The Role of Unimodal and Transmodal Cortex in Perceptually-Coupled and Decoupled Semantic Cognition: Evidence from fMRI

Charlotte Elizabeth Murphy

Doctor of Philosophy

University of York

Psychology

September 2017

Abstract

Semantic retrieval extends beyond the here-and-now, to draw on abstract knowledge that has been extracted across multiple experiences; for instance, we can easily bring to mind what a dog looks and sounds like, even when a dog is not present in our environment. However, a clear understanding of the neural substrates that support patterns of semantic retrieval that are not immediately driven by stimuli in the environment is lacking. This thesis sought to investigate the neural basis of semantic retrieval within unimodal and heteromodal networks, whilst manipulating the availability of information in the environment. Much of the empirical work takes inspiration from modern accounts of transmodal regions (Lambon Ralph et al. 2017; Margulies et al. 2016), which suggest the anterior temporal lobe (ATL) and default mode network (DMN) support both abstraction and perceptual decoupling. The first empirical chapter examines whether words and experiences activate common neural substrates in sensory regions and where, within the ATLs, representations are transmodal. The second empirical chapter investigates how perceptually-decoupled forms of semantic retrieval in imagination are represented across unimodal and transmodal regions. The third empirical chapter interrogates whether transmodal regions respond in a similar manner to conceptually-guided and perceptually-decoupled cognition, and whether these two factors interact. The data suggests ventrolateral ATL processes both abstract modality-invariant semantic representations (Chapter 3) and decoupled semantic processing during imagination (Chapter 4). In addition, this thesis found comparable networks recruited for both conceptual processing and perceptually-decoupled retrieval corresponding to the broader DMN (Chapter 5). Further interrogation of these sites confirmed lateral MTG and bilateral angular gyrus were pivotal in the combination of conceptual retrieval from memory. Collectively, this data suggests that brain regions situated farthest from sensory input systems in both functional and connectivity space are required for the most abstract forms of cognition.

List of Contents

List of Tables

List of Figures

Acknowledgements

Firstly, I am thankful to my supervisors, Prof. Beth Jefferies and Dr. Shirley-Ann Rueschemeyer, who have provided me with the theoretical and methodological building blocks for me to grow intellectually. Thank you for always ensuring that I produced the highest possible standard of work, and for your unquestionable patience and financial resources, which ensured I had the training needed to tackle these ambitious projects.

I would also like to express my deepest gratitude to Dr. Jonny Smallwood, whose door has always been open and has provided many insightful discussions that have shaped both my work and research interests.

I would also like to take this opportunity to thank Prof. Tim Andrews for your support throughout my PhD, as chair of my thesis advisory panel – these meetings were always a pleasure (even if I did forgot to bring cake!).

I could not have conducted this research without the help of David Watson who patiently showed me the ropes when it came to machine learning and Theo Karapanagiotidis who provided the technical support and guidance for the resting-state analyses. My sincere thanks go out to both of them.

Finally, I would like to acknowledge all my family and friends who have supported me throughout my PhD. In particular I would like to thank my beautiful mother, who always sent words of encouragement when I needed them most , and my wonderful partner, Dean, who believed in me, even when I did not. Your encouragement and helpful advice have kept me on track.

Author's Declaration

I, Charlotte Murphy, declare that this thesis is a presentation of original work and I am the sole author, under the joint supervision of Prof. Beth Jefferies and Dr. Shirley-Ann Rueschemeyer. This work has not previously been presented for an award at this, or any other, University. All sources are acknowledged as References.

The empirical work presented in this thesis has been published or is currently under review in the following peer-reviewed journals:

Murphy, C., Rueschemeyer, S. A., Watson, D., Karapanagiotidis, T., Smallwood, J., & Jefferies, E. (2017). Fractionating the anterior temporal lobe: MVPA reveals differential responses to input and conceptual modality. NeuroImage, 147, 19-31.

Murphy, C., Rueschemeyer, S. A.,Smallwood, J., & Jefferies, E. (*in review*). Imagining sounds and images: Decoding the contribution of unimodal and transmodal brain regions to semantic retrieval in the absence of meaningful input

Murphy, C., Jefferies, E., Rueschemeyer, S. A., Sormaz, M., Wang, H. T., Margulies, D., & Smallwood, J. (*in review*). Isolated from input: Transmodal cortex in the default mode network supports perceptually-decoupled and conceptually-guided cognition. bioRxiv, 150466.

Results from multiple empirical chapters have been presented at the following conferences:

Murphy, C., Rueschemeyer, S. A., Watson, D., Karapanagiotidis, T., Smallwood, J., & Jefferies, E. (May, 2015). Fractionating the anterior temporal lobe: MVPA reveals differential responses to input and conceptual modality. *Poster presented at the meeting of the Concepts, Actions, Objects and Scenes Conference, Trento, Italy.*

Murphy, C., Rueschemeyer, S. A., Watson, D., Karapanagiotidis, T., Smallwood, J., & Jefferies, E. (September, 2015). Fractionating the anterior temporal lobe: MVPA reveals differential responses to input and conceptual modality. *Paper presented at the meeting of British Association for Cognitive Neuroscience, Colchester, UK.*

Chapter 1 - Introduction and Review of Literature

"Whilst part of what we perceive comes through our senses from the object before us, another part (and it may be the larger part) always comes out of our own head."

William James (1890)

1.1. Introduction

As noted by William James in the 19th century not all cognitive functions are driven by perceptual experience (William James, 1983; p. 102). One ubiquitous phenomenon that occupies over half of waking thought is the retrieval of memories unrelated to the here-and-now (Klinger & Cox, 1987; Killingsworth & Gilbert, 2010; Poerio et al. 2013), such as imagining lying on a sandy beach during the daily commute. Such experiences are made possible by conceptual knowledge [1] (i.e., knowledge of what a sandy beach looks likes, how the sand feels on your feet, the sound of waves crashing). Conceptual knowledge puts meaning to our world and shapes our interaction with it; therefore it is fundamental for nearly all human behaviour including language processing, communication, perception, judgement, action, reasoning, remembering the past and thinking about

Glossary

Unimodal* - regions that typically receive input from a single sensory modality (such as vision or audition). These include primary and adjacent visual, auditory, somatosensory, olfactory and gustatory cortex.

Heteromodal* – receives convergent inputs from unimodal areas in more than one modality. Heteromodal regions typically reflect an area outside of unimodal sensory regions (i.e., association cortex). These areas include prefrontal cortex, posterior parietal lobule, lateral temporal cortex and parahippocampal gyrus.

Transmodal* – reflects a heteromodal brain region that does not show specificity for any single modality of sensory input. ***Following this definition a transmodal region is always a heteromodal region, but a heteromodal region is not always a transmodal region.*** Such regions are considered the most abstracted from sensory experience, as these regions receive information predominantly from downstream parts of unimodal areas and from other heteromodal/transmodal regions. These regions include limbic and paralimbic structures and the default mode network.

Perceptual-decoupling – cognition that is independent of perception, such as imagining lying on a sandy beach during the daily commute.

* Definitions taken from Mesulam (1998)

[1] Throughout the thesis, the terms *conceptual* and *semantic* will be used interchangeably.

the future (Binder et al. 2009; Binney, Embelton, Jefferies, Parker & Lambon Ralph, 2010; Hart et al. 2007; Irish, 2016; Jefferies & Lambon Ralph, 2006; Lambon Ralph & Patterson, 2008).

Although, for a healthy brain, retrieval of conceptual knowledge - both in the presence and absence of external sensory cues - appears effortless, it is composed of several interactive components, such as (i) a system that stores transmodal *conceptual representations* by linking concepts together according to meaning, independent of modality information (e.g., Binney et al. 2010; Damasio, 2008; Lambon Ralph et al. 2017; Meteyard et al. 2012; Mion et al. 2010; Patterson, Nestor & Rogers, 2007; Tranel, Damasio & Damasio, 1997), and (ii) a system that allows us to disengage or *perceptually decouple* from the external world in order to support internally generated retrieval of memories (Schooler et al. 2011; Smallwood et al. 2013). However, historically, these two systems have been investigated in relative isolation of one another (for exceptions see Binder et al. 2009; Humphreys et al. 2015); with the former focusing on semantic networks encompassing core sensory regions, bilateral anterior temporal lobe (ATL), temporal-parietal regions and lateral frontal cortex (Lambon Ralph et al. 2017; Noonan et al. 2010; 2013; Binder et al. 2009), while the later concentrates on a distributed network, whose core regions include angular gyrus, posterior cingulate cortex and medial prefrontal cortex, collectively known as the default mode network (DMN) (Fox et al. 2005; Greicius et al. 2003). Therefore, a clear understanding of the neural substrates that support patterns of semantic retrieval that are not immediately driven by stimuli in the environment is lacking.

This opening chapter will firstly discuss brain regions thought to contribute to *conceptual representations* by reviewing neuropsychological, functional neuroimaging and neuroanatomical evidence. It will then discuss brain regions thought to contribute to *perceptually decoupled* states. Finally, it will discuss a component process hypothesis, that postulates the retrieval of memories depends on operations of potentially independent, but nevertheless interactive, components or networks (Moscovitch, 1992), while drawing comparisons from modern accounts of anatomical organisation (e.g., Margulies et al. 2016) to provide an explanatory

framework for the multifaceted function of these brain regions with regard to perceptually-decoupled semantic retrieval.

1.2. Conceptual Representations

Semantic memory enables us to effortlessly understand the meaning of items encountered via any of our senses. This aspect of memory is distinct from our episodic memory for personal experience, which is dependent on a particular time or place (Tulving, 1972). That is, not all of us went on holiday to France last summer (episodic memory), but the majority of us could identify that Paris is the capital of France (semantic memory). Semantic memory is abstracted from specific similar experiences, allowing us to generalize the information without reference to an actual experience (Binder & Desai, 2011). It is therefore fundamental for nearly all human behaviour including language processing, communication, perception, judgement, action, reasoning, remembering the past and thinking about the future (Binney, Embelton, Jefferies, Parker & Lambon Ralph, 2010; Hart et al. 2007; Irish, 2016; Jefferies & Lambon Ralph, 2006; Lambon Ralph & Patterson, 2008). As a result, the manner in which this knowledge is represented and organized has been of great interest to researchers (Binder, Desai, Graves & Conant, 2009; Bonner & Price, 2013; Cree & McRae, 2003; Damasio, 1989; Martin, 2007; Meteyard et al. 2012; Lambon Ralph et al. 2017; Patterson et al. 2007; Pulvermüller, 2013; Sitnikova, West, Kuperburg & Holcomb, 2006; Visser, Jefferies & Lambon Ralph, 2010). Before discussing the theory and evidence for transmodal semantic representations, we will first briefly review both historical and contemporary views of semantic memory.

1.2.1. Distributed Theories of Semantic Memory

Embodied accounts of semantic processing postulate that neural regions generally used for perception and action, are also recruited during semantic processing (Barsalou, 1999; 2008; Humphrey & Forde, 2001; Martin, 2007; Patterson, et al. 2007; Pulvermüller, 2005). These regions are considered modality-specific as they represent modality-specific attributes and are located in areas proximal to and

reciprocally linked to primary sensory and motor regions. For instance, knowledge about the sound a flute makes is modally auditory so would be represented in auditory cortex, whereas knowledge about what a flute looks like is modally visual so would be represented in visual cortex. It is therefore suggested that retrieving a concept will engage neural pathways that encode the items distinct colour, shape, sound, motor properties and so on (see Figure 1.1; Patterson et al. 2007). Indeed, a plethora of functional neuroimaging studies have provided compelling evidence that these neural pathways are, to some extent, shared with perception and action systems (Martin et al. 1995; Goldberg, Perfetti Charles & Schneider Walter, 2006; Rueschemeyer et al. 2014). For instance, processing tool concepts activates regions associated with non-biological motion and action execution (Chao & Martin, 2000). Similarly, processing food concepts activated regions previously implicated in representation of taste and food reward properties (Simmons, Martin & Barsalou, 2005). Particularly noteworthy are findings that unimodal regions are activated by words denoting sensory or motor properties; for instance, words denoting actions (e.g., kick) and manipulable objects (e.g., hammer) elicit activation in the brain's motor system (Hauk & Pulvermüller, 2004), words associated with specific smells (e.g., cinnamon) elicit activation in olfactory cortex (Gonzalez et al. 2006) and words associated with colour overlapped with (Simmons et al. 2007) or were processed in regions proximal to (Chao & Martin, 1999) the neural substrates used for actual colour perception.

The umbrella term "embodied" encompasses many theories that fall on a continuum of embodiment; from theories that suggest distributed-modality specific regions form the *entire* neural basis of semantic memory (coined distributed-only views by Patterson et al. 2007) to theories that suggest, in addition to distributed-modality specific regions, that higher-level convergence "zones" or "hubs" are required to assist semantic processing (for review see Meteyard, Cuadrao, Bahrami & Vigliocco, 2012; Caramazza et al. 1990; Damasio, 1989; Geschwind, 1965; Lambon Ralph, Sage, Jones & Mayberry, 2010; Martin, 2007; Patterson et al. 2007; Pulvermüller, 2013; Riddoch et al. 1988; Rogers et al. 2004). This review is not intended to be exhaustive and given the majority of modern theorists acknowledge the role of higher-level integrative brain regions in semantic processing this chapter will therefore focus on evidence for the existence of convergence zones.

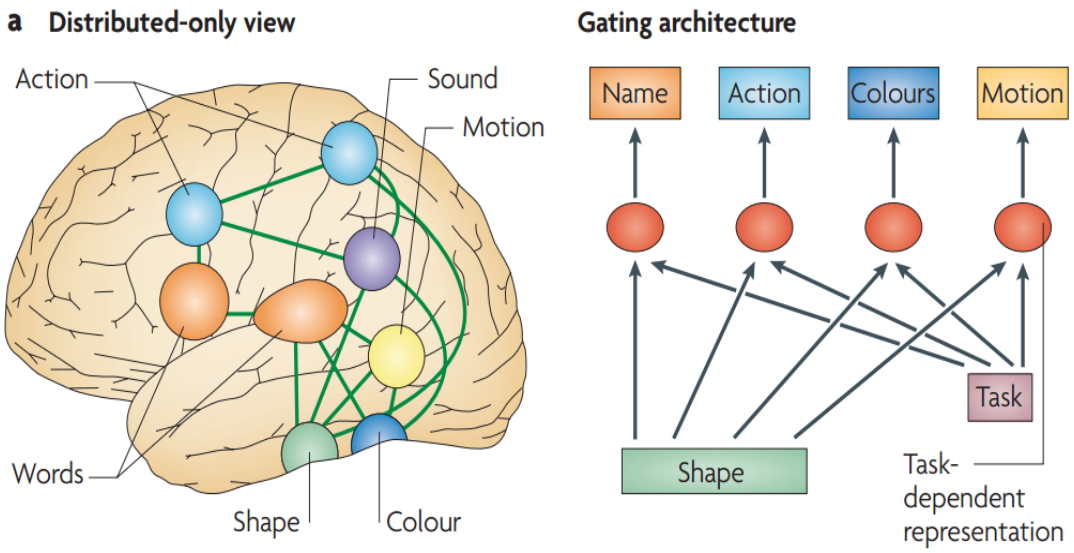

Figure 1.1. Distributed-only view of semantic processing suggests our entire semantic network is a product of these widely distributed regions, along with the connections between them (green lines; taken from Patterson, Nestor & Rogers, 2007).

1.2.2. The Role of Convergence Zones in Semantic Processing

A long standing argument, in favour of convergence zones, highlights that distributed-only views fail to fully explain how we are able to generalize across concepts that are semantically similar, despite having different sensory, motor and language attributes (Lambon Ralph et al. 2017; Patterson et al. 2007). Take the

comparison of a “flute” and “piano”; both have different shapes, sizes, names, descriptions and so on, but we can easily identify them as instruments because they have considerable conceptual overlap. Distributed-only views cannot explain our ability to abstract away from modality-specific properties and generalize across semantically similar concepts. Several theories have been reported that suggest this abstraction process is the function of a convergence zone, however these theories differ in their details. An early proposal suggested that a hierarchical set of convergence zones integrate modality-specific information across modalities (e.g., convergence zone theory, Damasio, 1989; Pulvermüller, 1999; Simmons & Barsalou, 2003; Martin, 2007). More recent integration proposals suggest that a single transmodal “hub” binds modality-specific information into a representation of a concept (Lambon Ralph et al. 2017; McClelland & Rogers, 2003; Patterson, et al. 2007; Rogers et al. 2004).

Although many recent theories agree about the need for a convergence zone or hub, there is still controversy regarding its neural basis (Binder & Desai, 2011; Gainotti, 2011; Martin, 2007; Visser et al. 2010). The regions that have been implicated as a transmodal convergence zone include the ATLs (Lambon Ralph et al. 2010; Patterson et al. 2007; Rogers & McClelland, 2004; Vandenberghe et al. 1996); posterior MTG (Fairhall & Caramazza, 2013); angular gyrus; and the lateral temporal cortex (for review see Bookheimer, 2002; Binder & Desai, 2011; Martin, 2007, Martin & Chao, 2001). Over the course of this chapter evidence from neuropsychological, neuroimaging and neuroanatomical fields will be reviewed to understand the specific contribution of each of these regions; firstly, due to a compelling body of evidence, this review will focus on the ATL as a candidate semantic hub with the later portion of this chapter discussing the evidence for a wider range of brain regions implicated in heteromodal processing.

1.2.2.1. Neuropsychology

Semantic Dementia. Neuropsychological studies of patients with semantic dementia (SD) have provided critical evidence that a heteromodal semantic hub is housed in ATL. SD is a progressive neurodegenerative disorder that results from bilateral atrophy and hypermetabolism of ATLs. As a result patients show

progressive deterioration of their conceptual knowledge, despite other aspects of cognition and memory being preserved (Butler et al. 2009; Hodges et al. 1992; Mummery et al. 2000; Patterson et al. 2007; Warrington, 1975; Snowden et al. 1989). The progressive nature of the disease is reflected in a specific-to-general decline in conceptual knowledge (Hodges, Graham & Patterson, 1995; Patterson et al. 2007); for example SD patients may not be able to identify specific-level information (e.g., camels have humps), however, as the disease progresses they may also lose general information (e.g., camels are animals). These deficits are evident in a wide range of semantic tasks across all modalities of testing suggests a pan-modal nature of the deficit (Bozeat, Lambon Ralph, Patterson, Garrard & Hodges, 2000; Coccia, Bartolini, Luzzi, Provincilali & Lambon Ralph, 2004).

Herpes Simplex Virus Encephalitis. Further evidence that a heteromodal hub is housed within the ATL comes from patients with herpes simplex virus encephalitis (HSVE). HSVE results in widespread damage to frontotemporal regions (usually bilaterally). HSVE patients with predominately frontal lobe damage often have no semantic deficits, but rather show executive deficits and poor control of memory retrieval. In contrast HSVE patients who show damage to bilateral ATL (overlapping with the damage in SD), have heteromodal semantic deficits on both production and comprehension tests; for instance when asked to select which of two photos (dog or goat) is the dog, they are significantly worse than healthy controls (Warrington & Shallice, 1984; Lambon Ralph, Lowe & Rogers, 2007; Noppeney et al. 2007). This deficit is specific to living categories as opposed to non-living categories; arguably due to the high number of shared features living concepts have in common. For example, 'dog' and 'goat' have high featural overlap (e.g., 4 legs, fur, mammals), whereas two vehicle concepts (such as car and aeroplane) share few featural properties (different shapes, sounds, colours, movements etc). Concepts that share more featural overlap are argued to place greater demands on semantic processing during both identification and differentiation (Ikeda et al. 2006; Moss et al. 2005; Tyler & Moss, 2001). Therefore damage to a semantic store (e.g., ATL) should result in more profound deficits in the living than non-living conditions. This category-specific impairment for living things has been attributed to more medial

ATL damage in HSVE compared to SD, affecting visual aspects of knowledge (Noppeney et al. 2006).

Further Evidence. Neuroanatomical research provides further evidence that the ATLs are the ideal site for forming transmodal semantic representations because of two key features: (i) it has extensive connections with sensorimotor cortical regions, allowing modality-specific information to converge at this point but (ii) its function is not dominated by a single sensorimotor input/output (Gloor, 1997; Plaut, 2002; Margulies et al. 2016). Consistent with these findings, computational models have shown that transmodal semantic representations can be constructed from input/output information from different modalities (verbal and nonverbal; Rogers et al. 2004). Convincingly, neural network models that incorporate intermediate transmodal hubs are best able to capture the semantic relationship between conceptually similar items (e.g., flute and piano) despite them having vastly different sensory-motor properties (McClelland et al. 2010). Moreover, when damaged, the model yielded accurate predictions of the performance of patients with semantic disruption (such as SD), across a wide variety of semantically demanding tasks.

Theory: The hub-and-spoke model of semantic processing draws on this aforementioned work and proposes that SD reflects damage to a transmodal semantic hub housed within the ATL bilaterally (Patterson et al. 2007; see Figure 2). This theory combines two important, existing ideas. First, in keeping with the embodied accounts discussed earlier, the hub-and-spoke model assumes that our distributed semantic network is comprised of modality-specific information that contributes only to concepts experienced in that modality – these representations are processed in areas proximal to and reciprocally linked to primary sensory cortices, referred to as “spokes”. Second, a transmodal “hub” housed within bilateral ATL integrates modality-specific information from the spokes (see Figure 1.2). As summarised by Reilly, Garcia & Binney (2016), ‘the additional representational layer afforded by the ATL hub provides a tertiary level of abstraction allowing for distillation of the highly complex, non-linear transmodal relationships between multi-modal features that comprise concepts’. Crucially, this account does not suggest that conceptualization can be achieved by the exclusive

activation of the representations performed by the hub. Rather, this model emphasises bi-directional connections between hub and spoke regions to permit the complete conceptualization of knowledge (Lambon Ralph et al. 2010; Pobric, Jefferies & Lambon Ralph, 2010; Reilly, Peelle, Garcia, & Crutch, 2016). Therefore, the hub-and-spoke model predicts that semantic processing will be supported by the joint activation of (i) transmodal representational cortex (e.g., ATL) and (ii) distributed unimodal regions.

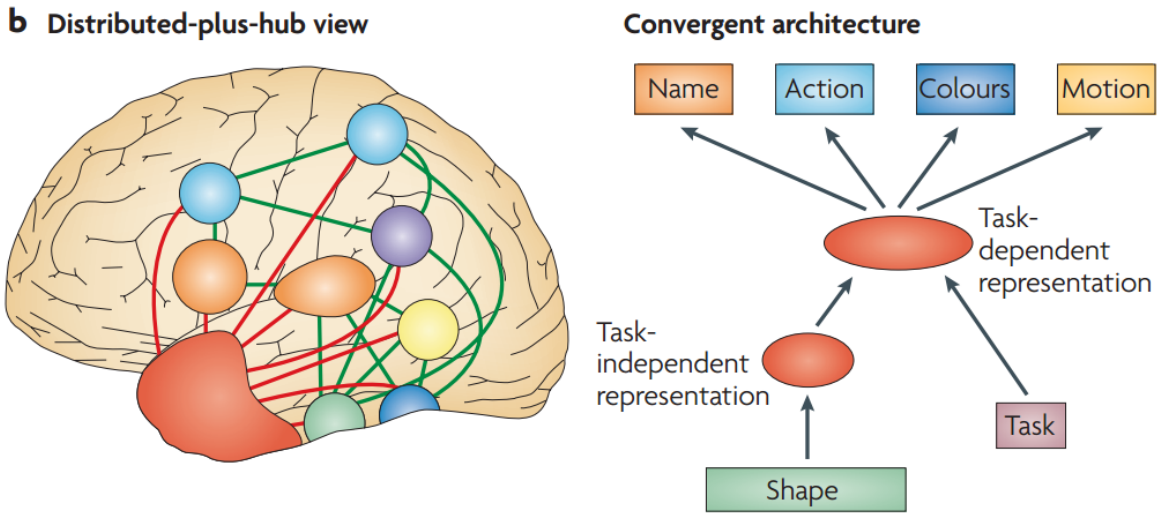

Figure 1.2. Distributed-plus-hub model of semantic processing (more commonly referred to as the hub-and-spoke model) suggests in addition to the widely distributed modality-specific regions (and their connections - green lines), a "convergence zone" or "hub" is required to allow for integration of information (red lines) across all distributed regions (transmodal hub housed within bilateral ATL; taken from Patterson, Nestor & Rogers, 2007).

1.2.2.2. Functional Neuroimaging Studies

Despite neuropsychological, neuroanatomical and computational evidence highlighting the crucial role of the ATL in semantic processing, as summarised by the hub-and-spoke model, neuroimaging studies have not been as conclusive. A number of functional neuroimaging studies have reported ATL activation for a variety of semantic tasks in healthy individuals (e.g., Baron & Osherson, 2011; Binder et al. 2009; Devlin et al. 2000; Rogers et al. 2006; Vandenberghe, Nobre &

Price, 2002). Yet, critically, many studies of semantic memory have implicated regions *outside* of the ATL: typically a combination of frontal, posterior temporal, and inferior parietal regions (for review see Binder et al. 2009; Martin, 2007; Patterson et al. 2007; Thompson-Schill, 2003). One possible explanation for the discrepancies between functional neuroimaging studies and patient findings relates to imaging artefacts. Functional neuroimaging studies typically employ functional magnetic resonance imaging (fMRI) to acquire data - although some use positron emission tomography (PET). Brain regions close to air-filled sinuses, such as ventral ATL, are affected by discrepancies in magnetic susceptibility across different tissue types (e.g. water, air, bone), resulting in loss of signal and distortion. This susceptibility artefact in fMRI studies could potentially explain why ATL activation is absent from many fMRI studies investigating semantic memory (Jezzard & Clare, 1999; Devlin et al. 2000). Support for this assumption comes from a meta-analysis which revealed PET studies were more likely to find ATL activation than fMRI; lending to the assumption that fMRI, but not PET, is sensitive to magnetic susceptibility leading to signal distortion (Visser et al. 2010).

Despite the inconsistencies in conventional fMRI studies, recent advances in neuroimaging methodologies such as distortion-corrected fMRI (Binney et al. 2010; Visser et al. 2010; Visser & Lambon Ralph, 2011); fMRI decoding (Correia et al. 2014; Coutanche & Caramazza, 2014; Peelen & Caramazza, 2012); MEG and EEG decoding (Chan et al. 2011; Chen et al. 2016); and repetitive transcranial magnetic stimulation (rTMS; Lambon Ralph, Pobric & Jefferies, 2009; Pobric, Jefferies & Lambon Ralph, 2007; 2009) have provided converging evidence that the ATLs play an important role in transmodal semantic processing. In brief, neuroimaging studies have implicated the ATLs in tasks that manipulate semantic specificity (Rogers et al. 2006; Tyler et al. 2004); studies involving semantic judgements of both auditory and visual input (Binney et al. 2010; Bright, Moss & Tyler, 2004; Lambon Ralph et al. 2009; Pobric et al. 2007; 2009; Visser & Lambon Ralph, 2012); comprehension of both spoken and written sentences (Hickok & Peoppel, 2007; Scott et al. 2000), decoding abstract conceptual properties of objects (Peelen & Caramazza, 2012; Coutanche & Thompson-Schill, 2014) and decoding language invariant concepts (Correia et al. 2014). However, the peak ATL activation reported across such studies

is often variable (for review see Visser et al. 2010; Lambon Ralph et al. 2017). Given the anatomical variability in previous findings, the next section will address which specific regions within the ATL contribute to transmodal semantic memory.

1.2.3. Differentiation of Function within ATL Structures – Which Region is Transmodal?

Historically the ATL has been discussed as one homogenous structure (Patterson et al. 2007). However, the ATLs are made up of multiple anatomically defined structures, each of which arguably plays a distinct role in the representation of conceptual knowledge (see Figure 1.3 adapted from Bonner & Price, 2013; Brodmann, 1909; Blaizot et al. 2010); this may explain why the peak activation is variable across semantic studies. Prior investigations of the ATL have revealed that anterior superior temporal gyrus (aSTG) is particularly involved in the processing of auditory and verbal stimuli (Scott et al. 2000; Spitsyna et al. 2006; Visser & Lambon Ralph, 2011), while anterior fusiform is particularly involved in the processing of pictures (Coutanche & Thompson-Schill, 2014; Peelen & Caramzza, 2012; Visser et al. 2010). Collectively, these findings lend to the idea that there is some degree of modality-specificity in the function of the ATL (Bajada et al. 2017; Jackson et al. 2015; Lambon Ralph et al. 2017; Visser et al. 2001). Meta-analytic results are consistent with this account; Visser et al (2010) found that although there were significant differences in the location of peaks for pictures and words (ventral ATL structures including inferior temporal gyrus (ITG) and fusiform cortex for pictures and superior temporal gyrus (STG) for words); the activation across pictures, spoken words and written words was highly overlapping. Moreover, the distribution between pictorial and word peaks gradually became more distinct in regions closer to input modalities (i.e., regions posterior the ATL).

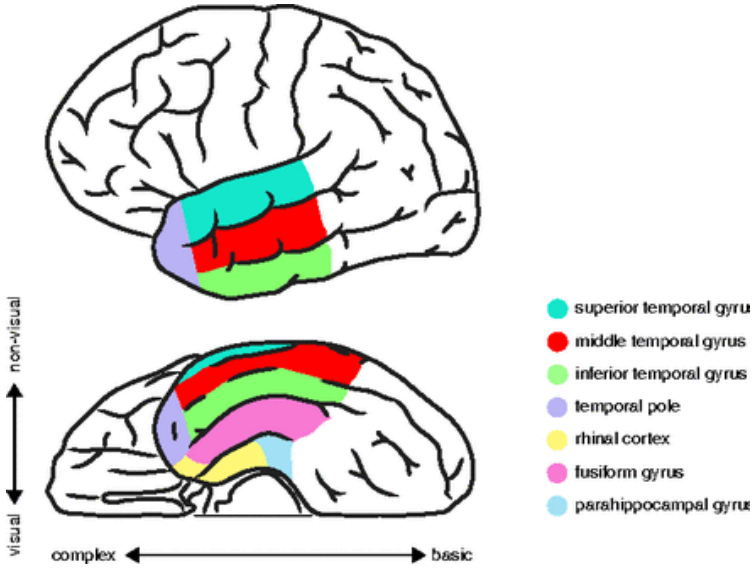

Figure 1.3. Structures of the anterior temporal lobe taken from Bonner and Price (2013). The y-axis refers to the contribution of regions in relation to visual and non-visual processing. The x-axis refers to the complexity of processing.

One suggestion for such findings is that information converges within heteromodal and transmodal regions with graded connectivity (Bajada et al. 2017; Buckner & Krienen, 2013; Lambon Ralph et al. 2017; Margulies et al. 2016; Mesulam, 1998; Plaut, 2002; Schapiro et al. 2013). For instance, Plaut's (2002) computational model proposed there is a posterior-to-anterior gradient in the degree of modality-specificity within our semantic system. This gradient reflects connection distance, that is, areas closest to sensory input systems (i.e., verbal and visual) are relatively more important for semantic tasks in that domain, while regions furthest from modality input (e.g., the ATL) detach from sensory information. Plaut's model also highlighted that there is no *absolute* distinction between semantic processing of different stimuli types in the ATLs, in line with a transmodal perspective. Consistent with this finding, Margulies et al. (2016) investigated representational gradients across the entire brain. This paper described a "principal gradient" of connectivity explaining the largest source of variance in temporal correlations at rest, which extends from unimodal sensory-motor regions at one end, to transmodal default mode regions at the other. Regions situated at the top of this representational hierarchy, in the default mode network

(including ventrolateral ATL) might support higher-order representations with predictive value across multiple situations and modalities, which maximally integrate features from diverse sensory-motor regions and other transmodal integration zones.

In line with these graded perspectives, more recent studies have placed an increasing emphasis on ventrolateral temporal regions - that are posterior to the temporal pole (Mion et al. 2010) extending laterally into middle and inferior aspects of the ATL (Binney et al. 2010) – as the core transmodal site. This region is consistent with the site of maximal atrophy in SD patients (Binney et al. 2010; Mion et al. 2010). Moreover, distortion-corrected fMRI studies have revealed peak activation in ventrolateral ATL region across a variety of semantic tasks including synonym judgements and verbal/picture semantic association tests (Binney et al. 2012; Visser et al. 2010; 2012). In addition, macroscale decomposition reveals that ventrolateral ATL is situated further from sensory input, compared to superior and ventral ATL regions, allowing it to support higher-order representations with predictive value across multiple situations and modalities (Bajada et al. 2017; Buckner & Krienen, 2013; Margulies et al. 2016; Mesulam, 1998; Plaut, 2002).

