examining manipulating the learning environment in rehabilitation whilst also allowing a background level of recovery to occur in parallel with learning. The goal of this chapter is to realistically simulate patient rehabilitation. Since spontaneous recovery can continue in patients during their rehabilitation (Howard, 1994), allowing a background level of recovery to occur in the model in parallel to rehabilitation learning offers a more realistic simulation of what happens in rehabilitation therapy. The multimodal versus unimodal comparison that was first explored for developmental and recovery learning in Chapter 1, and verified in a larger model in Chapter 3, is carried out for rehabilitation learning in this simulation whilst incorporating a background level of recovery learning delivered identically to the recovery learning described in Chapter 3.

Method

The model was trained in the manner described in Simulation 4.1 for the severe damage condition (i.e. 93% of connections removed, since that yielded a suitably low starting performance) with 10 items in the therapy and control sets. The following modifications were made: throughout rehabilitation, for each epoch, the model was re-exposed to the entire original training corpus at the same learning rate used during the model's recovery to stable baseline (0.005). Similarly, at each testing point, the therapy and control sets were assessed but alongside the entire corpus set in order to gain a measure of background spontaneous recovery as well as the targetted rehabilitation. This approach was taken to capture the fact that spontaneous recovery still occurs during therapy and to gain a measure of its effects.

This simulation compared a therapy (and control) set size, of 5 items with a set size of 10 and 20 items, in order to explore the effect that this has on therapy outcome. All other parameters and processes in this simulation are identical to those used in simulation 4.1. Analysis also now included an ANOVA comparison of multimodal and unimodal background recovery.

Results

The results for Simulation 4.2 are presented first as summary bar charts showing the mean performance accuracy for all 10 models (simulated patients) alongside the individual results for each model (see Figure 4.7). This is done to draw as much comparison as possible to the reporting of data from therapy studies. Then the breakdown graphs of accuracy, representational economy and log network error are presented .It is also worth noting, in advance, that unlike previous observations (from Chapters, 1,2 and 3), equated error comparisons are no longer approximately identical to unimodal relearning performance.

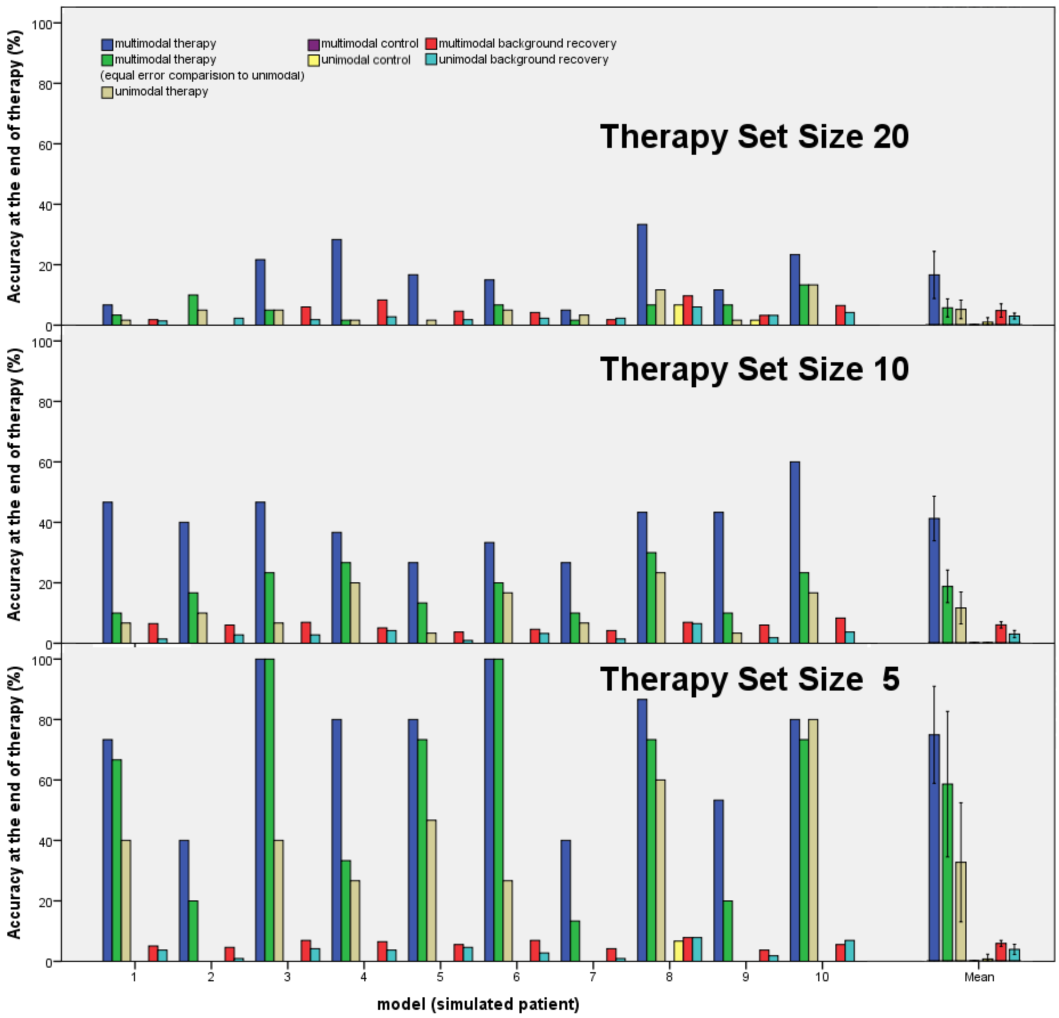

Figure 4.7 Rehabilitation performance for therapy and control sets for set sizes of 5, 10 and 20 items with background recovery. Performance shown for each of the 10 individual models then the mean performance is given on the far right of each chart

Clearly smaller therapy set sizes resulted in better performance with multimodal training outperforming unimodal training, this time even after equating for error unlike in Simulation 4.1. The control sets showed no improvement in performance as would be expected for untreated items. The level of background spontaneous recovery remains roughly identical regardless of changes in therapy set size.

Rehabilitation with a therapy set size of 5 and background spontaneous recovery after a severe lesion

After severe lesioning, and with a therapy set size of 5, there was a large improvement in terms of accuracy and network error, and in convergence amongst semantic representations measured as representational economy (as illustrated in Figure 4.8) . Multimodal training achieves higher accuracy, representational economy and lower network error than unimodal training on the therapy set and this difference is confirmed as statistically significant by the results of the ANOVA in Table 4.5. The control set exhibits no improvement in performance. The equated error comparison shows statistically significant differences in accuracy, network error and representational economy (see Table 4.5) between multimodal and unimodal training on the therapy set. The level of background recovery is greater for multimodal training and shows statistically significant differences in accuracy, network error and representational economy (see Table 4.5) compared to unimodal training.

Table 4.5 shows an advantage for multimodal therapy over unimodal therapy as well as an advantage for equated error multimodal therapy over unimodal therapy with as statistically significant difference across the board in accuracy, representational economy and network error. The multimodal control did slightly better than the unimodal control as show by the significant effect of training and a significant interaction between training and time for representational economy and an effect of training, time and a significant interaction between training and time in network error. Multimodal background recovery also did better than unimodal background recovery with a significant difference in accuracy, representational economy and network error but no effect of time on network error.

Representational Economy was strongly correlated with learning Accuracy for all conditions: Multimodal Therapy $r(20)=.964$, $p<.001$; Equated Error Multimodal Therapy $r(20)=.932$, $p<.001$; Unimodal Therapy $r(20)=.914$,$p<.001$: Multimodal Control $r(20)=-.115$,$p=.630$:Unimodal Control $r(20)=.284$,$p=.225$; Multimodal General $r(20)=.651$,$p=.002$:Unimodal General $(20)=-.173$,$p=.466$.

Representational Economy was strongly correlated with Network Error for all conditions: Multimodal Therapy $r(20)=-.954$, $p<.001$; Equated Error Multimodal Therapy $r(20)=-.958$, $p<.001$; Unimodal Therapy $r(20)=-.972$,$p<.001$: Multimodal Control $r(20)=-.734$,$p<.001$:Unimodal Control $r(20)=-.697$,$p<.001$; Multimodal General $r(20)=-.814$,$p<.001$;Unimodal General $(20)=--.881$,$p<.001$.

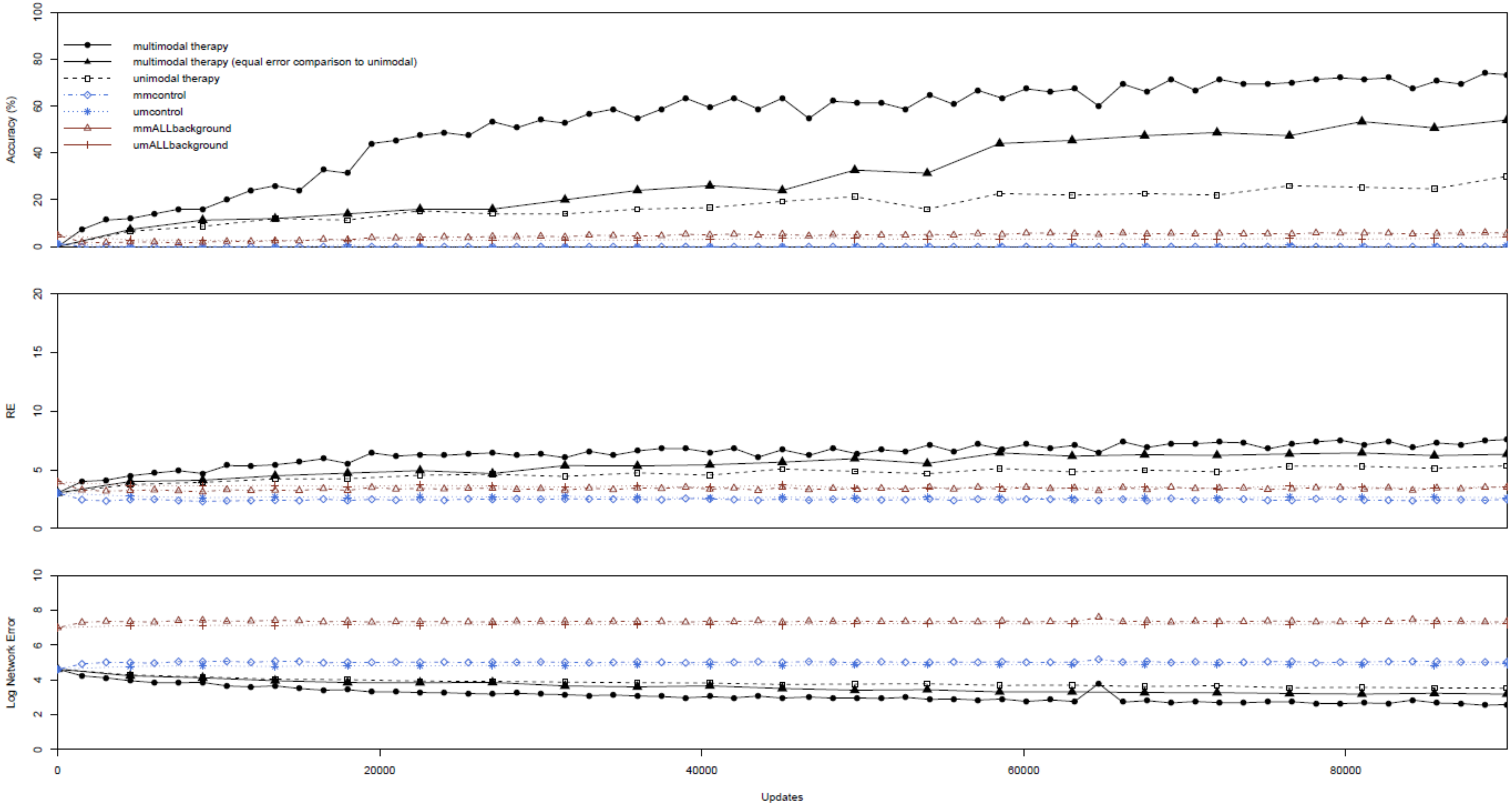

Figure 4.8 Mean variation in training performance for a therapy set of 5 with background recovery during rehabilitation after a severe lesion

Table 4.5 2x20 ANOVAs indicating significant performance differences between training regimes for a therapy set of 5 during rehabilitation after a severe lesion