Taken together, the functional segregation identified within the ATLs has been incorporated into a recently updated version of the hub-and-spoke account, coined the graded-hub account (Lambon Ralph et al. 2017; see Figure 1.4), which highlights that (i) superior ATL regions play a role in auditory processing due to its connection to auditory and language systems, (ii) portions of the inferior and ventral ATL play a role in visual processing due to their connections to visual cortices and (iii) the core site for the transmodal semantic hub can be assigned to ventrolateral ATL encompassing anterior MTG and ITG by virtue of its distance from both auditory and visual sensory input systems; as this region is transmodal in nature it should integrate conceptual information from all sensory input modalities, but does not exhibit preference for any single modality. That is, because the representation of a concept within the "hub" is now translational across information sources, the representation becomes modality invariant (e.g., we are able to know the visual properties of a flute given only its sound).

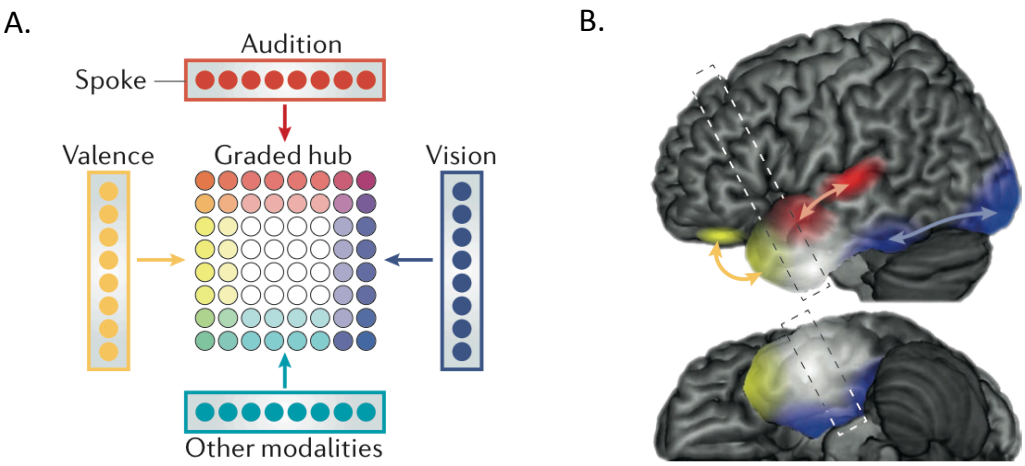

Figure 1.4. Graded hub account adapted from Lambon Ralph et al (2017). (A) Illustration of the computational framework of the graded hub-and-spoken model. The 8x8 grid represents the reciprocal connectivity to the modality-specific spoke layers (each spoke is depicted by a different colour). At the centre-point (denoted by the white coloured circles) there is equally weighted connectivity to all input, and thus the function of the unit remains evenly transmodal. (B) Illustration of the graded hub framework projects onto the cortical surface. This highlights that transmodal function is assigned to ventrolateral ATL (white cluster).

1.2.4. Organization of Conceptual Representations within Unimodal and Transmodal Regions

Despite our growing understanding of the functional role and anatomical location of both "hub" and "spoke' regions of the cortex, as captured by the graded-hub account (Lambon Ralph et al. 2017), several questions remain regarding how concepts are represented across our distributed-semantic network: (i) in "spoke" regions do words and experiences really activate common neural substrates? (ii) How is information represented in transmodal regions – for instance, are concepts organized by overlap in shared features and (iii) despite ventrolateral ATL activating equally across multiple modalities of testing it is still unclear whether the conceptual representations themselves are identical or unique; a truly transmodal representation should be the same regardless of how it is accessed it (i.e., the representation of a flute triggered by the sound it makes should map on to the representation of a flute when triggered by its visual appearance). Notably, a methodological advancement, known as multivoxel pattern analysis (MVPA), has

allowed researchers to address such questions. As opposed to standard univariate analyses, which are interested in the overall activity of individual voxels or regions of interest, MVPA is interested in the distributed pattern of activity across multiple voxels (these methods will be discussed in more detail in Chapter 2). Therefore a common brain region may be activated equally by two conditions (e.g., smelling cinnamon and reading the word cinnamon) but this does not mean the pattern of activity is the same for these two conditions. The following section will summarise recent MVPA findings that have helped shed light on how concepts are represented in both unimodal and heteromodal brain regions.

As alluded to earlier, embodied accounts assert that knowledge is represented in the neural systems used for perceiving and acting on sensory information in our environment (Allport, 1985; Barsalou, Simmons, Barbey & Wilson, 2003; Chao & Martin, 1999; 2000; Goldberg, Perfetti Charles & Schneider Walter, 2006; Gonzalez et al. 2006; Hauk & Pulvermüller, 2004; Humphrey & Forde, 2001; Martin, 2007; Pulvermüller, 1999; 2005; Simmons, Martin & Barsalou, 2005; Simmons, et al. 2007). Clearly, however, the two cognitive abilities are not the same phenomena, as such the relationship between perception and meaning remains unclear. For example, despite sensory and motor cortices *activating* to both perception and knowledge, it remains unclear whether the two *representations* are similar or whether they simply share a common neural region, but are coded in a unique format. Notably, a recent MVPA study investigated whether overlapping cortical activation for action execution and action-word comprehension reflected common neural sources or proximal, but distinct, sources (Rueschemeyer et al. 2014). MVPA was used to investigate whether the pattern of activity captured by distinguishing action execution vs. nonbiological motion, could also capture the distinction between action words (e.g., grasp) vs. nonbiological motion words (e.g., to snow). The results demonstrated that activation in unimodal "spoke" regions of the cortex (e.g., ventrolateral pre-motor cortex) was not comparable for action execution and action word comprehension; suggesting that although common neural regions are required for both processes, the underlying representations are distinct. However, it remains unclear whether this is true in other sensory modalities (e.g., visual and auditory).

Furthermore, the central assumption of the graded-hub account is that transmodal processing occurs in ventrolateral ATL. However, if this region is truly transmodal then the pattern of activity for a concept activated through the visual domain (i.e., seeing a picture of a flute) should map on to the same representation activated through the auditory domain (i.e., hearing a flute) or any other modality for that matter. Recent MVPA studies have provided tentative evidence along this line. For instance, Peelen and Caramazza (2012) found that bilateral ventral ATL encodes information about the abstract conceptual properties of objects. In their study, participants viewed images of objects that differed on two dimensions: location and action. Half of the objects were found in a kitchen (e.g., garlic press) the other half in a garage (e.g., screwdriver). Half were objects requiring a rotating movement to operate (e.g., screwdriver) and the other half a squeezing movement (e.g., garlic press). The authors reasoned, for a region to process abstract conceptual features it should have similar neural responses to objects that share a conceptual feature (i.e., both found in a kitchen), even if those objects differed on all other dimensions. MVPA revealed visual similarity was encoded in early visual processing regions of temporal-occipital cortex, whereas the strongest similarity between abstract concepts was revealed in bilateral anterior ventral temporal lobe. This region showed a similar effect for both the location and action features of concepts. These findings indicate a posterior-anterior gradient along the ventral visual stream, with higher-level conceptual associations processed in anterior ventral temporal lobe. This lends support to the idea that ventral ATL is a heteromodal region that integrates knowledge about motor and spatiotemporal properties. Similar results were found in a recent MVPA electrocorticogram (ECoG) study that utilised a picture-naming task and found that ventral ATL held patterns of activity regarding the categories of pictures being presented (Chen et al. 2016). Interestingly, both of these studies used picture stimuli which heavily relies on visual-feature knowledge, and has been shown to recruit ventral ATL regions in line with a graded effect of modality-input from ventral visual stream into ventral ATL (Plaut, 2002; Visser et al. 2010). These factors may explain why both Peelen and Caramazza (2012) and Chen and colleagues (2016) identified a more ventral cluster compared to previous findings - spanning neuropsychology, distortion-corrected

fMRI or macroscale decompositions - that pin point ventrolateral ATL as the site of transmodal processing (Bajada et al. 2017; Binney et al.2010; 2012; Buckner & Krienen, 2013; Margulies et al. 2016; Mesulam, 1998; Mion et al.2010; Plaut, 2002; Visser et al. 2010; 2012).

Consistent with notion that the presentation format influences the location of heteromodal processing within the ATLs, Correia et al (2014) identified superior ATL in a crossmodal study investigating semantic representations in Dutch-English bilinguals. In their study, MVPA was employed to investigate the neural mechanisms underlying the representation of language-independent concepts in the brain. Bilingual subjects were presented with spoken animal nouns in both English and Dutch. A classifier was trained to see whether patterns of activity trained to distinguish between spoken nouns in one language (e.g., "horse" vs. "duck" in English) could accurately predict the same distinction in the other language (e.g., "paard" vs. "eend" in Dutch). This cross-modal classifier revealed a significant cluster in the left superior ATL, lending to the idea that the superior ATL is a hub region that organizes conceptual information in a language-invariant manner (i.e., abstracted away from input modality). Taken together these findings suggest that studies that use MVPA to decode visual information retrieve clusters in the most ventral portion of the ATLs (Chen et al. 2016; Devereux, Clarke, Marouchos & Tyler, 2013; Peelen & Carmazza, 2012;), whereas those who decode auditory information retrieve clusters in the most superior portions of the ATL (Correia et al.2014). These results are consistent with graded effect of input modality (Bajada et al. 2017; Lambon Ralph et al. 2017; Margulies et al. 2016; Plaut, 2002; Visser et al. 2010), with language and auditory semantic processing recruiting superior portions of the ATL and visual and picture processing recruiting ventral portions of the ATL; however they do not provide compelling evidence for a 'transmodal' hub as these effects seem to be, at least in part, driven by the modality of presentation.

Therefore, although MVPA is a valuable tool for investigating representations within the ATLs, the precise location of the putative transmodal hub remains a hotly debated topic (Lambon Ralph, 2014). Moreover, in addition to the lack of consistency of MVPA results *within* the ATL, several MVPA studies

investigating the representation of conceptual knowledge only report findings *outside* of the ATL (e.g., Bonnici et al. 2016; Bruffaerts et al. 2013; Fairhall & Caramzza, 2013; Fernandino et al. 2016; Simanova et al. 2014); including posterior middle temporal gyrus (pMTG), angular gyrus and inferior frontal gyrus. These regions are compatible with the semantic control network (a distributed network made up of fronto-temporal-parietal brain regions), which permits relevant information to be brought to the fore, and are particularly important when we need to retrieve distance semantic associations or weakly activated features (Badre et al. 2005; Binder et al. 2009; Davey et al. 2016; Pobric et al. 2007; Lambon Ralph et al. 2017; Thompson-Schill et al. 1997; Wagner et al. 2001; Whitney et al. 2011). Given that only semantic control regions, and not the putative hub, are reported in several machine-learning studies, one could speculate that the experimental paradigms being tested are not tailored for revealing transmodal properties within the ATL. For instance, Fairhall and Caramazza's (2013) paradigm required participants to rate how typical each item was within its semantic category (e.g., rating the typicality of "apple" or "coconut" as a fruit). This type of judgement arguably manipulates the requirement of control processes; that is an "apple" is a highly typical fruit example and thus requires little/no input from control processes, whereas identifying a "coconut" as a fruit may require control process to facilitate the suppression of strongly activated features (e.g., coconut – tree) to activate distant associations (coconut – fruit). Thus patterns of activity may have been more informative in semantic control regions for decoding between different semantic concepts. Furthermore, as stated in their study, the signal-to-noise ratio in ventrolateral ATL was significantly weaker than the regions they found MVPA results in, lending to the notion that a lack of ATL finding could be due to the susceptibility artefacts. Finally, the lack of ATL results can also be attributed to a limited ROI analysis that neglects ATL contribution (e.g., Bonnici et al. 2016). Taken together, the MVPA literature sheds light on how semantic memories are represented across the cortex, but it remains unclear whether *transmodal* semantic concepts are represented within the ATL.

1.2.5. Summary of Conceptual Representation

The graded-hub account proposes that semantic processing is comprised of two components: (i) modality-specific regions of the cortex referred to as spokes and (ii) a transmodal hub (bilateral ventrolateral ATL) optimal for the abstraction of conceptual representations from multiple experiences and modalities (e.g., sound, visual and action attributes). A wealth of evidence from neuropsychological, neuroanatomical and functional neuroimaging studies supports this theory. However, several questions remain unclear, namely, which brain structures are important for forming conceptual representations and what role or type of information do they add. The primary goal of this thesis is to utilise MVPA to address the following questions (i) in "spoke" regions, do words and experiences really activate common neural substrates? That is, despite sensory and motor cortices responding to both perception and knowledge, it remains unclear whether the two *representations* are similar or whether they simply share a common neural region, but are coded in a unique manner. (ii) The current thesis also aims to elucidate where, within the ATLs, are representations truly transmodal by comparing representations across both spoken (auditory input) and written (visual input) presentations formats. This will overcome the limitations of 'sensory bias' results (the notion that the presentation format influences the location of heteromodal processing within the ATLs; e.g., Peelen & Caramazza, 2012; Chen et al. 2016; Correia et al. 2014) which are potentially mediated by the presentation format as well as the conceptual meaning.

1.3. Perceptually-decoupled Retrieval

While previous research has resulted in a number of theories regarding the broad organisation of semantic cognition in the brain (e.g., Barsalou et al. 2003; Binder et al. 2009; Binder & Deasi, 2011; Damasio, 2008; Koenig & Grossman, 2007; Lambon Ralph et al. 2017; Martin, 2007; Meteyard et al. 2012; Patterson et al. 2007; Pulvermüller, 2013; Tranel, Damasio & Damasio, 1997), most of this work has considered processes such as picture recognition or matching; less is known about

patterns of semantic retrieval that are perceptually-decoupled (i.e., not immediately driven by stimuli in the environment). This is an interesting avenue for investigation since semantic retrieval extends beyond the here-and-now, to draw on abstract knowledge that has been extracted across multiple experiences; for instance, we can easily bring to mind what a dog looks and sounds like, regardless of whether or not there is a dog present in our immediate environment. The ability for semantic regions to process information "off-line" enables us to make sense of past experiences and create effective plans regarding behaviour in the future (Binder et al. 2009; Binney, Embelton, Jefferies, Parker & Lambon Ralph, 2010; Hart et al. 2007; Irish, 2016; Jefferies & Lambon Ralph, 2006; Lambon Ralph & Patterson, 2008). Particularly noteworthy are suggestions that this uniquely human behaviour to perform high-level computations off-line is the main explanation for our ability to adapt, create culture, and even develop technology (Binder et al. 2009; Irish, 2016; Klein, 2013; Suddendorf & Corballis, 2007). However, a clear understanding of the neural substrates that support patterns of semantic retrieval that are not immediately driven by stimuli in the environment is lacking. Therefore, the second goal of this thesis is to better understand how we can retrieve semantic concepts from memory, in the absence of external input.

1.3.1. The Neural Correlates of Perceptually-decoupled Semantic Retrieval

Contemporary semantic accounts (e.g., Lambon Ralph et al. 2017; Patterson et al. 2007) suggest that memory retrieval relies on modality-specific components located in unimodal sensory cortex and abstract representations that are largely invariant to the input modality, in transmodal portions of the brain. In the absence of perceptual input, memory retrieval is likely to depend on abstract representations to a greater extent, suggesting that transmodal regions, such as the ventrolateral ATL, may be especially important for stimulus-independent cognition (Damasio, 1989; Margulies et al. 2016). In line with this perspective, the literature on perceptually-decoupled semantic retrieval can be broadly divided into two camps: (i) embodied accounts that focus on the role of unimodal sensory cortices during

internally-generated semantic retrieval tasks, such as mental imagery (e.g., Albers et al. 2013; Daselaar et al. 2010; de Borst et al. 2016; Halpern & Zatorre, 1999; Ishai et al. 2000; Reddy et al. 2010; Vetter et al. 2014; Zvyagintsev et al. 2013) and (ii) a growing body of work that investigates transmodal brain regions such as anterior and posterior temporal lobe, fronto-parietal and medial regions (e.g., Binder et al. 2009; Binder & Desai, 2011; Fox et al. 2005; Humphreys et al. 2015; Irish & Piguet, 2013; Jefferies, 2013; Margulies et al. 2017; McKiernan et al. 2006; Noonan et al. 2010; Smallwood & Schooler, 2006). Evidence from neuroimaging and neuroanatomical fields will be reviewed to understand the specific contributions of both of these bodies of research, to perceptually-decoupled semantic retrieval.

1.3.1.1. Unimodal Brain Regions

As discussed previously embodied accounts propose that knowledge of objects resides in the very cortices that process their features during perception or use (Barsalou, 2008; Barsalou et al. 2003; Kiefer & Pulvermüller, 2011; Martin, 2007; Meteyard et al. 2012; Pulvermüller, 2012; Wilson, 2002). Particularly noteworthy, with regard to perceptually-decoupled semantics, are studies that investigate how sensory cortices are recruited during mental imagery; which requires the retrieval and maintenance of information from memory, such as imagining what a dog looks or sounds like, without the supporting sensory experience (i.e., without seeing or hearing a dog). Imagery therefore relies on previously organised and stored semantic information about the features to be imagined (Kosslyn et al. 1997), therefore retrieval of memories – that is, semantic retrieval – is argued to underlie imagery (Barsalou, 1999). Consistent with this viewpoint, studies have shown that visual cortex is activated by imagined images (Albers et al. 2013; Coutanche & Thompson-Schill, 2014; de Borst et al. 2016; Ishai et al. 2000; Reddy et al. 2010; Vetter et al. 2014) and auditory cortex is activated by imagined sounds (Daselaar et al. 2010; de Borst et al. 2016; Halpern & Zatorre, 1999; Zvyagintsev et al. 2013). Notably, the majority of studies find recruitment of sensory association cortices as opposed to primary cortices during visual (Amedi et al. 2005; Ishai et al. 2000; Knauff et al. 2000), auditory (Bunzeck et al. 2005; Zatorre & Halpern, 2005) and action mental imagery (Willems et al. 2009; 2010).

One possible explanation for the recruitment of sensory association cortices is the 'anterior shift' noted by Thompson-Schill (2003). She found that areas engaged during semantic processing are not isomorphic to those areas used in direct experience, but rather are shifted anterior to those areas (for a wider review see Chatterjee, 2010; Binder & Desai, 2011; McNorgan et al. 2011; Meteyard et al. 2012). This anterior shift therefore suggests that information in modality-specific regions is abstracted from direct experience during retrieval of semantic concepts from memory. Such accounts are therefore consistent with a gradient of 'abstraction', where, as one moves way from primary sensory and motor cortex, more complex relationships are captured (Bajada et al. 2017; Buckner & Krienen, 2013; Lambon Ralph et al. 2017; Margulies et al. 2016; Mesulam, 1998; Meteyard et al. 2012; Patterson et al. 2007; Plaut, 2002; Schapiro et al. 2013; Visser et al. 2010).

Furthermore, as highlighted by Mahon and Caramazza (2008), embodied findings, similar to those discussed earlier, do not clarify whether unimodal regions are directly activated to facilitate the retrieval of a semantic concept from memory (bottom-up) or whether they are subsequently activated due to the retrieval of conceptual knowledge from a transmodal region (top-down). Convincingly, fMRI studies that utilise dynamic causal modelling (Friston et al. 2003; 2013) support the top-down recruitment of unimodal cortices for perceptually-decoupled semantic retrieval. For example, Mechelli, Price, Friston and Ishai (2004) found that conceptual representations in secondary visual cortices are facilitated by primary visual areas, in a bottom-up fashion during perception (e.g., viewing a picture of a DOG) while top-down activation from heteromodal regions to visual cortices are seen during imagery (e.g., imaging a DOG). This is consistent with a growing body of fMRI literature indicating top-down activation of sensory and motor cortices during perceptually-decoupled semantic retrieval (e.g., Coutanche & Thompson-Schill, 2014; Dijkstra et al. 2017; Ganis & Schendan, 2008; Ishai et al. 2000; Kalkstein et al. 2011; Kosslyn, 2005; Kosslyn & Thompson, 2003; Pearson et al. 2011; Stokes et al. 2009). Moreover, electrophysiological data showed greater top-down activation during visual imagery as compared to visual perception (Dentico et al. 2014; van Wijk et al. 2013). Taken together, this evidence suggests that (i) unimodal portions

of cortex, in particular secondary association regions, are recruited during perceptually-decoupled retrieval of semantic information and (ii) the activation in these regions are facilitated in a top-down fashion by higher-order transmodal cortices.

1.3.1.2. Transmodal Brain Regions

In an effort to elucidate how transmodal portions of the semantic network engage during perceptually-decoupled semantic retrieval, attention has also been given to imagination-based tasks that capture the deliberate retrieval of internally-generated information (e.g., Coutanche & Thompson-Schill, 2014; Daselaar, Porat, Huijbers & Pennartz, 2010; Golchert et al. 2017; McNorgan, 2012). For instance, a recent fMRI study in conjunction with MVPA found that left ATL could successfully decode the properties of an imagined object (Coutanche & Thompson-Schill, 2014). In this study, classification accuracy in early visual regions, related to the shape (in V1) and colour (in V4) of the object, predicted classifier accuracy for the specific object in ATL. This is consistent with the theory that information from sensory cortex is integrated in transmodal ATL to form modality-invariant conceptual representations that are critical for perceptually-decoupled semantic cognition, as well as for the comprehension of words and objects in the external environment.

In addition to specific task manipulations that help elucidate the role of the semantic network in retrieval of concepts from memory (i.e., imagination inducing experiments), focus has been drawn to “passive” or “rest” periods within a task, in which subjects are given either no task, or a minimally demanding task such as staring at a fixation cross (Binder et al. 1999; 2009). Although such conditions are routinely used across neuroimaging studies as low-level baselines, the majority of subjects report experiencing detailed and memorable thoughts and mental images during such conditions (Binder et al. 1999; James, 1890; McKiernan et al. 2006; Singer, 1993; Visser et al. 2010). Consequently, a growing body of researchers have argued that such “task-unrelated thought” is essentially underpinned by semantic processing (Binder et al. 1999; 2009; Irish, 2016; Irish & Piguet, 2013; McKiernan et al. 2006; Visser et al. 2010). It has subsequently been argued that semantic memory

is the root of imagination; that is, all forms of imagination, whether it is creative or autobiographical in nature, emerge from semantic memory (Abraham & Bubic, 2015; Irish & Piguet, 2013). Strong support for these arguments comes from the finding that during rest periods a network almost identical to the semantic network is recruited (Binder et al. 1999; 2009; Jackson et al. 2016; Pascual et al. 2015). For instance, using an intrinsic connectivity measure, Jackson and colleagues (2016) revealed the same semantic regions were active during rest and semantic task states, supporting the necessity for semantic cognition in internal processes occurring during rest.

An important, and often overlooked, observation - that may help shed light on how transmodal regions are recruited during retrieval of knowledge from memory - is that the network implicated in semantic tasks extends over multiple intrinsic networks, including "task-positive" networks such as the frontoparietal network, but also the default mode network (DMN), whose core regions are functionally connected during rest (see Figure 1.5). Given the anatomical similarity between the semantic and default mode networks, there may be common underlying processes that facilitate these two independent, but nevertheless interactive, components or networks (Moscovitch, 1992). Although investigations of the DMN and semantic network have been primarily independent of one another (for review see Binder et al. 2009; Humphreys et al. 2015), there are several reasons why direct comparison of the networks may be beneficial. First, during "rest" the brain is engaged in the activation of rich conceptual representations, and thus the DMN processing places strong demands on the semantic system (Binder et al. 1999; 2009; Irish, 2016; Irish & Piguet, 2013; Jackson et al. 2015; McKiernan et al. 2006). Second, the DMN and semantic network engage several common anatomical areas, including regions in parietal and lateral temporal lobes (Binder et al. 1999; 2009; Buckner, Andrews-Hanna & Schacter, 2008; Humphreys et al. 2015; Raichle, 2001). For instance, both ATL and angular gyrus are considered semantic "hubs" that help represent heteromodal semantic representations (Binder & Desai, 2011; Lambon Ralph, 2013). Consequently, a recent upsurge of interest has seen the simultaneous investigation of these two networks for the retrieval of semantic knowledge (e.g., Humphreys et al. 2015; Seghier & Price, 2012; Seghier, 2013; Wirth et al. 2011). The

following section will briefly review the literature surrounding the anatomical organisation and function of the DMN and then reflect on how the semantic and default mode networks may support a common cognitive function.

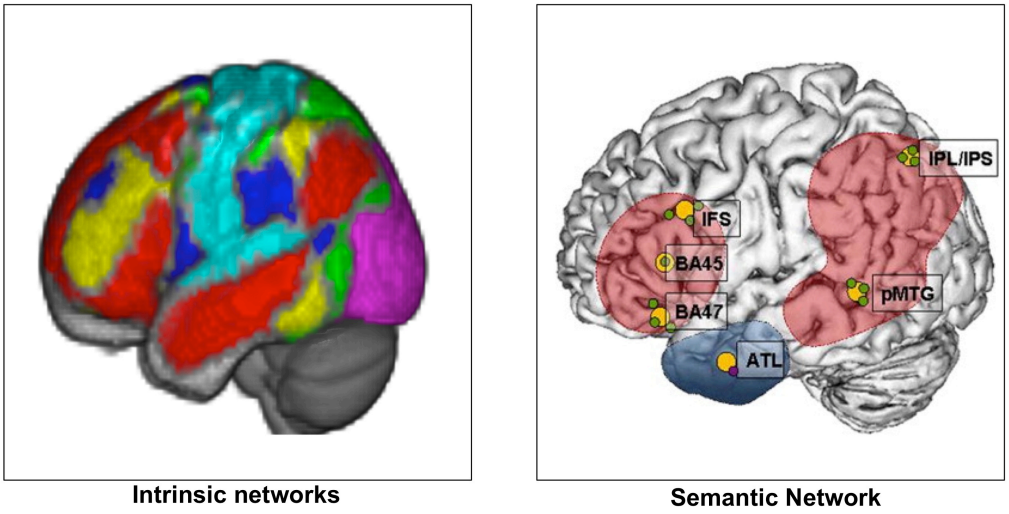

Figure 1.5. Anatomical similarity between the semantic network (Noonan et al. 2010) and other core intrinsic networks (Yeo et al. 2011). Notably, the red intrinsic network (which depicts the DMN) has considerable overlap with the semantic network.

The DMN is an anatomically defined network consisting of a constellation of transmodal brain regions including anterior and posterior temporal lobe, fronto-parietal and medial regions (Fox et al. 2005; Greicius et al. 2003; Raichle et al. 2001; Smallwood & Schooler, 2006). Early studies characterised the DMN as "task-negative", as this network was shown to be more active during "rest" periods (such as staring at a fixation cross) than during explicit task conditions (Shulman et al. 1997). The term 'default mode' caught on after the publication of 'A Default Mode of Brain Function' by Raichle et al (2001). This seminal paper supported Shulman's findings and concluded that there was a distributed network of brain regions that decreased their activity during goal-directed and attention-demanding tasks. Moreover, this paper showed that these regions were highly related at 'rest'. Since the discovery of a distributed 'default mode' network, research has moved away from this initial interpretation of a "task-negative" state. Notably, the understanding of the function and organisation of the DMN has been propelled by studying spatial coherence patterns in the spontaneous fluctuations in the fMRI

BOLD signal during the resting state (see Raichle 2010, 2011); commonly known as resting-state functional connectivity (this method will be discussed in more detail in Chapter 2). During such scans participants are required to lie still in an fMRI scanner for several minutes without an explicit task. Functional connectivity analysis revealed that the DMN can be reliably delineated based on the functional connectivity during such resting-state scans; that is, the activity within the DMN regions were functionally coupled when participants were not engaged in an explicit task. This led to a new wave of researchers exploring internally-oriented cognitive processes during these rest-periods, a phenomenon widely described as task-unrelated thought or mind-wandering (Buckner et al. 2008; Mason et al. 2007).

To-date an established body of work has routinely shown that the DMN actively supports several aspects of cognition (Spreng, 2012), including semantic processing (Binder et al. 2009; Humphreys et al. 2015; Irish & Piguet, 2013; Krieger-Redwood et al. 2016), autobiographical and episodic recollection (Andrews-Hanna, 2012; Buckner et al. 2008; Rugg & Vilberg, 2013), working memory (Konishi et al. 2015; Spreng et al. 2014; Vatansever et al. 2015), mental imagery (Hassabis et al. 2007), self-generation of emotion (Engen, Kanske & Singer, 2017) and imagining the future or recalling the past (Huijbers et al. 2009; Maguire, 2001; Schacter et al. 2007; Spreng et al. 2009; Svoboda et al. 2006; Szpunar et al. 2007). This evidence therefore goes against historical accounts of the DMN as "task-negative" by showing that the DMN does activate under a variety of task conditions. Notably, many of these situations involve memory retrieval – i.e., a requirement to focus cognition on previously-encoded knowledge as opposed to information in the external environment. The current thesis was therefore motivated by the hypothesis that there might be common neurocognitive processes underpinning perceptually-decoupled and conceptually-guided cognition in the DMN. During states of episodic recollection, we recreate past experiences that involve places, objects and people not currently present in the environment. Consequently, memory retrieval might necessitate a process of decoupling from sensorimotor systems, allowing cognition to be generated internally in a way that diverges from what is going on around us (Smallwood, 2013). These perceptually-decoupled states

might be easier in brain regions whose neural computations are functionally independent, or distant, from systems important for perceiving and acting.

In line with this observation, many regions within or allied to the DMN are considered to be transmodal 'hubs' for memory-related processes, including posterior cingulate cortex (Leech, Braga & Sharp, 2012; Leech & Sharp, 2014), angular gyrus (Binder & Desai, 2011; Bonnici et al. 2016; Seghier, 2013), hippocampus (Moscovitch et al. 2016) and anterior temporal lobes (Lambon Ralph et al. 2017; Patterson et al. 2007; Visser, Jefferies and Lambon Ralph, 2010). In addition to the ATLs, a region of particular importance to perceptually-decoupled semantic retrieval, is the angular gyrus. Using the NeuroSynth database (Yarkoni et al. 2011), a "reverse inference" of the left angular gyrus reveals key concepts that have been associated with this region in neuroimaging studies; the top five concepts comprise 'retrieval', 'default mode network', 'memory', 'semantic' and 'sentence' (Seghier, 2013). Semantic processing is therefore one of the most consistent functions that activates the angular gyrus (Binder et al. 2009; Price et al. 2015; Seghier, 2013; Seghier et al. 2010; Wang et al. 2010). Furthermore, the activity in bilateral angular gyrus during 'rest' is remarkably reliable (Buckner et al. 2008; Shehzad et al. 2009; Smith et al. 2009; Spreng et al. 2009). To account for the wide variety of tasks that the angular gyri have been associated with, several influential theories have been proposed.

One prominent hypothesis put forward to explain the varied cognitive functions assigned to the angular gyrus is the proposal that it is involved in the combination and manipulation of acquired knowledge about the world during rest, that is interpreted during demanding tasks, such as sentence comprehension, problem solving and planning (Binder et al. 1999; Seghier et al. 2010). A second, but not mutually exclusive proposal is that the angular gyri show a stronger response to a range of memory retrieval situations in which the retrieved representations are detailed, specific or precise (Binder et al. 2005; Price et al. 2015; Bonnici et al. 2016; Davey at al. 2015). Alternatively, Andrews-Hanna and colleagues (2010) suggest that the angular gyrus might be engaged in the construction of mental scenes based on memory during perceptually-decoupled retrieval, during rest. Finally, a fourth more general theory is that the inferior parietal cortex, including angular gyrus, is

responsible for focusing attention on memory (Cabeza et al. 2011; Humphreys et al. 2015). Clearly we lack an over-arching account of the functions of the angular gyrus, however the common denominator between these hypotheses is the engagement of the angular gyrus in the manipulation of increasingly abstract conceptual knowledge and mental representations.

1.3.2. Intrinsic Connectivity Characterizes Functional Organisation of the Brain

A complementary method that has helped shed light on the role of the angular gyrus (and other transmodal regions) in abstract conceptual knowledge and mental representations, focuses on understanding the underlying connectivity structures and intrinsic organisation of these networks (Braga et al. 2013; Buckner & Krienen, 2013; Leech et al. 2012; Margulies et al. 2016; Mesulam, 1998; Plaut, 2002). Recent macroscale decompositions of brain connectivity have helped characterize the neural regions that are likely to be important for heteromodal memory representations in a more formal manner; such data provides support for transmodal processing in ventrolateral ATL and angular gyrus, as well as other core DMN regions (see Figure 6). Margulies and colleagues (2016) described a "principal gradient" of connectivity with unimodal sensory regions at one end and transmodal regions including posterior cingulate cortex, medial prefrontal cortex and ventrolateral ATL at the other – regions that are allied to or fall within the default mode network (DMN; Raichle et al. 2001). These regions are maximally distant in functional and structural space from primary landmarks of unimodal function such as the calcarine sulcus or the central sulcus (Margulies et al. 2016). This topographic architecture suggests regions, such as ventrolateral ATL and angular gyrus, are situated at the top of a representational hierarchy allowing these regions to integrate information across systems. Such regions are therefore able to support higher-order representations with predictive value across multiple situations and modalities, which integrate features from diverse sensorimotor regions. In other words, increasingly abstract and complex representations might be formed at greater distances along the gradient, where the influence of specific features linked

to stimuli in the immediate environment is reduced (Buckner & Krienen, 2013; Margulies et al. 2016; Mesulam, 1998; Plaut, 2002).

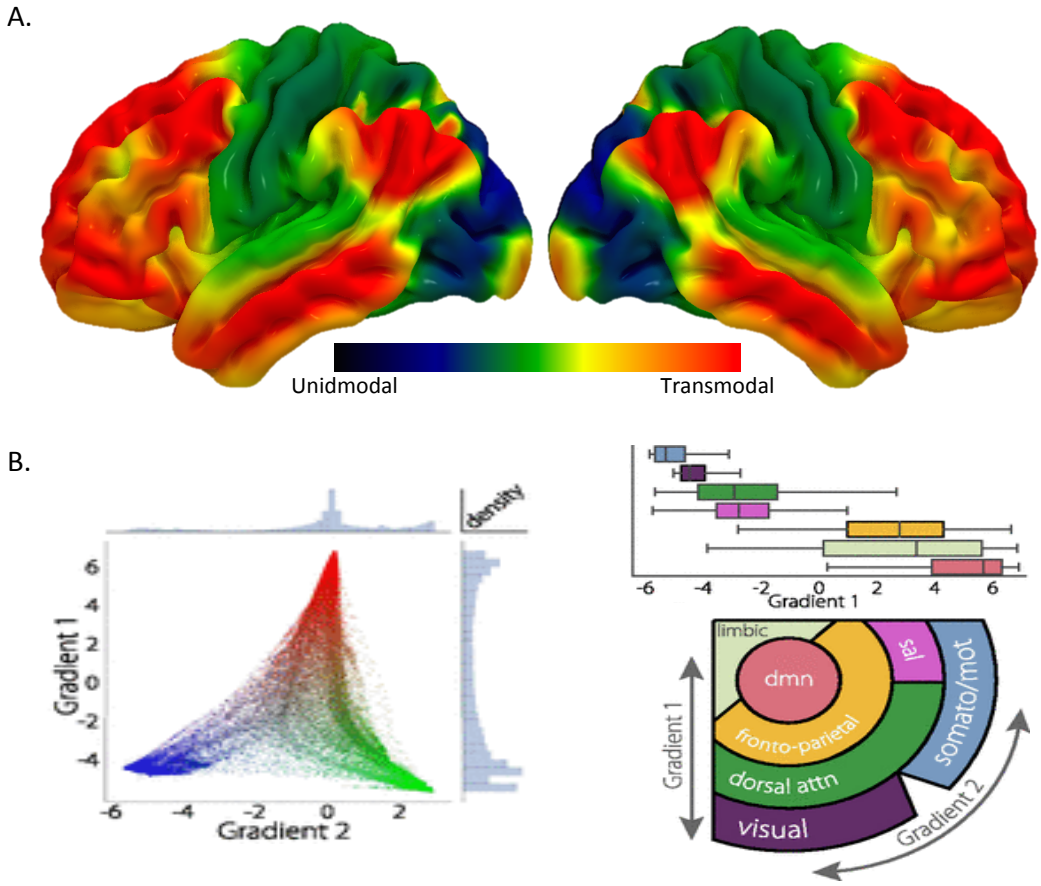

Figure 1.6. Principal gradient adapted from Margulies et al (2016). (A) The principal gradient of connectivity in humans, which captures the functional spectrum from perception, and action in unimodal regions (dark blue) to more abstract cognitive functions in transmodal cortex (red). (B) Gradient 1 extends between primary sensorimotor (blue-green) and transmodal regions (red). Gradient 2 separates somatomotor and auditory cortex (green) from visual cortex (blue). The right hand panel summarises the order of networks in the gradient and provides a schematic of the spatial relationships of intrinsic resting-state networks. dmn, default-mode network; dorsal attn, dorsal attention network; sal, salience network; somato/mot, somatosensory/motor network.

This principal gradient account suggests that aspects of cognition that are less related to input (e.g., perceptually-decoupled semantic retrieval) might be easier in regions whose neural computations are functionally independent, or distant, from systems important for perceiving and acting in the here and now (Buckner et al. 2009; Bullmore & Sporns, 2009; Hagmann et al. 2008; Margulies et al. 2016). As such, it has been suggested that the DMN integrates information from

multiple sources in memory to create states that are not reducible to the moment they take place. This decoupling hypothesis therefore provides a framework for how internally generated representations may be supported. That is, thinking about concepts not present in the environment, such as imagining lying on a beach during the daily commute, is made possible in regions that are functionally isolated from input (as this sensory information is not necessary for thinking about a beach). Therefore, the explanatory power of this gradient account is emphasised in its ability to explain common neurocognitive processes underpinning perceptually-decoupled and conceptually-guided cognition in both the DMN and semantic network. For instance, the gradient account explains why regions, such as ventrolateral ATL, would be optimal for the abstraction of conceptual knowledge away from sensory-features into transmodal representations as proposed by several semantic theories (Lambon Ralph et al. 2017; Patterson et al. 2007; Rogers et al. 2004) – this region sits close to the top of the representational hierarchy and receives convergent inputs from all other sensory networks. In addition, this account situates the angular gyrus, as well as other core DMN regions such as lateral middle temporal gyrus (MTG), at the furthest end of the gradient, suggesting they play an integral role in the most abstract forms of cognition, including internally-oriented retrieval and manipulation of knowledge (Binder et al. 2009; Cabeza et al. 2011).

Another striking observation is the gradient's ability to align with other influential explanations regarding the functional architecture of the DMN and transmodal cortices. Braga and colleagues (2013) suggest that if heteromodal regions, captured by the DMN and semantic network, converge information to form coherent cognition, perception and behaviour then they should show traces or "echoes" of information from the intrinsic networks that they are converging. Using fMRI they were able to show that unimodal portions of the cortex, such as primary sensory and motor cortices contain traces of few or even single networks, while the existence of multiple "echoes" were identified within transmodal regions. Therefore the presence of multiple echoes within transmodal regions provides an explanatory framework for why these regions are involved in multiple different cognitive states (see also Leech et al. 2012; de Pasquale et al. 2012; Smith et al. 2009). This is

consistent with the principal gradient account, which suggests that transmodal regions sit at the top of a representational hierarchy and thus receive convergent input from all other networks beneath them. Consequently, both of these bodies of work highlight that the unique macroscale organisation of transmodal regions permits them to mediate cross-talk between the brain's functional networks.

Taken together this established body of research outlines how transmodal regions of the brain are pivotal for processing abstract conceptual knowledge and perceptually-decoupled mental representations by virtue of their macroscale organisation and intrinsic connectivity (Buckner & Krienen, 2013; Braga et al. 2013; Leech et al. 2012; Margulies et al. 2016; Mesulam, 1998; Plaut, 2002). However, there is little empirical evidence regarding the pattern of activity within these regions during different forms of perceptually-decoupled semantic retrieval. Therefore several important questions remain unclear. First, if transmodal regions contain 'echoes' of the systems they integrate, for example in the form of differential connectivity to these systems across each cortical region contributing to this network, transmodal regions of cortex might still contain information about the sensory properties of internally driven experiences (Leech et al. 2012). In the case of semantically-driven imagery, one might expect that the ATL hub would play a crucial role in the generation of different types of features in a perceptually-decoupled way (e.g., thinking about what a dog looks like versus what it sounds like); however, if there is graded convergence of information in this region, different forms of imagery might not maximally evoke the same portions of ATL. Both machine learning methods, such as MVPA, and investigation of the intrinsic connectivity of transmodal regions will be integral to addressing these outstanding questions.

Second, the framework suggests that transmodal regions located farthest from input systems would be optimal for (i) abstract conceptual processing and (ii) perceptually-decoupled retrieval of knowledge from memory. However, it is not well specified if this network shows preference for either of these processes, or alternatively whether it prefers the combination of abstraction and perceptually-decoupled states. Moreover, it is unclear how different components of transmodal cortex are recruited during perceptually-decoupled retrieval of abstracted

conceptual knowledge. For instance, given that both the angular gyrus and ATL are two prominent regions across DMN and semantic networks, it is not clear what their precise contributions are. According to Margulies and colleagues (2016), the angular gyrus sits slightly higher in the representational hierarchy than ventrolateral ATL. As it is situated farthest from unimodal systems, this framework suggests it may play a role in the most abstract forms of cognition; one could therefore predict that ventrolateral ATL may be involved in transmodal conceptual processing both in the presence and absence of input, whereas the angular gyrus is more influential when cognition is orientated inwardly to memory.

Third, although the DMN has been shown to activate to several aspects of cognition (Spreng, 2012), including semantic processing (Binder et al. 2009; Davey et al. 2015; 2016; Humphreys et al. 2015; Irish & Piguet, 2013; Krieger-Redwood et al. 2016), autobiographical and episodic recollection (Andrews-Hanna, 2012; Buckner et al. 2008; Rugg & Vilberg, 2013), working memory (Konishi et al. 2015; Spreng et al. 2014; Vatansever et al. 2015), mental imagery (Hassabis et al. 2007), self-generation of emotion (Engen, Kanske & Singer, 2017) and imagining the future or recalling the past (Huijbers et al. 2009; Maguire, 2001; Schacter et al. 2007; Spreng et al. 2009; Svoboda et al. 2006; Szpunar et al. 2007), it remains unclear the specific circumstances in which regions of the transmodal DMN shows above baseline activation during semantic tasks. For instance, is it the general retrieval of knowledge from memory that activates the DMN or the richness of the concept being retrieved?

1.3.3. Summary of Perceptually-decoupled Semantic Retrieval

Semantic retrieval extends beyond the here-and-now, to draw on abstract knowledge that has been extracted across multiple experiences; for instance, we can easily bring to mind what a dog looks and sounds like, regardless of whether or not there is a dog present in our immediate environment. However, a clear understanding of the neural substrates that support patterns of semantic retrieval that are not immediately driven by stimuli in the environment is lacking. Given the anatomical and functional similarity between the semantic network and DMN

(Binder et al. 2009; Humphreys et al. 2015; Margulies et al. 2016), this thesis is motivated by the hypothesis that there might be common neurocognitive processes underpinning perceptually-decoupled and conceptually-guided cognition in the DMN. Firstly, perceptually-decoupled states might recruit brain regions whose neural computations are functionally independent, or distant, from systems important for perceiving and acting. Secondly, DMN regions might support higher-order representations with predictive value across multiple situations and modalities, which integrate features from diverse sensory-motor regions. This view is consistent with the observation that the DMN and many portions of the semantic network lie at the extreme end of a gradient from heteromodal to unimodal cortex (Margulies et al. 2016), since increasingly abstract and complex representations might be formed at greater distances along the gradient, as the influence of specific features and modalities is reduced (Buckner & Krienen, 2013; Margulies et al. 2016; Mesulam, 1998). However, it is not clear if these aspects of processing recruit identical regions and if they interact. Therefore, the second goal of this thesis is to better understand how we can retrieve semantic concepts from memory, in the absence of external input. The combination of machine learning methods, intrinsic connectivity and univariate activation will be integral to addressing the following questions: (i) Which portions of unimodal sensory cortex are necessary for the processing of perceptually-decoupled semantic features (e.g., thinking about what a dog sounds or looks like)? (ii) How are transmodal regions recruited during perceptually-decoupled retrieval of knowledge from memory; for instance, do they show global increases in activation to internally-generated semantic content, or is their contribution reflected in the multivariate pattern of activity? and (iii) given transmodal cortex is theoretically optimal for abstract and complex representations from memory, it remains unclear whether transmodal regions respond in a similar manner to conceptually-guided and perceptually-decoupled cognition, but it is unclear whether these two factors interact.

1.4. Conclusion and Research Aims

Perceptually-decoupled semantic cognition requires a store of transmodal conceptual knowledge and an ability to disengage from the external world and orientate cognition inwardly on previously encoded information. Our store of conceptual knowledge is arranged such that we have (i) "spokes" (sensory-motor regions of the cortex) and (ii) a transmodal convergence zone or "hub" (ventrolateral ATL). The role of the hub is to abstract away from modality-specific attributes and generalize across semantically similar concepts. Despite a large body of evidence - spanning neuroanatomical, neuropsychological and neuroimaging fields - supporting the graded-hub account of semantic processing, it remains unclear *how* conceptual knowledge is represented within this semantic network; such as, (i) the exact portion of ATL that is transmodal, (ii) whether traces of sensory systems are still seen even in the most transmodal parts and (iii) how the transmodal ATL is supported by other transmodal regions in the DMN. Furthermore, perceptually-decoupled retrieval of memory is thought to rely on transmodal regions allied to or within the DMN. These regions are thought to be critical for higher-order representations with predictive value across multiple situations and modalities, which integrate features from diverse sensorimotor regions. Moreover, aspects of cognition that are less related to input (e.g., perceptually-decoupled semantic retrieval) might be easier in regions, such as the DMN, whose neural computations are functionally independent, or distant, from systems important for perceiving and acting in the here and now. Nevertheless, the precise mechanisms behind perceptually-decoupled retrieval of semantic memories remain elusive. This thesis will utilise a combination of MVPA, intrinsic connectivity measures and standard univariate analysis in an attempt to understand (i) how concepts are represented in the brain and (ii) how transmodal mechanisms adapt for perceptually-decoupled memory states. The specific aims of this thesis were:

- To explore the role of unimodal sensory-motor cortex in semantic representations. Using MVPA, chapters 3 and 4 will explore whether the patterns of activity in unimodal cortex represent modality of presentation (spoken vs. written; Chapter 3), modality of word meaning (loud vs. bright;

Chapter 3) and modality of perceptually-decoupled retrieved memory (thinking about what a dog sounds like vs. what it looks like; Chapter 4).

- To localise where within the ATL abstract heteromodal semantic representations are supported. Using a searchlight MVPA, Chapter 3 investigates where, within the ATLs, patterns of activity for a concept activated through the visual domain map on to the same concept activated through the auditory domain. This cross-classification will permit the localisation of a transmodal 'hub' region that captures abstract meaning irrespective of presentation modality. Moreover, analysis of the intrinsic connectivity of this region will provide evidence of whether this region is embedded within a network that facilitates abstract transmodal processing (e.g., default mode network).
- To investigate perceptually-decoupled semantic retrieval states such as imagination (Chapter 4) and judgements from memory (Chapter 5). Chapter 4 examines perceptually-decoupled forms of semantic retrieval and establishes which brain regions can decode sensory features in imagination, in the absence of input. Chapter 5 assesses the conjunction of cognitive processes that require (i) multi-featural abstract concepts and (ii) decoupling. Notably, this later study combines the two dominant features of chapter 3 (abstract conceptual representations) and perceptual-decoupling (retrieval of knowledge from memory in the absence of input). Both of these experiments will interrogate whether unimodal and/or transmodal regions are necessitated by perceptually-decoupled semantic retrieval and measure the intrinsic connectivity of these brain regions.
- To identify circumstances in which regions of the transmodal default mode network shows above baseline activation during semantic tasks. This is addressed in Chapter 5 by comparing semantic conditions (multi-featural concepts) with perceptual conditions (simple colour patches) in a univariate fMRI analysis.

Chapter 2 – fMRI Methods Review

This thesis adopts multiple functional magnetic resonance imaging (fMRI) analysis methods to investigate how semantic information is represented and processed in the brain. fMRI measures neural responses, *in vivo,* by tracking changes in the blood flow associated with increased neural activity. As neurons increase their firing rate they consume energy reserves. To replenish these reserves oxygenated blood is transferred to those cells via the bloodstream, thereby increasing the local supply of oxygenated blood in that region. fMRI provides a measure of this oxygenation using a contrast called Blood Oxygenation Level-Dependent signal (BOLD; Ogawa et al. 1990). The BOLD signal therefore provides an indirect measure for investigating the underlying neural activity. This method is one of the most widely used techniques for studying the function of the brain, and as such will be utilized in the current thesis. This chapter will outline the statistical fMRI methods applied within this thesis that broadly fall under two categories: (i) detection of functional activation and (ii) detection of functional connectivity. The final section of this chapter will discuss additional approaches proposed to merge across functional-activity and connectivity measures.

2.1. Functional Activity

2.1.1. Univariate Analysis

The most popular statistical approach for analysing fMRI data is referred to as a univariate general linear model (GLM) approach (Friston et al. 1994). This method aims to detect voxels that are activated during a specific stimulus condition by analysing each voxel's time series independently (known as a "univariate analysis": Huettel, Song & McCarthy, 2004; 2009).

2.1.1.1. Univariate Analysis Statistics

The GLM approach sets up a model derived from the timing of the task in the MRI scanner, which creates the general pattern you would expect to see in the data (i.e.,

periods when the stimulus condition were present and periods when it was absent) (Monti, 2011). This model is referred to as a box-car model. Next, because the hemodynamic response is a delayed process, each of the regressors is convolved with a hemodynamic response function (HRF), which mimics the effect of the brains neurophysiology and when convolved with the model produces an expected time-series of response. This convolved model can then be regressed against the original fMRI BOLD signal on a voxel-by-voxel basis, which produces a whole-brain statistical map of parameter estimates (regression coefficients). These parameter estimates reflect how well the model fitted the fMRI data; a large coefficient value will be given if a particular voxel responds strongly to the stimulus condition as it will be predicted well by the model, whilst a smaller coefficient value will be given if a voxel is not responsive to the stimulus condition and thus is predicted poorly by the model (Monti, 2011). This pipeline is illustrated in Figure 2.1. Many stimulus conditions can be modeled in parallel, and thus univariate analysis allows for conditions to not only be contrasted against rest periods but also directly contrasted against one another (e.g. condition A > condition B). From here, statistical maps can be combined across scan sessions and / or subjects (Holmes & Friston, 1998; Worsley et al. 2002). In order for datasets to be comparable within and across subjects, individual results are warped into a common reference space (such as MNI152). This allows corresponding structures to be aligned across subjects with differing brain anatomy (Brett et al. 2002). The final step of the univariate analysis is to threshold the statistical maps in order to determine which parts of the brain were activated, above a given threshold (e.g. z-statistic or p-value) (Huettel, Song & McCarthy, 2004). These thresholded statistical maps therefore reveal brain regions whose BOLD signal correlated with the stimulus condition under examination.

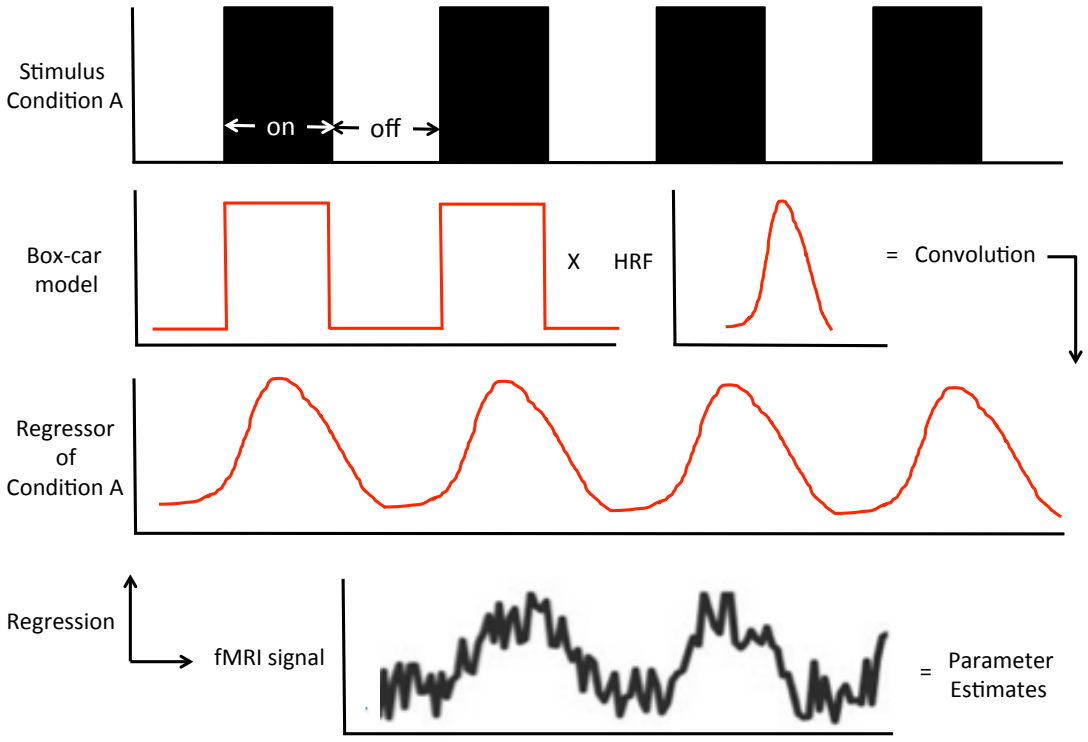

Figure 2.1. Example of univariate general linear model (GLM) analysis of functional MRI data. Top panel indicates the general pattern you would expect to see in the data for stimulus condition A (i.e., periods when the stimulus condition were present and periods when it was absent). A box-car function is then defined that corresponds to these periods of stimulus presentation. A hemodynamic response function (HRF) is then convolved with the box-car model to produce a hemodynamic regressor of stimulus condition A. Finally, this is regressed against the fMRI signal for each voxel independently. This results in a parameter estimate (e.g., statistical map) that indicates the fit of the regressors to the fMRI BOLD signal at each voxel.

An additional process that is often undertaken is a region of interest (ROI) analysis (Friston et al. 1994; Poldrack, 2007; Poldrack et al. 2011). This method adopts the same univariate principles as discussed above, however it looks at the average activity of all voxels within a pre-defined region. There are several reasons why one might perform an ROI analysis, for instance adopting an ROI analysis can control for type 1 error by reducing the number of statistical tests to a limited

number of ROIs rather than all voxels in the entire cortex. Moreover, researchers can explore their data in pre-defined regions from the literature or a separate functional localizer to help guide their analysis and shape interpretation (Poldrack, 2007). This is particularly useful when you have a complex design (e.g., factorial design) where there are multiple conditions and it is unclear which is driving the effect. Using an ROI based approach you can extract the signal for each condition to depict the pattern of signal across all conditions. This univariate GLM method has been considered the 'gold standard' in fMRI research (Mahmoudi et al. 2012; Norman, Polyn, Detre & Haxby, 2006) and has therefore been widely used for identifying brain regions associated with semantic processing and internally generated thought.

2.1.1.2. Limitations

Despite the wealth of knowledge that conventional univariate GLM-based analyses have provided, they are not without limitations. The first draw back to this method is the assumption that covariance across voxels does not convey information about the cognitive function under investigation. As univariate analysis focuses on voxel-by-voxel activity it neglects the information that may arise from the pattern of distributed neighbouring voxels. In traditional univariate methods this information is considered uncorrelated noise and is normally dealt with by using spatial filters that smooth BOLD signal across neighbouring voxels. Secondly, because these methods spatially average across neighbouring voxels that respond significantly to that stimulus condition only responses that differ significantly from zero reflect stimulus-related neural activity. Therefore univariate analyses assume that sub-threshold voxels or voxels that only marginally deviate from zero are non-informative. Furthermore, this spatial averaging blurs out potential fine-grained spatial patterns present across neighbouring voxels that may be informative for discriminating between stimulus conditions. This limits the understanding of dimensional structure that may be present across subthreshold voxels and / or multiple spatially distributed voxels (Norman et al. 2006). For instance, research employing standard univariate methods has shown that the anterior temporal lobes (ATL) respond to both written forms of an object (e.g. APPLE) and picture forms of

the same object (Visser et al. 2012). However, it is unclear whether the pattern of activity is the same for these two modalities. To overcome the limitations of univariate methods and address questions regarding the patterns of activity that represent content, many researchers have begun implementing multivariate methods within fMRI research (as well as EEG and MEG; Chan et al. 2011; Murphy et al. 2011; Simanova et al. 2010; Wang et al. 2004).

2.1.2. Multivariate Pattern Analysis (MVPA)

Whilst univariate methods are sensitive to the mean activity of individual voxels (Friston et al. 1994; Poldrack et al. 2011), multivariate approaches are sensitive to the distributed patterns of BOLD fMRI across multiple voxels (Cox & Savoy, 2003; Haxby et al. 2001; Haynes & Rees, 2006, Norman et al. 2006). These methods are often referred to as Multi-Voxel or Multivariate Pattern Analysis (MVPA). Because MVPA methods are sensitive to distributed patterns of activity, and thus able to capture multidimensional effects (i.e., different voxels within a region carry non-identical information about stimulus conditions (Diedrichsen et al. 2013)), they can overcome some of the limitations of conventional univariate approaches. For instance, weak information available at each voxel that would be considered sub-threshold in a standard univariate analysis can be gathered in an efficient way across many voxels using MVPA. Moreover, even if two voxels or ROIs are not informative to the stimulus condition in a univariate analysis, they might do so when analysed together (for review see Haynes & Rees, 2006). The following sections will provide a brief summary of the different varieties of MVPA available, and conclude which method will be adopted in the current thesis.

2.1.2.1. Correlation-based methods

Haxby and colleagues (2001) first applied MVPA to fMRI data, where it was used to demonstrate that distributed neural patterns of activity could distinguish between object categories. In this seminal paper, the authors used a simple correlation-based MVPA approach to identify reliable neural profiles in the ventral temporal (VT) cortex evoked by each object category. There are three basic steps undertaken

in this correlation analysis. Firstly, MVPA is performed across a subset of voxels (i.e., a pre-defined ROI) rather than the whole-brain volume. This step is known as *feature selection,* where features refer to brain voxels. The second step, *pattern assembly,* involves organizing the fMRI data into discrete 'brain patterns'. Parameter estimates are therefore calculated for each stimulus condition for each experimental run, (as per the univariate analysis). An advantage of using the GLM to produce parameter estimates is the ability to include motion predictors in the model in order to obtain better estimates of the pattern of activity (Mur et al. 2009).

To ensure that patterns of activity are not contaminated by a high degree of shared variance across the stimulus conditions, such as global and local variations in baseline hemodynamic response, the data is typically normalized. Normalization of the data can be achieved by de-meaning (subtracting the mean across all stimulus conditions per-voxel from each stimulus condition) or z-scored normalization (in addition to de-meaning each stimulus condition is divided by the per-voxel estimate of the standard deviation across conditions). Both procedures remove the shared variance across conditions, and therefore better isolate the unique variance in each stimulus condition (although z-scored normalization is more commonly used; Misaki, Bandettini & Kriegeskorte, 2010). As a result, unlike univariate methods where the main component of interest is the subject-specific variability in mean activation across an ROI, MVPA captures the latent multidimensional code across multiple voxels within a brain volume, and is therefore insensitive to the mean activation (Davis et al. 2014).

Next, to determine if these patterns of activity are reliable, it is vital to compare stimulus conditions across independent estimates of the neural response (i.e., across independent runs). This process is known as cross-validation. To ensure that each split of the cross-validation process remains independent, normalization must be computed within each split independently (Kriegeskorte et al. 2009). Once the normalized patterns of activity have been generated, the third step is to conduct the correlation MVPA. For each subject, the first-half patterns are correlated with the second-half for all conditions. This was the first study to show

that reliable distributed patterns of activity capture multidimensional information about stimulus conditions.

2.1.2.2. Classification methods

Since its advent, more sophisticated methods for pattern recognition have emerged. Machine learning techniques originally developed for pattern classification of a variety of domains, such as handwriting recognition, have now been successfully applied to fMRI data analysis (Cox & Savoy, 2003; Hanson et al. 2004; Haynes & Rees, 2006; Norman et al. 2006, Pereira, Mitchell & Botvinick, 2009). As these algorithms are more sophisticated than the original correlation method they have the potential to provide greater sensitivity to capturing multidimensional effects of stimulus conditions. Classification methods apply a function that takes the values of multiple features (i.e., voxels) in a set of independent variables known as an example and predicts the class (i.e., stimulus condition) that the example belongs to (the dependent variable). The initial stages of a classification analysis are similar to those for a correlation method. Firstly, *feature selection* is undertaken where the voxels included in the classification analysis need to be determined (see Figure 2.2a).

Secondly, *pattern assembly* is necessary to produce estimates of the patterns of response for independent splits of the data. Here parameter estimates are created in line with univariate analysis approach, these are then normalized, as per the correlation method. Furthermore, cross-validation is necessary in order to examine how well the model, derived from the classification algorithm, will generalize to new stimulus examples. This results in independent training and testing sets (Figure 2.2b). Since Haxby and colleagues (2001) use of split-half cross-validation more powerful methods for cross validating have been proposed. The most popular of which is known as leave-one-run-out cross-validation, where the original data set is partitioned into individual runs and a single run is retained for validating the model. The model is therefore trained on N – 1 subsamples of the data (where N = number of runs). The cross-validation method is then repeated N times, where each time a different run is removed from the training sample, so that each run is the independent left-out run exactly once. The N results can then be

averaged into a single estimate of classifier accuracy. This method is considered superior to split-half as it is generally recommended to choose a larger training set in order to enhance classifier convergence (Kreigeskorte et al. 2009; Mahmoudi et al. 2012). Once pattern assembly has been achieved the following stages of the classifier analysis differ substantially from the correlation approach.

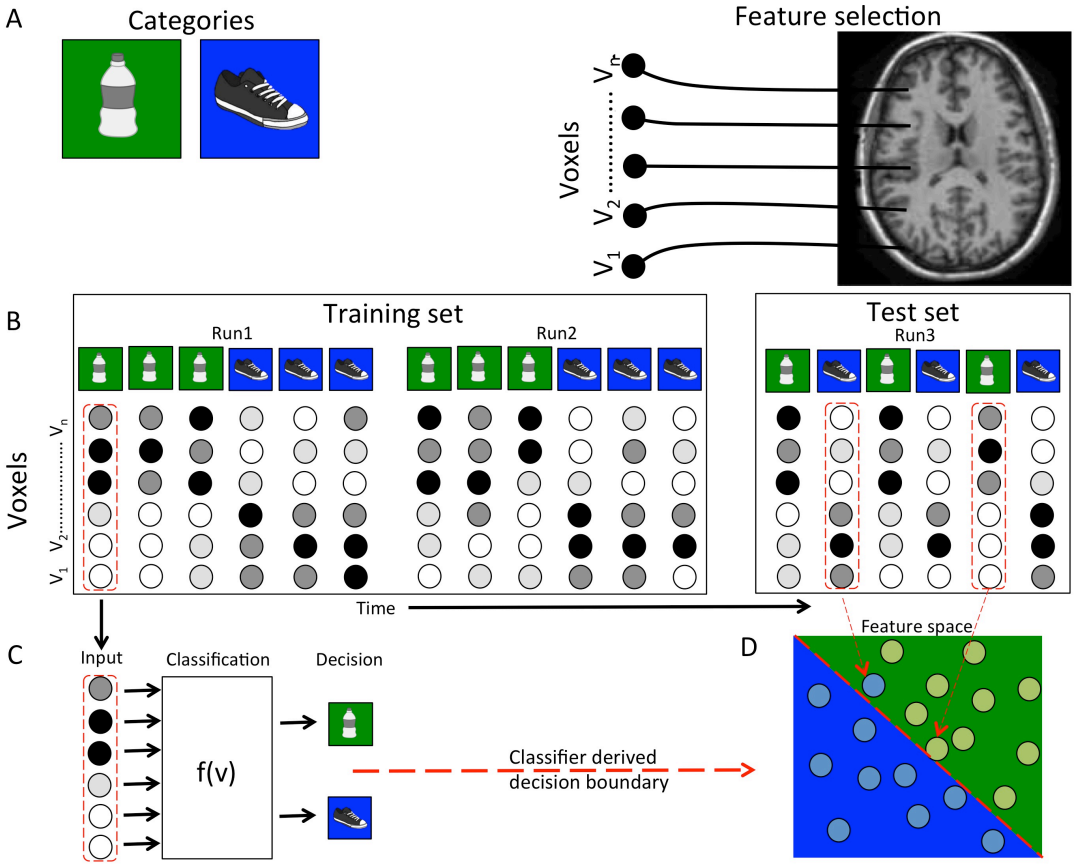

Figure 2.2. Illustration of classification-based MVPA adapted from Norman et al. (2006). (a) Participant views blocks of objects from two conditions (bottles and shoes). *Feature selection* is then used to identify which voxels will be included in the classification analysis. (b) *Pattern assembly:* The BOLD fMRI time series is decomposed into discrete patterns of activity, which produces estimates of the patterns of response for independent splits of the data. Here parameter estimates are created in line with univariate analysis approach. Each pattern of brain activity is labelled with its corresponding stimulus condition (e.g., bottle or shoe). These patterns are then divided into independent training and testing sets. (c) *Classifier training:* a classifier function is then trained on the training set to map between brain patterns and stimulus conditions. (d) Red-dashed line represents the classifier

derived decision boundary in the high-dimensional space of voxel patterns (depicted here in 2-D for illustrative purposes). Each dot denotes the associated pattern of activity and the colour of the dot indicates its stimulus condition (shoe or bottle). The background colour corresponds to the guess the classifier makes for patterns in that region (blue = shoe; green = bottle). The trained classifier is then used to predict category membership for patterns from the independent test set. The figure shows (i) a successful classification where a bottle pattern (green dot) was accurately classified as a bottle, and (ii) an unsuccessful classification where a shoe pattern (blue dot) was misclassified as a bottle.