Training Regime Comparison	Dependent Variable Measure	Independent Variables	df	F	p	Effect size (η^2)
Multimodal Therapy Vs Unimodal Therapy	Accuracy	Training	1,9	54.106	<0.001*	0.857
		Time	19,171	37.104	<0.001*	0.805
		Training*Time	19,171	9.209	<0.001*	0.506
	Representational Economy	Training	1,9	181.149	<0.001*	0.953
		Time	19,171	27.442	<0.001*	0.753
		Training*Time	19,171	4.964	<0.001*	0.355
	Network Error	Training	1,9	277.138	<0.001*	0.969
		Time	19,171	121.535	<0.001*	0.931
		Training*Time	19,171	1.997	0.011*	0.182
Equated Error Multimodal Therapy Vs Unimodal Therapy	Accuracy	Training	1,9	20.585	0.001*	0.696
		Time	19,171	18.486	<0.001*	0.673
		Training*Time	19,171	7.979	<0.001*	0.470
	Representational Economy	Training	1,9	346.459	<0.001*	0.975
		Time	19,171	37.463	<0.001*	0.806
		Training*Time	19,171	3.723	<0.001*	0.293
	Network Error	Training	1,9	381.414	<0.001*	0.977
		Time	19,171	65.456	<0.001*	0.879
		Training*Time	19,171	3.004	<0.001*	0.250
Multimodal Control Vs Unimodal Control	Accuracy	Training	1,9	2.087	0.182	0.188
		Time	19,171	0.851	0.643	0.086
		Training*Time	19,171	0.552	0.934	0.058
	Representational Economy	Training	1,9	29.736	<0.001*	0.768
		Time	19,171	1.281	0.202	0.125
		Training*Time	19,171	1.593	0.063	0.150
	Network Error	Training	1,9	133.423	<0.001*	0.937
		Time	19,171	1.650	0.050	0.155
		Training*Time	19,171	2.148	0.005*	0.193

Training Regime Comparison [continued]	Dependent Variable Measure	Independent Variables	df	F	p	Effect size (η^2)
Multimodal Background Recovery Vs Unimodal Background Recovery	Accuracy	Training	1,9	14.219	0.004*	0.612
		Time	19,171	20.214	<0.001*	0.692
		Training*Time	19,171	6.406	<0.001*	0.416
	Representational Economy	Training	1,9	56.926	<0.001*	0.863
		Time	19,171	2.200	0.004*	0.196
		Training*Time	19,171	3.183	<0.001*	0.261
	Network Error	Training	1,9	192.158	<0.001*	0.955
		Time	19,171	1.396	0.134	0.134
		Training*Time	19,171	2.440	0.001*	0.213

Rehabilitation with a therapy set size of 10 and background spontaneous recovery after a severe lesion

After severe lesioning with a therapy set size of 10 there was a large improvement in terms of accuracy and network error, and in convergence amongst semantic representations, measured in terms of representational economy (as illustrated in Figure 4.9) . This improvement was less than that observed for a therapy set size of 5. Multimodal training achieves higher accuracy, representational economy and lower network error than unimodal training on the therapy set and these differences were confirmed as statistically significant by the results of the ANOVA in Table 4.6. The control set makes no improvement in performance. The equated error comparison shows statistically significant differences in accuracy, network error and representational economy (see Table 4.6) between multimodal and unimodal training on the therapy set. The level of background recovery is greater for multimodal training and again shows statistically significant difference in accuracy, network error and representational economy (see Table 4.6) compared to unimodal training.

In Table 4.6 the advantage for multimodal therapy over unimodal therapy can be seen in the statistically significant difference in accuracy, representational economy and network error between the two regimes. There is also a statistically significant difference in accuracy, representational economy and network error between equated error multimodal and unimodal therapy. The multimodal control performed better than the unimodal in terms of training and time for representational economy and network error. Multimodal background recovery also showed an advantage over unimodal background recovery with a significant difference in accuracy, there was also a significant effect of training on representational economy and network error, a significant interaction between training and time for representational economy and a significant effect of time on network error. Representational Economy was strongly correlated with learning Accuracy for all conditions: Multimodal Therapy $r(20)=.958$, $p<.001$; Equated Error Multimodal Therapy $r(20)=.894$, $p<.001$; Unimodal Therapy $r(20)=.903$,$p<.001$: Multimodal Control $r(20)=-.096$,$p=.686$:Unimodal Control $r(20)=.033$,$p=.891$; Multimodal General $r(20)=.466$,$p=.039$:Unimodal General $(20)=.298$,$p=.201$. Representational Economy was strongly correlated with Network Error for all conditions: Multimodal Therapy $r(20)=-.987$, $p<.001$; Equated Error Multimodal Therapy $r(20)=-.981$, $p<.001$; Unimodal Therapy $r(20)=-.977$,$p<.001$; Multimodal Control $r(20)=-.912$,$p<.001$;Unimodal Control $r(20)=-.954$,$p<.001$; Multimodal General $r(20)=-.466$,$p=.039$;Unimodal General $(20)=-.908$,$p<.001$.

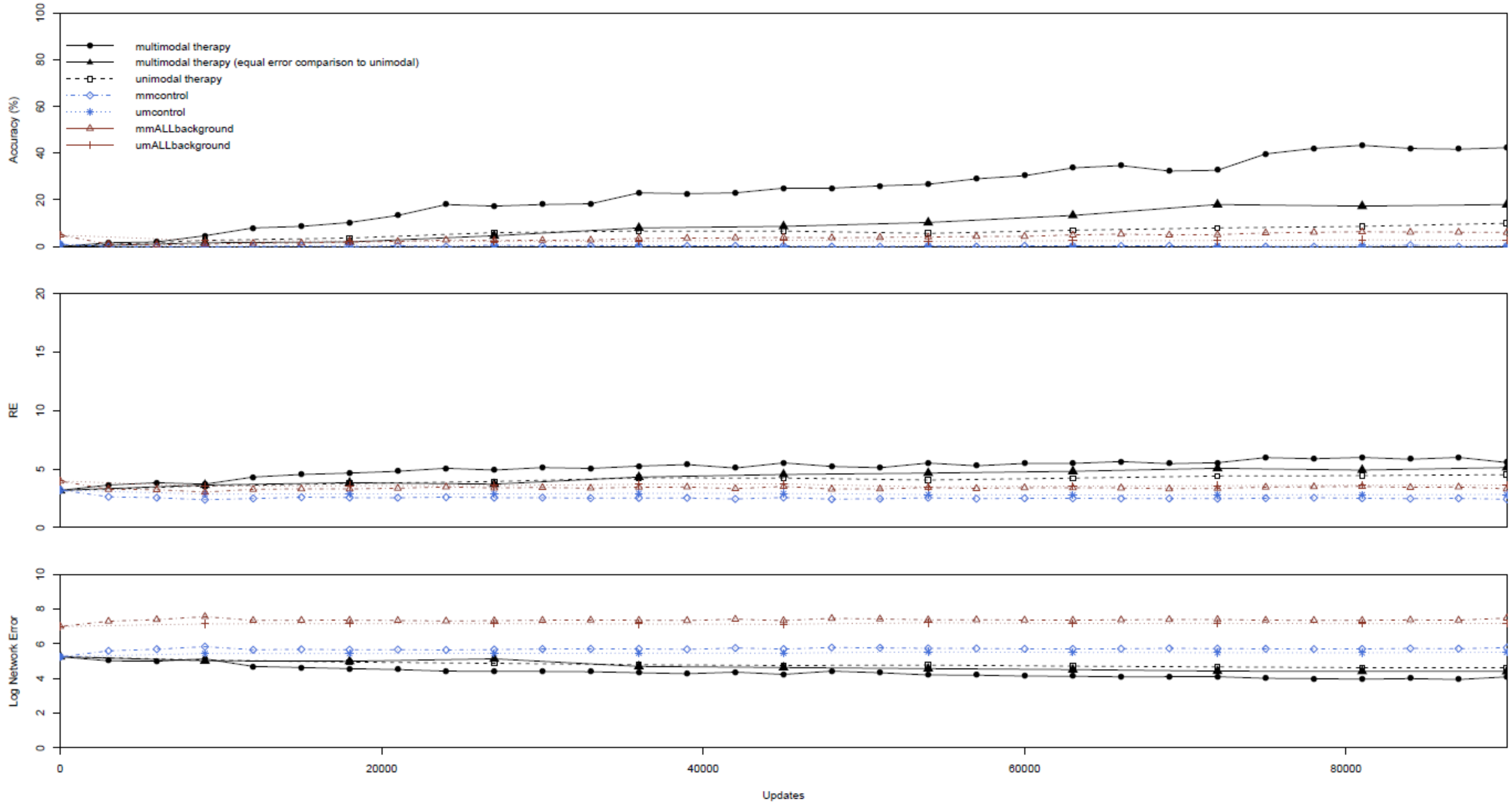

Figure 4.9 Mean variation in training performance for a therapy set of 10 with background recovery during rehabilitation after a severe lesion

Table 4.6 2x20 ANOVAs indicating significant performance differences between training regimes for a therapy set of 10 during rehabilitation after a severe lesion

Training Regime Comparison	Dependent Variable Measure	Independent Variables	df	F	p	Effect size (η^2)
Multimodal Therapy Vs Unimodal Therapy	Accuracy	Training	1,9	260.153	<0.001*	0.967
		Time	19,171	57.630	<0.001*	0.865
		Training*Time	19,171	26.107	<0.001*	0.744
	Representational Economy	Training	1,9	519.540	<0.001*	0.983
		Time	19,171	20.641	<0.001*	0.696
		Training*Time	19,171	4.065	<0.001*	0.311
	Network Error	Training	1,9	593.567	<0.001*	0.985
		Time	19,171	84.490	<0.001*	0.904
		Training*Time	19,171	3.280	<0.001*	0.267
Equated Error Multimodal Therapy Vs Unimodal Therapy	Accuracy	Training	1,9	42.960	<0.001*	0.827
		Time	19,171	18.460	<0.001*	0.672
		Training*Time	19,171	6.824	<0.001*	0.431
	Representational Economy	Training	1,9	59.778	<0.001*	0.869
		Time	19,171	34.062	<0.001*	0.791
		Training*Time	19,171	4.547	<0.001*	0.336
	Network Error	Training	1,9	88.910	<0.001*	0.908
		Time	19,171	56.932	<0.001*	0.863
		Training*Time	19,171	6.257	<0.001*	0.410
Multimodal Control Vs Unimodal Control	Accuracy	Training	1,9	1.374	0.271	0.132
		Time	19,171	1.137	0.318	0.112
		Training*Time	19,171	1.201	0.262	0.118
	Representational Economy	Training	1,9	141.360	<0.001*	0.940
		Time	19,171	2.866	<0.001*	0.242
		Training*Time	19,171	0.937	0.538	0.094
	Network Error	Training	1,9	134.815	<0.001*	0.937
		Time	19,171	6.684	<0.001*	0.426
		Training*Time	19,171	0.755	0.757	0.077

Training Regime Comparison [continued]	Dependent Variable Measure	Independent Variables	df	F	p	Effect size (η^2)
Multimodal Background Recovery Vs Unimodal Background Recovery	Accuracy	Training	1,9	59.296	<0.001*	0.868
		Time	19,171	19.959	<0.001*	0.689
		Training*Time	19,171	20.681	<0.001*	0.697
	Representational Economy	Training	1,9	44.943	<0.001*	0.833
		Time	19,171	1.458	0.107	0.139
		Training*Time	19,171	2.275	0.003	0.202
	Network Error	Training	1,9	175.416	<0.001*	0.951
		Time	19,171	3.953	<0.001*	0.305
		Training*Time	19,171	1.169	0.289	0.115

Rehabilitation with a therapy set size of 20 and background spontaneous recovery after a severe lesion

After severe lesioning with a therapy set size of 20 there is only a small improvement in terms of accuracy but not network error, or representational economy (as illustrated in Figure 4.10) . This improvement is much less than that observed for a therapy set size of 5 or 10. Multimodal training achieves higher accuracy, representational economy and lower network error than unimodal training on the therapy set and these differences were confirmed as statistically significant by the results of the ANOVA in Table 4.7. The control set exhibits no improvement in performance. The equated error comparison shows statistically significant differences in accuracy and network error but not representational economy (see Table 4.7) between multimodal and unimodal training on the therapy set. The level of background recovery is only slightly greater for multimodal training and shows statistically significant difference in network error and representational economy (see Table 4.7) compared to unimodal training.

In Table 4.7 the advantage of multimodal therapy over unimodal therapy can be seen in the statistically significant difference between the two in accuracy, representational economy and network error. There is also a significant difference between equated error multimodal therapy and unimodal therapy in training on network error and a significant effect of time and a significant interaction between training and time for representational economy, accuracy and network error. There is a significant difference between multimodal and unimodal controls for representational economy and network error. Finally there is a significant difference between multimodal and unimodal background recovery in training for representational economy and network error as well as a significant effect of time and a significant interaction between training and time in accuracy. Representational Economy was strongly correlated with learning Accuracy for all conditions: Multimodal Therapy $r(20)=.963$, $p<.001$; Equated Error Multimodal Therapy $r(20)=.901$, $p<.001$; Unimodal Therapy $r(20)=.884$,$p<.001$: Multimodal Control $r(20)=.509$,$p=.019$:Unimodal Control $r(20)=.603$,$p=.001$; Multimodal General $r(20)=.736$,$p<.001$:Unimodal General $(20)=-.028$,$p=.890$. Representational Economy was strongly correlated with Network Error for all conditions: Multimodal Therapy $r(20)=-.987$, $p<.001$; Equated Error Multimodal Therapy $r(20)=-.985$, $p<.001$; Unimodal Therapy $r(20)=-.988$,$p<.001$: Multimodal Control $r(20)=-.923$,$p<.001$:Unimodal Control $r(20)=-.955$,$p<.001$; Multimodal General $r(20)=-.466$,$p=.021$:Unimodal General $(20)=-.827$,$p<.001$.

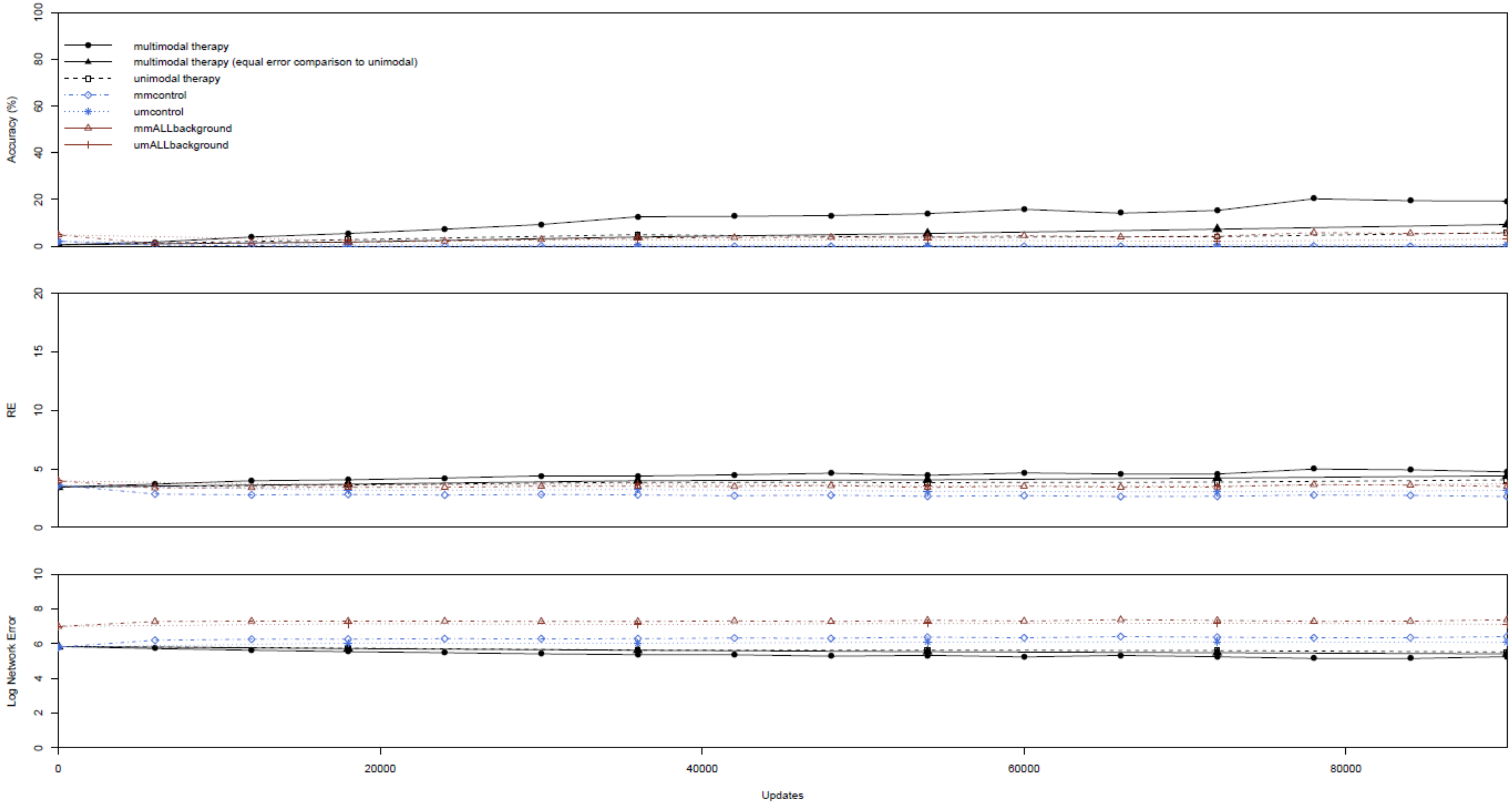

Figure 4.10 Mean variation in training performance for a therapy set of 20 with background recovery during rehabilitation after a severe lesion

Table 4.7 2x20 ANOVAs indicating significant performance differences between training regimes for a therapy set of 20 with background recovery during rehabilitation after a severe lesion

Training Regime Comparison	Dependent Variable Measure	Independent Variables	df	F	p	Effect size (η^2)
Multimodal Therapy Vs Unimodal Therapy	Accuracy	Training	1,9	33.605	<0.001*	0.789
		Time	19,171	21.245	<0.001*	0.702
		Training*Time	19,171	10.138	<0.001*	0.530
	Representational Economy	Training	1,9	375.967	<0.001*	0.977
		Time	19,171	21.420	<0.001*	0.704
		Training*Time	19,171	3.031	<0.001*	0.252
	Network Error	Training	1,9	719.723	<0.001*	0.988
		Time	19,171	35.894	<0.001*	0.800
		Training*Time	19,171	3.341	<0.001*	0.271
Equated Error Multimodal Therapy Vs Unimodal Therapy	Accuracy	Training	1,9	6.781	0.029*	0.430
		Time	19,171	11.158	<0.001*	0.554
		Training*Time	19,171	2.311	0.002*	0.204
	Representational Economy	Training	1,9	0.062	0.809	0.007
		Time	19,171	23.854	<0.001*	0.726
		Training*Time	19,171	3.678	<0.001*	0.290
	Network Error	Training	1,9	6.538	0.031*	0.421
		Time	19,171	39.059	<0.001*	0.813
		Training*Time	19,171	6.277	<0.001*	0.411
Multimodal Control Vs Unimodal Control	Accuracy	Training	1,9	2.975	0.119	0.248
		Time	19,171	1.204	0.259	0.118
		Training*Time	19,171	0.287	0.999	0.031
	Representational Economy	Training	1,9	866.123	<0.001*	0.990
		Time	19,171	4.381	<0.001*	0.327
		Training*Time	19,171	1.216	0.250	0.119
	Network Error	Training	1,9	461.251	<0.001*	0.981
		Time	19,171	4.790	<0.001*	0.347
		Training*Time	19,171	1.132	0.324	0.112

Training Regime Comparison [continued]	Dependent Variable Measure	Independent Variables	df	F	p	Effect size (η^2)
Multimodal Background Recovery Vs Unimodal Background Recovery	Accuracy	Training	1,9	3.012	0.117	0.251
		Time	19,171	8.348	<0.001*	0.481
		Training*Time	19,171	10.946	<0.001*	0.549
	Representational Economy	Training	1,9	149.186	<0.001*	0.943
		Time	19,171	1.625	0.055*	0.153
		Training*Time	19,171	1.727	0.036*	0.161
	Network Error	Training	1,9	787.446	<0.001*	0.989
		Time	19,171	0.963	0.507	0.097
		Training*Time	19,171	0.836	0.661	0.085

Simulation 4.3:

Manipulating the learning environment, therapy set size and lesion severity in rehabilitation from stable baseline recovery that incorporates background spontaneous recovery.

This simulation extends simulations 4.1 and 4.2 by examining manipulating the learning environment in rehabilitation but examing the variation in performance at different severity levels when different numbers of items (here described as therapy set size) are used for rehabilitation whilst maintaining a background level of recovery to occur in parallel with learning. Having firmly established a multimodal advantage in all previous simulations documented in this thesis, this final simulation explores the effect of increasing therapy set size on learning performance at different levels of severity.

Method

In order to look for the possibility of severity by set size interactions the model was trained in the same way as Simulation 4.2 but this time from the stable baseline recovery point it achieved after a moderate lesion (i.e. 92% of connections removed) and only with therapy set sizes of 5 and 10 items. The results were combined with data from the model trained in Simulation 4.2 for the severe damage condition (i.e. 93% of connections removed) for therapy set sizes of 5 and 10 item. This gave a 2x2 design with which to explore the effect of manipulating therapy set size and damage severity (each explored at two levels) for both the multimodal and unimodal learning environment training regimes. All other parameters and processes in this simulation are identical to those used in simulation 4.2.

Results

As expected the difference in rehabilitation performance between therapy sets of 5 or 10 items are more pronounced for the severe lesion explored in Simulation 4.2 than for the moderate lesion explored in the current simulation.

Other qualities observed in Simulation 4.2 are replicated here. Specifically, relearning performance of background recovery and control sets in both multimodal and unimodal conditions is virtually identical. Multimodal learning achieves better rehabilitation performance in all conditions. This multimodal advantage is more pronounced when therapy sets contain smaller numbers of items. The equated error multimodal condition tends to outperform the unimodal condition for smaller therapy set sizes and more severe lesions.

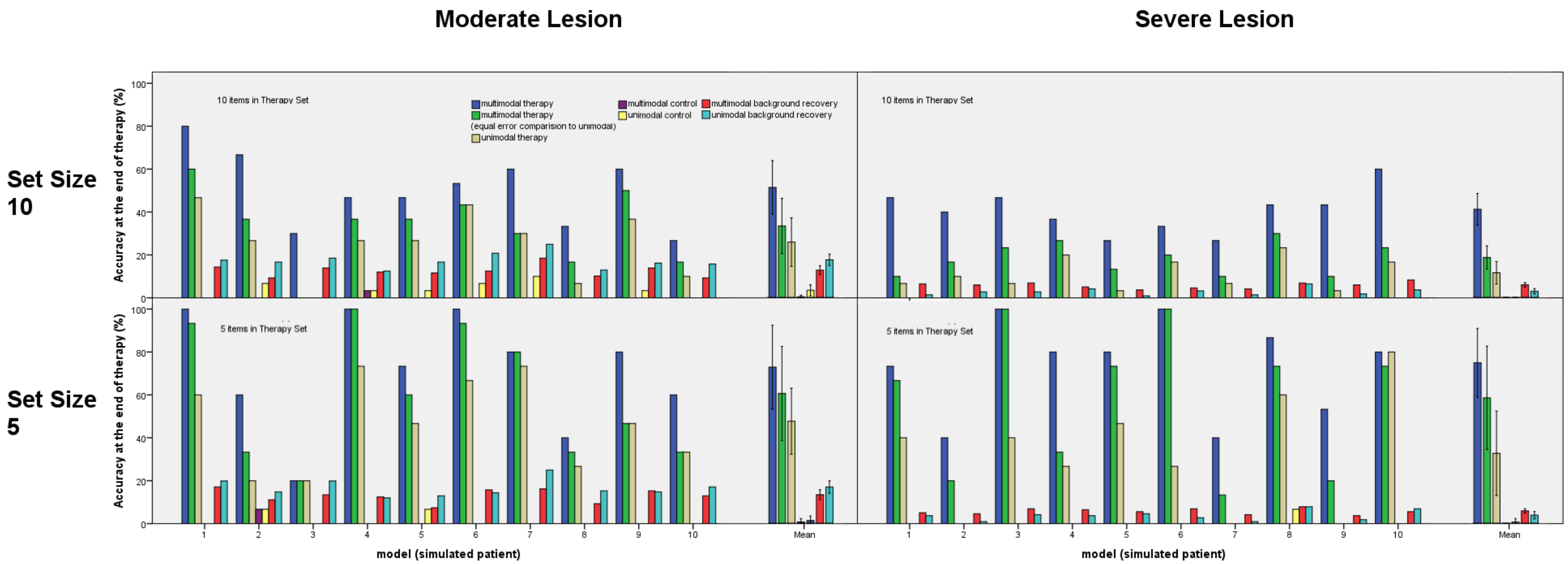

Figure 4.11 Mean variation in rehabilitation performance contrasting therapy sets of 5 or 10 items with background recovery after moderate and severe lesions

Rehabilitation with a therapy set size of 5 and background spontaneous recovery after a moderate lesion

After moderate lesioning with a therapy set size of 5 there is a large improvement in terms of accuracy and smaller improvement for network error, and in convergence amongst semantic representations measured as representational economy (as illustrated in Figure 4.12) . This improvement is similar to that observed for a therapy set size of 5 with a severe lesion. Multimodal training achieves higher accuracy **,** representational economy and lower network error than unimodal training on the therapy set and this difference is confirmed as statistically significant by the results of the ANOVAs in Table 4.8. The control set makes no improvement in performance. The equated error comparison shows statistically significant differences in accuracy, network error and representational economy (see Table 4.8) between multimodal and unimodal training on the therapy set. The level of background recovery is now slightly greater for unimodal training, unlike after a severe lesion and shows statistically significant difference in accuracy, network error and representational economy (see Table 4.8) compared to multimodal training.