The third step in the classifier analysis is known as *classifier training,* which involves passing a set of labelled patterns into a MVPA classification algorithm (known as the training set). This algorithm learns a function that maps between the stimulus conditions and the voxel activity patterns. Within this step, the most pivotal decision is choosing which classifier to train. There are many different classification algorithms available (for review see Duda et al. 2012). In general these classifiers can be split into two branches: (i) linear classifiers, such as support vector machine (SVMs) (Correia et al. 2014; LaConte et al. 2005; Kamitani and Tong, 2005) and Gaussian Naïve Bayes classifiers (GNB) (Mitchell et al. 2004) and (ii) nonlinear classifiers, such as non-linear SVMs (Cox & Savoy, 2003; Kamitani & Tong, 2005; Kragel, Carter & Huetel, 2012). The main difference between these two branches is that nonlinear classifiers can capture multidimensional conjunctions between features in a way that differs from their response to individual features. Consequently, nonlinear classifiers are considered potentially more powerful than linear classifiers, as they are more flexible in the types of mappings they can learn (Norman et al. 2006). However, a growing body of research has shown that nonlinear classifiers perform no better (and occasionally worse) than linear classifiers (Cox & Savoy, 2003; LaConte et al. 2005; Misaki et al. 2010). This poorer performance has been attributed to overfitting, as the decision boundary is more flexible and therefore is likely to capture the idiosyncrasies of the noise in the training data (Duda et al. 2000; Kriegeskorte et al. 2009). Furthermore, it has been suggested that a 'good performance' in a nonlinear classifier is more difficult to

interpret than a 'good performance' in a linear classifier (for discussion see Norman et al. 2006; Kamitani & Tong, 2005). Therefore, due to their simplicity and sensitivity, linear classification is the most commonly used and successful pattern classifier analysis in neuroimaging so far (Corriea et al. 2014; Countanche & Thompson-Schill, 2014; Cox & Savoy, 2003; LaConte et al. 2005; Misaki et al. 2010; Mur, Bandettini & Kriegeskorte, 2009; Norman et al. 2006). In light of this evidence, the current thesis will focus on using linear classifiers.

Once trained, all linear classifiers place a separating decision boundary (i.e., hyperplane) between different conditions in the multidimensional space. In Figure 2.2d the multivariate pattern of activity for two experimental conditions are represented by blue (shoe condition) and green (bottle condition) dots. The red dashed line separating these two conditions is the classifier derived decision boundary. However, where the boundary is placed differs subtly across the different linear classification methods (Mur et al. 2009). Two of the most commonly used linear classifiers are SVMs and GNBs (for extensive discussion on the mathematics of a wide variety of pattern classification techniques see Duda et al. 2001).

Linear SVMs weight each voxels activity and then project these patterns onto a linear discrimination dimension in order to maximize the margin (i.e., distance of the closest data point to the hyperplane) (see Figure 2.3 adapted from Mur et al. 2009). Linear SVM classifiers do not assume multivariate normality. Instead, the margins are widened on either side of the decision boundary, until the margin cannot be widened anymore without including more than one stimulus conditions data points (i.e., blue dots should be on one side of the decision boundary and red on the other). The triangles indicate the patterns of activity closest to the decision boundary. These points (known as "support vectors") define the margins and the decision boundary is placed in the middle of these. Whichever side of the decision boundary a data point is closet to will be the stimulus condition to which that point will be assigned.

In contrast, GNB classifiers model each voxels activity as Gaussians (Ng & Jordan, 2002; Raizada & Lee, 2013). Firstly this method calculates the z-score distances for each stimulus condition; therefore the model takes into account not

only the distance of each voxel from the mean but also how it compares to the stimulus condition variance (see Figure 2.3). The "Gaussian" part of GNB is highlighted in the assumption that the stimulus conditions have Gaussian normal distributions. Each z-score distance is then converted into a p-value; probability of observing a given data point if that data point were drawn from the distribution of a particular stimulus condition. However, the purpose of a classifier is to calculate the probability of a stimulus condition, given our observed data. Therefore the "Bayes" part of the GNB classifier adopts Bayes theorem to derive each from one another (predict stimulus condition from data *and* predict data from stimulus condition). Finally, the "Naive" part of GNB treats all of the input dimensions as independent from each other. As a result GNB classifiers do not model the covariance of the dimensions. In sum, the GNB classifier calculates distances from stimulus-condition-centers and positions a decision boundary halfway between the two centers. Whichever center a data point is closet to will be the stimulus condition to which that point will be allocated.

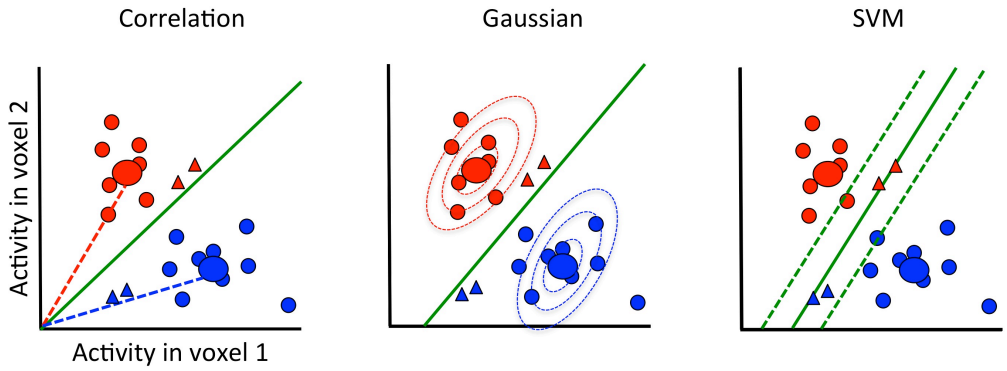

Figure 2.3. Hypothetical illustrations of classification by different classifiers adapted from Mur et al. (2009).

To determine the most powerful linear classification method, several studies have directly compared the sensitivity of linear SVMs and GNBs, as well as several other classifiers (e.g., fishers linear discriminant analysis (LDA) and k-nearest-neighbour classifiers (KNN); Duda et al. 2001). Mitchell et al. (2004) present three studies where classifiers were used to distinguish cognitive states between (i) pictures or sentences, (ii) ambiguous or non-ambiguous sentences and (iii) category

of words (e.g., food, buildings etc.). They found that, despite having a large number of features (> 100,000) and less than a dozen noisy training examples per stimulus condition, linear SVMs out performed both GNB and KNN classifiers (with KNN coming off as the worst performing across all three studies). Notably, they illustrated that SVMs performance increases when the number of features is reduced, and the number of training examples increases. Similar results have been shown in high-resolution imaging of BOLD signals in monkeys (Ku et al. 2008). Here, the authors compared classifier performance across both single trials and averaged across numerous stimulus presentations. In general, GNB classifiers performed significantly worse than all other classifiers across all modes of testing, while SVMs and LDAs consistently out-performed correlation and GNB classifiers. However, Ledoit & Wolf (2003) illustrated that linear SVMs handle limited data in high-dimensional space more effectively than LDAs (which may require an additional regularized covariance estimate). As fMRI data typically produces many more features than examples, SVMs are potentially better equipped at dealing with this limited data than LDA. Taken together, these findings suggest that linear SVMs are arguably the most powerful classification method; it is therefore not surprising that linear SVMs are the most popular classification method (Mur et al. 2009).

However, to perform a linear SVM classification successfully it is critical to effectively reduce the number of features used in the classification analysis. This is particularly important given the findings that a classifier performance is known to degrade in the presence of many irrelevant features, known as over-fitting, particularly when the number of training sets is somewhat limited, as in typical fMRI studies (Guyon & Elisseef, 2003, Kohavi & John, 1997, Norman et al. 2006). A common method for reducing the number of features (i.e., voxels) is known as a searchlight-analysis (Kriegeskorte et al. 2006). Here, instead of analysing distributed fMRI activity patterns contained within all the voxels in a predefined ROI - that could potentially contain hundreds of thousands of voxels (Cox & Savoy, 2003; Haxby et al. 2001) - a spherical “searchlight” can be moved throughout a functional volume to continuously map informative regions. The searchlight is centred on each voxel in turn. The size of the searchlight can vary (see Kriegeskorte et al. 2006) but typically a 6-mm radius is used, which is comprised of 123 2 mm voxels. At each

location, a multivariate classifier can then be computed on the pattern of activity across neighbouring voxels within the searchlight sphere. The accuracy of each sphere is then allocated to the central voxel of that sphere. Resulting in a whole-volume accuracy map that depicts, for each voxel in the volume, how well the pattern of activity in the local spherical neighbourhood decodes between the stimulus conditions. Moreover, as pattern classification has been shown to be most sensitive to information in contiguous macroscopic regions, the use of a searchlight is optimal for sampling of informative voxels (Kriegeskorte et al. 2006).

So far this chapter has discussed the first three stages of the classification method: *feature selection, pattern assembly* and *classifier training*. The fourth step in the classification method is known as *generalization testing.* Here the classification model is fed a new unseen pattern of brain activity and the classifier has to determine which stimulus condition is associated with that pattern. In the case of leave-one-run-out cross-validation the 'unseen' data refers to the independent left out run. Accuracy scores are averaged across N folds of the cross-validation (N = number of runs), resulting in whole-volume accuracy maps that depict, for each voxel in the volume, how well the pattern of activity in the local spherical neighbourhood decodes between the stimulus conditions.

The final step in the classification method, *statistical thresholding,* is to determine which results are statistically significant. Typically, this is done across all participants in order to make group-level inferences (Raizada & Lee, 2013). First, for each individual searchlight map, the chance-level score (50% for binary classification) is subtracted from the classification accuracy score (e.g., 61%) assigned to each voxel. This adjusted-score map can then be submitted to a standard random effects analysis. Here searchlight maps across all participants can be thresholded and corrected for multiple comparisons using family wise error (FWE) and / or false discovery rate (FDR) (Kragel et al. 2012; LaConte, 2005; Raizada & Lee, 2013). The resulting maps indicate where, within the brain volume, voxels could decode between stimulus conditions at significantly above chance-level (e.g., 50%).

2.1.2.3. Limitations

Despite its superior ability to detect distributed patterns of activity across multiple voxels, it is worth briefly noting that pattern classification is not without limitations. One drawback of the MVPA method is voxel resolution, as the spatial information about neural populations codes, provided by fMRI, is limited. For instance, the BOLD signal in a 2mm voxel (common voxel dimension for 3T fMRI studies) carries information about the underlying neural activity of *many* voxels. The fine-grained activity patterns across these voxels may not reflect precise fine-grained patterns because of hemodynamic blurring and distortion (Mur et al. 2009). This caveat alters the interpretation that can be drawn from fMRI MVPA, specifically, that a change of activity patterns across stimulus conditions can only be interpreted as a change of neural *population* activity. The use of high-resolution MRI scanners (7T) has been proposed as a way to reduce this limitation, however this is beyond the resources available within the current thesis.

2.1.2.4. Summary of MVPA methods

In light of the methods outlined above it is clear that there are many different forms of MVPA, the following section will briefly outline the specific classification methods that will be adopted in the current thesis. Firstly, after reviewing the strengths and weaknesses of different MVPA methods it is evident that linear SVM classifiers are superior at decoding underlying patterns of activity associated with stimulus conditions, particularly so when feature selection is restricted (Mitchell et al. 2004; Ku et al. 2008). Linear classifiers are often praised for their simplicity (Pereira et al. 2009) and over the past decade an increase in software packages have emerged such as LIBSVM (Chang & Lin, 2011) and PyMVPA (Hanke et al. 2009), making linear SVM classifiers more standardized across the neuroimaging field. Secondly, a powerful and widely used method for reducing the number of features fed into the classifier is known as a searchlight analysis (Kriegeskorte et al. 2006). This method iterates a spherical "searchlight" across the entire brain volume, centring this searchlight on each voxel in turn. All neighbouring voxels that fall within this searchlight are included as features within the classification. As pattern classification has been shown to be most sensitive to information in contiguous macroscopic regions, the use of a searchlight has been commended for its capacity

to optimize sampling of informative voxels (Kriegeskorte et al. 2006). Finally, to improve classification performance a leave-one-out-cross-validation method is suggested. This method maximizes the amount of data the classifier trains on, improving the classifiers chance at learning the latent multidimensional patterns of activity associated with different stimulus conditions, whilst reducing the chances of over-fitting (Kriegeskorte et al. 2009). In light of this, the current thesis will adopt a searchlight-analysis, to reduce feature selection, with a leave-one-run-out cross validation method in conjunction with a linear SVM.

2.1.3. Combining MVPA and Univariate Methods

As discussed, MVPA has been successfully employed to detect a broader, and often more subtle, class of task-related effects compared to traditional univariate methods. Notably, MVPA has been successfully used to investigate the neural mechanisms underlying the representation of language-independent concepts in the brain. For instance, Correia et al. (2014) presented bilingual subjects with animal nouns in both English and Dutch. A classifier was trained to see whether patterns of activity trained to distinguish between spoken nouns in one language (e.g., “horse” vs. “duck” in English) could accurately predict the same distinction in the other language (e.g., “paard” vs. “eend” in Dutch). This cross-modal classifier revealed a significant cluster in the left superior ATL, lending to the idea that the ATL is a region that organizes conceptual information in a language-invariant manner (i.e., abstracted away from input modality). Furthermore, Coutanche & Thomspon-Schill (2014) successfully employed MVPA to reveal stable patterns of activity across several cortical regions (e.g., ATL and V4) that captured information about objects that participants were merely imagining (e.g., thinking about an orange). Notably, both of these papers highlight how these findings would not have been achieved using univariate approaches.

Oftentimes, however, in addition to exploiting MVPAs superior ability to detect stimulus-conditions researchers are also interested in comparing between MVPA and univariate results to draw inferences about how task-related effects are coded in the brain (Correia et al. 2014; Coutanche, 2013; Fairhall & Caramazza,

2013; Jimura & Poldrack, 2012; Peelen & Caramazza, 2012; Tusche, Smallwood, Bernhardt & Singer, 2014), as MVPA captures distributed coding of information whereas univariate analysis is sensitive to global engagement in ongoing tasks – all voxels changing their activity in the same direction. Therefore the results from both methods will be reported in the current thesis. Specifically, in Chapter 3 MVPA will used to address research aims 1 and 2 of this thesis: (i) understand what the patterns of activity in unimodal cortex represent (e.g., modality of input; spoken versus written words or modality meaning; light vs. bright), and (ii) localise where in the ATL representations are truly decoupled from sensory input (e.g., transmodal). While univariate analysis will be used to investigate the unimodal/transmodal properties of our MVPA findings. In Chapter 4 MVPA will examine perceptually-decoupled forms of semantic retrieval (research aim 3) and establish which brain regions can decode sensory features in imagination, in the absence of input (i.e., thinking about what something sounds like versus what it looks like), while univariate analysis will be performed to interrogate which regions are crucial for producing internally generated thoughts across all experimental conditions. Finally, to identify circumstances in which regions of the transmodal default mode network shows above baseline activation during semantic tasks (research aim 4), Chapter 5 will use univariate analysis to detect brain regions involved in tasks when information is either present or not present in the external world, and compare this to brain regions that are interested in conceptually complex (Dalmatian) versus simple (colour patch).

2.2. Functional connectivity

Since its discovery in 1990 by Ogawa and colleagues BOLD fMRI has been widely used as a tool for detecting functional activation: univariate methods capture global activity and MVPA approaches capture distributed patterns of activity. Whilst these forms of analysis have been integral for understanding the function of specific brain regions they fail to fully capture the interactions between spatially distributed regions that arguably underpin these brain activations. It therefore remains unclear whether there is a core network within the brain that underlies conceptual

knowledge across an array of complex stimulus domains (e.g., perceptually-coupled versus perceptually-decoupled). As our brain consists of spatially distributed, but functionally associated brain regions that continuously share information with each other (Van Den Huevel & Pol, 2010), recent methodological advances are now moving beyond the localization of task-related responses, as outlined in the previous sections, to capturing connectivity of remote brain areas. These methods provide conceptually complimentary evidence to the inferences made from task-fMRI data and as such are increasingly being applied across a variety of fields of neuroscience to further inform our knowledge of the fundamental organisation of processing systems in the human brain.

2.2.1. Resting-state connectivity

fMRI can be used as a means for studying the dynamics of neural networks by tracking BOLD response characteristics across spatial and temporal scales (Aertsen et al. 1989; Logothetis et al. 2001; Lowe et al. 2000). This technique, known as resting-state fMRI or functional connectivity, captures the co-activation (i.e., temporal dependence) of neuronal activation in distributed brain regions across the fMRI time series, during rest (Lowe et al. 2000). This co-activation is taken to reflect "functional communication" between brain regions (Damoiseaux et al. 2006; Salvador et al. 2005; Scholvinck, et al. 2010). Investigating the resting-state connectivity highlights the ongoing spontaneous brain activity in the absence of cognitive or sensory stimulation (Scholvinck, et al. 2010). Since the seminal paper by Biswal et al. (1995) where functional connectivity was first used to show that BOLD signal time course within motor cortex strongly correlated with the contralateral motor region during rest, many researchers have questioned what exactly functional connectivity is capturing and, more specifically, whether it provides (i) indirect information about anatomical connectivity, (ii) correlation patterns that emerge from common inputs (i.e., 'information about neuromodulatory input from ascending neurotransmitter systems or thalamo-cortical afferents' (Van Dijk et al. 2010)) or (iii) both. The following section will briefly outline the current literature surrounding this topic (for extensive review see Damoiseaux & Grecius, 2009).

Several lines of evidence suggest that BOLD fluctuations captured by functional connectivity are constrained by anatomic connectivity. For example, Margulies and colleagues (2009) revealed that in monkeys functional connectivity corresponded to tracer injections in several distinct pathways. However, it has also been shown that functional connectivity exists between regions that do not have direct anatomical connectivity (e.g., bilateral primary visual cortex; Vincent et al. 2007). Several studies investigating cerebro-cerebellar circuits have demonstrated that functional connectivity also reflects long-range polysnaptic connections (Krienen & Buckner, 2009; O'Reilly et al. 2010). The cerebellum is connected to other cerebral regions, not via direct anatomical projections, but via long-range polysynaptic circuits. These functional connectivity studies showed that, despite a lack of anatomical connectivity, the cerebellum was correlated with activity in the motor cortex. In light of this, functional connectivity is thought to capture *both* anatomical connectivity and polysynaptic connectivity (i.e., correlation patterns that emerge from common inputs).

Over the past 15 years this method has gained popularity and has revealed several distinct networks of correlated activity in the brain (for review see Fox & Raichle, 2007; Cole, Smith & Beckmannn, 2010; Yeo et al. 2011) including visual and auditory networks (Bianciardi et al. 2009; Damoiseaux et al. 2006; Hunter et al. 2006), the dorsal attentional network (Fox et al. 2005; 2006), the language system (Hamspon et al. 2002), medial temporal lobe memory systems and the default-mode network (DMN; Buckner et al. 2008; Fox et al. 2005; Greicius et al. 2003; 2004) and the frontoparietal control network (Vincent et al. 2008). For illustration of these networks see Figure 2.4 adapted from Yeo et al. (2011). These studies emphasize that during rest the brain is not idle, but instead shows a large amount of spontaneous activity that is highly correlated between brain regions. Importantly, these networks reflect brain regions that have been shown to typically work together during task-based fMRI experiments.

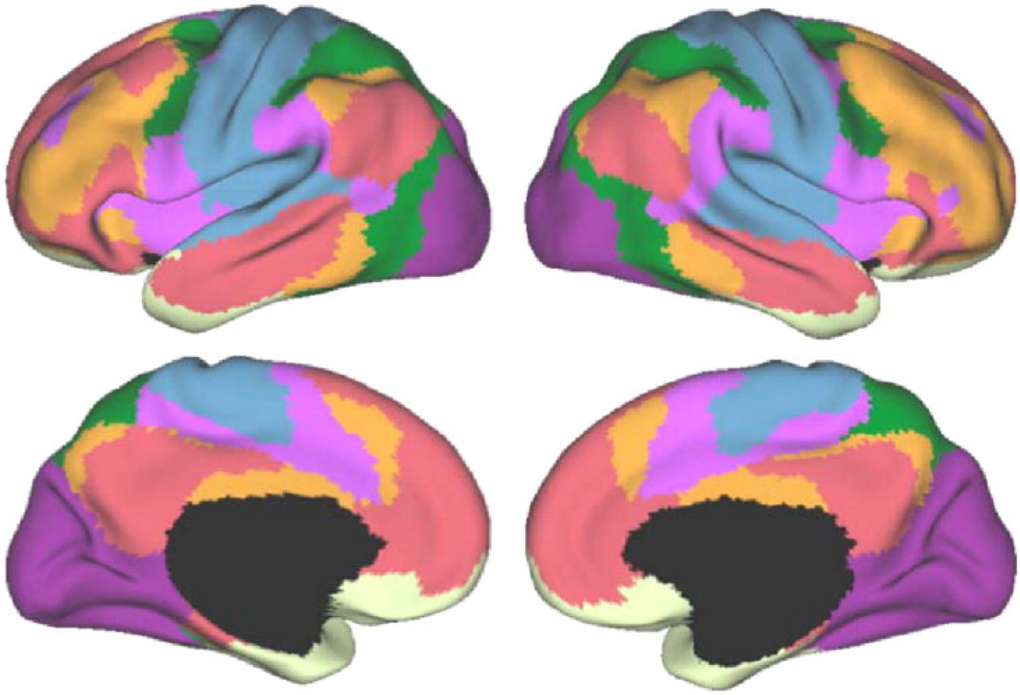

Figure 2.4. The 7-network adapted from Yeo et al. (2011). Large-scale cortical networks identified through resting-state connectivity analysis. Each of the 7 networks are depicted by a unique colour: visual (dark purple), dorsal attentional (green), ventral attentional (violet), frontoparietal control (orange), somatomotor (blue), default mode (red), limbic (cream).

Despite an increase in the number of functional connectivity studies, deciphering the appropriate methods to analyse this data remains a controversial topic (Cole, Smith & Beckmannn, 2010). Current debates in the literature include (i) how to successfully remove the effect of nuisance covariates (such as respiration-induced fMRI signal and motion-related artifacts: Birn et al. 2006; Murphy et al. 2009) to improve the quality of resting-state data and (ii) which statistical analyses are most appropriate for functional connectivity (for review see Cole, Smith & Beckmann, 2010). As such, this field has seen a rapid growth in both pre-processing and statistical analysis methods. The following section will review these methods, outlining the strengths and caveats of each, and conclude with which approach will be adopted in the current thesis.

2.2.1.1. Acquisition

Resting-state connectivity involves participants lying in an MRI scanner for an allotted time (typically 5-10 minutes) without an explicit stimulus or task present. During this time they are asked to think about nothing in particular, without falling asleep. Measurement during rest is advantageous as it minimizes task-related BOLD fMRI fluctuations (Van Dijk et al. 2010). Just as in conventional task-based fMRI, the BOLD fMRI signal is measured throughout the resting-state scan.

2.2.1.2. Functional connectivity pre-processing

For functional connectivity the initial pre-processing steps are similar to those routinely applied to traditional task-based BOLD fMRI data (Beckmann et al. 2005); including slice-timing correction for slice dependent time shifts / intensity differences and spatial smoothing. However, unlike univariate and MVPA methods, which employ a high pass filtering, functional connectivity typically applies low pass filtering to retain frequencies less than 0.1 Hz (Biswal. 1995; Lowe et al. 2000). Low pass filtering is necessary to remove signal from cardiac and respiratory oscillations that typically occur above 0.3 Hz (Corders et al. 2001). The resulting low-pass filtered data is therefore thought to reflect a neuronal basis of functional connectivity (as opposed to signal explained by physiological responses such as cardiac and respiratory noise). Removal of this noise is pivotal as physiological sources of noise produce time dependent changes in the magnetic field due to subject's chest movements (Brosch et al. 2002), moreover it is has been demonstrated that variations in the static magnetic field within brain tissue occur due to respiration (Raj et al. 2001).

Although temporal filtering reduces cardiac and respiratory related-noise, it has been shown that these physiological noises alias in the low frequency range (~ 0.1Hz) due to the long TR (~ 2-3 s) typically used in standard resting-state BOLD sequences (Bhattacharyya & Lowe, 2004; Lowe et al. 1998). Utilizing external recordings of physiological noise (e.g., recording cardiac, respiratory and blood pressure traces) allows researchers to model these responses and remove these estimates from each voxels time series. However, there are many reasons why obtaining physiological recordings is not routinely done in the field of functional

connectivity: for instance, lack of equipment, data corruption and non-compliant/uncomfortable subjects (Murphy et al. 2009).

Therefore, more advanced statistical methods have been developed to reduce physiological causes of noise in functional connectivity data. This is particularly important in resting-state analysis as, unlike traditional task-based BOLD activity - where the timing of the task is known *a priori,* and many events are averaged together to increase the signal to noise ratio – functional connectivity uses some metric (typically correlation coefficient) to determine the temporal similarity of the BOLD time series. In the latter, trials are not averaged together and as such functional connectivity is more susceptible to spurious sources of noise. For example, in the seminal paper by Biswal and colleagues (1995), the correlation coefficient was calculated between the BOLD time series of a voxel in the left motor cortex and all other voxels in the brain. Voxels were considered functionally connected if their correlation coefficient passed a statistical threshold. This resulted in common 'spontaneous fluctuations' between bilateral motor cortices. However, as the time series of the two voxels are measured simultaneously, any non-neural activity affecting one or both time series may influence the functional connectivity, thus producing a spurious result. By introducing false similarities these confounds may increase the apparent functional connectivity between two time series (Bright and Murphy, 2013; Murphy et al. 2011; Van Dijk et al. 2012).

One proposed method to remove such noise is known as global signal regression (GSReg), which subtracts the global signal (i.e., mean BOLD signal computed across all voxels) from each voxel in the brain (Greicius et al. 2003; Macey et al. 2004). This method proposed that any activity that globally influences the BOLD signal must be unrelated to neural activity and, as such, must be a confound. This method was popular in the advent of functional connectivity studies as it produced reliable patterns of functional connectivity across the brain (Fox et al. 2005; 2009). However, many studies now question the effect of global signal removal on resting state maps. Firstly, this technique only removes the average time-series; therefore voxel-specific phase differences in the noise produced by physiological fluctuations are not accounted for (Behzadi et al. 2007). Secondly, GSReg consistently produces anti-correlations between the DMN and the extended

dorsal attentional network. This negative correlation is hard to interpret given the fact that artifactual deactivations are produced in task-based fMRI after global normalizations (Desjardins et al. 2001; Macey et al. 2004). More poignantly, it has been mathematically demonstrated that when GSReg was not performed the anti-correlated network disappears (Murphy et al. 2009). In light of this evidence, it has been argued that GSReg produces unreliable negative correlations making interpretation of results after GSReg difficult and, therefore, should be avoided (Murphy, Birn & Bandettini, 2013).

To circumvent this problem, alternative methods have been devised. One popular method, known as CompCor, defines regions unlikely to be associated with neural activity (e.g., cerebrospinal fluid and white matter; Behzadi et al. 2007), as any signal here is likely to reflect physiological noise (Dagli et al. 1999; Van Dijk et al. 2010). CompCor then applies Principal Component analysis (PCA) to the signal extracted from these "noise ROIs" to obtain a subset of nuisance regressors from the larger set of voxel-wise signals of no interest. These nuisance regressors can then be included as nuisance parameters within a standard GLM for BOLD fMRI time-series data. This method overcomes the issues present in GSReg as it obtains more locally-specific signals that many not be well-reflected in the mean signal. Moreover, the use of nuisance regressors derived from these noise ROIs have been shown to improve specificity of functional connectivity maps (Bright & Murphy, 2013; Weissenbacher et al. 2009).

In addition to physiological noise, functional connectivity is also susceptible to motion artifacts. Over the last few years, several independent research groups have reported that motion adds spurious variance that is typically more similar for nearby voxels than at distance voxels (Power et al. 2012; Satterthwaite et al.. 2012; Van Dijk et al. 2012). This is problematic in functional connectivity as these differences in the distributed variance causes distant-dependent modulation of signal correlations. For instance, all things being equal, a 'higher-motion' group would have comparatively stronger short-distance correlations than a 'lower-motion' group (Power et al. 2015). Pre-processing steps therefore also need to include motion correction regressors, to remove this confound from the functional connectivity BOLD signal.

Motion artifacts can be measured directly from the fMRI datasets via realignment parameters (as the fMRI data contains information required for motion artifact removal). A standard whole-brain approach for modelling head motion is through FSL's MCFLIRT function (Jenkinson et al. 2002). This method produces a motion transformation matrix for each time point relative to the reference volume (middle time point of the time series). In total, MCFLIRT derives six motion parameters consisting of three rotations and three translations (Jenkinson et al. 2002). These six parameters can then be entered as confound regressors within a standard GLM for BOLD fMRI time-series data. Despite its popularity and ability to significantly reduce motion variance (Lund et al. 2005) more recent studies have suggested that these regressors are inadequate at removing subtle motion artifacts (Kundu et al. 2013; Power et al. 2014).

An alternative approach for motion denoising is to utilize the signal in non-grey matter regions (e.g., white matter and CSF). Any signal here is likely to reflect non-neural activity. As discussed previously, CompCor defines voxels in noise ROIs (non-grey matter locations) and uses PCA to obtain a small set of nuisance regressors. Note physiological sources of noise can also manifest as subtle motion artifacts such as time dependent changes in the magnetic field due to subject's chest movements (Brosch et al. 2002). As such, it has been demonstrated that CompCor not only tackles physiological noise issues but also reduces subtle motion artifacts (Muschelli et al. 2014). Collectivity speaking, both physiological noise and motion-artifacts must be appropriately modeled in order to make reliable interpretations about resting-state data. Therefore both (i) the top subset of nuisance regressors derived from the CompCor analysis and (ii) the six motion parameters derived from MCFLIRT can be simultaneously entered into the GLM model as confound regressors. Multiple linear regression can then be performed to remove these confounds from the BOLD time-series.

2.2.1.3. Functional connectivity statistical analysis

The most common statistical method for detecting functional connectivity is known as a seed-based correlation analysis (SCA; Andrews-Hanna et al. 2007; Biswal et al. 1995; Fox et al. 2005; Fransson, 2005; Jiang et al. 2004; Larson-Prior et al. 2009;

Song et al. 2008). SCA is a model-dependent method that requires the selection of a voxel, cluster or atlas region, guided by *a priori* knowledge (typically from previous literature or from a localizer experiment). Once selected the time series data is extracted from the seed region and used as a regressor in a GLM analysis (along with confound regressors of no interest outlined in the previous section). This calculates whole-brain, voxel-wise functional connectivity maps of co-variance with the seed region. This technique reveals reliable connectivity properties of many seed areas and, as such, has been utilized by many prolific research groups in the resting-state field (e.g., Greicius et al. 2003; Fox et al. 2005; Margulies et al. 2007).

The primary advantage of SCA is the simplicity in interpreting the resting-state findings (Buckner & Vincent, 2007). This method asks a direct question – which regions most strongly functionally connect with the seed region - and as such provides a direct answer. Moreover, assessments of test-retest reliability have shown that resting-state connectivity can be identified by SCA with moderate to high reliability (Shezad et al. 2009). One potential weakness of this method reflects the influence that structured confounds (e.g., rigid head movements) may have on the functional connectivity, as the presence of residual confounds can negatively impact the correlation maps by also including voxels that describe the artifact and not the true functional connectivity to the seed region. Furthermore, this method is limited to investigating the connectivity of pre-determined seed regions; as such it is not designed to examine functional connection patterns on a whole-brain scale.

An alternative approach, known as independent component analysis (ICA), has been proposed to avoid the influence of noise attached to the seed and capture whole-brain connectivity patterns (Beckmann et al. 2005; De Luca et al. 2006; Kiviniemi et al. 2003; 2009). This model-free approach aims to capture underlying sources that can explain resting-state patterns by identifying spatial sources of resting-state signals that are maximally independent from each other. This method has gained popularity in the resting-state field due to its ability to capture whole-brain connectivity patterns (Beckman et al. 2005; De Luca et al. 2006; Kiviniemi et al. 2003; 2009; Smith et al. 2009). However, the components produced from ICA methods are often perceived as more difficult to understand, compared to traditional seed-based maps, as they compare a more complex representation of

the data (Fox & Raichle, 2007). Nevertheless, although both methods have their advantages and disadvantages, the connectivity maps derived from both ICA and SCA analyses show strong overlap, supporting the notion that robust functionally linked networks exist in the brain at rest (Beckmann et al. 2005).

In light of this evidence, SCA will be adopted in the current thesis due to its elegance and ease in both implementing and interpreting the data. Moreover, SCA links regions of activity identified in task-based fMRI studies (e.g., functional activation) with functional connectivity. The combination of task-based activation and seed-based functional connectivity allows us to simultaneously address two complimentary aspects of semantic retrieval. Task-based fMRI can interrogate direct hypotheses regarding discrete activity of a small number of regions whereas experimentally guided seed-based functional connectivity explains the true distributed nature and complexity of these neural functions and how they are embedded in large-scale functional networks. Therefore, SCA will be adopted in Chapters 3-5 to further interrogate regions identified from task-based fMRI.