Table 4.8 presents results showing a statistically significant difference between multimodal and unimodal therapy for accuracy, representational economy and network error except for non-significant interactions between training and time for accuracy and network error. The equated error multimodal therapy also performed better than unimodal therapy for accuracy, representational economy and network error. The multimodal control also performed better showing a significant effect of training and a significant interaction between training and time for both representational economy and network error. Mu;timodal background recovery also outperformed unimodal background recovery as illustrated in the significant differences between accuracy, representational economy and network error. Representational Economy was strongly correlated with learning Accuracy for all conditions: Multimodal Therapy $r(20)=.961$, $p<.001$; Equated Error Multimodal Therapy $r(20)=.980$, $p<.001$; Unimodal Therapy $r(20)=.977$,$p<.001$: Multimodal Control $r(20)=-.272$,$p=.245$:Unimodal Control $r(20)=.456$,$p=.043$; Multimodal General $r(20)=.891$,$p<.001$:Unimodal General $(20)=.835$,$p<.001$.Representational Economy was strongly correlated with Network Error for all conditions: Multimodal Therapy $r(20)=-.963$, $p<.001$; Equated Error Multimodal Therapy $r(20)=-.965$, $p<.001$; Unimodal Therapy $r(20)=-.971$,$p<.001$: Multimodal Control $r(20)=-.904$,$p<.001$:Unimodal Control $r(20)=-.796$,$p<.001$; Multimodal General $r(20)=-.960$,$p<.001$:Unimodal General $(20)=-.964$,$p<.001$.

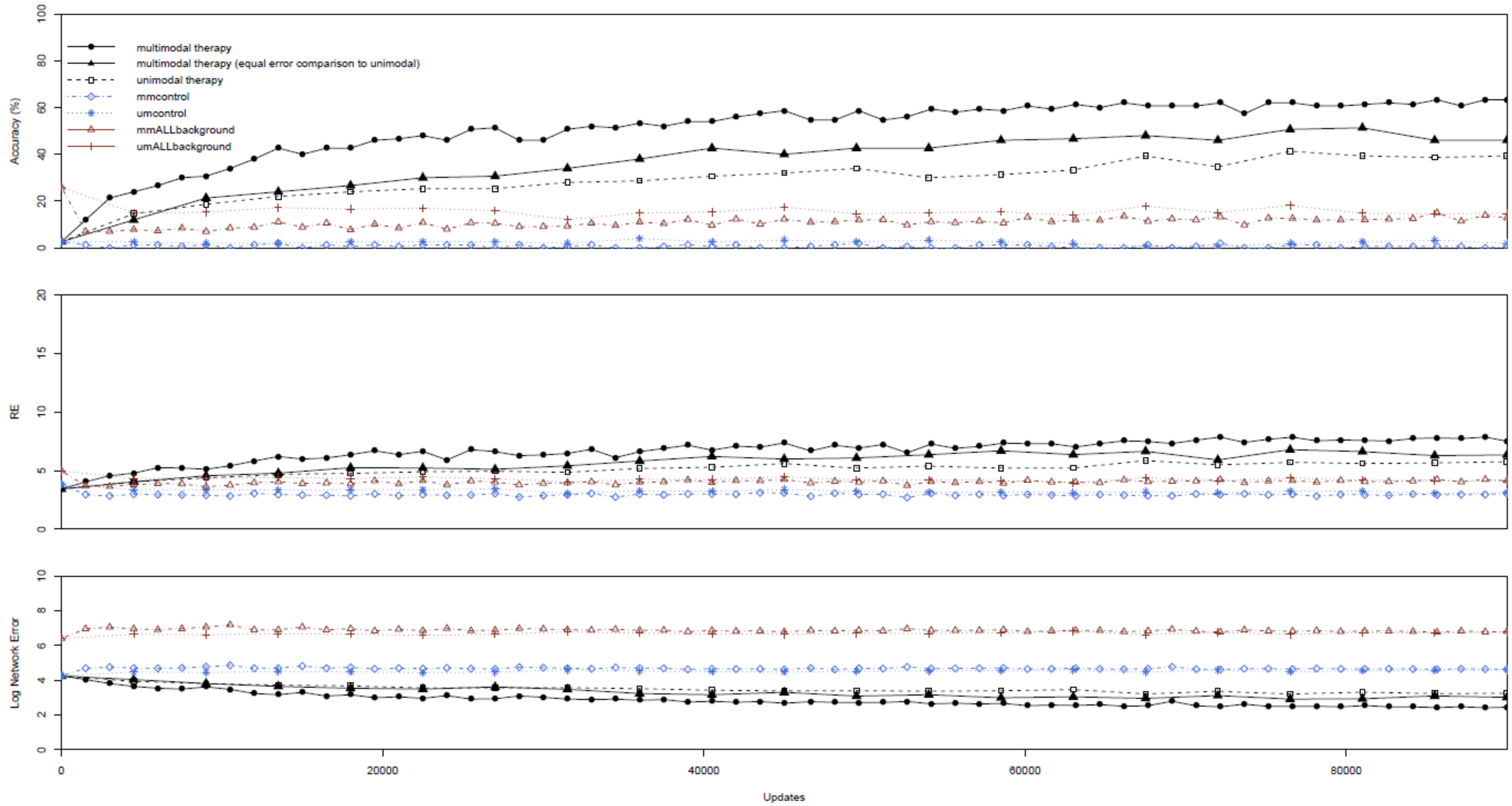

Figure 4.12 Mean variation in training performanc for a therapy set of 5 with background recovery during rehabilitation after a moderate lesion

Table 4.8 2x20 ANOVAs indicating significant performance differences during rehabilitation between training regimes for a therapy set of 5 with background recovery after a moderate lesion

Training Regime Comparison	Dependent Variable Measure	Independent Variables	df	F	p	Effect size (η^2)
Multimodal Therapy Vs Unimodal Therapy	Accuracy	Training	1,9	60.923	<0.001*	0.871
		Time	19,171	31.629	<0.001*	0.778
		Training*Time	19,171	0.953	0.520	0.096
	Representational Economy	Training	1,9	153.346	<0.001*	0.945
		Time	19,171	33.290	<0.001*	0.787
		Training*Time	19,171	2.251	0.003	0.200
	Network Error	Training	1,9	157.523	<0.001*	0.946
		Time	19,171	95.854	<0.001*	0.914
		Training*Time	19,171	1.493	0.093	0.142
Equated Error Multimodal Therapy Vs Unimodal Therapy	Accuracy	Training	1,9	24.889	0.001*	0.734
		Time	19,171	17.176	<0.001*	0.656
		Training*Time	19,171	1.911	0.016*	0.175
	Representational Economy	Training	1,9	91.612	<0.001*	0.911
		Time	19,171	38.972	<0.001*	0.812
		Training*Time	19,171	3.380	<0.001*	0.273
	Network Error	Training	1,9	100.340	<0.001*	0.918
		Time	19,171	74.803	<0.001*	0.983
		Training*Time	19,171	6.096	<0.001*	0.404
Multimodal Control Vs Unimodal Control	Accuracy	Training	1,9	4.047	0.075	0.310
		Time	19,171	1.407	0.129	0.135
		Training*Time	19,171	1.386	0.139	0.133
	Representational Economy	Training	1,9	41.422	<0.001*	0.822
		Time	19,171	0.814	0.688	0.083
		Training*Time	19,171	1.855	0.021*	0.171
	Network Error	Training	1,9	139.141	<0.001*	0.939
		Time	19,171	0.980	0.487	0.098
		Training*Time	19,171	1.802	0.026*	0.167

Training Regime Comparison [continued]	Dependent Variable Measure	Independent Variables	df	F	p	Effect size (η^2)
Multimodal Background Recovery Vs Unimodal Background Recovery	Accuracy	Training	1,9	57.254	<0.001*	0.864
		Time	19,171	4.793	<0.001*	0.348
		Training*Time	19,171	5.551	<0.001*	0.382
	Representational Economy	Training	1,9	44.157	<0.001*	0.831
		Time	19,171	2.700	<0.001*	0.231
		Training*Time	19,171	4.137	<0.001*	0.315
	Network Error	Training	1,9	138.771	<0.001*	0.939
		Time	19,171	1.690	0.042*	0.158
		Training*Time	19,171	3.218	<0.001*	0.263

Rehabilitation with a therapy set size of 10 and background spontaneous recovery after a moderate lesion

After moderate lesioning with a therapy set size of 10 there is a medium improvement in terms of accuracy and smaller improvement for network error, and in convergence amongst semantic representations measured as representational economy (as illustrated in Figure 4.13) . This improvement is similar to that observed for a therapy set size of 5 with a severe lesion. Multimodal training achieves higher accuracy , representational economy and lower network error than unimodal training on the therapy set and this difference is confirmed as statistically significant by the results of the ANOVAs in Table 4.9. The control set makes no improvement in performance. The equated error comparison shows statistically significant difference in accuracy, and representational economy but not network error (see Table 4.9) between multimodal and unimodal training on the therapy set. The level of background recovery is now slightly greater for unimodal training, unlike after a severe lesion and shows statistically significant difference in accuracy, network error and representational economy (see Table 4.9) compared to multimodal training.

Table 4.9 shows the multimodal advantage as a statistically significant difference between multimodal and unimodal therapy in accuracy, representational economy and network error as well as between equated error multimodal and unimodal therapy in accuracy, representational economy and network error with the exception of no significant effect of training on network error. The multimodal control performs better than the unimodal with a significant effect of training and time in accuracy, representational economy and network error. Similarly the multimodal background recovery outperforms the unimodal in accuracy, representational economy and network error. Representational Economy was strongly correlated with learning Accuracy for all conditions: Multimodal Therapy r(20)=.966, p<.001; Equated Error Multimodal Therapy r(20)=.946, p<.001; Unimodal Therapy r(20)=.949,p<.001: Multimodal Control r(20)=.104,p=.663:Unimodal Control r(20)=.483,p=.031; Multimodal General r(20)=.821,p<.001:Unimodal General (20)=.876,p<.001. Representational Economy was strongly correlated with Network Error for all conditions: Multimodal Therapy r(20)=-.984, p<.001; Equated Error Multimodal Therapy r(20)=-.956, p<.001; Unimodal Therapy r(20)=-.973,p<.001: Multimodal Control r(20)=-.919,p<.001:Unimodal Control r(20)=-.929,p<.001; Multimodal General r(20)=-.884,p<.001:Unimodal General (20)=-.957,p<.001.