2.3. Methods that combine functional-activity and functional-connectivity data

Thus far, this chapter has outlined methods to statistically analyse (i) functional activation (univariate and MVPA) and (ii) functional-connectivity (functional connectivity) separately. The final section of this chapter will discuss new and emerging approaches that attempt to bridge the gap between these two domains by simultaneously taking into consideration task-based activation and resting-state networks.

2.3.1. Neurosynth Decoder

Neurosynth is an online platform for large scaled automated synthesis of fMRI data (Yarkoni et al. 2011). Neurosynth pools results from published fMRI studies (including univariate, MVPA and functional connectivity studies) and stores them in

a large database. From this database, each published article is examined for key-terms using a text-mining technique. Key terms are extracted if they are reported in the article more than once, along with their frequency. Next, peak coordinates are extracted from tables in the published article. Both key-terms and coordinates are saved in the database allowing investigators to interrogate either key words of interest or peak coordinates. For key words of interest, the user enters a search string (e.g., 'semantic memory') and peak coordinates from articles with this key term in are extracted from tables. A meta-analysis is automatically performed on each peak-coordination, producing a whole-brain map of the probability of the term producing activation at each point. The user can therefore see regions of activity related to the search string and the probability of that region being activated by that term. Alternatively, a user can interrogate the function of a specific coordinate by entering peak coordinates of interest. In return, NeuroSynth will produce a list of key terms associated with that coordinate and their probability. In addition to searching specific coordinates, and arguably the most interesting property of this tool is its ability to handle whole-brain statistical maps. For this users can upload whole-brain maps (e.g., z-stat or t-stat map) and NeuroSynth will decode the list of key terms associated with the entire map (and their probabilities). This method therefore allows for both forward and reverse inferences to be made on fMRI data: given a known cognitive phenomenon (e.g., day dreaming) one can quantify the corresponding neural activity and generate forward inferences, whereas given an observed pattern of activity, Neurosynth can draw a reverse inference about associated cognitive states using probability algorithms.

There are many advantages for utilizing a tool like Neurosynth. Firstly, given an observed pattern of activity (e.g., an unthresholded z-stat map created from univariate, MVPA or functional connectivity analysis) it provides confirmation of cognitive function from a large independent data set. Secondly, it provides a novel framework that simultaneously aggregates and synthesizes both task-based and resting-state data. This is particularly important, given the fact that many neuroimaging studies are often underpowered and exhibit relatively high false positive rate (Yarkoni, 2009), thus Neurosynth provides a mechanism through which results from multiple methods can be combined in order for a consensus to be

achieved regarding the relationship between brain and cognitive function. Finally, as researchers are besieged by an increasing number of fMRI publications each year, these tools provide an automated platform to synthesis and compare these findings. Overall, although this tool is not a classic statistical analysis (like those discussed in previous sections of this chapter), it does provide for automated meta-analyses across a wide range of recent publication that tackles many prominent issues in the cognitive neuroimaging field. This tool will therefore be used as a complimentary approach to more conventional statistical analyses to quantify inferences drawn on the neural activity identified in all Chapters of this thesis.

2.3.2. Principal Gradient

The final method discussed in this chapter explores the topographical organisation of functional brain networks. In addition to linking functional connectivity with functional activation, recent studies have shown that the topology of these networks is organized in such a way as to produce highly efficient and cost-effective processing: by integrating across different sub-systems of the brain network, this topology optimizes towards a high level of information processing (Margulies et al. 2016; Van Den Heuvel & Pol, 2010). A recent account of topographical organisation suggests that the cortex is arranged along a macroscale principal gradient that describes gradual transitions in functional connectivity, repeated in multiple regions (Margulies et al. 2016). At one end, this functional gradient is anchored by unimodal regions that have well-described roles in perception and action (sensory/motor areas), while the other end is characterized by transmodal association regions, within the so-called default mode network (DMN) that are thought to represent complex and abstract functions (see Figure 2.5). The following section will briefly outline how Margulies et al (2016) computed the principal gradient and then discuss the potential benefits of using such a measure to explain task-based fMRI activation.

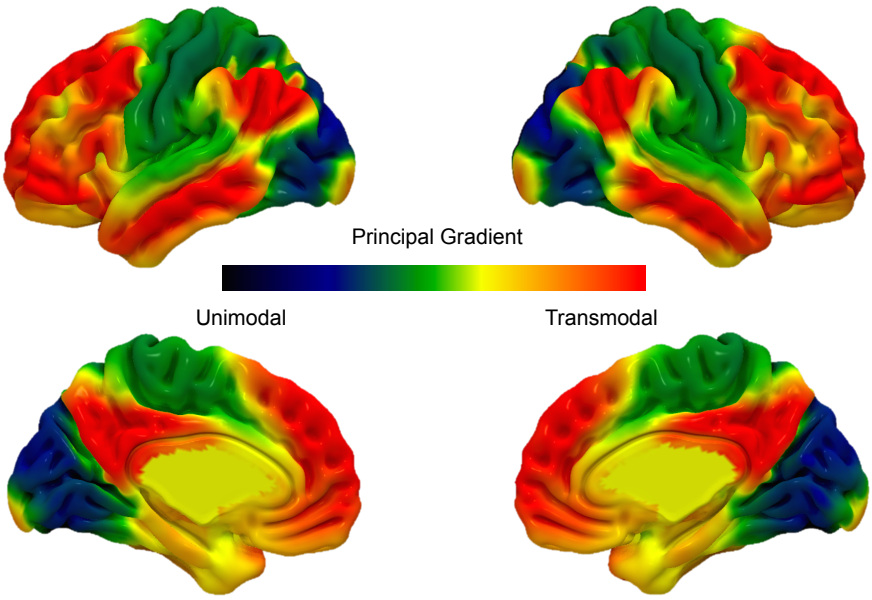

Figure 2.5. The principal gradient of connectivity in human cortices adapted from Margulies et al (2016). The principal gradient shows a spectrum between unimodal regions (dark blue) and transmodal regions (red), which in the human cortex, peaks in regions corresponding to the DMN. The proximity of colours can be interpreted as greater similarity of connectivity patterns.

The principal gradient was generated using a non-linear dimensionality reduction technique, known as diffusion map embedding. This method takes high-dimensional data (e.g., functional connectivity data) and returns parameters that describe the lower-dimensional structures of which it is comprised (Coifman et al. 2005; de la Porte et al. 2008). In higher-dimensions, distances between points are often too large for linear techniques (such a Euclidean distance) to handle, given the sparsely populated feature space. The key to the non-linear embedding technique is the assumption that data lies on (embedded) a lower-dimensional structure or manifold; as we measure distance on the manifold itself, rather than in Euclidean space, this approach is able to characterize neural data and the relationship between individual points using fewer dimensions. Importantly, dimensionality reduction produces meaningful parameters that preserve the important relationships between data points.

In Margulies et al. (2016) the resulting diffusion map represents the global connectivity structure as a distribution of cortical points on an axis (e.g., embedding space). The position of cortical points on this axis represents connectivity similarities between regions such as points that are strongly connected by either many connections or fewer strong connections are close together, whereas points with weak or no connections are far apart. This revealed a principal gradient axis that explained the most variance in resting-state connectivity when primary sensory and motor regions were positioned on one end and higher-order association regions (DMN) at the other end (see Figure 2.5). In addition to the principal gradient derived from diffusion map embedding, the authors showed that a regions position on the principal gradient axis were predicted by its geodesic distance from the DMN: the further a region was from the DMN in geodesic space the further away it was positioned on the gradient.

There are many advantages to explaining links between neural function and cognition in this manner. For example, describing the macrolevel organization of connectivity patterns from unimodal regions to transmodal association cortex (Krienen & Sherwood, 2016) highlights that features of higher order cognition maybe organized into a successions of functional gradients that are present in the topographical organization of visual processing (Rosa, 2002). This gradient perspective is important because it highlights, not only the topography of these distinct distributed networks but also the *reason* for their particular spatial relationship and how this constrains their function. Following this view information is converged across modalities into progressively more abstract representations and situates the DMN at the top of this representational hierarchy in order to describe the current cognitive landscape in the most abstract terms. Notably, regions tied to a specific modality (e.g., sensory and motor regions) have the greatest geodesic distance from the DMN supporting the notion that abstract processing requires abstraction away from modality input.

Given the principle gradient's potential explanatory power in capturing the organizing principle of human cognition, it could be used to help interpret task-based fMRI data. In the section that follows, two novel applications of the principal gradient - as an analysis method for task-based fMRI - will be described. The first

will focus on identifying where ROIs fall on the principal gradient and the second will aim to capture whole-brain neural shifts in activity.

2.3.2.1. Principal Gradient Analysis

For an ROI analysis, first regions need to be identified, typically from previous literature or clusters of activity derived from traditional univariate analyses. Following Margulies et al (2016) the original principal gradient map can then be divided into five-percentile bins; yielding twenty bins in total (bin 1 = 0 – 5, bin 2 = 5 – 10 etc.). Next, for each ROI the number of voxels in each of the twenty principal gradient bins is calculated. This is calculated for each individual participant and then group-averaged to calculate the group-average gradient for each ROI (i.e., the position on the gradient where the greatest number of voxels fell). To evaluate whether these values differed from chance, gradient values can be compared to those obtained based on random permutation labelling. Random permutation labelling is a nonparametric test: within each subject a random label is assigned to each gradient bin, in a way no label recurs. Next, the number of voxels in that bin is recalculated with these random labels for each subject. The gradient values obtained can then compared with the distribution of the values obtained with these random labels. This method therefore quantifies where on the gradient ROIs fall and the statistical likelihood of identifying a cluster at that position.

Alternatively, to capture more general whole-brain patterns of cortical activity the average neural signal can be extracted from each of the 20 bins, for unthresholded zstat maps. These zstat maps are created as part of a standard univariate analysis and contrast task condition > rest. The purpose of using unthresholded z-stat maps is to capture changes that may occur at the level of whole-brain, that manifest as distributed shifts in activity as opposed to discrete clusters of activity that surpass a statistical threshold. Once extracted, the group-averaged activity in each bin can be plotted across the principle gradient separately for each condition. This approach will reflect general shifts in the patterns of cortical activity at the whole-brain level. Follow up statistics can interrogate the linear relationships between the loading of each stimulus-condition across the

principal gradient (such as analysis of variance (ANOVA) or principal component analysis (PCA)).

Taken together, the principal gradient analyses offer a complementary approach to functional activation statistics (univariate and MVPA) that provides additional evidence that the topographical location of a neural region (e.g., position on the principal component) is related to its functional characterization. Moreover, as the far extreme of the gradient is tied to higher-order association zones - regions that have been implicated in spontaneous thought, daydreaming and semantic processing of multi-modal concepts - this method can also highlight the *reason* for their particular spatial relationship and how this constrains their function.

2.4. Additional caveats

In addition to identifying appropriate statistical analyses, careful consideration regarding the fMRI scanning acquisition is needed. This is particularly important as the current thesis is interested in regions close to air-filled sinuses, such as ventral anterior temporal lobes. Such regions are affected by discrepancies in magnetic susceptibility across different tissue types (e.g. water, air, bone), resulting in loss of signal and distortion. Emerging methods have been proposed to overcome these distortions such as dual-echo EPI sequences (coined distortion-corrected fMRI). Such methods consecutively collect two EPI read-outs at a short and longer echo time (TE), which has been shown to reduce signal loss due to spin dephasing (Halai et al. 2014; Visser et al. 2010). However, due to the scanner resources available for data collection, this method is unable to be implemented.

Alternative, and easily implemented methods for reducing signal distortion have also been proposed. First, optimization of slice orientation (e.g., orientating slices with the temporal lobe) has been shown to reduce signal distortion in susceptible regions (De Panfilis & Schwarzbauer, 2005; Deichmann et al. 2003; Halai et al. 2014). Second, reducing the TE (e.g., selecting the shortest possible TE) can effectively reduce fMRI signal distortions (Hutton et al. 2002; Weiskopf et al. 2007). As distortion-corrected fMRI is not feasible, the current thesis will optimize slice

orientation and select a short TE to reduce the effects of signal distortion, and improve the likelihood of detecting activation in crucial regions of interest (e.g., vATL).

2.5. Methods Overview

Simultaneous analysis that focuses on different dimensions of the same data will provide a rich explanation for how complex representations decoupled from perceptual input are represented within the brain. In all cases, the experiments proposed in this thesis make use of univariate and MVPA methods to provide a clear understanding of the activation (univariate) and representational content (MVPA) of regions associated with complex abstract concepts that are processed independently of sensory input. Furthermore, all experiments will include seed-based functional connectivity that aims to further interrogate the clusters of functional activity identified from either univariate or MVPA approaches, to understand the neural networks in which these different clusters of activity are embedded. In addition, all chapters will include the search tool Neurosynth to decode the most common interpretations of patterns of functional connectivity, in the broader neuroimaging literature. Finally, Chapter 5 will also employ a principal gradient analysis to investigate the underlying function of the transmodal end of the gradient (e.g., regions that have been implicated in spontaneous thought, daydreaming and semantic processing of multi-modal concepts) as this method can highlight the reason for their particular spatial relationship *and* how this constrains their function.

Chapter 3 - Fractionating the anterior temporal lobe: MVPA reveals differential responses to input and conceptual modality

This chapter is adapted from: Murphy, C., Rueschemeyer, S. A., Watson, D., Karapanagiotidis, T., Smallwood, J., & Jefferies, E. (2017). Fractionating the anterior temporal lobe: MVPA reveals differential responses to input and conceptual modality. NeuroImage, 147, 19-31.[2]

3.1. Abstract

Words activate cortical regions in accordance with their modality of presentation (i.e., written vs. spoken), yet there is a long-standing debate about whether patterns of activity in any specific brain region capture modality-invariant conceptual information. Deficits in patients with semantic dementia highlight the anterior temporal lobe (ATL) as a transmodal store of semantic knowledge but these studies do not permit precise localisation of this function. The current investigation used multiple imaging methods in healthy participants to examine functional dissociations within ATL. Multi-voxel pattern analysis identified spatially segregated regions: a response to input modality in anterior superior temporal gyrus (aSTG) and a response to meaning in more ventral anterior temporal lobe (vATL). This functional dissociation was supported by resting-state connectivity that found greater coupling for aSTG with primary auditory cortex and vATL with the default mode network. A meta-analytic decoding of these connectivity patterns implicated aSTG in processes closely tied to auditory processing (such as phonology

[2] The author, Charlotte Murphy, designed the experiment, analysed the results and wrote the article under the supervision of Prof. Beth Jefferies and Dr. Shirley-Ann Rueschemeyer. Dr. David Watson and Theo Karappanagiotidis provided technical support for the analysis. Dr. Jonathan Smallwood helped to supervise the analysis and provided comments on the manuscript.

and language) and vATL in meaning-based tasks (such as comprehension or social cognition). Thus we provide converging evidence for the segregation of meaning and input modality in the ATL.

3.2. Introduction

Current neurocognitive models propose that concepts are represented in a large-scale distributed network comprising (1) sensory and motor 'spoke' regions that store knowledge of physical features and (2) convergence zones that integrate across multiple modalities (e.g., visual vs. auditory) to form abstract transmodal representations (Damasio, 1989; Patterson, Nestor & Rogers, 2007). For example, the hub-and-spoke model of Patterson and colleagues (2007) proposes that information from modality-specific spoke regions is integrated in a transmodal 'hub' region within the anterior temporal lobes (ATL), allowing the conceptual similarity of items that are semantically similar yet share few surface features, such as 'flute' and 'violin', to be represented, and making it possible to map between modalities so that we can picture a flute and imagine the sound that it makes from only its name (e.g., Damasio, 1989; Lambon Ralph, Sage, Jones & Mayberry, 2010; Patterson et al. 2007; Rogers et al. 2004). This hub-and-spoke model proposes that both the ATL and modality-specific spokes make a crucial contribution to conceptual representation, and these elements are mutually-constraining through a pattern of interactive-activation.

The spokes are hypothesized to represent the contributions of sensory and motor cortex to conceptual knowledge, as words associated with specific sensorimotor attributes activate corresponding sensorimotor cortex. For example, words denoting actions (e.g., kick) activate the motor system (Postle, McMahon, Ashton, Meredith, & de Zubicaray, 2008; Rueschemeyer, Brass, & Friederici, 2007; Rueschemeyer, van Rooij, Lindemann, Willems, & Bekkering, 2010), while words associated with specific smells (e.g., cinnamon) elicit activation in olfactory cortex (Cerf-Ducastel & Murphy, 2004; Gonzalez et al. 2006). Although these neural regions are important for perception and action, they are also recruited during

semantic processing to provide meaning to words (Barsalou, 1999; 2008; Martin, 2007; Patterson, et al. 2007; Kiefer & Pulvermüller, 2012).

The proposal that the ATL forms a key semantic "hub" capturing knowledge across different input modalities was initially put forward to account for the pattern of impairment in semantic dementia (SD), in which relatively focal atrophy centered on ATL leads to progressive conceptual degradation across modalities and tasks (e.g., Patterson et al. 2007; Rogers, Patterson, Jefferies & Lambon Ralph, 2015). SD patients are highly consistent in the knowledge they can demonstrate when the same concepts are probed in different ways, suggesting central semantic representations degrade in this condition. Patients with SD have atrophy which increasingly affects inferior frontal and posterior temporal areas, as well as ATL, making it difficult to draw strong conclusions about the location of the "hub" from neuropsychology alone; however, the severity of the semantic impairment correlates most strongly with the degree of hypometabolism in ventrolateral ATL (Mion et al. 2010). The crucial role of ATL is also supported by functional neuroimaging studies of healthy participants that show transmodal conceptual processing in ATL (Rice et al. 2015; Visser et al. 2010). For example, Visser et al. (2011) characterized the degree of modality convergence in STG, MTG, ITG and fusiform cortex comparing posterior and anterior parts of the temporal lobe. Both STG and fusiform were modality-sensitive along the temporal lobe, showing stronger activation for spoken words and pictures respectively. MTG showed a multimodal response in both anterior and posterior regions. ITG uniquely showed a pattern consistent with the increasing integration of information from different inputs, namely sensitivity to modality in posterior but not anterior regions. Moreover, Spitzyna et al. (2006) showed that, despite originating from different sensory inputs, there is considerable activation overlap for spoken and written processing in ATL regions. Thus, emerging evidence from both patients with SD and healthy participants suggests that the semantic hub may be located in ventrolateral ATL.

These observations raise the possibility of functional dissociations with ATL. Jackson et al. (2016) recently observed different patterns of functional connectivity within superior and ventral regions of the ATL, with anterior STG showing stronger

connectivity to language, auditory and motor regions, while ventrolateral ATL showed connectivity to other heteromodal semantic regions including inferior frontal gyrus, angular gyrus and posterior middle temporal gyrus. These parallel the pattern of white-matter connections found by Binney et al. (2012) and Jung et al. (2016). Consistent with these findings it has been proposed that superior regions of the ATL are important in lexical and auditory processing, while ventrolateral regions support conceptual processing across all sensory modalities (Rice et al. 2015; Visser et al. 2010; Visser & Lambon Ralph, 2011). Ventral and ventrolateral ATL regions have been found to respond to meaningful inputs across multiple modalities by studies employing convergent methods; including fMRI and transcranial magnetic stimulation (Binney et al. 2010; Visser et al. 2011; 2012; Hoffman et al. 2015) and representational similarity analysis (RSA) of ECoG data (Chen et al. 2016).

The current study used multiple imaging methodologies to simultaneously investigate the organization of knowledge in the ATL (hub) and auditory and visual regions (as potential spokes). In a functional experiment we manipulated the format in which words were presented (i.e., spoken, written) and the modality-specific features associated with the word's meaning (e.g., auditory features: "loud" vs. visual features: "shiny"). We used Multi Voxel Pattern Analysis (MVPA) to decode how these different features (modality of presentation and underlying meaning) are represented. Based on the hub-and-spoke model, we expected this analysis to reveal regions that are distributed across the cortex that responded to the meaning of the stimulus regardless of the input modality. In this experiment we were particularly interested in identifying regions in ATL where the meaning of words is represented that are independent of input modality. The transmodal hub regions should be able to code the meaning of a stimulus regardless of the presentation format (e.g., auditory feature words should elicit similar patterns of activation even when spoken and written words are compared). In addition, this region should represent the meaning of words tied to different sensory modalities (i.e., it should represent words with auditory meanings like 'loud' and words with visual meanings like 'shiny'). In contrast, the spokes should represent particular semantic features in regions of sensory cortex (i.e., words with an auditory meaning, such as loud, should be represented in auditory cortex regardless of how

they are presented (written or spoken). However, spoke regions are not expected to represent meaning that is tied to a different sensory modality (i.e., auditory cortex may not contribute to semantic representation for words with a visual meaning, such as shiny).

Next we used the regions identified in our MVPA analysis as regions of interest in a seed based resting-state connectivity analysis to understand the neural networks in which these different regions of the ATL are embedded. We expected the transmodal region of ATL to show functional connectivity with regions of cortex that are important in more abstract forms of cognition, e.g., the default mode network, rather than regions important in unimodal sensory processing, such as the auditory and visual cortex. Finally, we used the search tool Neurosynth to decode the most common interpretations of this pattern of functional connectivity in the broader neuroimaging literature.

3.3. Materials and Methods

3.3.1. Functional Experiment

3.3.1.1. *Participants*

Twenty participants were recruited from the University of York. One participant's data was excluded due to excessive motion artefacts, leaving nineteen subjects in the final analysis (10 female; mean age 24.55, range 18-36 years). Participants were native British speakers, right handed and had normal or corrected-to-normal vision. Participants gave written informed consent to take part and were reimbursed for their time. The study was approved by the York Neuroimaging Centre Ethics Committee at the University of York.

3.3.1.2. Stimuli

Participants were presented with blocks of spoken and written items from three conditions: AUD words denoted auditory features (e.g., loud), VIS words denoted visual features (e.g., shiny) and NON stimuli were meaningless non-words (e.g.,

brodic). A block consisted of a sequence of items; participants were asked to pay attention to the meaning of each item, and respond with their left index finger when an out-of-category item was presented (see Figure 3.1). For VIS and AUD blocks, half of the out-of-category items were taken from the non-presented feature condition, while the other half were taken from a separate list of taste words (e.g., spicy). Participants could not predict the category of the out-of-category item and therefore had to focus on the AUD or VIS feature specified in the instructions. In the NON condition, participants were asked to respond to any item that was a word. All stimuli were presented in both spoken and a written format. Spoken words were recorded digitally and then normalized for volume and power. Written words were presented centrally as white letters on a black background. The combination of item meaning (AUD, VIS, NON) and presentation format (Spoken, Written) yielded 6 experimental conditions (Spoken-AUD, Spoken-VIS, Spoken-NON, Written-AUD, Written-VIS, Written-NON).

The selection of AUD and VIS words was validated in a behavioural study with twelve participants who did not take part in the fMRI session. Participants were asked to rate a subset of modality-specific words (n=220), according to how much each one related to four sensory categories; auditory, visual, haptic and taste. Participants also provided ratings of familiarity and emotional valence. All ratings were given on a 5-point likert-scale. We selected adjectives with strong auditory or visual associations. Each set contained 8 items, which were matched for key psycholingusitic variables such as frequency and length (see Table 3.1; Wilcoxon signed rank tests revealed all $p > .05$). AUD words (such as 'loud') were selected if they scored significantly higher on the auditory than visual, haptic or taste modalities (all $p < .001$). Likewise VIS words (such as 'shiny') were selected if they scored significantly higher on the visual than the auditory, haptic or taste modalities (all $p < .001$).

A set of 8 taste-features were used in out-of-category catch trials. These items scored significantly higher on the taste modality than auditory, visual and haptic ($p < .001$). These items were also matched to AUD or VIS words on the variables in Table 3.1 (all $p > .05$). Finally, NON words were made by recombining the phonemes from the AUD and VIS conditions to create 8 pseudo-words. The non-

word condition matched AUD and VIS conditions on number of letters, syllable and Levenshtein distance (Levenshtein, 1965), which quantifies the number of phoneme insertions, deletions and/or substitutions required to change one word into another, (all p > .05). The use of a small number of items is consistent with other MVPA studies into semantic representation (Corriea et al. 2014; Peelen & Caramazza, 2012).

Table 3.1 - Mean psycholinguistic properties of stimuli (SD in parentheses)

Property	**Auditory feature words**	**Visual feature words**	**Non-words**
Example	"loud"	"shiny"	"brodic"
Log frequency	2.27 (1.05)	2.54 (.82)	N/A
Length	5.25 (.76)	5.50 (.80)	5.88 (1.17)
Syllables	1.88 (.45)	1.63 (.49)	2.00 (.50)
Age of acquisition	7.17 (2.70)	6.85 (2.76)	N/A
Familiarity	4.43 (.63)	4.40 (.51)	N/A
Emotional Valence	3.18 (.70)	3.3 (.67)	N/A
Levehnstein distance	5.11 (.94)	6.00 (1.25)	5.89 (.86)
Behavioural feature-rating (auditory)	4.45 (.61)*	1.15 (.04)*	N/A
Behavioural feature-rating (visual)	1.65 (.32)*	4.77 (.19)*	N/A
Behavioural feature-rating (haptic)	1.5 (.39)	1.76 (.72)	N/A
Behavioural feature-rating (taste)	1.19 (.07)	1.21 (.09)	N/A

Footnote: Log frequency = log-transformed lemma frequencies from the SUBTLEX database (Brysbaert, New & Keuleers, 2012; http://expsy.ugent.be/subtlexus). Length = number of letters. Age of acquisition (AoA norms; Kuperman et al. 2012). Part of speech also taken from SUBTLEX database. Familiarity, emotional valence and behavioural feature rating (auditory; visual; haptic; taste) were obtained from a behavioural experiment with a separate cohort of participants from the fMRI study. These were scored on a Likert-scale (1-5). * Wilcoxon signed rank tests revealed a significant difference between auditory-feature and visual-feature conditions (p < .001).

Stimulus presentation was controlled by a PC running Neurobehavioural System Presentation® software (Version 0.07, www.neurobs.com). Stimuli were projected onto a screen viewed though a mirror mounted on the head coil. Spoken stimuli were presented binaurally using MR-compatible headphones.

3.3.1.3. Task Procedure

Prior to being scanned, participants were shown a printed copy of all stimuli (8 AUD, 8 VIS, 8 NON) to familiarize them with the items. They also performed a practice session consisting of 12 blocks, identical to one scanning run.

In the scanner there were 4 runs of 12 blocks. The choice of 4 functional runs is consistent with many MVPA studies that also presented trials within 4 runs that were each 5-10 minutes long (Countanche & Thompson-Schill, 2012; Fairhall & Caramazza, 2013; Li et al. 2011; Peelen & Caramzza, 2012). Within each run, there were two blocks related to each of the 6 experimental conditions (spoken and written words combined with three meaning conditions: AUD, VIS and NON). These were presented in a pseudo-random order, with no immediate repetition of conditions. Blocks were separated by a jittered gap (4-8s) during which a red fixation cross was presented. A block consisted of 17 stimuli: eight stimuli related to that experimental condition presented twice in a pseudo-random order, with no immediate repetition, plus one out-of-category catch trial. Written stimuli were presented for 600ms; spoken stimuli were presented on average for 633.57ms (SD = 71.57ms). Words within each block were separated by a 500ms inter-stimulus interval.

Block transitions were marked with a written task instruction, which indicated (i) the aspect of meaning that participants needed to focus on and (ii) the presentation format presented in parentheses. The task instructions were presented for 3500ms (followed by 500ms fixation). A grey fixation cross against a black background was used to minimize eye movements during the duration of a block. Each block (including task instruction and jittered rest period) lasted on average 28.7 seconds.

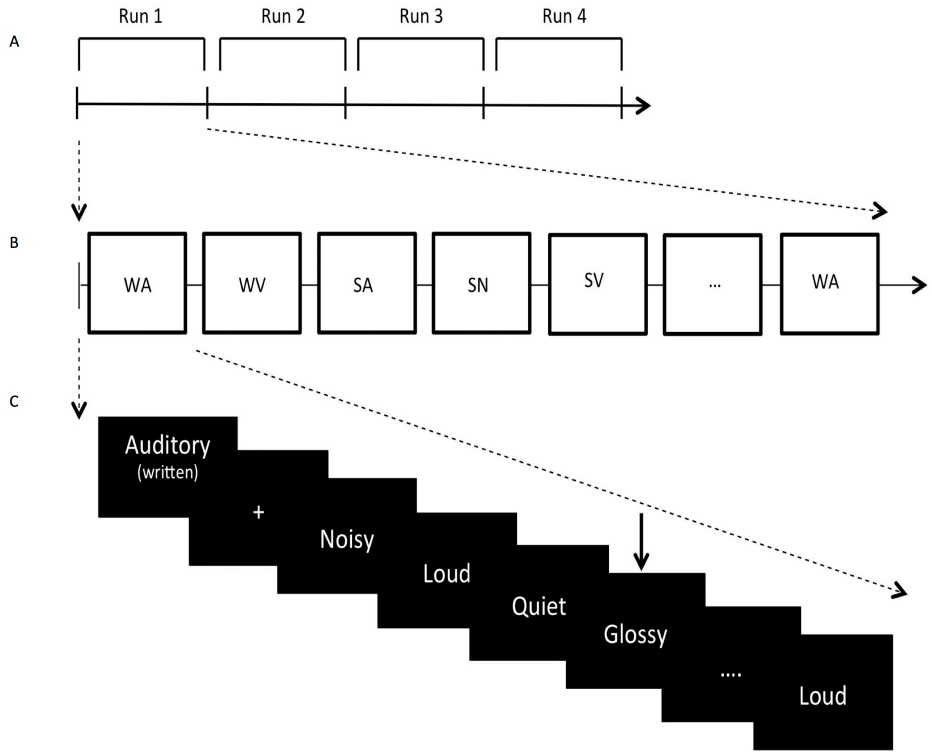

Figure 3.1. Experimental design. (A) Four runs across the fMRI session. Each run lasted no longer than 6 min 19 s. (B) Block organization across each run. WA = written-Aud, WV = written-VIS, WN = written-NON, SA = spoken-AUD, SV = spoken-VIS and SN = spoken-NON. Only 6 are depicted for illustration (from a total of 12 blocks). Each of the 6 conditions were randomly presented twice, with no immediate repetition. Written blocks lasted 22.7 seconds, spoken blocks lasted no longer than 23.2 seconds. (C) Each block began with written instructions stating the semantic feature type and presentation format, for 3500ms (followed by 500ms fixation). The 8 items from the condition were then presented twice in a random order, with no immediate repetition. Only 5 are depicted for illustration (from a total of 16 items). The arrow represents an out-of-category item (e.g., visual feature 'glossy' in a block of auditory features). In total, 17 words were presented within each block (16 targets and 1 catch trial).

3.3.1.4. Acquisition

Data were acquired using a GE 3T HD Excite MRI scanner at the York Neuroimaging Centre, University of York. A Magnex head-dedicated gradient insert coil was used

in conjunction with a birdcage, radio-frequency coil tuned to 127.4 MHz. A gradient-echo EPI sequence was used to collect data from 38 bottom-up axial slices aligned with the temporal lobe (TR = 2s, TE = 18 ms, FOV = 192 × 192 mm, matrix size = 64 × 64, slice thickness = 3 mm, slice-gap 1mm, flip-angle = 90°). Voxel size was 3 × 3 × 3 mm. Functional images were co-registered onto a T1-weighted anatomical image from each participant (TR = 7.8 s, TE = 3 ms, FOV = 290 mm x 290 mm, matrix size = 256 mm x 256 mm, voxel size = 1.13 mm x 1.13 mm x 1 mm) using linear registration (FLIRT, FSL).

3.3.1.5. Pre-processing

Imaging data were pre-processed using the FSL toolbox (http://www.fmrib.ox.ac.uk/fsl). Images were skull-stripped using a brain extraction tool (BET, Smith, 2002) to remove non-brain tissue from the image. The first five volumes (10s) of each scan were removed to minimize the effects of magnetic saturation, and slice-timing correction was applied. Motion correction (MCFLIRT, Jenkinson et al. 2002) was followed by temporal high-pass filtering (cutoff = 0.01 Hz). Individual participant data were first registered to their high-resolution T1-anatomical image, and then into a standard space (Montreal Neurological Institute (MNI152); this process included tri-linear interpolation of voxel sizes to 2 × 2 × 2 mm. For univariate analyses, data were additionally smoothed (Gaussian full width half maximum 6 mm).

3.3.1.6. Univariate Analysis

The condition onset and duration were taken from the first item in each block (after the initial instructions) to the end of the last item. The response to each of the 6 conditions was contrasted against rest. Box-car regressors for each condition, for each run, in the general linear model were convolved with a double gamma hemodynamic response function (FEAT, FSL). Regressors of no interest were also included to account for head motion within scans. A fixed effect design (FLAME, http://www.fmrib.ox.ac.uk/fsl) was then conducted to average across the four runs, within each individual. Finally, individual participant data were entered into a

higher-level group analysis using a mixed effects design (FLAME, http://www.fmrib.ox.ac.uk/fsl) whole-brain analysis.

3.3.1.7. Multivariate Pattern Analysis

Parameter estimates were calculated in the same manner as for univariate analyses, for each condition and for each run: in this way, the spatial pattern information entered into the classifier from each condition represented the average response to the 8 exemplars. This method is consistent with previous literature investigating semantic representations (Countanche & Thompson-Schill, 2012; Fairhall & Caramazza, 2013; Peelen & Caramagzza, 2012): it allows us to make inferences that a particular region is able to discriminate between words referring to auditory and visual features, for example, but not the meanings of these individual words. MVPA was conducted on spatially unsmoothed data to preserve local voxel information.

As we had a priori knowledge of strong selectivity for the classes in particular brain regions (ATL, primary auditory cortex and primary visual cortex), we opted for a ROI-based MVPA method rather than whole-brain analysis. This reduced the number of voxels used for classification (and therefore the number of free parameters which can lead to over-fitting; for similar approaches see Kamitani and Tong, 2005) and Kuhl, Rissman, Chun and Wagner, 2011). The following masks were used; primary visual cortex (taken from FSL Juelich Atlas; http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/Atlases), primary auditory cortex (taken from FSL Juelich Atlas; http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/Atlases) and ATL (anterior to Y = -22; Lambon Ralph et al. 2015). The size of these masks are as follows; primary visual cortex, 12662 voxels; primary auditory cortex, 2372 voxels; ATL, 18523 voxels.

To ensure that our ROIs had sufficient signal to detect reliable fMRI activation, the temporal signal-to-noise ratio (tSNR) for each participant was calculated for the first run of the experiment by dividing the mean signal in each voxel by the standard deviation of the residual error time series in that voxel (Friedman and Glover, 2006). tSNR values were averaged across the voxels of each ROI. Mean tSNR values, averaged across participants, were as follows: ATL, 76.74;

primary auditory cortex (PAC), 93.61; primary visual cortex (PVC), 102.96. The percentage of voxels in each ROI that had "good" tSNR values (>20; Binder et al. 2011) was above 85% for all ROIs: ATL, 86.17%; PAC, 99.87%; PVC, 94.58%. These values indicate that, although mean tSNR was lower in anterior temporal cortex than in sensory regions, the tSNR was sufficient to detect reliable fMRI activation in all ROIs (Binder et al. 2011). Moreover, to determine whether tSNR was sufficient in each sub-region of the ATL (as signal drop out is most prominent in ventral anterior regions), the tSNR was calculated for the following regions: aSTG, 85.97; aMTG, 89.00; aITG, 69.79; anterior fusiform gyrus, 69.74; anterior parahippocampal gyrus, 67.13; temporal pole, 63.27. These values suggest that, again, although mean tSNR was lower in more ventral anterior regions, it was still sufficient to detect reliable fMRI activation (Binder et al. 2011).

For each voxel in our three ROI masks, we computed a linear support vector machine (LIBSVM; with fixed regularization hyper-parameter C = 1) and a 4-fold cross-validation (leave-one-run-out) classification, implemented in custom python scripts using the pyMVPA software package (Hanke et al. 2009). A support vector machine was chosen as this aims to combat over-fitting by limiting the complexity of the classifier (Lewis-Peacock & Norman, 2013). The classifier was trained on three runs and tested on the independent fourth run; the testing set was then alternated for each of four iterations. Classifiers were trained and tested on individual subject data transformed into MNI standard space. The functional data were first z-scored per voxel within each run. The searchlight analysis was implemented by extracting the z-scored β-values from spheres (6mm radius) centered on each voxel in the masks. This sized sphere included ~ 123 3mm voxels (Kriegeskorte et al. 2006). Classification accuracy (proportion of correctly classified trials) for each sphere was assigned to the sphere's central voxel, in order to produce accuracy maps. The resulting accuracy maps were then smoothed with a Gaussian kernel (6mm FWHM). To determine whether accuracy maps were above chance-levels (50%), individual accuracy maps were entered into a higher-level group analysis (mixed effects, FLAME; http://www.fmrib.ox.ac.uk/fsl), testing the

accuracy values across subjects against chance for each voxel. Voxel inclusion was set at z = 2.3 with a cluster significance threshold at FWE $p < .05$.

The following classification tests were performed: (1) *Semantic feature classifier*: this examined whether patterns of activity conveyed information regarding the meanings of words, by training a classifier to discriminate between words referring to auditory features (e.g. loud) and visual features (e.g., shiny). This classifier was truly format-independent in the sense that it was trained on this semantic distinction using spoken words and tested using written words (and vice versa). The advantage of performing the classification in this manner is only semantic information common to both presentation formats was informative to the classifier (see Figure 3.2.A). The results from the two classifications were averaged to produce a single estimate of classification accuracy. (2) *Perceptual classifier*: here a classifier was trained to discriminate between spoken and written non-words and was tested on these two presentation formats for words. In this way only the presentation format that was general to both non-words and words was informative to the classifier (see Figure 3.2.B).

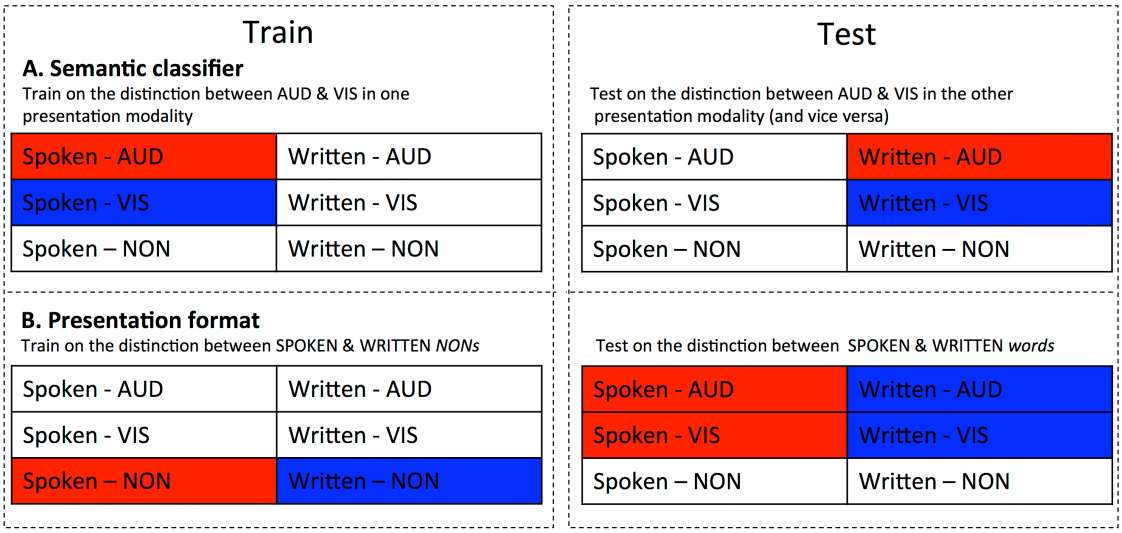

Figure 3.2. Schematic illustration of the MVPA searchlight classifiers performed. Each box includes the six experimental conditions. Classifiers were trained to distinguish between two conditions (red and blue). The classifiers were then tested on independent trials that differed in the same way. (A) Classifiers were trained and tested based on semantic content (trained on Spoken-AUD vs. Spoken-VIS, tested

on Written-AUD vs. Written-VIS – and vice versa). The results from both comparisons were then averaged. (B) Classifiers were trained and tested based on presentation format (trained on Spoken-NON vs. Written-NON, tested on Spoken words vs. Written words). (C) Classifiers were trained on one distinction and tested on the other (trained on Spoken vs. Written items, tested on AUD vs. VIS; trained on AUD vs. VIS items, tested on Spoken vs. Written items).

3.3.2. Resting state fMRI

3.3.2.1. Participants

This analysis was performed on a separate cohort of 42 healthy participants at York Neuroimaging Centre (13 male; mean age 20.31, range 18-25 years). Subjects completed a 9-minute functional connectivity MRI scan during which they were asked to rest in the scanner with their eyes open. Using these data we examined the resting-state fMRI connectivity of ATL regions that were informative to the semantic feature (aITG) and perceptual classifiers (aSTG) to investigate whether these regions fell within similar or distinct networks. In addition, we investigated the rs-fMRI connectivity of semantic regions within primary sensory cortices that showed significant decoding by the semantic classifiers to examine whether these regions overlap with the connectivity maps of the ATL seeds.

3.3.2.2. Acquisition

As with the functional experiment, a Magnex head-dedicated gradient insert coil was used in conjunction with a birdcage, radio-frequency coil tuned to 127.4 MHz. For the resting-state data, a gradient-echo EPI sequence was used to collect data from 60 axial slices with an interleaved (bottom-up) acquisition order with the following parameters: TR = 3s, TE = minimum full, volumes = 180, flip angle = 90°, matrix size = 64 × 64, FOV = 192 × 192 mm, voxel size = 3x3x3 mm. A minimum full TE was selected to optimise image quality (as opposed to selecting a value less than minimum full which, for instance, would be beneficial for obtaining more slices per TR). Functional images were co-registered onto a T1-weighted anatomical image

from each participant (TR = 7.8 s, TE = 3 ms, FOV = 290 mm x 290 mm, matrix size = 256 mm x 256 mm, voxel size = 1 mm x 1 mm x 1 mm).

3.3.2.3. Pre-processing

Data were preprocessed using the FSL toolbox (http://www.fmrib.ox.ac.uk/fsl). Prior to conducting the functional connectivity analysis, the following pre-statistics processing was applied to the resting state data; motion correction using MCFLIRT to safeguard against motion-related spurious correlations (Baker et al. 2015; Smallwood et al. 2016; Kreiger-Redwood et al. 2016; Davey et al. 2016); slice-timing correction using Fourier-space time-series phase-shifting; non-brain removal using BET; spatial smoothing using a Gaussian kernel of FWHM 6 mm; grand-mean intensity normalisation of the entire 4D dataset by a single multiplicative factor; high-pass temporal filtering (Gaussian-weighted least-squares straight line fitting, with sigma = 100 s); Gaussian lowpass temporal filtering, with sigma = 2.8 s.

3.3.2.4. Low-level Analysis

For our ATL sites we created two spherical seed ROIs, 6 mm in diameter, centered on the co-ordinates of the central voxel in the highest performing spheres in our presentation and semantic searchlight analyses; left aSTG [-54 2 -10] and aITG [-50 -10 -26] respectively (see Table 3.2). For our sensory semantic regions we created two spherical seed ROIS centered on intracalcarine cortex [-18 -84 4] and planum polare [-48 -12 -4] from the best performing spheres in our semantic searchlight analysis; as these regions showed high performance accuracy on the semantic classifier *and* fall within primary sensory regions.

The time series of these regions were extracted and used as explanatory variables in a separate subject level functional connectivity analysis for each seed. Subject specific nuisance regressors were determined using a component based noise correction (CompCor) approach (Behzadi et al. 2007). This method applies principal component analysis (PCA) to the fMRI signal from subject specific white matter and CSF ROIS. In total there were 11 nuisance regressors, five regressors from the CompCorr and a further 6 nuisance regressors were identified using the motion correction MCFLIRT. These principle components are then removed from

the fMRI data through linear regression. The WM and CSF covariates were generated by segmenting each individual's high-resolution structural image (using FAST in FSL; Zhang et al. 2001). The default tissue probability maps, referred to as Prior Probability Maps (PPM), were registered to each individual's high-resolution structural image (T1 space) and the overlap between these PPM and the corresponding CSF and WM maps was identified. These maps were then thresholded (40% for the SCF and 66% for the WM), binarized and combined. The six motion parameters were calculated in the motion-correction step during pre-processing. Movement in each of the three Cartesian directions (x, y, z) and rotational movement around three axes (pitch, yaw, roll) were included for each individual.

3.3.2.5. High-level Analysis

At the group-level the data were processed using FEAT version 5.98 part of FSL (FMRIB's Software Library,www.fmrib.ox.ac.uk/fsl) and the analyses were carried out using FMRIB's Local Analysis of Mixed Effects (FLAME) stage 1 with automatic outlier detection. The z statistic images were then thresholded using clusters determined by $z > 2.3$ and a (corrected) cluster significance threshold of $p = 0.05$ (Worsley, 2001). No global signal regression was performed.

To investigate the differences between the connectivity maps a fixed effect design (FLAME, http://www.fmrib.ox.ac.uk/fsl) was conducted for each participant to investigate four contrasts; (i) aSTG > aITG seed, (ii) aITG > aSTG seed, (iii) auditory semantic > visual semantic seed and (iv) visual semantic > auditory semantic seed. Individual participant data were then entered into a higher-level group analysis using a mixed effects design (FLAME, http://www.fmrib.ox.ac.uk/fsl) whole-brain analysis. Finally, to determine whether our ATL seeds connectivity maps overlap with the connectivity maps of the sensory semantic seeds we calculated the number of overlapping voxels for our two ATL sites and the sensory semantic connectivity maps.

3.3.3. Resting state decoder

To allow quantitative inferences to be drawn on the functional neural activity identified through our seed based correlational analyses we performed an automated meta-analysis using NeuroSynth (http://neurosynth.org/decode; Yarkoni et al. 2011). This software computed the spatial correlation between each ATL component mask and every other meta-analytic map (*n* = 11406) for each term/concept stored in the database (e.g., semantic, language, memory, sensory). The 15 meta-analytic maps exhibiting the highest positive correlation and negative correlation for each sub-system mask were extracted, and the term corresponding to each of these meta-analyses is shown in Figure 3.4. The font size reflects the size of the correlation (ranging from *r* = 0.10 to 0.45 for positive correlations and *r* = -0.05 to -0.2 for negative correlations, in increments of 0.05). This allows us to quantify the most likely reverse inferences that would be drawn from these functional maps by the larger neuroimaging community.

3.4. Results

3.4.1. Behavioural Results

Accuracy and reaction times (RT) were calculated for each participant (n=19) for the catch trials in each experimental condition. Results showed that all participants paid attention to the words as indicated by a mean accuracy above 80% for all experimental conditions (spoken AUD = 80.63% ± 15.33, spoken VIS = 88.12% ± 4.86, spoken NON = 85.62% ± 11.47, written AUD = 83.12% ± 19.01, written VIS = 86.25% ± 13.52, written NON = 88.75% ± 5.45). A chi-square test of independence revealed that accuracy did not significantly differ across the six experimental conditions (x^2(5) = 6.09, p = .303) or across spoken and written input (x^2(1) = .301, ns). RTs differed significantly between modality-input (*t*(59) = 7.36, p < .001), but not semantic-category within each modality (spoken: F(2,38) = .92, ns; written: F(2,38) = 0.074, ns). In line with previous findings (Booth et al. 2002; Cohen et al. 2004), participants were significantly faster at responding to written than spoken stimuli. Furthermore, there was no difference in RT between AUD, VIS and NON items within each presentation modality, suggesting that the

experimental conditions were well matched at the behavioural level within our stimuli subset.

3.4.2. Searchlight Analysis

3.4.2.1. Semantic-feature Classifier

The format-independent searchlight classifier, trained on the distinction between visual and auditory features in one presentation modality and tested on this distinction in the other modality, was run in three separate masks (ATL; primary auditory cortex and primary visual cortex). All results reported are above chance levels (50%, cluster corrected $p < .05$). The searchlight analysis within the ATL mask revealed a left hemisphere cluster that could decode semantic information across modalities in ventrolateral regions, namely aMTG and aITG (see Figure 3.3, Table 3.2). Additionally, right hemisphere clusters were revealed in anterior parahippocampal gyrus and temporal pole (TP). The searchlight analysis within the primary auditory mask revealed a cluster in planum polare (see Figure 3.4, Table 3.2). Finally, the primary visual cortex mask revealed a cluster in intracalcarine cortex that could decode semantic content (see Figure 3.5, Table 3.2).

3.4.2.2. Perceptual Classifier

The classifier that was trained on the distinction between spoken and written non-words and tested on the distinction between these presentation modalities for words, was also run in three separate masks (ATL; primary auditory cortex and primary visual cortex). All results reported are above chance levels (50%, cluster corrected $p < .05$). Within the ATL, anterior portions of STG, extending into temporal pole, were able to decode between presentation formats (see Figure 3.3; Table 3.2). The classifier results for the primary auditory cortex mask revealed an extensive cluster of voxels that could classify perceptual information in Heschl's Gyrus, planum temporale and superior temporal gyrus (see Figure 3.4.; Table 2). The classifier results for the primary visual cortex mask revealed an extensive cluster of voxels in occipital pole (see Figure 3.5; Table 3.2).

To explicitly determine whether the ventrolateral site (aITG) and aSTG were differentially able to classify the modality of presentation and the meaning of the stimulus, we conducted a 2 X 2 repeated-measures ANOVA in which we compared the prediction accuracies for each classifier output for each significant cluster. This revealed three significant effects. First, a main effect for classifier type (presentation format vs. semantic classifier; $F(1,18) = 36.76$, $p < .001$). Second, a significant main effect of region (aSTG vs. aITG; $F(1,18) = 79.71$, $p < .001$). Critically, we also found a significant interaction between classifier type and ATL region ($F(1,18) = 1087.51$, $p < .001$). Post-hoc tests revealed a significant difference between aSTG and aITG for the presentation format classifier, with aSTG performing significantly better than aITG ($t(18) = 29.04$, $p < .001$). There was also a significant difference between aITG and aSTG for the semantic feature classifier, with aITG performing significantly better than aSTG ($t(18) = 28.30$, $p < .001$). Collectively, these analyses show a dissociation between ATL regions: aSTG classification accuracy was higher for presentation modality than word meaning, while the reverse pattern was obtained for aITG.

In addition to our ROI-based MVPA results, a whole-brain searchlight analysis was computed for both the semantic feature classifier and perceptual classifier, using the same analysis pipeline outlined for our ROI analysis. Results from the whole-brain searchlight reveal similar clusters across primary auditory cortex, primary visual cortex and anterior temporal lobe. In addition, the whole-brain analysis revealed clusters in occipital-parietal cortex and clusters extending along the temporal lobe. The unthresholded maps from the whole-brain searchlight analysis have been uploaded to the neurovault database and can be found here http://neurovault.org/collections/1970/."

Table 3.2
Centre Voxel Coordinates of Highest Decoding Sphere in the Searchlight Analyses.

Condition	Mask	Cluster Peak	Extended Cluster Regions	Cluster Extent	Z-score	Acc (%)	x	y	z
Semantic Feature									
	ATL	L Anterior ITG/MTG	L Heschls gyrus, L putamen	478	4.91	61.22	-50	-10	-26
	ATL	R Temporal pole	R Anterior parahippocampal gyrus, R Anterior MTG, R Anterior STG.	416	4.58	61.05	42	12	-24
	Auditory	L Planum polare	L Heschls gyrus, L Planum temporale	88	3.92	59.53	-48	-12	-4
	Visual	L Intracalcarine cortex	L Lingual gyrus	81	4.26	61.18	-18	-84	4
Presentation format									
	Visual	L Occipital pole	L Occipital fusiform gyrus, L Inferior lateral occipital cortex.	607	4.3	58.57	-16	-92	0
	Auditory	L Planum temporale	L Heschl's gyrus, R Planum Temporale, R Heschl's gyrus,	581	4.97	59.85	-58	-24	8
	ATL	L Anterior STG	L Temporal pole, R Anterior STG	66	2.8	58.36	-58	-10	-2

Footnote: Highest decoding accuracy clusters for semantic feature (AUD vs. VIS) and presentation format (spoken vs. written words) analysed separately. Semantic feature classifier was trained on the distinction between spoken AUD vs. spoken VIS and tested on written AUD vs. written VIS (and vice versa). Presentation format classifier was trained on the distinction between written non-words vs. spoken non-words and tested on spoken words vs. written words. Results are thresholded at $p < .05$ (cluster corrected). L=left, R=right. As well as peak accuracy (reported under the 'Cluster Peak' column), the 'Extended Cluster Regions' includes all significant regions within each ROI. In addition to the searchlight analyses reported in the table, a further searchlight analysis was run on the distinction between all spoken vs. all written items. This revealed accuracies as high as 99.6% in primary sensory regions and 93.2% in ATL. The unthresholded MVPA maps for each searchlight have been uploaded to the Neurovault database and can be found here http://neurovault.org/collections/1970/.

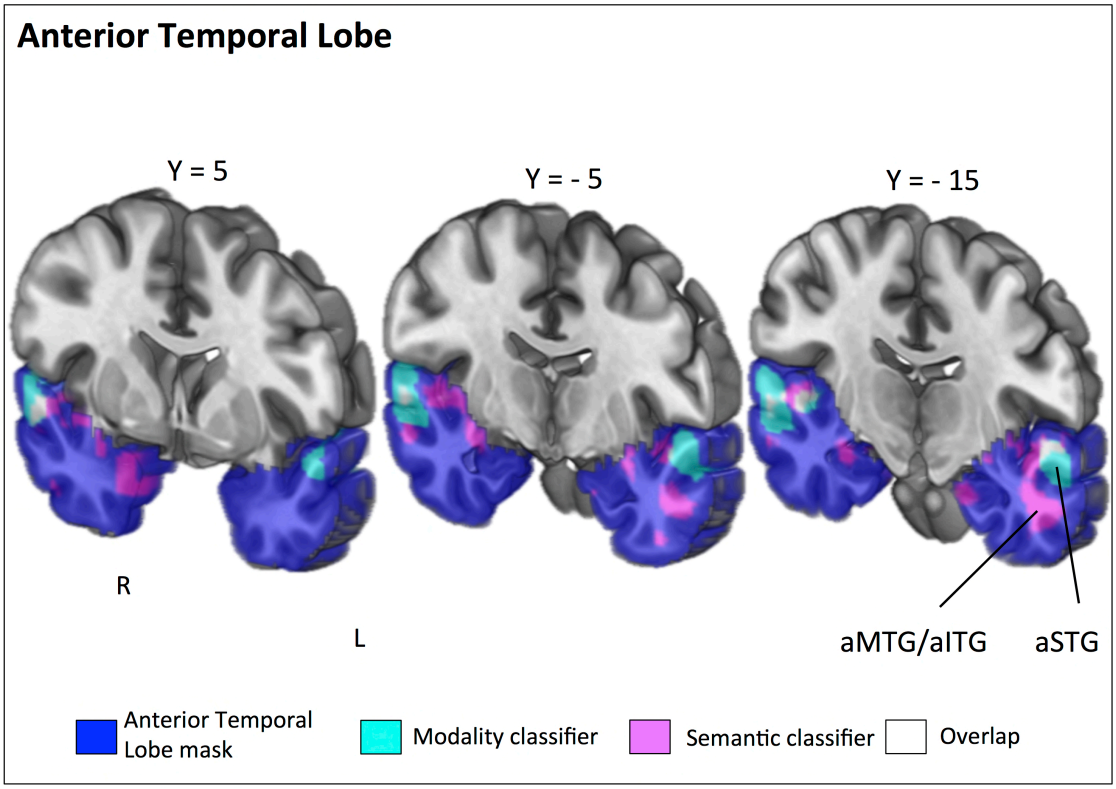

Figure 3.3. Coronal slices taken at Y = 5, Y = -5 and Y = -15. Anterior temporal lobe mask shows all regions of the temporal lobe anterior to Y = -22 in line with Lambon Ralph et al. (2015) projected in blue. Results of the group-level searchlight analysis for *semantic feature classification* (AUD vs. VIS) projected in magenta (cluster-corrected $p < .01$). Results for *perceptual classifier* (spoken vs. written) projected in cyan (cluster-corrected $p < .01$). Overlap of the two searchlight analyses in white. In total 47 voxels overlapped across the two searchlight analyses in aSTG (right hemisphere, 38 voxels; left hemisphere, 9 voxels). aSTG = anterior superior temporal gyrus; aMTG/aITG = anterior middle temporal gyrus/inferior temporal gyrus.

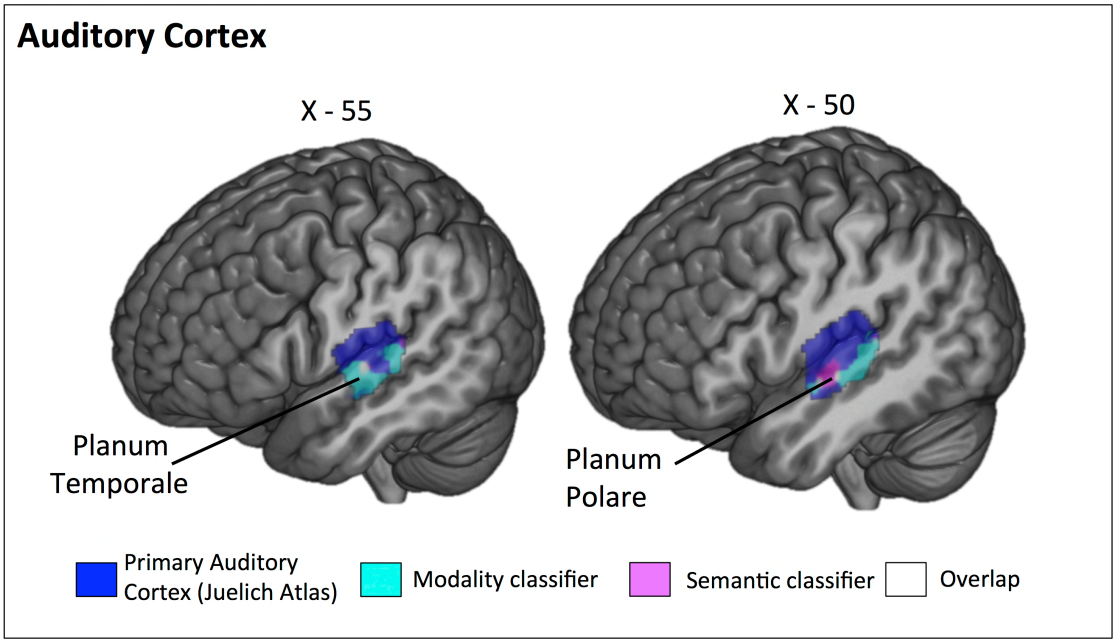

Figure 3.4. Left hemisphere sagittal slices taken at X = -55 and X = -50. Primary auditory ROI taken from Juelich histological atlases projected in blue. Results of the group-level searchlight analysis for *semantic feature classification* (AUD vs. VIS) projected in magenta (cluster-corrected $p < .01$). Results for *perceptual classifier* (spoken vs. written) projected in cyan (cluster-corrected $p < .01$). Overlap of the two searchlight analyses in white.

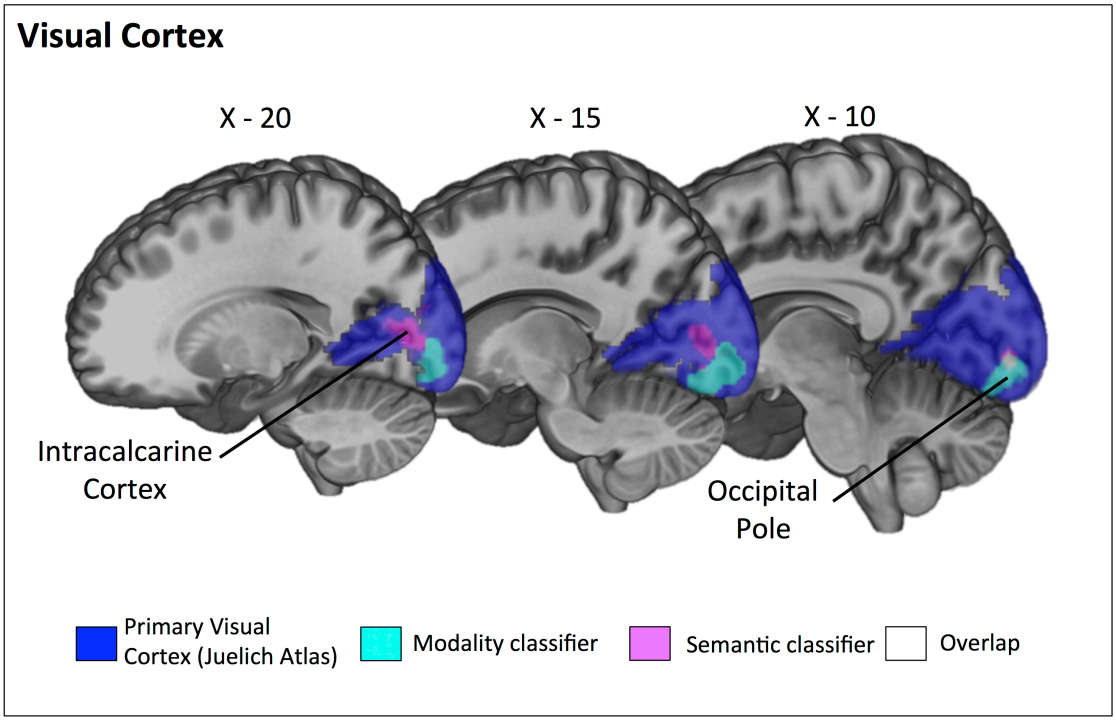

Figure 3.5. Left hemisphere sagittal slices taken at X = -20, X = -15 and X = -10. Primary visual ROI taken from Juelich histological atlases projected in blue. Results of the group-level searchlight analysis for *semantic feature classification* (AUD vs. VIS) projected in magenta (cluster-corrected $p < .01$). Results for *perceptual classifier* (spoken vs. written) projected in cyan (cluster-corrected $p < .01$). Overlap of the two searchlight analyses in white.

3.4.3. Univariate Analysis

The searchlight results revealed that in ATL, primary auditory cortex and visual cortex, distinct regions were able to decode semantic feature type and presentation modality. As an additional complementary analysis, the percentage signal change was extracted for each condition from the pairs of clusters that were able to decode semantic feature type and modality of presentation in ATL, visual cortex and auditory cortex (generating six analyses; see Figure 3.6). A 6mm sphere was centered at the peak MVPA accuracy in each of these sites (see Table 3.2). The ventrolateral ATL region (encompassing aITG and aMTG, decoding feature type) showed deactivation across all four conditions, and the degree of deactivation was sensitive to meaning (auditory > visual features) but not input modality (spoken = written words). In contrast,

aSTG (which decoded presentation modality) was sensitive to modality (spoken > written) but not meaning (auditory = visual features). Thus, univariate analyses also revealed a functional dissociation within ATL. We also examined regions that could decode modality of presentation and semantic feature type within primary auditory cortex (planum temporale and planum polare respectively) and primary visual cortex (occipital pole and intracalcarine cortex). All four sites showed strong effects of input modality in univariate analyses across both feature types. In addition, the intracalcarine cortex showed greater activity to words that denoted a visual property (e.g., bright) whereas planum polare showed greater activation to words that denoted an auditory property (e.g., loud). This effect of meaning in primary visual and auditory areas was only seen when the words were presented in the complementary input modality: primary visual cortex responded more to visual features when written words were presented, while primary auditory cortex responded more to auditory features when spoken words were presented. Thus, aITG was unique in showing a pattern across both multivariate and univariate analyses consistent with the predictions for a transmodal 'hub': i.e., sensitivity to meaning and insensitivity to presentation modality.