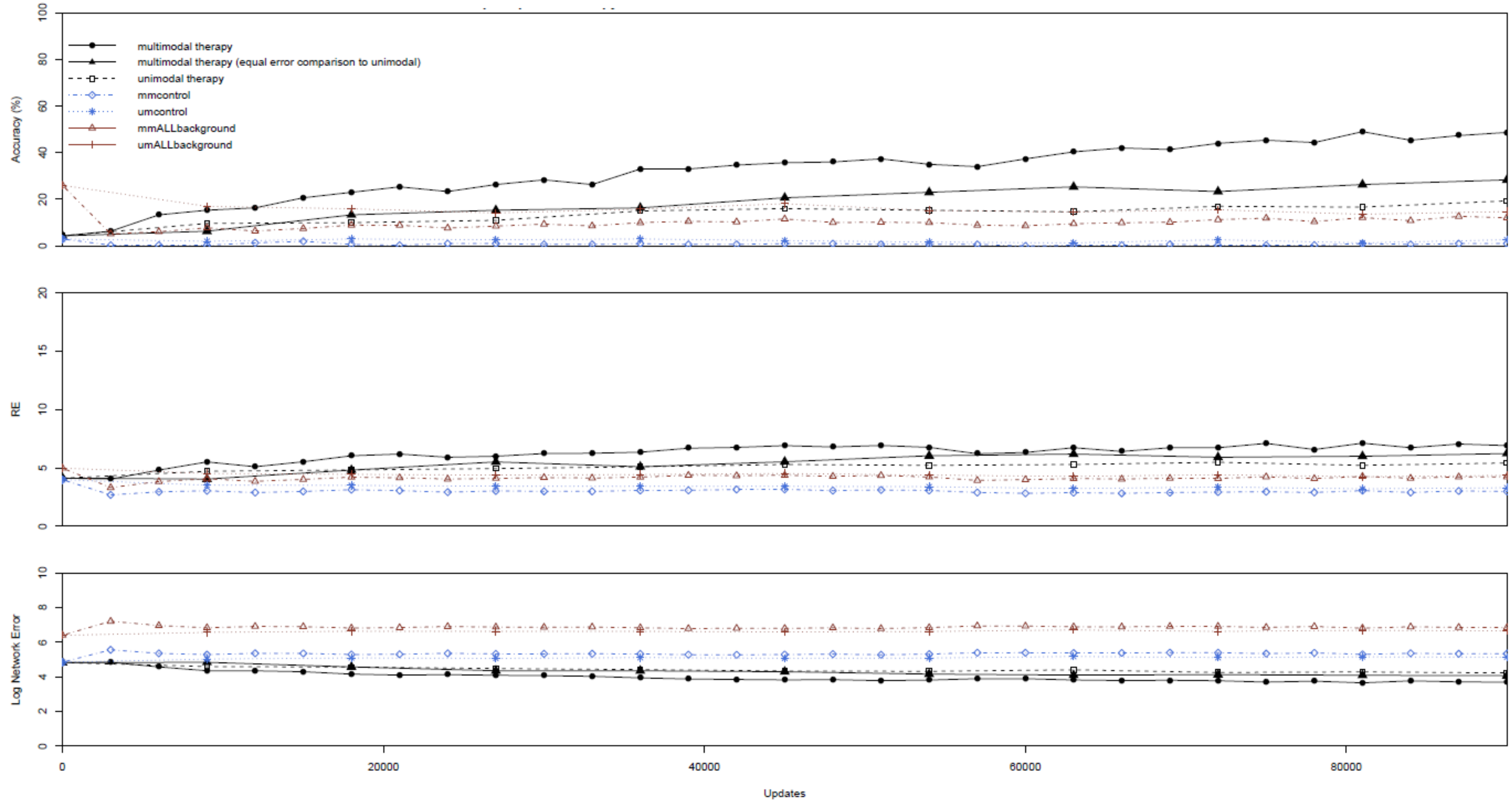

Figure 4.13 Mean variation in training performance for a therapy set of 10 with background recovery during rehabilitation after a moderate lesion

Table 4.9 2x20 ANOVAs indicating significant performance differences between training regimes for a therapy set of 10 with background recovery during rehabilitation after a moderate lesion

Training Regime Comparison	Dependent Variable Measure	Independent Variables	df	F	p	Effect size (η^2)
Multimodal Therapy Vs Unimodal Therapy	Accuracy	Training	1,9	102.348	<0.001*	0.919
		Time	19,171	29.392	<0.001*	0.766
		Training*Time	19,171	5.885	<0.001*	0.395
	Representational Economy	Training	1,9	713.347	<0.001*	0.988
		Time	19,171	23.475	<0.001*	0.723
		Training*Time	19,171	3.301	<0.001*	0.268
	Network Error	Training	1,9	520.454	<0.001*	0.983
		Time	19,171	63.352	<0.001*	0.876
		Training*Time	19,171	2.467	0.001*	0.215
Equated Error Multimodal Therapy Vs Unimodal Therapy	Accuracy	Training	1,9	23.354	0.001*	0.722
		Time	19,171	15.857	<0.001*	0.638
		Training*Time	19,171	4.511	<0.001*	0.334
	Representational Economy	Training	1,9	208.567	<0.001*	0.959
		Time	19,171	21.507	<0.001*	0.705
		Training*Time	19,171	3.974	<0.001*	0.306
	Network Error	Training	1,9	0.151	0.707	0.016
		Time	19,171	12.628	<0.001*	0.584
		Training*Time	19,171	2.719	<0.001*	0.232
Multimodal Control Vs Unimodal Control	Accuracy	Training	1,9	9.455	0.013*	0.512
		Time	19,171	2.816	<0.001*	0.238
		Training*Time	19,171	1.234	0.235	0.121
	Representational Economy	Training	1,9	171.143	<0.001*	0.950
		Time	19,171	2.121	0.006*	0.191
		Training*Time	19,171	1.431	0.118	0.137
	Network Error	Training	1,9	192.947	<0.001*	0.955
		Time	19,171	2.831	<0.001*	0.239
		Training*Time	19,171	1.466	0.103	0.140

Training Regime Comparison [continued]	Dependent Variable Measure	Independent Variables	df	F	p	Effect size (η^2)
Multimodal Background Recovery Vs Unimodal Background Recovery	Accuracy	Training	1,9	137.304	<0.001*	0.938
		Time	19,171	2.083	<0.001*	0.188
		Training*Time	19,171	4.913	<0.001*	0.353
	Representational Economy	Training	1,9	57.388	<0.001*	0.864
		Time	19,171	1.961	0.013*	0.179
		Training*Time	19,171	2.903	<0.001*	0.244
	Network Error	Training	1,9	164.700	<0.001*	0.948
		Time	19,171	1.946	0.014*	0.178
		Training*Time	19,171	1.868	0.019*	0.172

Discussion

Summary of Results

Simulation 4.1

As lesion severity increases greater improvements in therapy sets were observed. Multimodal training was always more efficient than unimodal training. The equated error comparison would again suggest multimodal efficiency is a function of its higher error signal (as observed in the simulations reported in Chapters 1,2 and 3). It is also worth noting that greater improvement in therapy sets were accompanied by higher measures of representational economy in all cases. There was no improvement in performance on control sets.

Simulation 4.2

The performance of background recovery learning in both multimodal and unimodal conditions is virtually identical, though there is a slight multimodal advantage. Multimodal learning achieves better rehabilitation performance in all conditions. This multimodal advantage is more pronounced when therapy sets contain smaller numbers of items. The equated error comparison would suggest that multimodal efficiency is no longer just a function of its higher error signal but actually has a greater effect in pushing the network's weights towards a correct solution for a given item. This effect is lost for larger numbers of items as shown by the bar chart comparison in Figure 4.7, and the identical results for the multimodal equated error comparison and unimodal training for a therapy set size of 20.

Simulation 4.3

Simulation 4.3 echoes some of the findings in Simulation 4.2. There is a difference in the relearning performance of background recovery learning, which is greater for unimodal training after a moderate lesion, and greater for multimodal training after a severe lesion. Aside from these differences, however, it was again found that multimodal learning achieves better rehabilitation performance in all conditions. The multimodal advantage was consistently more pronounced when therapy sets contained smaller numbers of items. The equated error comparison would suggest, that multimodal efficiency is no longer just a function of its higher error signal for relearning small subsets of the original training corpus when accompanied by background recovery. It is likely that for small numbers of items, where learning occurs both at the original background recovery rate and at the higher (double) rehabilitation rate, simultaneous multimodal presentation results in network weights being pushed so substantially in the optimal direction of a correct solution and very few therapy items to conflict with, even the equated error comparison outperforms the unimodal condition.

Analysis of simulated rehabilitation in the context of real therapy

The purpose of simulating rehabilitation was to attempt to gain a greater understanding of the cognitive mechanisms that support learning behaviour during rehabilitation. The advantage of considering learning within a simulation is the complete control that is possible over every aspect of the situation in which that learning takes place. Such degrees of control are impossible in clinical studies. Whilst therapists can administer therapeutic tasks in a near identical manner for a number of patients, it would be impossible for all patients to have brain damage of exactly identical severity and, perhaps more interestingly, for patients to be involved in relearning items of equal difficulty.

The model of semantic memory used in this study employs a prototype system (see methods section of Simulation 4.1) to generate learning items for the model. These items are of identical semantic difficulty, something which is unlike the real world. It is possible then that variations in item difficulty within therapy may play a role in the efficiency of that therapy, which may account for the findings within this study that suggest therapy set size affects therapy outcome, since patients studies (see Snell, Sage and Lambon Ralph, 2010) do not confirm this finding. The greater proportion gains for smaller therapy set sizes observed in the above simulations do not reflect existing research (see Figure 4.11 and Snell et al., 2010) where meta analysis and clinical studies were used to understand the effect of varying therapy set size. This study suggests that the advantages observed in proportion gain for small therapy set sizes are related to the degree of representational economy that develops during learning as a result of convergence processes. Snell, Sage and Lambon Ralph (2010) found no difference in performance improvement between using small and large therapy sets. However both their small and large sets contained a very small proportion of the total number of items humans are expected to be familiar with.

In contrast in Simulation 4.2 a therapy set size of 20 represents relearning 27.8% (20/72 items) of pre-damage knowledge.

This study set out predominantly with an interest in the effect of the learning environment (in terms of targeted modalities during therapy) on rehabilitation efficiency. All simulations showed a statistically significant advantage for multimodal learning regimes. Furthermore when multimodality was combined with small therapy set sizes in simulations 4.2 and 4.3, a multimodal advantage was determined even if the number of presentations of each item for learning was identical (i.e., the equated error comparison). This finding suggests that for small therapy set sizes, rehabilitation learning benefits much more from multimodality than for large therapy set sizes.

The results also show a difference in background recovery learning depending upon the degree of lesion severity: unimodal training appears to benefit background recovery more than multimodal training for less severe lesions (the moderate lesion case of simulation 4.3). This finding is consistent for different therapy set sizes, suggesting that unimodal therapy perhaps drives improvements in recovery more than any rehabilitation effect it may have.

Conclusion

Multimodal learning was observed to be more efficient than unimodal learning in rehabilitation. Representational economy appeared to be a guide to the efficiency advantages for multimodal learning . During rehabilitation learning, the observed differences between multimodal and unimodal learning seem to be more accentuated in the case of more severe lesions, and this is reflected in the representational economy measure, in other words driven by a need for convergent internal representations supporting learning. Simulation 4.1 established a basic principle of simulating rehabilitation and simulations 4.2 and 4.3 offer small refinements towards addressing real-world therapy issues. The idea that it is possible to explore manipulations of damage severity and therapy set size in the context of manipulating the learning environment, yields a powerful approach to thinking about therapy, and predicting patient behaviour. Any insights that can be gained through simulation generate hypotheses for the manner in which rehabilitation occurs as a result of considering the connectionist mechanisms supporting rehabilitation. Thinking about multimodality in therapy makes intuitive sense when it is contrasted against findings of multimodal learning in children (Quinlan, 2003) that are supported by modelling. The current study used a model of semantic memory that supports the idea that multimodal learning is more efficient due to simultaneous perceptions in multiple modalities offering a higher error signal as the brain attempts to develop internal representations that support and associate these perceptions. The findings from this study offer an appealing, and somewhat intuitive, case for adopting a greater degree of multimodality within rehabilitation therapy tasks and that multimodality is advantageous irrespective of therapy set size or damage severity, but can display greater advantage when considered in the context of these other factors.

General Discussion

This concluding chapter of the thesis focuses upon summarising and reviewing the findings of the simulations described in the preceding chapters in the context of the two themes which this thesis set out to explore. At this point it is perhaps worth reiterating that the goal of this thesis was to consider simulations of learning during different life stages and thus simulate rehabilitation in the context of exploring those issues likely to be major factors in the success of patient-related work. The starting point for thinking about simulating rehabilitation, and consequently learning during different life stages, was the role of the learning environment in terms of the sensory modalities targeted during therapy. The evidence, both from PDP modelling and from experimental studies on children's learning, that developmental learning is multimodal (as discussed in the introduction to Chapter 1) formed the starting point for thinking about simulating rehabilitation: Any PDP model that was to undergo rehabilitation learning must first necessarily undergo developmental learning with subsequent damage and recovery learning as a prelude to rehabilitation. The two themes that govern the explorations detailed in this thesis form an attempt at a logical structure with which to approach the problem of simulating rehabilitation whilst prioritising the inclusion of some facility to manipulate the learning environment and measure the impact of any manipulations upon learning efficiency.