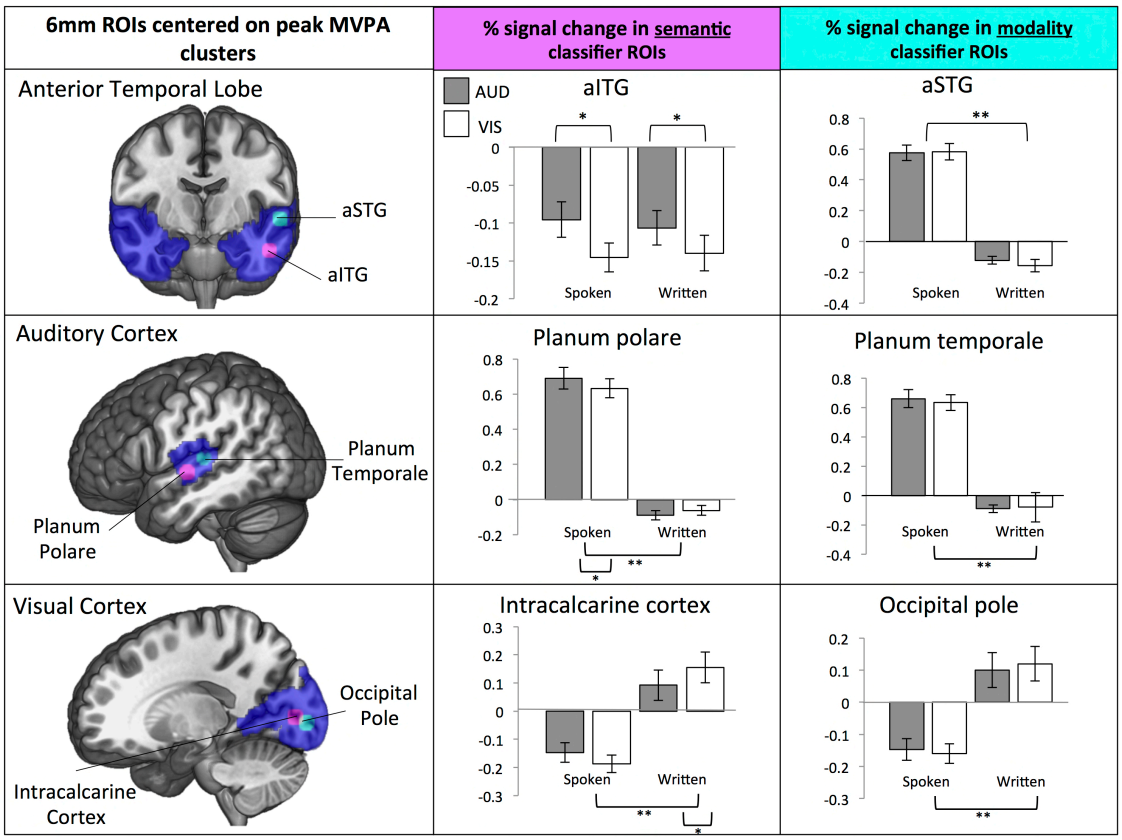

Figure 3.6. The first column shows 6mm ROIs centered on the peak MVPA results from the searchlight analyses (shown in figure 3.3-3.5) for semantic classifier in magenta and modality classifier in cyan, for each of our three masks (anterior temporal lobe, primary auditory cortex and primary visual cortex) projected in blue. The centre for these ROIs are as follows; aITG seed [-50 -10 -26], aSTG seed [-58 -10 -2], planum polare [-48 -12 -4], planum temporale [-58 -24 8], intracalcarine cortex [-16 -84 4] and occipital pole [-16 -92 0]. The second column shows the univariate percent signal change for each of our four conditions within the semantic (magenta) ROI. The third column shows the univariate percent signal change for each of our four conditions with the modality (cyan) ROI. Grey bars show the results for auditory-feature words (e.g., 'loud') and white bars show the results for visual-feature words (e.g., 'bright). * indicates a significant difference between auditory-features and visual-features within a modality (i.e., spoken auditory-features and spoken visual-features; $p < .05$). ** indicates a significant difference between spoken and written presentation format ($p < .001$). The unthresholded univariate maps for each

condition have been uploaded to the Neurovault database and can be found here http://neurovault.org/collections/1970/

3.4.4. Resting-state fMRI

To provide a better understanding of the neural architecture that supported the functional distinction between aSTG (effect of input modality) and aITG (effect of semantic feature type), we explored the connectivity of these regions in functional fMRI (see Figure 3.7) by placing spherical ROIs at peaks in the MVPA analysis. The aSTG seed showed significant positive connectivity across the entire length of STG through primary auditory cortex and into supramarginal gyrus (SMG). It coupled with posterior and anterior regions of MTG, pre- and post-central gyrus, supplementary motor cortex and anterior cingulate gyrus and deactivation with visual regions, including lateral occipital cortex, intracalcarine cortex, occipital fusiform gyrus (OFG) and temporal occipital fusiform gyrus, as well as posterior cingulate and precuneous. In contrast, the aITG site showed connectivity with core parts of the default mode network and multimodal semantic regions, including angular gyrus, posterior parts of MTG and ITG, temporal pole extending medially to include hippocampus and anterior parahippocampal gyrus, and anterior and inferior prefrontal regions, including orbital cortex and left inferior frontal gyrus (LIFG). This seed also coupled with lateral visual regions (e.g., LOC and occipital fusiform gyrus). Table 3.3 presents location and size of each of these clusters.

Table 3.3. - Coordinates of Peak Clusters in the Resting-state Connectivity Analyses

Seed Region	Cluster	Cluster Extent	Z-score	x	y	z
aSTG	*Increased Correlation*					
	L. aSTG	15745	12.3	-54	2	-10
	R. Temporal pole	12970	9.24	52	8	-14
	Cingulate Gyrus	7618	7.02	-4	12	32
	Reduced Correlation					
	L. Cuneal cortex	26667	6.19	-20	-74	32
	R. Superior frontal gyrus	4128	4.69	20	12	52

	L. Middle frontal gyrus	2259	4.53	-32	10	50
	L. Lateral occipital cortex, inferior	1457	5.46	-46	-70	-12
aITG	*Increased Correlation*					
	L. aITG/MTG	20324	13.1	-50	-10	-26
	L. Frontal pole	2899	7.22	-10	50	32
	L. Occipital fusiform gyrus	1981	4.49	-26	-82	-8
	Reduced Correlation					
	Postcentral gyrus	3725	4.44	0	-54	74
	R. Frontal pole	2717	5.07	42	54	12
	L. IFG, pars triangularis	2118	5.17	-46	35	16
	R. Cingulate gyrus	1276	4.44	12	32	16
	L. Angular gyrus	783	4.39	-40	-50	42
	L. Superior parietal lobule	769	3.94	-30	-48	-56
	L. Middle frontal gyrus	724	4.72	-28	8	60
	R. Middle frontal gyrus	626	4.16	30	12	56

Footnote: The table shows peak clusters in the resting-state connectivity analysis from two seed regions; aSTG and aITG. Results are thresholded at $p < .01$ (cluster corrected). L = left, R = right.

To investigate the differences between these two ATL maps a difference analysis was performed (Figure 3.7.B). The contrast of aSTG > aITG identified bilateral superior temporal and frontal polar regions. The contrast aITG > aSTG revealed bilateral inferior and middle portions of the temporal lobe and multimodal semantic sites including angular gyrus, pMTG and LIFG. These differences resemble resting state differences for aSTG and vATL reported by Jackson et al. (2016), helping to validate the functional dissociation we observed using MVPA.

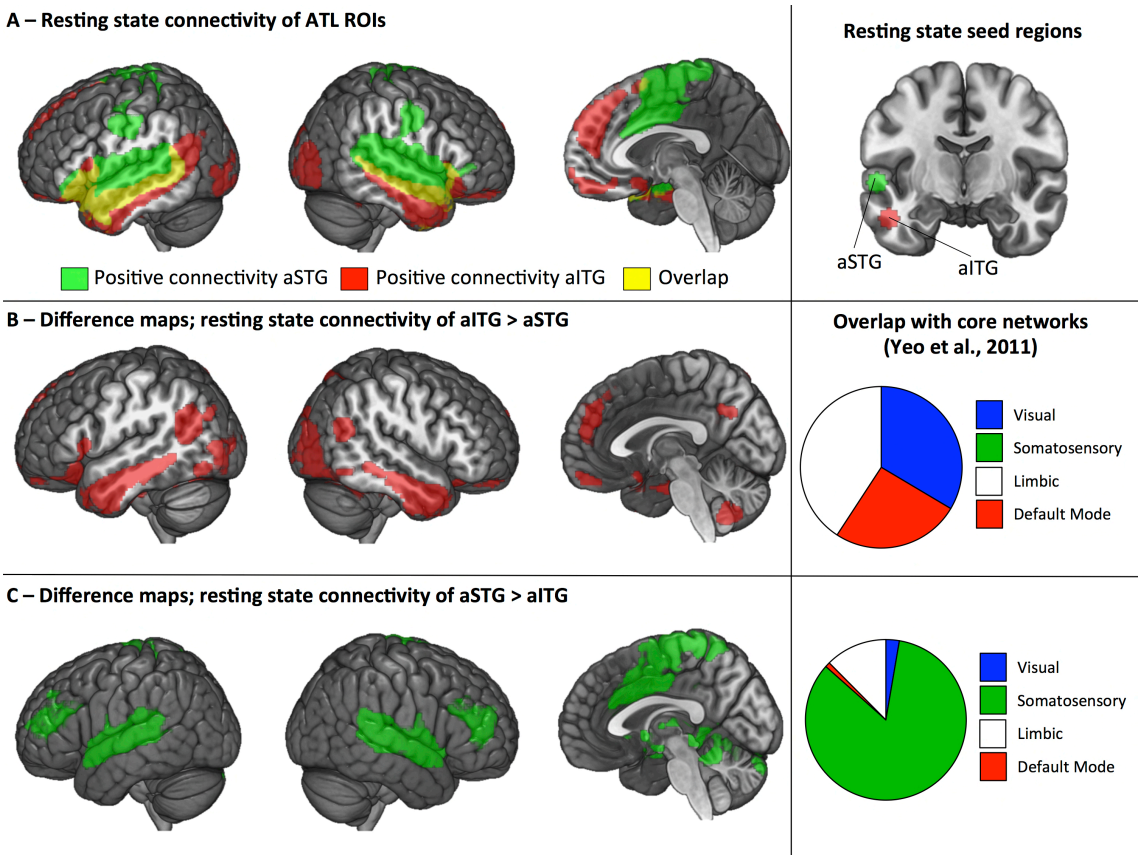

Figure 3.7. Resting state connectivity maps projected on rendered brain, displaying (from left-to-right) left hemisphere, right hemisphere, medial view. Maps thresholded at z = 2.3, cluster corrected p < .01. (A) Resting state connectivity from two ATL regions connectivity maps; green seed = aSTG (taken from peak accuracy for modality classifier within anterior temporal lobe) and red seed = aITG (taken from peak accuracy for semantic classifier within anterior temporal lobe) - the seed locations are highlighted on the right. (B) Subtraction analysis from two ATL connectivity maps; red = aITG > aSTG. Pie chart on the right shows proportion of overlapping voxels for this difference map with core networks taken from Yeo et al. (2011). These four networks include two sensory maps (Visual, Somatosensory), Limbic and Default Mode Network. (C) Subtraction analysis from two ATL connectivity maps; green = aSTG > aITG. Pie chart on the right shows proportion of overlapping voxels for this difference map with core networks taken from Yeo et al. (2011).

To further interrogate the assumption that aITG exhibits a connectivity profile consistent with an transmodal region, whereas aSTG is connected to sensory regions, we looked at the similarity between our two ATL difference maps (see Figure 3.7.B and C) and that of four core networks taken from Yeo et al. (2014). These included two networks sensitive to sensory input (visual, somatosensory) and two networks thought to be crucial in the generation of cognitive states that do not rely on sensory inputs for their mental content (limbic and default mode network) (for a review see Andrews-Hanna, Smallwood and Spreng, 2014). The results, outlined in Figures 3.7.B and 3.7.C, indicated substantial overlap between the sensory networks (namely somatosensory) and aSTG. In contrast, aITG showed substantial overlap with limbic and DMN networks.

3.5. Discussion

The current study used multiple imaging methods to identify regions in the anterior temporal lobe (ATL) and primary sensory regions that showed the pattern expected for the semantic hub of the hub-and-spokes model (Patterson et al. 2007). In an fMRI study, participants listened to or viewed words that referred to either visual or auditory features (e.g., BRIGHT or LOUD). Multivoxel pattern analysis (MVPA) revealed a dissociation between (i) ventrolateral ATL (aMTG / aITG), which could classify semantic categories relating to feature type (e.g., auditory features like “loud” as being different from visual features like “bright”) across auditory and visual inputs and (ii) anterior superior temporal gyrus (aSTG), which was sensitive to input modality across meaningful and meaningless items. This dissociation within ATL was further supported by univariate contrasts and patterns of functional connectivity: aSTG showed a stronger response to spoken than written inputs and was functionally coupled to an auditory-motor network (somatosensory network; Yeo et al. 2014), while ventrolateral ATL was insensitive to input modality and showed substantial connectivity with regions in the default mode network and limbic network, plus

some overlap with visual regions (see Jackson et al. 2016, for similar findings).

Our findings make an important contribution to our understanding of the neural basis of semantic cognition in three ways: (1) We provide evidence that conceptual knowledge, extracted from different modalities of input across many learning experiences, is represented within ventrolateral portions of ATL which act as a 'hub' (Patterson et al. 2007; Rogers et al. 2004). (2) Across converging methods, we observe a functional dissociation between ventrolateral and superior portions of ATL and provide evidence that these regions are situated within distinct large-scale cortical networks. (3) Responses in primary visual and auditory cortex confirm the contribution of these 'spoke' regions to semantic processing.

According to the hub-and-spoke model (Patterson et al. 2007), conceptual knowledge depends on the co-activation of spoke regions that convey information about specific unimodal and multimodal features of concepts, and an ATL hub, which integrates these features to form transmodal conceptual representations that are independent of specific sensory input. Studies of patients with semantic dementia (SD) provided the original motivation for this proposal yet neuropsychological methods are not especially well-suited to the precise localization of transmodal conceptual representations given the widespread atrophy in this condition. Nevertheless, the degree of semantic impairment correlates with hypometabolism in ventral rather than superior portions of ATL across patients (Mion et al. 2010), suggesting that ventrolateral ATL could be the critical substrate for transmodal knowledge. Relevant evidence is also provided by univariate fMRI analyses of the ATL response to verbal comprehension tasks in healthy participants, which show multiple peak responses in both ventral ATL and aSTG, often to the same contrasts (Binney et al. 2010; Hoffman et al. 2015; Visser & Lambon Ralph, 2011). Semantic matching and naming tasks have also shown multiple peak responses in the ATL with the more superior ATL region being involved in object naming and the more ventrolateral region in semantic matching (Sanjuán et al. 2015). Furthermore, the differential patterns of functional connectivity across ATL

regions have been observed by both Jackson et al. (2016) and Pascual et al. (2015).

Our findings therefore add to existing knowledge by showing a dissociable response in these two regions: only the ventrolateral ATL site showed a pattern consistent with the representation of conceptual information, since it was able to classify responses according to semantic category (i.e., feature type, not input modality). In univariate analyses, this site also showed deactivation (arguably due to the use of rest rather than an active baseline) for both auditory and visual feature types, irrespective of whether these words were spoken or written – and the magnitude of this deactivation was greater for visual than auditory features. Finally, this site showed stronger functional connectivity at rest with the default mode and limbic systems, as expected for a region implicated in transmodal conceptual processing. Therefore, our combination of functional and resting state methods provides novel converging evidence that anterior ventrolateral temporal areas allow different sensory representations to be integrated to form 'transmodal' conceptual representations (particularly for auditory features, see limitations below).

Previous studies have used MVPA to explore the neural basis of semantic processing, and have identified a conceptual response in ATL using classification of stimuli within a single presentation modality (Coutanche & Thompson-Schill, 2014; Peelen & Caramazza, 2012). Other studies, examining semantic cognition across modalities of presentation (Devereux et al. 2013; Fairhall & Caramazza, 2013; Man et al. 2015), have largely not observed effects in ATL. An exception is a recent crossmodal MVPA study, investigating Dutch-English bilinguals (Correia et al. 2014). The research tested whether patterns of activity related to the distinction between spoken nouns in one language (e.g., "horse" vs. "duck" in English) could accurately predict the same distinction in the other language (e.g., "paard" vs. "eend" in Dutch). Consistent with our findings, the cross-language classifier revealed a significant cluster in the left ATL. This largely fell within mid-superior temporal pole rather than the more ventrolateral regions we identified in our analysis, perhaps because aSTG is an important interface between semantic processing and other aspects of language.

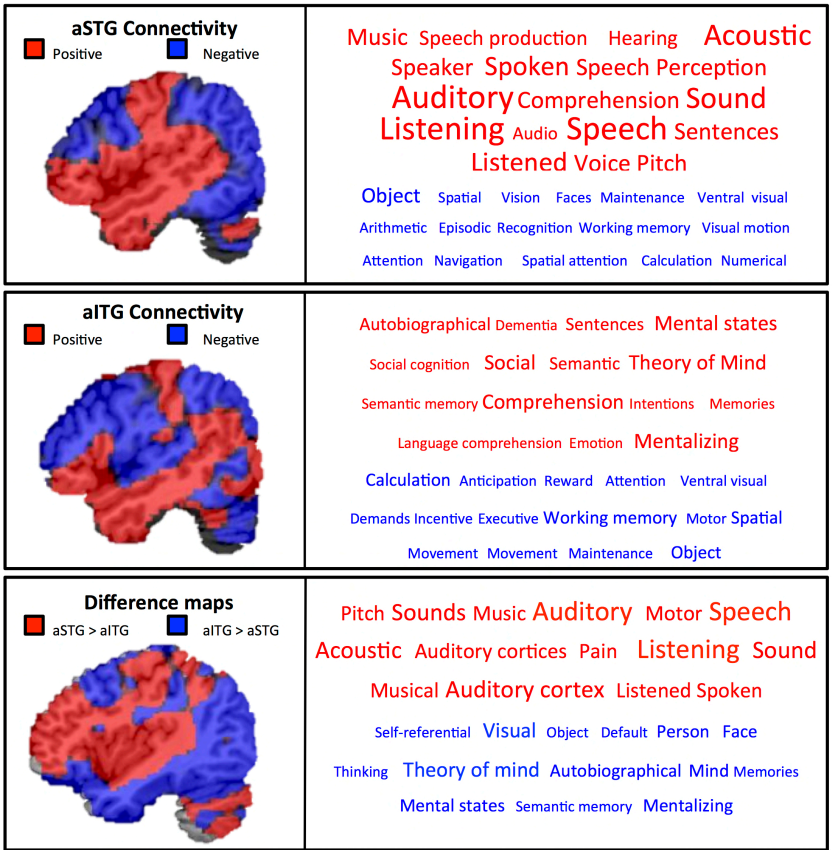

Figure 3.8. Decoding the functions of two ATL components (aSTG and aITG) using automated fMRI meta-analyses (NeuroSynth, Yarkoni et al. 2011). This software computed the spatial correlation between each ATL component unthresholded zstat mask (shown on the left; red = positive correlation and blue = negative correlation) and every other meta-analytic map (*n* = 11406) for each term/concept stored in the database (e.g., semantic, language, memory and sensory). The 15 meta-analytic maps exhibiting the highest positive correlation (red words) and negative correlation (blue words) for each sub-system mask were extracted, and the term corresponding to each of these meta-analyses is shown in the respective box (shown on the right). The font size reflects the size of the correlation (ranging from *r* = 0.10 to 0.45 for positive correlations (red) and *r* = -0.05 to -0.2 for negative correlations (blue), in increments of 0.05). This allows us to quantify the most likely reverse inferences that would be drawn from these functional maps by the larger neuroimaging community.

1

Analyses of functional connectivity from the ATL regions that were able to classify input modality (aSTG) and semantic feature type (aITG) revealed that these two sites lie within distinct large-scale functional networks. A similar dissociation between the resting state connectivity of ventrolateral ATL and anterior STG was recently reported by Jackson et al. (2016), providing further evidence for the validity of the functional dissociation in ATL that we observed using MVPA. To quantify the interpretation of the functional connectivity of the aSTG and aITG connectivity maps, we performed a decoding analysis using automated fMRI meta-analytic software NeuroSynth (see Figure 3.8). Meta-analytic decoding of these spatial maps revealed that our aSTG connectivity map correlated with terms related to language (e.g., sentence, comprehension) and auditory processing (e.g., speech, sound) whilst anti-correlating with other modality information (e.g., visual, spatial) and memory (e.g., working memory, episodic). In contrast, the aITG connectivity map correlated with terms related to memory (e.g., semantic, autobiographical) and social processes (e.g., theory of mind, social cognition) terms, whilst anti-correlating with modality-specific (e.g., ventral visual, motor, spatial) and executive terms (e.g., maintenance, demands). This is consistent with previous findings that relate aSTG to speech comprehension, language and sensory processing (Patterson & Lambon Ralph, 1999; Jobard, Vigneau, Mazoyer & Tzourio-Mazoyer, 2007; Scott & Johnsrude, 2003; Scott, Leff & Wise, 2003; Scott, Blank, Rosen & Wise, 2000; Spitsyna et al. 2006) and ventrolateral ATL to semantic processing but not sensory experience (Margulies et al. 2016; Patterson et al. 2007; Visser et al. 2010). Furthermore, the differences in function across temporal areas as revealed by the Neurosynth database seem to align with differences in the white-matter terminations (see Bajada et al. 2016). These findings confirmed associations between (i) the network anchored in the aSTG and auditory processing and speech perception, plus (ii) the aITG network and more abstract domains (such as social cognition, theory of mind, or mental states).

Thus, the putative semantic 'hub' in ventrolateral ATL was functionally coupled to aspects of cortex that specialize in forms of stimulus-independent higher order cognition, including angular gyrus (AG) and posterior and anterior areas on the medial surface that correspond to the midline core of the so-called default mode network (DMN)(see also Hurley et al. 2015). This network is known to be deactivated by input

(Raichle et al. 2001) and is thought to be crucial in the generation of cognitive states that do not rely on sensory information for their mental content (for a review see Andrews-Hanna, Smallwood and Spreng, 2014). Tasks which are associated with the default mode network include those that depend on episodic memory, semantic processing, mental state attribution as well as states of spontaneous thought studied under the rubric of mind-wandering / daydreaming (Spreng, Mar & Kim, 2009; Raichle, 2015). Although previous literature has shown that connectivity to the AG may not be due to shared semantic processing (Humphreys et al. 2015). Therefore, as many cognitive states that involve the DMN are stimulus-independent in nature, their association with ventrolateral ATL both in terms of functional connectivity and their meta-analytic decoding is consistent with the view that this region supports semantic processing across different input modalities and may form conceptual representations that are not tied to a specific input modality. In contrast, aSTG showed greater functional connectivity with auditory and motor regions and this spatial map was associated with auditory processing and language tasks, as opposed to transmodal tasks, in the meta-analytic decoding. Therefore, our combination of functional and resting state methods provides novel converging evidence that anterior ventrolateral temporal areas allow different sensory representations to be integrated to form 'transmodal' conceptual representations.

As discussed, the hub and spoke model (Lambon Ralph, Sage, Jones & Mayberry, 2010; Patterson et al. 2007; Rogers et al. 2004) makes novel predictions about the contribution of the ATL to transmodal conceptual knowledge, but it also anticipates an important role for modality-specific 'spoke' regions in visual and auditory cortex, in line with many influential accounts of semantic processing (Damasio, 1989; Martin, 2007; Meteyard, Cuadrado, Bahrami, & Vigliocco 2012; Pulvermüller, 2013). Furthermore, the involvement of both hub and spoke regions in semantic representations has been shown using TMS (Pobric et al. 2010). In line with this view, MVPA revealed regions that responded to meaning in both ventrolateral parts of ATL (putative 'hub') and in primary visual and auditory regions (putative 'spokes'). In addition, even though the putative 'spoke' regions (i.e., voxels sensitive to meaning) were adjacent to areas that coded for input modality, the specific voxels that could classify meaning and input modality were largely different. These findings do not

readily support traditional 'strong' embodied accounts that equate semantic representations with traces of perceptual/motor experience (for a review, see Meteyard et al. 2012) since this would suggest a greater degree of overlap between the results of these two classifiers. While our data suggests that sensory systems appear to play a critical role in the representation of meaning, they also suggest that perceptual experience and imagery generated as part of semantic retrieval may be distinguishable on the basis of differences in the patterns of activity in sensory cortex.

One potential limitation of our study is that we did not observe evidence that aITG responds to both auditory and visual semantic features in the univariate contrasts: this site showed deactivation for both feature types that was greater for visual features. Thus, the strongest evidence for the aITG as an transmodal hub is provided by the MVPA results and our meta-analytic decoding of this region's pattern of distinct functional connectivity, and not the univariate analyses. Our design was optimized for decoding rather than univariate effects – as we focused on obtaining the maximum number of blocks for MVPA and did not employ a high-level non-semantic baseline which would have allowed us to recover semantic activation in ATL for both auditory and visual features from a contrast (Humphreys et al. 2015). Since we found that ventrolateral responds more to auditory features (words such as "loud") than visual features (words such as "bright"), it remains unclear whether this region reflects the meanings of auditory features alone, or both feature types equally. Future studies might allow these possibilities to be disentangled using a high-level baseline with which both feature types can be compared (e.g. Jackson et al. 2015).

3.6. Conclusion

Collectively, our findings from both pattern classification and functional connectivity provide converging evidence that sub-regions of the ATL support different aspects of semantic processing. Anterior ITG and MTG capture meaning independent of input modality, consistent with the fact that semantic dementia patients (who have multimodal semantic impairment) have considerable atrophy in this same region of ATL (Binney et al. 2010; Galton et al. 2001). In contrast, aSTG exhibited a degree of modality specificity: this structure, which is known to be important for understanding

speech and environmental sounds, does not fulfil the criteria for an transmodal semantic hub. Finally, the current results provide evidence for modality-specific spokes regions within the vicinity of primary auditory and visual cortex (intracalcarine cortex and planum polare respectively). However, the specific voxels that could classify between each condition (presentation format and semantic feature) were largely different. These findings challenge traditional embodied accounts (Pulvermüller, 2005) that attempt to equate semantic representations with traces of perceptual/motor experience, and instead support the view that the richness of semantic cognition arises at least in part from abstraction away from specific input modalities in ventrolateral regions of the anterior temporal lobe.

Chapter 4 – Imagining sounds and images: Decoding the contribution of unimodal and transmodal brain regions to semantic retrieval in the absence of meaningful input

This chapter is adapted from: Murphy, C., Rueschemeyer, S. A., Smallwood, J., & Jefferies, E. (in review). Imagining sounds and images: Decoding the contribution of unimodal and transmodal brain regions to semantic retrieval in the absence of meaningful input.[3]

4.1. Abstract

In the absence of sensory information, we can generate meaningful images and sounds from representations in memory. However, it remains unclear which neural systems underpin this process, and whether different types of imagery recruit similar or different neural networks. We asked people to imagine the visual and auditory features of objects, either in isolation (car, dog) or in specific meaning-based contexts (car/dog race). Using an fMRI decoding approach, in conjunction with functional connectivity analysis, we examined the role of primary auditory/visual cortex and transmodal brain regions. Conceptual retrieval in the absence of external input recruited sensory and transmodal cortex. The response in transmodal regions – including anterior middle temporal gyrus – was of equal magnitude for visual and auditory features, yet nevertheless captured modality information in the pattern of response across voxels. In contrast, sensory regions showed greater activation for

[3] The author, Charlotte Murphy, designed the experiment, analysed the results and wrote the article under the supervision of Prof. Beth Jefferies, Dr. Shirley-Ann Rueschemeyer and Dr. Jonathan Smallwood

modality-relevant features in imagination (even when external inputs did not differ). These data are consistent with the view that transmodal regions support internally-generated experiences through the integration of stored perceptual information encoded in memory.

4.2. Introduction

In the absence of sensory information, the mind produces experiences with rich sensorimotor features through the retrieval of information from memory (Singer, 1966; Antrobus, Singer & Greenberg, 1966; Mason et al. 2007). For instance, in everyday life we regularly hear voices and music in the mind's ear when no sound is delivered (e.g., Alderson & Fernyhough, 2015; Halpern, 2001) and studies suggest more than one third of our time is spent engaged in thoughts and experiences that are unrelated to the ongoing environment (Kane et al. 2007; Killingsworth & Gilbert, 2010). Although attempts have been made to understand how the brain retrieves memories in the absence of input (Albers et al. 2013; Daselaar, Porat, Huijbers & Pennartz, 2010; Vetter, Smith & Muckli, 2014), we lack an account of the component neurocognitive processes critical for mental imagery, whether these vary with respect to the modality of the memories being retrieved, and how these processes combine to support more complex multi-dimensional aspects of cognition. Studies of imagination have almost entirely focused on a constrained regions-of-interest analysis, which may not adequately represent the rich involvement of multiple brain regions distributed across the cortex. Moreover, they have seldom attempted to differentiate between different forms of imagery, with the majority of studies focusing solely on visual imagery (Albers et al. 2013; Countanche & Thompson-Schill, 2014; Dijkstra et al. 2017; Ishai, Ungerleider & Haxby, 2000; Lee, Kravitz & Baker, 2012; Reddy, Tsuchiya & Serre, 2010; Stokes, Thompson, Cusack & Duncan, 2009; Vetter et al. 2014). As such, there is limited understanding of the neural signature of different modalities (e.g., visual versus auditory), and whether different forms of imagination share similar or unique neural representations. Notably, studies that have compared visual and auditory

imagery within the same experiment have been criticized for not employing comparable task conditions (see Daselaar et al. 2010; Halpern et al. 2004).

We addressed these issues by applying multivoxel pattern analysis (MVPA) and resting-state functional magnetic resonance imaging to identify neural patterns that support different aspects of imagination at the whole-brain level. Using a constant source of visual and auditory noise as a baseline, participants were asked to imagine information under three different conditions: visual (e.g. what a dog looks like), auditory (what a dog sounds like) and contextual (e.g. imagining a dog in a specific context, such as a race dog). This latter condition combines features from multiple modalities in a complex way (e.g., imagining a race dog may involve the visual properties of a greyhound and race track, as well as the auditory properties of dogs panting and crowds cheering). Figure 4.1 presents a schematic description of the experimental design used in our experiment. We compared the time points during which participants imagined a given concept whilst observing visual and auditory noise to those in which participants only observed visual and auditory noise (baseline). Our paradigm, therefore, permitted us to investigate the mechanisms involved in imagery whilst controlling for sensory input across our conditions.

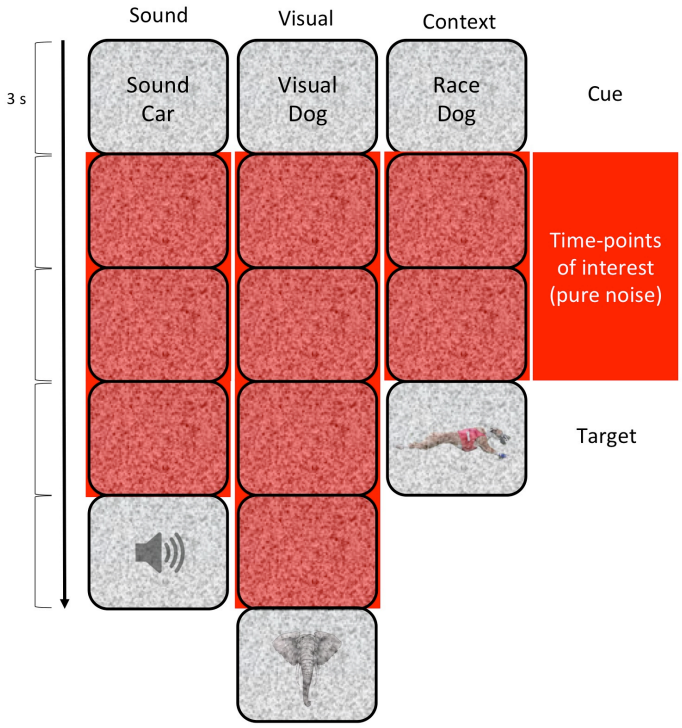

Figure 4.1. Experimental design. Participants were presented with written cues embedded in visual and auditory noise that referred to items they must detect. Cues referred to one of three tasks (Thinking about the *sound* of a concept; Thinking about the *visual* properties of a concept; Thinking about a concept in a particular complex *context* i.e., at the races) for one of two concepts (Dogs; Cars). This yielded six experimental conditions (Sound Car; Sound Dog; Visual Car; Visual Dog; Context Car (e.g., Race Car); Context Dog (e.g., Race Dog)). Cues were followed by blocks of pure noise that lasted 6-12 s. Each block ended with either an image or a sound embedded in noise, that was either congruent to the cue (e.g., greyhound for the context cue 'Race Dog') or incongruent (e.g., elephant trunk for the visual cue 'Visual Dog'). Participants responded with a yes/no response to whether the target trial matched the cue. Time points of interest are highlighted in red, these refer to pure noise trials where participants were thinking about the relevant cue (e.g., thinking about what a sound looked like). Cues, each pure-noise image and targets were shown for 3 s each.

A wealth of evidence supports the view that regions of unimodal sensory cortex are important for modality-specific elements of memory retrieval during imagination. Visual cortex is activated by mental images (Albers et al. 2013; de Borst & de Gelder, 2016; Ishai et al. 2000; Reddy et al. 2010; Vetter et al. 2014) and auditory cortex is activated by imagined sounds (Daselaar et al. 2010; de Borst & de Gelder, 2016; Halpern & Zatorre, 1999; Zvyagintsev et al. 2013). These findings are consistent with embodied cognition accounts, which propose that sensory regions important for perception and action also support mental processes such as comprehension and imagery (for discussion, see Barsalou, 1999; 2008; Patterson, Nestor & Rogers, 2007; Kiefer & Pulvermüller, 2012). Notably, the majority of studies find recruitment of sensory association cortices during visual (Amedi et al. 2005; Ishai et al. 2000; Knauff et al. 2000) and auditory mental imagery (Bunzeck et al. 2005; Zatorre & Halpern, 2005). Moreover, a recent fMRI study showed that both secondary sensory regions and top-down mechanisms are necessary in visual imagery for enhancing the relevant representations in early sensory areas (Dijkstra et al. 2017). However, some studies have also found imagery-induced activation in primary sensory cortex (Kosslyn et al. 1999; 2001; Slotnick, Thompson & Kosslyn, 2005), and the extent to which primary and/or secondary sensory regions are recruited during different modalities of imagery remains a source of contention (Daselaar et al. 2010; Kosslyn et al. 2001). By directly

comparing visual and auditory imagery under equivalent conditions in the same experiment, the present study can elucidate the role of primary and secondary sensory cortex in mental imagery.