Summarising the findings and their implications in terms of thesis theme 1

Theme 1 deals with the problem of attempting to understand how PDP models extract efficient and effective representations from the learning environment. As discussed in the introduction the starting point for both themes was a replication of the Rogers et al. (2004)

model of semantic memory. This model represented the most appropriate established model that possessed the necessary architecture and theoretical implications to explore rehabilitation. Chapters 1 and 2 tackle theme 1: Chapter 1 described the results of simulations that compared multimodal and unimodal learning using an original quantitative measure of efficiency (representational economy). These results indicated that the efficiency of representation development was dependent upon the structuring of the learning environment. Throughout all simulations in the thesis representational economy was strongly correlated with accuracy and network error whenever learning took place in the model suggesting it took be an accurate measure of convergence in the model as a function of its developing structure as a result of learning. The multimodal and unimodal learning episodes yielded similar levels of accuracy after developmental and recovery learning, however the time course of learning showed significant efficiency advantages for multimodal learning that were proportional to the degree of multimodality found in the learning environment. When learning was considered at the item level, multimodal learning was found to be more efficient in terms of speed (accelerated early development) and convergence (representational efficiency) in development and recovery learning. It was also noted that multimodal learning yielded representations that were more robust to damage. Chapter 2 took this finding of an advantage for multimodal learning as a starting point to explore whether it was possible to rearrange unimodal learning tasks in such a way that would approximate the efficiency advantages of multimodal learning tasks by targeting every modality sequentially for each item in turn. The findings from Chapter 2 indicate that this manipulation, while theoretically appealing, due to its potential consequences for patient-related work, does not in fact achieve the benefits of multimodal learning, nor does it display greater efficiency than traditional task-focused unimodal learning. Robustness to damage, speed of recovery and representational economy were also again examined as in Chapter 1, but there were no significant improvements on any of these measures. Chapter 2 thus cleared up the question of whether reorganised unimodal learning could approximate multimodal learning, suggesting that it is the simultaneity of multimodal item presentation that is important, as opposed to merely presentation order. The conclusion of Chapters 1 and 2 therefore is simply that the higher error signal available during simultaneous multimodal presentation of learning items is responsible for the multimodal advantage.

Chapter 3 was a technical report on scaling up the Rogers et al. (2004) model so it was capable of learning twice the number of items (i.e 72 instead of 36). This was done so that there would be enough items to model rehabilitation therapy. Although rescaling a model is a non-trivial task the activity itself has no implications for this conclusion beyond it having been accomplished successfully and suggesting future research could benefit from further scaling of this model or work with other larger models.

Summarising the findings and their implications in terms of thesis theme 2

Theme 2 deals with how PDP models can be used to understand recovery and rehabilitation for impairments in the context of patient-related work. Following Welbourne and Lambon Ralph (2005b) Chapter 4 addresses this by modelling rehabilitation as intense exposure to a small number of items from the learning environment. Simulation 4.1 showed that as lesion severity increases greater performance improvements in therapy sets were observed. Multimodal learning was always found to be more efficient than unimodal learning across all severities. Simulation 4.2 manipulated therapy set size and incorporated background recovery learning into the simulation of rehabilitation by alternating trials of recovery and rehabilitation. Incorporating recovery like this offered greater realism within the simulation since any patient will still be exposed to a background learning environment during their rehabilitation therapy sessions. The role of set size has been explored in clinical rehabilitation studies: Snell, Sage and Lambon Ralph (2010) recruited a case-series of aphasic patients with varying severity and gave all of them the same naming therapy which varied in two levels of therapy set size. They found that patients learnt the same proportion of items irrespective of set size. Contrary to this clinical finding, when expressed as a proportion of set size the model showed a decreasing effect of therapy as set size increased. The model basically re-learned a roughly fixed number of items irrespective of set size and so the proportion varies accordingly.

In contrasting the multimodal and unimodal learning conditions the performance of background recovery was virtually identical, though a slight multimodal advantage was evident. Multimodal learning achieved better rehabilitation performance in all conditions. This multimodal advantage was also more evident for therapy sets with smaller numbers of items. Simulation 4.3 extended Simulation 4.2 by adding a manipulation of lesion severity (at two levels - moderate and severe) to the manipulation of therapy set size with background recovery. A greater relearning performance of background recovery was found for unimodal training after a moderate lesion, and greater for multimodal training after a severe lesion. In general though and corresponding to findings from the other simulations in Chapter 4, multimodal learning achieves better rehabilitation performance in all conditions.

Review of the findings in the context of thesis theme 1

Theme 1 deals with the problem of attempting to understand how PDP models extract efficient and effective representations from the learning environment.
Most work on representations in PDP models has concentrated upon realism within input and output representations. One particularly notable example of this focus on improving input and output representations is Plaut et al.'s (1996) refinement of the representations in Seidenberg and McClelland's (1989) triangle model of reading, which improved model performance and generated better non-word reading. Similarly, Chang et al. (2012a) extended the model by moving the input representations from orthography to vision thus allowing the orthographic representations to emerge naturally. This resulted in improvements in accounting for a number of serial effects most notably the length effect in nonword reading. This thesis answers the question of the degree to which hidden layer internal representations are responsible for the efficiency with which conceptual knowledge is acquired, re-established after damage, and maintained over time.

PDP models have three distinct types of representation: Input and output representations attempt to capture salient features of the external world whose relations can be learnt by the model as the simulation runs. The third type of representations are those internal representations that emerge within the model's hidden layers as a consequence of learning, eventually developing into stable attractor states which support the model's emergent acquired knowledge of the relationship between input and output representations. Much modelling work concentrates upon the nature of input/output representations as they relate to, and capture through encoding, aspects of the external world which the model wishes to explore. Such research considers learned mappings and how they may account for the relationship between input and output representations in the context of how the brain may acquire and store such a relation within its neural network. The findings of this thesis contribute towards understanding the manner in which internal representations develop to support learning and the degree to which these representations are dependent upon the model's learning environment. Analysis of internal representations is rarely done in PDP modelling work, two examples of how such analyses can be employed are the work of Rogers et al. (2004) and Chang et al. (2012b). This thesis used the Rogers et al. (2004) model of semantic memory as a basis for the simulations conducted. Rogers et al.'s (2004) cluster analysis of internal representations gives an account of semantics where semantic representations generated for unrelated objects become increasingly differentiated, whilst the representations of the same item elicited from different domains (i.e. names, verbal descriptions and visual features) become more similar, thus semantics is accounted for as an emergent categorising structure that develops as the model learns. Chang et al. (2012b) developed a model of letter recognition that can deal with the problem of size and shape invariance. In order to try and understand the internal representations that support these properties they used a principal components analysis to check that the internal representations of different sizes and cases of the same letters their model learnt were similar, and thus that the model had learnt that variations of the same letter are related. The degree to which representations that PDP models extract from the learning environment are effective can be measured purely in terms of learning performance. So if a model has learnt to correctly map between all its input and output representations the model is clearly developing effective internal representations to support its learning. However efficiency is not so clear cut. Considering efficiency within internal representations means considering the performance measures of the model in terms of speed and accuracy of learning and what qualities of the internal representations may be responsible for influencing these factors. This thesis has focused on comparing the internal representations that arise when the model is trained unimodally (i.e. on single inputs to single outputs) or multimodally (i.e. on single inputs to multiple outputs). Learning efficiency is dependent upon the structure that develops amongst the internal representations in terms of the degree to which a single internal representation can support multiple input output mappings between different domains for the same item.

The thesis findings show consistent multimodal advantage in all simulations (see Chapters 1,2 and 4) of learning at different life stages. One question this thesis addresses is what constitutes multimodality in models and how that may translate to patient-related work. In theory multimodal learning can be the presentation of multimodal inputs that require single or multiple outputs or it can be the presentation of a single input that requires the production of multimodal outputs. This thesis takes the view that requiring the production of multimodal outputs is critical to producing efficient representations and is analogous to the manner in which developmental learning occurs. This view is supported in PDP modelling by Rogers et al. (2004) and developmental studies such as Gogate et al. (2006), as discussed in the Introduction and Chapter 1. This thesis was concerned with understanding learning during different life stages (i.e. developmental vs recovery/rehabilitation learning after brain damage) in order to question whether recovery/rehabilitation learning would benefit from the type of multimodal environment that drives developmental learning. Research in aphasia rehabilitation had suggested relearning after brain damage in adults may benefit from a multimodal learning environment (e.g. Howard et al, 1985), though this had focused on multimodal inputs. Focusing on the effect of multimodal outputs however seems potentially more beneficial specifically because developmental learning is very efficient and occurs with multimodal outputs. Chapter 1 trains the models using this multimodal outputs approach and contrasts it with the performance of the same model trained unimodally (i.e. single input to single outputs). Findings show multimodal learning to be more efficient than unimodal so Chapter 2 considers whether rearranging the unimodal training order so that all unimodal mappings for the same item are presented together before moving on to the next item. The findings of this thesis show that multimodality is only advantageous in efficiency terms when it consists of simultaneous presentation to multiple output modalities, in other words when the production of multimodal outputs are required. Since Chapter 2 shows no advantage for a re-arranged unimodal training regime, this thesis concluded that the higher error signal available when the model is trained using simultaneous multimodal outputs drives learning efficiency. This type of multimodal training is then explored in Chapter 4 in the context of rehabilitation learning where findings show the use of simultaneous multimodal outputs yields greater efficiency than unimodal outputs. The full implications of this are discussed in terms of patient-related work in the following section, in the context of thesis theme 2.

Representational economy is defined in Chapter 1 as the average ratio between the similarity of semantic representation for unrelated examples in the same modality and the similarity of semantic representations for the same example from different modalities. The findings of this thesis show representational economy to be a novel, effective measure of efficiency in the replication of the Rogers et al. (2004) model used in the simulations described in Chapters 1 and 2, as well as in the scaled up version of the model described in Chapter 3 and used for the rehabilitation simulations in Chapter 4. This means that representational economy is a statistic that measures efficiency in learning at different life stages (i.e. developmental, recovery and rehabilitation learning) that this thesis intended to investigate. One question that arises is how representational economy could be theoretically transferred for application to other types of neural network model. When considering transposing this measure to other types of model architectures it is necessary to solve the problem of identifying the area of the model where stable attractor states are established and then derive measures of the convergence that capture efficiency. In PDP models without the convergence zone architecture of the model used in this thesis there tends to be a single input modality, in this case efficiency would be measured in terms of the degree of convergence in the representation of related items generated by input in the same modality. Chang et al.'s (2012b) model of letter recognition is an example of using inputs of related items (size and shape variations in the letter form) in the same modality and analysing the similarity of internal hidden layer representations generated. In other words looking at similarity within internal representations of related input items. The measure of representational economy could be extended to neural network models possessing this type of structure: For example, Kohonen Self-Organizing Maps (Callan, 1999) have single input and single hidden layers, so representational economy would look at similarity of related items in the map generated at the hidden layer. Similarly Deep Belief networks (Hinton et al., 2006) have single input and multiple hidden layers. In this case representational economy could be measured as similarity at each hidden layer and it would be possible to look at variations in representational economy across layers to see where efficiency was most important during the model's training. Whilst such models are not generally used in considering rehabilitation they still represent possible accounts of other cognitive processes so could probably be applied to more abstract theoretical considerations of rehabilitation learning at the neuronal level.

Review of the findings in the context of thesis theme 2

Theme 2 deals with how PDP models can be used to understand recovery and rehabilitation for impairments in the context of patient-related work.

As discussed in the previous section the findings of this thesis show that multimodality is only advantageous in efficiency terms when it consists of simultaneous presentation to multiple output modalities - although exactly how this requirement for simultaneity translates to human learning has yet to be fully established. One of the goals of this thesis was to simulate rehabilitation and consider those factors in the learning environment that may be responsible for rehabilitation efficiency. The intention of such simulations was to evaluate the degree to which PDP models could be used to design cognitive rehabilitation therapy. The simulations in Chapter 4 consider the role that the learning environment, therapy set size and lesion severity play in shaping the outcome of therapy; illustrating how therapies could be designed and evaluated using models. However, it is clearly important that simulation results are grounded by comparison with results from therapy studies thus providing a clinical benchmark for accessing the simulations. Unfortunately, to the best of our knowledge, the use of multimodal outputs has not been explicitly explored in clinical populations. Instead therapy that is considered multimodal consists of multimodal inputs with a single output target. For example the use of multimodal cueing hierarchies (Nickels, 2002) that present semantic (picture) and phonological (verbal) cues to elicit word production, and there are many other similar examples (e.g. Wambaugh , 2003; Abel et al., 2005 or Wright et al., 2008). However, as discussed in the previous section developmental learning does often demand multimodal outputs (Gogate et al., 2000; Messer, 1978). The results of the simulations in Chapter 1 show that learning with multimodal outputs is more efficient that unimodal learning. Chapter 4 confirms that learning with multimodal outputs is more efficient in all learning scenarios at different life stages. This work therefore predicts that recovery and rehabilitation learning after brain damage will be most efficiently accomplished in a multimodal context that is similar to the context for developmental learning.

The simulations in Chapter 4 select therapy items by rank-ordering items by the highest error at stable baseline and then selecting the worst performing items for the therapy set. For therapy studies it is not possible to select items based on a continuous measure of error, but this method is approximated by selecting items that are consistently named incorrectly across multiple baselines. In clinical practice therapy items can be chosen in a variety of ways often depending on the nature of impairment and so this suggests future simulation work could usefully explore simulating different methods of selecting items for therapy. Renvall et al. (2013) provide an overview of the selection of aphasia therapy items that states the selection can be accomplished by considering the functional relevance of items selected. Renvall et al. (2013) offer the notion of two types of items; Personally chosen items and generally frequent items: Personally chosen items are words that a person with aphasia, their significant other and their therapist have identified as necessary for the patient to successfully communicate. These words will thus relate more directly to the type of activities the patient wishes to engage in in daily life. Generally frequent items are words that unimpaired adult speakers use often in their everyday conversations. These words are identified through objective counts of word frequency from large samples of spoken language. It is worth considering how these types of item selection could be re-created in simulation? The important thing is that personally chosen items have no requirement to be generally frequent. For example specific language relating to hobbies or professional interests. This thesis uses a PDP model that uses prototypes so emulating generally frequent items would require modifying prototypes so they could encode general frequency in their representation. This would make it possible to train on generally frequent items and personally chosen items could be implemented by random selection (i.e. without consideration of frequency) from the training corpus. The simulations in Chapter 4 use items identified by their continual failure to be named (i.e. high error value on naming) by the model at stable baseline. This method of item selection does not take into account functional relevance of items however this would need to be considered in future research

Although not one of the original questions, the results generated by the simulations in Chapter 4 allow for consideration of issues surrounding generalisation to untreated items during rehabilitation simulation. Plaut (1996) describes generalization in PDP models as weight changes during relearning that are in the right direction to re-establish the model's overall knowledge. Essentially this means that since the model's knowledge of all items is encoded in the same set of weights any training that pushes the weights in the right direction also offers the possibility that the weights will be able to handle other items it was initially trained on, but which are only present in the untreated control set for retraining during simulated rehabilitation. The simulations in Chapter 4 show no evidence of generalization to untreated items, this concurs with Best et al.'s (2013) clinical findings on generalisation and perhaps is accounted for by her observation that only patients with intact semantics showed evidence of generalization. Given that this thesis uses a model of semantic memory for its simulations, and that rehabilitation is explored in the context of increasingly severe damage to semantics it is not surprising that there is no real evidence of generalization. In a semantic system such as this model mappings are generally not systematic. Other systems like phonology will have much more systematic mappings and it may be this that determines whether or not generalisation is possible. Minor generalization can be seen amongst untreated items when including background spontaneous recovery (see Figure 4.12) in Chapter 4 simulation 4.3 for small therapy and control set of 5 after a moderate lesion. In this case it would seem likely that generalization is driven by background spontaneous recovery in conjunction with rehabilitation, but is not achieved with rehabilitation alone.

This thesis set out to simulate rehabilitation with a view to using the simulation to answer clinical questions about particular factors that affect therapy outcome. Therapy set size was chosen for the later simulation in Chapter 4 as there are some clinical findings regarding the effect of set size that the simulation outcome could be compared to in order to examine whether the simulation captured the real-world consequences of varying therapy set size. Two explanations seem possible for the discrepancy between clinical findings on therapy set size (i.e. patients learn the same proportion of items regardless of set size - Snell, Sage and Lambon Ralph, 2010) and the findings described in the simulations in Chapter 4 (i.e. the model learned the same number of items regardless of set size so the proportion learnt varies with size). In the model used in this thesis very few items are learnt compared to the number of items in average human vocabulary. So the number of therapy items used in simulation is a much higher proportion of the total number of items contained in the model's knowledge than in clinical therapy, essentially in the model the capacity limitations are more acute than in the patients. The second issue for clinical comparison is item difficulty. The model's use of prototypes means that item difficulty is balanced in a way that doesn't happen in the real world. For instance training could be accomplished using a mixture of known and unknown items which would bring out greater differences in a model with limited knowledge. Overall this thesis suggests a way of testing theories regarding various approaches to item presentation in therapy. Developing novel therapy theories requires precise formulation of clinical questions, and such questions can be usefully explored outside of clinical settings by employing computational frameworks that really learn and observing the consequences of certain therapy choices through simulating their outcome.

Future directions

The most important finding of this thesis is that simultaneous multimodal presentation of items for learning is the most efficient way to achieve learning. This finding makes an important prediction about patient-related work. All the simulations conducted testify to this finding however simultaneous multimodal presentation is likely to be hard to achieve in patient-related work and would require the development of new therapies based upon this principle. The clinical rehabilitation literature has called for more multimodality in therapy (e.g. Howard, 1985), but this has been limited to a focus on multimodal inputs generating single (i.e. unimodal) outputs as oppose to multiple (i.e. multimodal) outputs. It will be challenging to construct a learning environment that involves simultaneous multimodal inputs and outputs. It is likely that one possible source of inspiration for creating simultaneous multimodal learning environment may come from those observations of multimodal learning recorded in literature on developmental research such as the observations of Gogate et al. (2006). It is likely that structuring rehabilitation learning environments according to the environments in which young children learn multimodally may be one way to test the findings of this thesis in clinical research. Certainly the evidence from the rehabilitation simulations detailed in Chapter 4 suggests this would be beneficial as would some consideration of the interaction between lesion severity and the number of items used in therapy.

Future simulation work would necessarily benefit from larger models with vocabulary sizes that are more similar to those recorded in human knowledge. Such models would however require substantial computational resources and so pursuing such a line of enquiry would be hardware dependent to a certain degree. Further work with smaller models would likely still yield interesting results. For example it would be of interest to know how the nature of the mapping task changes the learning results. In the Rogers et al. (2004) model all the mappings are arbitrary, but in other domains there is considerably more systematicity in the mappings (e.g. reading). Similarly, models that incorporate the ability to simulate a wider range of rehabilitation tasks would also be of interest such as the inclusion of a motor domain to allow for simulating the role of gesture. Aphasia therapy sessions often contain gestural responses as clients attempt to communicate by all possible means. Simulating gesture in the context of the above conclusions regarding simultaneous multimodality represents one interesting next step that could arise from the work described in this thesis.

Conclusion

The thesis sought to answer three questions concerning learning occurring in different life stages: development; spontaneous recovery and rehabilitation. Firstly, is multimodal learning more efficient than unimodal learning? This thesis finds multimodal learning more efficient in all learning scenarios across different life stages. Secondly, can multimodality be approximated as sequential presentations of the same item in multiple modalities or must it be simultaneous presentations in multiple modalities? This thesis finds that the benefit of multimodality in the model is dependent upon simultaneous input in multiple modalities to show any advantage over unimodal learning. Finally, for rehabilitation learning what effect does varying therapy set size have on learning efficiency in the context of the efficiency of the learning environment and the severity of damage? This thesis found that the proportion of items learnt varies with therapy set size. The same total number of items were learnt regardless of variation in set size. Overall the findings suggest a conclusion that rehabilitation therapy can be designed in terms of measurable gains in efficiency using increasingly realistic PDP models, and that the greatest challenge is converting the theoretical implications of model results to genuine advances in clinical practice.

References

Abel, S. , Schultz, A. , Radermacher, I. , Willmes, K. & Huber, W. (2005) Decreasing and increasing cues in naming therapy for aphasia, *Aphasiology*, 19:9, 831-848,

Abel, S., Willmes, K. and Huber, W. (2007). Model-oriented naming therapy: Testing predictions of a connectionist model, *Aphasiology*, 21(5), 411-447

Ashcraft, M. (1989). *Human Memory and Cognition*, New York: HarperCollinsPublishers.

Barsalou, L.W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, 22(4), 637–660

Besner, D., Twilley, L., McCann, R. S., & Seergobin, K. (1990). On the connection between connectionism and data: Are a few words necessary? *Psychological Review,* 97, 432-446.

Best, W, Grassly, J., Greenwood, A., Herbert, R., Hickin, J., Howard, D. (2013) Aphasia rehabilitation: Does generalisation from anomia therapy occur and is it predictable? A case series study. *Cortex*, 49 (9) 2345 - 2357.

Best, W. and Nickels, L. (2000). From theory to therapy in aphasia: Where are we now and where to next?, *Neuropsychological Rehabilitation*, 10(3), 231-247

Broca, P. (1861). Remarques sur le siège de la faculté du langage articulé. *Bulletins de la Société Anatomique,* 6, 330-357

Buonomano, D. V., & Merzenich, M. M. (1998). Cortical plasticity: From synapses to maps. *Annual Review of Neuroscience*, 21, 149-186.

Chang, Y., Furber, S & Welbourne, S. (2012a). "Serial" Effects in Parallel Models of Reading. *Cognitive Psychology*, 64(4), 267-291.

Chang, Y, Furber, S & Welbourne,S. ((2012b). Modelling normal and impaired letter recognition: Implications for understanding pure alexic reading. *Neuropsychologia*, 50(12), 2773-2788.

Conroy, P., Sage, K., & Lambon Ralph, M. A. (2009). Errorless and errorful therapy for verb and noun naming in aphasia. Aphasiology, 23(11), 1311-1337.

Callan, R. (1999). *The Essence of Neural Networks*, London: Prentice Hall Europe

Conroy, P., Sage, K.E., and Lambon Ralph, M.A. (2006). Towards theory-driven therapies for aphasic verb impairments: A review of current theory and practice. *Aphasiology*, 20 (12), 1159-1185

Cree, G.S., McRae, K., and McNorgan, C. (1999). An Attractor Model of Lexical Conceptual Processing: Simulating Semantic Priming, *Cognitive Science*, 23(3), 371-414

Damasio, A. R. (1989). The brain binds entities and events by multiregional activation from convergence zones. *Neural Computation, 1,* 123–132.

Damasio, H., Grabowski, T. J., Tranel, D., and Hichwa, R. D. (1996). A neural basis for lexical retrieval. *Nature,* 380, 499 –505.

Dell, G.S., Martin, N., and 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, 490-520

Dell, G.S., Schwartz, M.F., Martin, N., Saffran, E.M., and Gagnon, D.A. (1997). Lexical access in aphasic and nonaphasic speakers. *Psychological Review*, 104(4), 801-838

Devlin, J.T., Gonnerman, L.M., Andersen, E.S. and Seidenberg, M.S. (1998). Category-Specific Semantic Deficits in Focal and Widespread Brain Damage: A Computational Account, *Journal of Cognitive Neuroscience*, 10(1), 77-94

Dilkina, K., McClelland, J.L. and Plaut, D.C. (2008). A single-system account of semantic and lexical deficits in five semantic dementia patients, *Cognitive Neuropsychology*, 25(2), 136-164

Elio, R.E., and Reutener, D.B. (1970) Colour context as a factor in encoding and as an organization device for retrieval of word lists, *Journal of General Psychology,* 99, 223-32.

Ellis, A.W., and Young, A.W. (1988). *Human cognitive neuropsychology*. Hove, UK: Lawrence Erlbaum Associates Ltd.

Eggert, G. H. (1977). *Wernicke's works on aphasia: A sourcebook and review* (Vol. 1). The Hague, the Netherlands: Mouton.

Farah, M.J., and McClelland, J.L. (1991). A computational model of semantic memory impairment: Modality-specificity and emergent category-specificity. *Journal of*

Experimental Psychology: General, 120, 339-357

Foygel, D., and Dell, G.S. (2000). Models of impaired lexical access in speech production. *Journal of Memory and Language*, 43, 182-216

Gallese, V. and Lakoff, G. (2005). The brain's concepts: the role of the sensory-motor system in conceptual knowledge. *Cognitive Neuropsychology*, 22(3-4), 455–479

Garrard, P. and Carroll, E. (2006). Lost in semantic space: a multi-modal, non-verbal assessment of feature knowledge in semantic dementia, *Brain*, 129, 1152-1163

Garrard, P., Lambon Ralph, M. A., Hodges, J. R., & Patterson, K. (2001). Prototypicality, distinctiveness and intercorrelation: Analyses of the semantic attributes of living and nonliving concepts. *Journal of Cognitive Neuroscience*, 18, 125–174.

Geva, S., Jones, P.S., Crinion, J.T., Price, C.J., Baron, J.C., and Warburton, E.A. (2011) The neural correlates of inner speech defined by voxel-based lesion-symptom mapping, *Brain* , 134 3071-3082.

Gogate, L. J., Bahrick, L. E., & Watson, J. D. (2000). A study of multimodal motherese: The role of temporal synchrony between verbal labels and gestures. *Child Development*, 71, 876–892.

Gogate, L.J., Bolzani, L.H. and Betancourt, E.A. (2006) Attention to Maternal Multimodal Naming by 6- to 8-Month-Old Infants and Learning of Word–Object Relations. *Infancy* 9(3),259-288.

Gotts, S.J. and Plaut, D.C. (2002). The impact of synaptic depression following brain damage: A connectionist account of “access/refractory” and “degraded-store” semantic impairments, *Cognitive, Affective & Behavioural Neuroscience*, 2(3), 187-213

Harm, M. W., McCandliss, B. D., & Seidenberg, M. S. (2003). Modeling the successes and failures of interventions for disabled readers. *Scientific Studies of Reading*, *7*, 155-182.

Harm, M.W., and Seidenberg, M. S. (1999). Phonology, reading acquisition, and dyslexia: Insights from connectionist models. *Psychological Review,* 106(3), 491-528

Harm, M.W., and Seidenberg, M. S. (2004). Computing the meanings of words in reading: Cooperative division of labor between visual and phonological processes. *Psychological Review*, 111(3), 662-720.

Harris, M. (1998). Can Connectionism Model Developmental Change? *Mind & Language*, 13(4), 576-581

Hinton, G. E., Osindero, S., and Teh, Y. (2006) A fast learning algorithm for deep belief nets. *Neural Computation 18,*1527-1554.

Hinton, G. E., Plaut, D. C., and Shallice, T. (1993). Simulating brain damage. *Scientific American*, *269*, 76-82.

Hinton, G. E., and Sejnowski, T. J. (1986). Learning and relearning in Boltzmann machines. In D. Rumelhart, J. McClelland, and The PDPResearch Group (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition,* (Vol. 1: Foundations, pp. 282-317). Cambridge, MA: MITPress.

Hinton, G.E., and Shallice, T. (1991). Lesioning an attractor network: Investigations of acquired dyslexia. *Psychological Review*, 98(1). 74-95

Howard, D. (1994). The Treatment of Acquired Aphasia, *Philosophical Transactions of the Royal Society London Biological Sciences*, 29th October 1994 vol.346 no.1315 113-120

Howard, D. , and Hatfield, M.H. (1987). *Aphasia therapy*. Hove , UK: Lawrence Erlbaum Associates Ltd.

Howard, D., Patterson, K., Franklin, S., Orchard Lisle, V., and Morton, J. (1985). The facilitation of picture naming in aphasia. *Cognitive Neuropsychology*, 2(1), 49-80

Hunt, R.R. and McDaniel, M.A. (1993) 'The enigma of organization and distinctiveness', *Journal of Memory and Language*, vol.32, pp.421–45.

Joanisse, M.F. and Seidenberg, M.S.(1999). Impairments in verb morphology after brain injury: A connectionist model. *Proceedings of the National Academy of Sciences, USA*, 96(13), 7592-7597.

Kay, J., Lesser, R., and Coltheart, M. (1992). *Psycholinguistic Assessment of Language Processing in Aphasia (PALPA).* Hove, UK: Psychology Press.

Kiran, S., and Thompson, C. (2003). The Role of Semantic Complexity in Treatment of Naming Deficits: Training Semantic Categories in Fluent Aphasia by Controlling Exemplar Typicality. *Journal of Speech, Language, and Hearing Research, 46*(4), 773-787.

Lambon Ralph, M.A., and Howard, D. (2000). Gogi-aphasia or semantic dementia? Simulating and assessing poor verbal comprehension in a case of progressive fluent aphasia. *Cognitive Neuropsychology*, 17, 437-465

Lambon Ralph, M.A., McClelland, J.L., Patterson, K., Galton, C.J. and Hodges, J.R. (2001). No Right to Speak? The Relationship between Object Naming and Semantic Impairment: Neuropsychological Evidence and a Computational Model, *Journal of Cognitive Neuroscience*, 13(3), 341-356

Lambon Ralph M.A., Patterson, K. (2008). Generalization and Differentiation in Semantic Memory: Insights from Semantic Dementia. *Annals of the New York Academy of Sciences*, 1124,51-76.

Lambon Ralph M.A., Sage, K., Jones, R.W., and Mayberry, E.J. (2010). Coherent concepts are computed in the anterior temporal lobes. *Proceedings of the National Academy of*

Sciences of the United States of America, 107(6), 2717-2722.

Mandler, G. (1979) ‘Organization and repetition: organizational principles with special reference to rote learning’ in Nilsson, L-G. (ed.) *Perspectives on Memory Research*, Hillsdale, NJ, LEA.

Mareschal, D., Johnson, M.H. and Grayson, A. (2004) ‘Brain and cognitive development’, in Oates, J and Grayson, A., (eds) *Cognitive and Language Development in Children*, Oxford, Blackwell/The Open University

Mayall, K. and Humphreys, G. (1996). A connectionist model of alexia: Covert recognition and case mixing effects, *British Journal of Psychology*, 87, 355-402

McClelland, J.L., Rumelhart, D.E. and The PDP Research Group (Eds.) (1986). *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*. Cambridge, MA: MIT Press.

McClelland, J.L. , St. John, M. and Taraban, R. (1989). Sentence comprehension: A parallel distributed processing approach, *Language and Cognitive Processes*, 4(3), 287 - 335

McCloskey, M. and Cohen, N. (1989) Catastrophic Interference in connectionist networks: The sequential learning problem. In G. H. Bower (Ed.), *The Psychology of Learning and Motivation, 24, 109-164*

McLeod, P., Plunkett, K. and Rolls, E.T. (1998) *Introduction to Connectionist Modelling of Cognitive Processes*. Oxford, UK: Oxford University Press.

McLeod, P., Shallice, T. and Plaut, D.C. (2000). Attractor dynamics in word recognition: converging evidence from errors by normal subjects, dyslexic patients and a connectionist model, *Cognition*, 74, 91-113

Messer, D. (1978). The integration of mothers' referential speech with joint play. *Child Development*, 49, 781–787.

Meyer, K and Damasio, A. (2009) Convergence and divergence in a neural architecture for recognition and memory. *Trends in Neurosciences,* 32(7), 376-382.

Moll, M. and Miikkulainen, R. (1997). Convergence-Zone Episodic Memory: Analysis and Simulations. *Neural Networks*, 10(6), 1017-1036.

Monaghan, P., Shillcock, R. & McDonald, S. (2004) Hemispheric asymmetries in the split-fovea model of semantic processing. *Brain and Language*, 88 (2004) 339–354

Munro, P.W. (1986). State- Dependent Factors Influencing Neural Plasticity: A Partial Account of the Critical Period. In J.L. McClelland, D.E. Rumelhart, and The PDP Research Group (Eds.), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition* (Vol. 2, pp. 471-502). Cambridge, MA: MIT Press.

Nickels, L. (2002) Therapy for naming disorders: Revisiting, revising, and reviewing, *Aphasiology*, 16(10-11), 935-979,

O'Reilly, R.C. and Munakata, Y. (2000). *Computational Explorations in Cognitive Neuroscience: Understanding the Mind by Simulating the Brain*. Cambridge, MA: MIT Press.

Patterson, K., Nestor P.J., and Rogers T.T. (2007) Where do you know what you know? The representation of semantic knowledge in the human brain. *Nature Reviews Neuroscience*, 8(12), 976-987

Patterson, K., Seidenberg, M. S., and McClelland, J. L. (1989). Connections and disconnections: Acquired dyslexia in a computational model of reading processes. In R. G. M. Morris (Ed.), *Parallel distributed processing: Implications for psychology and neuroscience* (pp. 131-181). London: Oxford University Press.

Plaut, D.C. (1996). Relearning after damage in connectionist networks: Toward a theory of rehabilitation. *Brain and Language*, 52(1), 25-82.

Plaut, D.C. (1997). Structure and Function in the Lexical System: Insights from Distributed Models of Word Reading and Lexical Decision, *Language and Cognitive Processes,* 12(5), 765-806.

Plaut, D.C. (2002). Graded modality-specific specialization in semantics: A computational account of optic aphasia. *Cognitive Neuropsychology*, 19, 603-639.

Plaut, D.C., McClelland, J.L., Seidenberg, M.S., and Patterson, K. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review*, 103(1), 56-115.

Plaut, D.C. and Shallice, T. (1993). Deep dyslexia: A case study of connectionist neuropsychology. *Cognitive Neuropsychology*, 10(5), 377-500

Plunkett, K., Karmiloff-Smith, A., Bates, E., Elman, J.L. and Johnson M.H. (1997). Connectionism and Developmental Psychology. *Journal of Child Psychology and Psychiatry*, 38(1), 53-80.

Quinlan, P.T. (2003). *Connectionist models of development: Developmental process in real and artificial neural networks.* Brighton, UK: Psychology Press

Renvall, K. , Nickels, L. & Davidson, B. (2013) Functionally relevant items in the treatment of aphasia (part I): Challenges for current practice, *Aphasiology*, 27(6), 636-650

Rogers, T.T., Lambon Ralph, M.A., Garrard, P., Bozeat, S., McClelland, J.L., Hodges, J.R. and Patterson, K. (2004). Structure and deterioration of semantic memory: A neuropsychological and computational investigation, *Psychological Review*, 111(1), 205-235.

Rogers, T.T, and McClelland, J.L. (2004). *Semantic Cognition: a parallel distributed processing approach*. Cambridge, MA: MIT Press.

Rohde, D. (2000). LENS, the light efficient network simulator.

http://tedlab.mit.edu/~dr/Lens/

Rumelhart, D.E., and McClelland, J.L. (1986). On Learning the Past Tenses of English Verbs. In J.L. McClelland, D.E. Rumelhart, and The PDP Research Group (Eds.), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition* (Vol. 2, pp. 216-271). Cambridge, MA: MIT Press.

Seidenberg, M.S., and McClelland, J.L. (1989). A Distributed, Developmental Model of Word Recognition and Naming. *Psychological Review*, 96(4), 523-568.

Sejnowski, T.J. and Rosenberg, C.R. (1987). Parallel networks that learn to pronounce English Text, *Complex Systems* 1,145-168.

Smith, A.C., Monaghan, P., & Huettig, F. (2014). Literacy effects on language and vision: Emergent effects from an amodal shared resource (ASR) computational model. *Cognitive Psychology, 75,* 28-54.

Snell, C. , Sage, K and Lambon Ralph, M.A. (2010): How many words should we provide in anomia therapy? A meta-analysis and a case series study, Aphasiology, 24(9), 1064-1094

Snowden, J.S., Goulding, P.J., and Neary, D (1989). Semantic dementia: a form of circumscribed cerebral atrophy. *Behavioral Neurology*, 2, 167–82.

Squire, L.R. (1992) ‘Declarative and nondeclarative memory: multiple brain systems supporting learning and memory', *Journal of CognitiveNeuroscience*, vol.4, pp.232–43.

Sullivan, J., & Horowitz, F. D. (1983). Infant intermodal perception and maternal multimodal stimulation: Implications for language development. In C. Rovee-Collier & L. P. Lipsitt (Eds.), *Advances in infancy research* (Vol. 2, pp. 183–239). Norwood, NJ: Ablex.

Tesak, J. and Code, C. (2008). *Milestones in the History of Aphasia: Theories and protagonists*. Hove, UK: Psychology Press.

Tranel, D., Damasio, H., and Damasio, A. (1997). A neural basis for the retrieval of conceptual knowledge. *Neuropsychologia, 35,* 1319 –1327.

Tulving, E. (1962) 'Subjective organization in free recall of ''unrelated'' words', *Psychological Review*, vol.69, pp.344–54.

Wambaugh, J. (2003) A comparison of the relative effects of phonologic and semantic cueing treatments, *Aphasiology*, 17(5), 433-441,

Welbourne, S. R., and Lambon Ralph, M. A. (2005a). Exploring the impact of plasticity-related recovery after brain damage in a connectionist model of single-word reading. *Cognitive, Affective & Behavioral Neuroscience*, 5(1), 77-92.

Welbourne, S.R. and Lambon Ralph, M.A. (2005b). Using computational, parallel distributed processing networks to model rehabilitation in patients with acquired dyslexia: An initial investigation. *Aphasiology*, 19 (9), 789-806.

Welbourne, S.R. and Lambon Ralph, M.A. (2007). Using PDP Models to Simulate Phonological Dyslexia: The Key Role of Plasticity-Related Recovery. *Journal of Cognitive Neuroscience*, 19 (7), 1125-1139.

Wernicke, C. (1874). *Die aphasische Symptomen-Complex*. Breslau: Cohn &Weigart.

Wilshire, C.E. (2008). Cognitive neuropsychological approaches to word production in aphasia: Beyond boxes and arrows. *Aphasiology*, 22(10), 1019-1053

Wright, H.H., Marshall, R.C. , Wilson, K.B. & Page, J.L. (2008) Using a written cueing hierarchy to improve verbal naming in aphasia , *Aphasiology*, 22(5), 522-536,

Zevin, J.D., and Seidenberg, M.S. (2002). Age of acquisition effects in word reading and other tasks. *Journal of Memory and Language*, 47(1), 1-29