Contemporary accounts of semantic cognition suggest that memory retrieval also relies on abstract representations that are largely invariant to the input modality. A prominent theory of conceptual representation, known as the hub-and-spoke account, suggests that input-invariant concepts draw on a convergence zone in the ventrolateral anterior temporal lobes (ATL), which extracts deep semantic similarities across multiple unimodal features (Lambon Ralph, Jefferies, Patterson & Rogers, 2017; Patterson et al. 2007). Support for this account comes from a recent fMRI study utilizing MVPA, which demonstrated that ventrolateral portions of the ATL (anterior inferior and middle temporal gyrus) supported modality-invariant patterns of activity corresponding to meaning, whereas superior temporal voxels held patterns of activity that reflected sensory input modality (Murphy et al. 2017). Therefore, if ventrolateral ATL represents abstract conceptual representations, as expected for a transmodal brain region (Margulies et al. 2016; Mesulam, 2012), it may be critical for stimulus-independent cognition regardless of the modality of what is being imagined.

In line with this broad perspective, studies have revealed ATL activation during the retrieval of concepts across modalities (Coutanche & Thompson-Schill, 2014; Gabrieli et al. 1997; Murphy et al. 2017; Reilly, Garcia & Binney, 2016; Rice et al. 2015; Van Ackeren & Rueschemeyer, 2014; Visser, Jefferies & Lambon Ralph, 2010). Coutanche and Thompson-Schill (2014) found that left ATL could successfully decode the properties of an imagined object. In this study, classification accuracy in early visual regions, related to the shape (in V1) and colour (in V4) of the object, predicted classifier accuracy for the specific object in ATL. This is consistent with the hypothesis that information from sensory cortex is integrated in ATL to form modality-invariant conceptual representations that are critical for perceptually-decoupled semantic cognition, as well as for the comprehension of words and objects in the external environment. Recent intrinsic connectivity accounts of ATL function have suggested that the role of this region in abstract conceptual processes, such as those occur in imagination, emerges from its location as a convergence zone for information from different sensory modalities (Visser et al., 2010; Lambon Ralph et al., 2017).

Consequently, it is assumed that this region retains some degree of differential connectivity to auditory and visual 'spoke' regions. A key question, therefore, is whether transmodal portions of ATL play a common or distinct role in the representation of information about specific modalities in imagination.

Furthermore, our context condition (race + dog) permits us to investigate brain regions recruited during more complex multi modal imagery (e.g., imagining a race dog may involve the visual properties of a greyhound and race track, as well as the auditory properties of dogs panting and crowds cheering). Baron and Osheron (2011) found that conceptual combination (young + man) was represented in left ATL, indicating a role in broader conceptual thinking. However, recent studies have shown that complex mental events are more associated with activation within a broader transmodal network including medial prefrontal cortex (Hartung et al. 2015) and both unimodal and attentional brain mechanisms (Berger, 2016). Taken together this literature suggests that when imagination is more complex and potentially involves multiple modalities, additional heteromodal regions may come into play – e.g., attention and transmodal regions beyond ATL. Therefore, a pertinent research question is to what extent additional heteromodal regions are recruited during context-based multimodal mental imagery?

The present study used a combination of imaging methods to understand patterns of common and distinct neural activity that are important for different forms of mental imagery (sound vs. visual imagery; sound vs. context; etc.). First, we used MVPA to identify regions that coded for each of our three conditions (visual, auditory and context). Second, we performed conjunctions of these MVPA maps to identify distinct regions recruited during the decoding of different forms of mental imagery. Third, we interrogated the univariate activation of our conjunction maps to identify the level of bold activity in each region. Fourth, we seeded these maps in an independent resting-state session to identify the intrinsic networks that these fall within. This allowed us to identify the large-scale networks that were identified in our experiment are allied with. Finally, we performed a conjunction of these resting-state maps to identify potential common regions within the large-scale networks necessary for all forms of imagery. To complement these resting state analyses, we performed a meta-analysis of the spatial maps generated through these analyses to provide a

quantitative description of the types of cognitive processes these patterns are most frequently interpreted as.

Using this analysis pipeline, the present study examined three questions that emerge from a common and distinct account of imagination. First, we examined whether different types of sensory cortex play a specific role in imagination. For example, auditory cortex should be recruited more for thinking about what a dog sounds like than what it looks like, moreover the patterns of activity in this region should be able to decode between thinking about what a dog sounds like and other forms of imagery (e.g., visual or context conditions). Given that the majority of literature highlights the recruitment of sensory association cortex, we also predicted that secondary sensory regions would be recruited more extensively than primary sensory regions during imagery. Second, we investigated the contribution of transmodal regions, including ATL, to different forms of imagery. If these regions combine information from different modalities in a *graded* fashion, differential connectivity might allow these regions to classify imagined visual and auditory features. Finally, using resting-state fMRI, we characterized the intrinsic connectivity of regions identified in our MVPA analysis to understand the neural networks they are embedded in. We anticipated that these regions would show functional connectivity to regions of transmodal cortex important in abstract forms of cognition, as well as to relevant portions of sensory cortex (i.e. visual cortex during visual imagery). Together these different analytic approaches permit the investigation of both similarities and differences in the networks recruited when semantic retrieval is internally-generated.

4.3. Materials and Methods

4.3.1. Functional Experiment

4.3.1.1 Participants

Twenty participants were recruited from the University of York. One participant's data was excluded due to excessive motion artifacts, leaving nineteen subjects in the final analysis (11 female; mean age 23.67, range 18-37 years). Participants were native British speakers, right handed and had normal or corrected-to-normal vision.

Participants gave written informed consent to take part and were reimbursed for their time. The study was approved by the York Neuroimaging Centre Ethics Committee at the University of York.

4.3.1.2. Design

The functional experiment contained six experimental conditions, in a 2 (concepts; dog, car) x 3 (type of imagery; auditory, visual and conceptually-complex context) design (see supplementary table A.2.1 for full list of experimental conditions).

4.3.1.3. Stimuli

Experimental stimuli consisted of (i) six verbal conceptual prompts that referred to each of our six experimental conditions (e.g., Dog Sound, which cued participants to imagine what a dog sounded like), (ii) visual and auditory noise which was presented throughout experimental conditions and rest periods. For this, Gaussian visual noise was generated through Psychopy (Pyschopy, 2.7), and auditory white noise was generated through Audacity software (Audacity Version 2.0.0), and (iii) target images/sounds. The targets used in this paradigm were piloted prior to fMRI scanning, on a separate group of participants (n=24) to determine the average length of time taken to detect a target (image or sound) emerging through noise. For the visual and context trials (car visual, car context, dog visual, dog context), a pictorial target was used (e.g., a picture of a car tyre for the car visual condition). For auditory trials, a sound target was used (e.g., a dog barking). On each trial of this behavioral pilot, participants were presented with both visual and auditory noise. One of two target types were then superimposed over the visual and auditory noise: (i) image targets and (ii) sound targets. For image targets, 150 different images were presented centrally to participants. There were 30 images for each of the following experimental conditions: Dog Visual-Features (e.g., dog paw), Cars Visual-Features (e.g., car tyre), Dog Contexts (e.g., race dog) and Car Contexts (e.g., race car) and an additional 30 catch-trials (that did not represent any of the experimental conditions). Each item emerged through the noise by adjusting the opacity of the image from 0 (transparent) to 1 (opaque) in increments of 0.025 every 150ms. For sound trials, 90 different sounds were presented binaurally to participants. There were 30 sounds for each of

following sound experimental conditions: Dog Sounds (e.g., barking), Car Sounds (e.g., breaks screeching) and an additional 30 catch-trials (that did not represent the other experimental conditions). All sound trials were modified to have the same average amplitude. Each sound emerged through noise by adjusting the volume from 0 to 1 in increments of 0.10. Each sound was played in full before the volume increased (the maximum length of any of the sound trials was 600ms).

For this pilot test, participants were instructed to respond with a button-press when they could identify the image or sound emerging through the noise. Images were presented first (for all image-based conditions), followed by sound trials. The order of presentation of individual image and sound trials was randomized across participants. To ensure that participants were accurately identifying the images and sounds, on 10% of trials participants were also required to type what they had seen or heard. The average detection time across all participants was calculated for every image and sound trial. Ten images were then selected for each of our six experimental conditions (Dog Visual-Features, Car Visual-Features, Dog Sound, Car Sound, Dog Context and Car Context) based on statistically similar reaction times (RTs) for detecting the item emerging through noise. Images were detected on average at 2861ms and sounds at 2912ms (see Table 4.1). These timings were used in the fMRI experiment to ensure that the in-scan detection task would be challenging enough to engage all participants. The fMRI scan therefore allowed 3000ms for participants to detect an item emerging through noise.

4.3.1.4. Task Procedure

Prior to being scanned participants completed a practice session, identical to one scanning run. For the in-scanner task stimuli were presented in four independent runs. Within each scanning run participants were instructed to respond with a button-press (yes/no) whether a target item emerging through visual and auditory noise was related to the cue word. Trials began with a word cue to indicate which experimental conditions they should be focusing on (e.g., “Dog Sound”). Task instructions were presented for 3s. A variable number of images then followed, each displaying visual and auditory noise (see Figure 4.1). Within the blocks, the pure-noise images were each shown for 3s. Following a variable length of time (between 6 and 12s after the

initial cue), a target image or sound began to emerge through the noise (at the rate outlined in the pilot experiment described above). Participants were given 3000ms to respond to this item. The block automatically ended after this image. This design afforded us the high signal sensitivity found with block designs, combined with unpredictability to keep participants cognitively engaged.

Each experimental condition (e.g., "Dog Sound") occurred two times in a run (giving 8 blocks for each experimental condition across the experiment). Blocks were presented in a pseudo-randomized order so the same cue did not immediately repeat, and blocks were separated by 12s fixation. During the fixation period the visual and auditory noise were also presented, to create an active baseline. 50% of the items emerging through noise contained an item that did not match the preceding cue (i.e., 4 of 8 were foils) in order to focus participants on detecting the specific target. To encourage participants to search for the cued target from the very start of every block, an additional short block was included in each run, in which an item emerged through noise after only 3s, followed by 12s of fixation. These blocks were disregarded in the analysis.

4.3.1.5. Acquisition

Data were acquired using a GE 3T HD Excite MRI scanner at the York Neuroimaging Centre, University of York. A Magnex head-dedicated gradient insert coil was used in conjunction with a birdcage, radio-frequency coil tuned to 127.4MHz. A gradient-echo EPI sequence was used to collect data from 38 bottom-up axial slices aligned with the temporal lobe (TR = 2s, TE = 18 ms, FOV = 192 × 192 mm, matrix size = 64 × 64, slice thickness = 3 mm, slice-gap 1mm, flip-angle = 90°). Voxel size was 3 × 3 × 3 mm. Functional images were co-registered onto a T1-weighted anatomical image from each participant (TR = 7.8 s, TE = 3 ms, FOV = 290 mm x 290 mm, matrix size = 256 mm x 256 mm, voxel size = 1.13 mm x 1.13 mm x 1 mm) using linear registration (FLIRT, FSL). This sequence was chosen as previous studies employing this sequence have produced an adequate signal-to-noise ratio in regions prone to signal dropout, such as ATL (e.g., Coutanche & Thompson-Schill, 2014; Murphy et al. 2017).

To ensure that our ROIs had sufficient signal to detect reliable fMRI activation, the temporal signal-to-noise ratio (tSNR) for each participant was calculated by

dividing the mean signal in each voxel by the standard deviation of the residual error time series in that voxel (Friedman et al., 2006). tSNR values were averaged across the voxels in both anterior temporal lobe (ATL) and medial prefrontal cortex (mPFC); regions that suffer from signal loss and distortion due to their proximity to air-filled sinuses (Jezzard & Clare, 1999). Mean tSNR values, averaged across participants, were as follows: ATL, 82.85; mPFC, 97.14. The percentage of voxels in each ROI that had "good" tSNR values (>20; Binder et al., 2011) was above 97% for all ROIs: ATL, 97.19%; mPFC, 99.24%. These values indicate that the tSNR was sufficient to detect reliable fMRI activation in all ROIs (Binder et al., 2011).

4.3.1.6. Pre-processing

Imaging data were preprocessed using the FSL toolbox (http://www.fmrib.ox.ac.uk/fsl). Images were skull-stripped using a brain extraction tool (BET, Smith, 2002) to remove non-brain tissue from the image. The first five volumes (10s) of each scan were removed to minimize the effects of magnetic saturation, and slice-timing correction was applied. Motion correction (MCFLIRT, Jenkinson et al. 2002) was followed by temporal high-pass filtering (cutoff = 0.01 Hz). Individual participant data were first registered to their high-resolution T1-anatomical image, and then into a standard space (Montreal Neurological Institute (MNI152); this process included tri-linear interpolation of voxel sizes to 2 × 2 × 2 mm. For univariate analyses, data were additionally smoothed (Gaussian full width half maximum 6 mm).

4.3.1.7. Multivariate Pattern Analysis

Analysis was focused on the moments when participants were imagining the target cues (e.g., thinking about what a dog looked like, or what a car sounded like). The condition onset and duration were taken from the first pure noise trial in each block (after the initial cue) to the end of the last pure noise trial (before the item began to emerge through the noise). The response to each of the 6 conditions was contrasted against the active rest baseline (periods of auditory and visual noise where participants were not cued to imagine concepts). Box-car regressors for each condition, for each run, in the general linear model were convolved with a double gamma hemodynamic response function (FEAT, FSL). Regressors of no interest were also included to account

for head motion within scans. MVPA was conducted on spatially unsmoothed data to preserve local voxel information. For each voxel in the brain, we computed a linear support vector machine (LIBSVM; with fixed regularization hyper-parameter C = 1) and a 4-fold cross-validation (leave-one-run-out) classification, implemented in custom python scripts using the pyMVPA software package (Hanke et al. 2009). A support vector machine was chosen to combat over-fitting by limiting the complexity of the classifier (Lewis-Peacock & Norman, 2013). The classifier was trained on three runs and tested on the independent fourth run; the testing set was then alternated for each of four iterations. Classifiers were trained and tested on individual subject data transformed into MNI standard space. The functional data were first z-scored per voxel within each run. The searchlight analysis was implemented by extracting the z-scored β-values from spheres (6mm radius) centered on each voxel in the masks. This sized sphere included ~ 123 3mm voxels (when not restricted by the brain's boundary; Kriegeskorte et al. 2006). Classification accuracy (proportion of correctly classified trials) for each sphere was assigned to the sphere's central voxel, in order to produce accuracy maps. The resulting accuracy maps were then smoothed with a Gaussian kernel (6mm FWHM). To determine whether accuracy maps were above chance-levels (50%), individual accuracy maps were entered into a higher-level group analysis (mixed effects, FLAME; http://www.fmrib.ox.ac.uk/fsl), testing the accuracy values across subjects against chance for each voxel. Voxel inclusion was set at $z = 2.3$ with a cluster significance threshold at FWE $p < .01$.

The following classification tests were performed: (1) Car vs. Dog classifier: this examined whether patterns of activity conveyed information about conceptual identity, by training a classifier to discriminate between periods of noise where participants were thinking about a dog and periods of noise where participants were thinking about a car. We were not able to successfully classify the semantic class (dog vs. car) in our dataset at the whole-brain level. This finding is broadly consistent with previous decoding studies of internally generated thought, which have shown that specific-level concepts (e.g., lime vs. celery) can be decoded; however categorical-level concepts (e.g., fruit vs. vegetable) were not successfully classified (Coutanche & Thompson-Schill, 2014). This may reflect the dynamic nature of conceptually driven

internally-generated thought; for instance, on one trial, subjects may have been thinking about the exterior look of a car and on the next trial imagining the interior decor. As this analysis revealed no regions across the cortex could successfully decode this information, the remaining classification tests combine car and dog trials. (2) Auditory vs. visual classifier: this examined whether patterns of activity conveyed information regarding the modality of imagery, by training a classifier to discriminate between periods of noise where participants were thinking about the visual properties of objects and periods of noise where participants were thinking about the auditory properties of objects. (3) Visual vs. context classifier: here a classifier was trained to discriminate between periods of noise where participants were thinking about the visual properties of objects and periods of time when participants were thinking about objects in more complex conceptual contexts. (4) Auditory vs. context classifier: here a classifier was trained to discriminate between periods of noise where participants were thinking about the auditory properties of objects and period of time when participants were thinking about objects in complex contexts. Unthresholded maps from all analyses are uploaded on Neurovault: http://neurovault.org/collections/2671/.

Next, we identified regions where patterns of activity consistently informed the classifier for each of our three tasks (visual, auditory and context) by running a formal conjunction on the uncorrected searchlight maps (using the FSL easythresh command). For visual patterns we looked at the conjunction of the two searchlight maps that decoded visual properties (visual vs. auditory and visual vs. context). Since regions that contributed to both of these searchlight maps were able to decode simple visual features in imagination, relative to both auditory features and more complex contexts, we reasoned that their pattern of activation related to simple visual features. Next, we looked at the conjunction of the two searchlight maps that decoded the auditory condition (auditory vs. visual and auditory vs. context), to identify brain regions containing patterns of activation relating to simple auditory properties in imagination. Finally, we looked at the conjunction of the two searchlight maps that decoded context properties (context vs. visual and context vs. auditory). This identified brain regions containing activation patterns relating to complex conceptual contexts, as distinct from both simple visual and auditory features. All analyses were cluster

corrected using a z-statistic threshold of 2.3 to define contiguous clusters. Multiple comparisons were controlled using a Gaussian Random Field Theory at a threshold of $p < .01$.

4.3.1.8 Univariate Analysis

We examined univariate activation to further characterise the response within our unimodal and transmodal regions defined by MVPA. The percentage signal change was extracted for each condition from regions of interest (ROIs) defined by the MVPA conjunctions (see above).

4.3.2 Resting state fMRI

4.3.2.1. Participants

This analysis was performed on a separate cohort of 157 healthy participants at York Neuroimaging Centre (89 female; mean age 20.31, range 18–31 years). Subjects completed a 9-minute functional connectivity MRI scan during which they were asked to rest in the scanner with their eyes open. Using these data, we examined the resting-state fMRI (rs-fMRI) connectivity of our conjunction regions that were informative to decoding visual imagery, auditory imagery and contextual imagery, to investigate whether these regions fell within similar or distinct networks.

4.3.2.2. Acquisition

As with the functional experiment, a Magnex head-dedicated gradient insert coil was used in conjunction with a birdcage, radio-frequency coil tuned to 127.4 MHz. For the resting-state data, a gradient-echo EPI sequence was used to collect data from 60 axial slices with an interleaved (bottom-up) acquisition order with the following parameters: TR=3 s, TE=minimum full, volumes=180, flip angle=90°, matrix size=64×64, FOV=192×192 mm, voxel size=3x3×3 mm. A minimum full TE was selected to optimise image quality (as opposed to selecting a value less than minimum full which, for instance, would be beneficial for obtaining more slices per TR). Functional images were co-registered onto a T1-weighted anatomical image from each participant (TR=7.8 s,

TE=3 ms, FOV=290 mmx290 mm, matrix size=256 mm x256 mm, voxel size=1 mm x 1 mm x 1 mm).

4.3.2.3. Pre-processing

Data were pre-processed using the FSL toolbox (http://www.fmrib.ox.ac.uk/fsl). Prior to conducting the functional connectivity analysis, the following pre-statistics processing was applied to the resting state data; motion correction using MCFLIRT to safeguard against motion-related spurious correlations slice-timing correction using Fourier-space time-series phase-shifting; non-brain removal using BET; spatial smoothing using a Gaussian kernel of FWHM 6 mm; grand-mean intensity normalisation of the entire 4D dataset by a single multiplicative factor; high-pass temporal filtering (Gaussian-weighted least-squares straight line fitting, with sigma=100s); Gaussian low-pass temporal filtering, with sigma=2.8s.

4.3.2.4. Low-level analysis

For each conjunction site we created spherical seed ROIs, 6mm in diameter, centered on the peak conjunction voxel; visual conjunction site in left inferior lateral occipital cortex (LOC) [-48 -70 -2], auditory conjunction site in left superior temporal gyrus [-48 -12 -10] and context conjunction site in left LOC [-48 -60 0] respectively (see supplementary table A.2.2). This ensured that we assessed the functional connectivity of a key site when the searchlight conjunction revealed a large cluster or multiple clusters. The time series of these regions were extracted and used as explanatory variables in a separate subject level functional connectivity analysis for each seed. Subject specific nuisance regressors were determined using a component based noise correction (CompCor) approach (Behzadi et al. 2007). This method applies principal component analysis (PCA) to the fMRI signal from subject specific white matter and CSF ROIs. In total there were 11 nuisance regressors, five regressors from the CompCor and a further 6 nuisance regressors were identified using the motion correction MCFLIRT. These principle components were then removed from the fMRI data through linear regression. The WM and CSF covariates were generated by segmenting each individual's high-resolution structural image (using FAST in FSL; Zhang et al. 2001). The default tissue probability maps, referred to as Prior Probability Maps (PPM), were

registered to each individual's high-resolution structural image (T1 space) and the overlap between these PPM and the corresponding CSF and WM maps was identified. These maps were then thresholded (40% for the SCF and 66% for the WM), binarized and combined. The six motion parameters were calculated in the motion-correction step during pre-processing. Movement in each of the three Cartesian directions (x, y, z) and rotational movement around three axes (pitch, yaw, roll) were included for each individual.

4.3.2.5. High-level analysis

At the group-level the data were processed using FEAT version 5.98 within FSL (www.fmrib.ox.ac.uk/fsl) and the analyses were carried out using FMRIB's Local Analysis of Mixed Effects (FLAME) stage 1 with automatic outlier detection. No global signal regression was performed. The z statistic images were then thresholded using clusters determined by $z > 2.3$ and a cluster-corrected significance threshold of $p = 0.05$. Finally, to determine whether our connectivity maps overlapped with one another we calculated the number of overlapping voxels for our three conjunction site connectivity maps.

4.4. Results

4.4.1. Behavioural Results

To determine whether our experimental conditions were well matched at the behavioural level, accuracy and reaction times (RT) for the fMRI session were calculated for each participant (n=19). All participants were engaged in the correct task (e.g., thinking about the sound of a dog) as indicated by a mean accuracy score above 75% for all experimental conditions (Table 4.1). A 2 (semantic category; car, dog) by 3 (imagery type; auditory, visual, context) repeated-measures ANOVA revealed no differences in accuracy across the three types of imagery (auditory, visual, conceptually-complex context; $F(2,36) = 2.32$, $p = .11$) and no effect of concept (car, dog; $F(1,18) = 1.95$, $p = .66$). RT scores were also well matched across our experimental conditions (Table 4.1). A 2 x 3 repeated measures ANOVA revealed there was no

difference in RT between the three experimental tasks (auditory, visual, conceptually-complex context; F(2,36) = 0.46, p=.64), no effect of concept (car, dog; F(1,18) = 2.61, p=.09) and no interaction between imagery types and concept (F(2,36) = 1.17, p = .37). Furthermore, the in-scan RT data were close to the RT in our pilot study, suggesting that participants required the same amount of time to detect stimuli both in and out of the scanner (mean RT for images = 2660 ms, SD = 233 ms, mean RT for sounds = 2763 ms, SD = 616 ms).

Table 4.1. Behavioural scores across pilot and fMRI experiments

Condition	Pilot Experiment		fMRI Experiment	
	RT	Acc	RT	Acc
Car Sound	2873 (635)	N/A	2748 (713)	82.11 (16.53)
Dog Sound	2951 (876)	N/A	2753 (552)	76.84 (12.04)
Car Visual	2886 (367)	N/A	2704 (204)	83.68 (11.64)
Dog Visual	2812 (402)	N/A	2620 (241)	82.63 (9.91)
Car Context	2994 (355)	N/A	2754 (211)	76.76 (12.62)
Dog Context	2752 (398)	N/A	2569 (250)	79.61 (14.71)

Footnote: RT = reaction time in milliseconds, ACC = percentage accuracy. Standard deviation in parentheses.

4.4.2. MVPA Decoding Results

To test which brain regions held patterns of activity related to the type of internally-generated conceptual retrieval, we examined brain regions that could classify imagery conditions during the presentation of auditory and visual noise. For example, the auditory vs. visual classifier was trained on the distinction between thinking about auditory and visual properties of concepts (collapsed across both cars and dogs) and tested on the same distinction in unseen data using a cross-validated approach. All results reported are above chance levels (50%, cluster-corrected, $p < .01$).

The whole-brain searchlight analysis for the distinction between visual and auditory imagery revealed an extensive network of brain regions including sensory regions, such as bilateral inferior lateral occipital cortex (LOC), left fusiform and left

auditory cortex (encompassing planum polare and Heschl's gyrus extending more broadly into superior temporal gyrus), as well as transmodal brain regions that have been implicated in semantic processing, such as middle temporal gyrus, ATL (middle, inferior, fusiform and parahippocampal portions) and on the medial surface, anterior cingulate gyrus and thalamus (see Figure 4.2A; Table 4.2).

Table 4.2. Centre Voxel Coordinates of Highest Decoding Sphere in the Searchlight Analyses

Condition	**Cluster Peak**	**Extended Cluster Regions**	**Cluster Extent**	**Z-Score**	**Acc (%)**	**x**	**y**	**z**
Auditory vs. Visual								
	L Lateral occipital cortex, superior division	L Lateral occipital cortex, inferior division, L Occipital pole, L Occipital fusiform gyrus.	975	4.13	68.32%	-36	-86	10
	L Thalamus	R Thalamus	599	4.18	61.09%	-12	-26	2
	R Lateral occipital cortex, inferior division	R Middle temporal gyrus, temporooccipital part.	431	4.43	62.76%	54	-66	10
	L Planum polare	L Superior temporal gyrus, posterior division , Insular cortex, L Heschl's gyrus, Anterior superior temporal gyrus.	226	3.77	60.63%	-40	-16	-8
	L Supramarginal gyrus, posterior division	L Planum temporale, Posterior superior temporal gyrus.	178	3.52	61.59%	-60	-42	16
	R Frontal operculum cortex	R Frontal orbital cortex, R Insular	156	3.37	58.24%	40	22	4

		cortex.						
	L Anterior parahippocampal gyrus	L Temporal fusiform gyrus, ,	75	4.34	59.83%	-36	-18	-18
	L Anterior middle temporal gyrus	L Anterior inferior temporal gyrus,	67	4.12	61.07%	-56	-6	-18
	L Anterior cingulate gyrus		49	3.82	58.16%	-4	34	-2
Visual vs. Context								
	L Lateral occipital cortex, inferior division	L Middle temporal gyrus, temporooccipital part, L Occipital Pole.	733	4.16	68.05%	-46	-72	0
Auditory vs. Context								
	L Lateral occipital cortex, inferior division	L Temporal occipital fusiform cortex, L inferior temporal gyrus, temporooccipital part.	312	3.81	66.49%	48	-62	-6
	R Temporal occipital fusiform gyrus	R Lateral occipital cortex, inferior division, R Inferior temporal gyrus, temporooccipital part, R Middle temporal gyrus. temporooccipital part	118	3.17	60.35%	34	-56	-20
	R Posterior middle temporal gyrus	R Posterior superior temporal gyrus, R Supramarginal	90	2.92	58.76%	56	-34	-2

	gyrus, R Anterior superior temporal gyrus						
R Posterior superior temporal gyrus	R Middle temporal gyrus, R Planum polare, R Planum Temporale	81	3.15	59.09%	60	-22	0

Footnote: Highest decoding accuracy clusters for each of our three classifiers analysed separately. The Auditory vs. Visual classifier was trained on the distinction between thinking about the sound of a concept versus thinking about what a concept looked like. The Visual vs. Context classifier was trained on the distinction between thinking about what a concept looked like versus thinking about it in a specific meaning-based context. The Sound vs. Context classifier was trained on the distinction between thinking about what a concept sounded like and thinking about it in a specific meaning-based context. All analyses were cluster corrected using a z-statistic threshold of 2.3 to define contiguous clusters. Multiple comparisons were controlled using a Gaussian Random Field Theory at a threshold of $p < .01$. L = left, R = right. As well as peak accuracy (reported under the 'Cluster Peak' column), the 'Extended Cluster Regions' includes all significant regions within each ROI. The unthresholded MVPA maps for each searchlight have been uploaded to the Neurovault database and can be found here http://neurovault.org/collections/2671/.

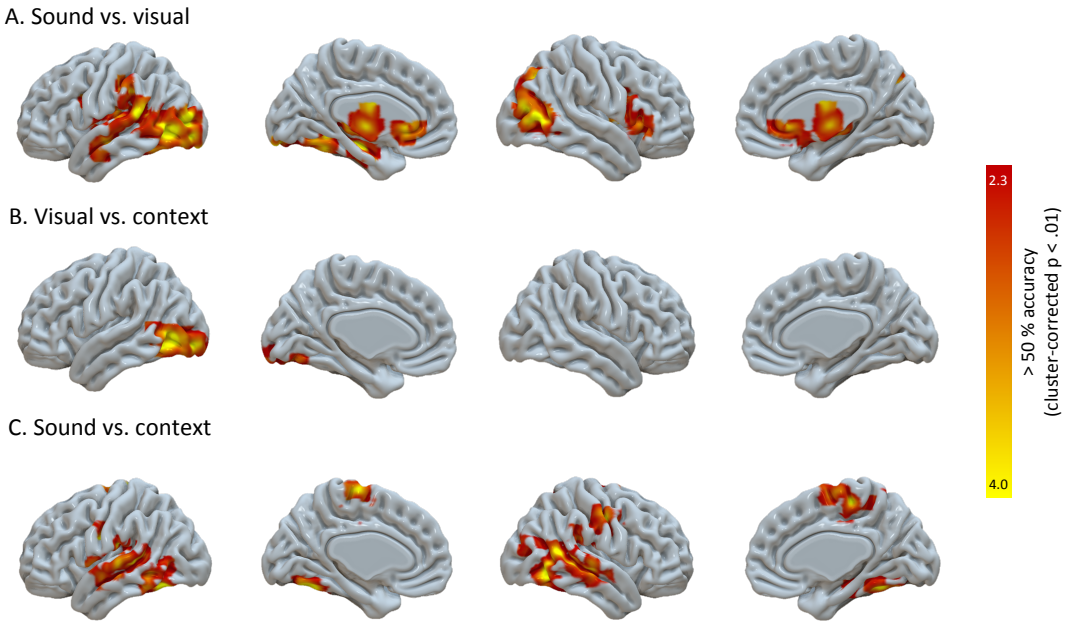

Figure 4.2. Results of the group-level whole-brain searchlight analysis with above-chance (50%) decoding projected in red ($z > 2.3$, cluster-corrected $p < .01$). All panels reveal results from binary choice searchlight analyses decoding the content of thought while participant's viewed

visual and auditory noise. (A) Location of searchlights that could decode between thinking about the sound and thinking about the visual properties of concepts. (B) Location of searchlights that could decode between thinking about the visual properties of concepts and thinking about the same concepts in more complex contexts. (C) Location of searchlights that could decode between thinking about the sound of concepts and thinking about the same concepts in more complex contexts.

Next, we examined a visual vs. context classifier, which identified regions that could classify the difference between thinking about the visual properties of concepts and thinking about the same concepts in complex conceptual contexts. This whole-brain searchlight analysis revealed a large region in the left occipital lobe that could decode between visual and context conditions at above chance levels (50%, cluster-corrected $p < .01$) (Figure 4.2B; Table 4.2). Finally, we tested whether auditory vs. context conditions could be decoded. This whole-brain searchlight analysis revealed a set of clusters in bilateral auditory cortex extending along the superior temporal gyrus (STG) into ATL and posterior occipital-temporal cortex that could decode between auditory and context conditions (50%, cluster-corrected $p < .01$) (Figure 4.2C; Table 4.2).

To identify regions that could consistently decode visual, auditory and context conditions, conjunction analyses were performed across the searchlight maps outlined in Figures 4.2A-C. The results of these conjunctions are presented in Figure 4.3A. For visual imagery, we looked at the conjunction of the two searchlight maps that involved decoding simple visual features (visual vs. auditory and visual vs. context). This revealed a left lateralized cluster in occipital pole extending into lateral occipital cortex, which reliably decoded the distinction between simple visual feature imagery and both of the other conditions. For auditory imagery, we looked at the conjunction of the two searchlight maps that involved decoding auditory properties (auditory vs. visual and auditory vs. context). This analysis revealed left hemisphere regions, including primary auditory cortex, STG, pMTG and occipital fusiform, that reliably decoded the distinction between simple auditory feature imagery and both of the other conditions. For imagery driven by complex conceptual contexts, we looked at the conjunction of the two searchlight maps that involved decoding context (visual vs.

context and auditory vs. context), which produced a cluster in left lateral occipital cortex.

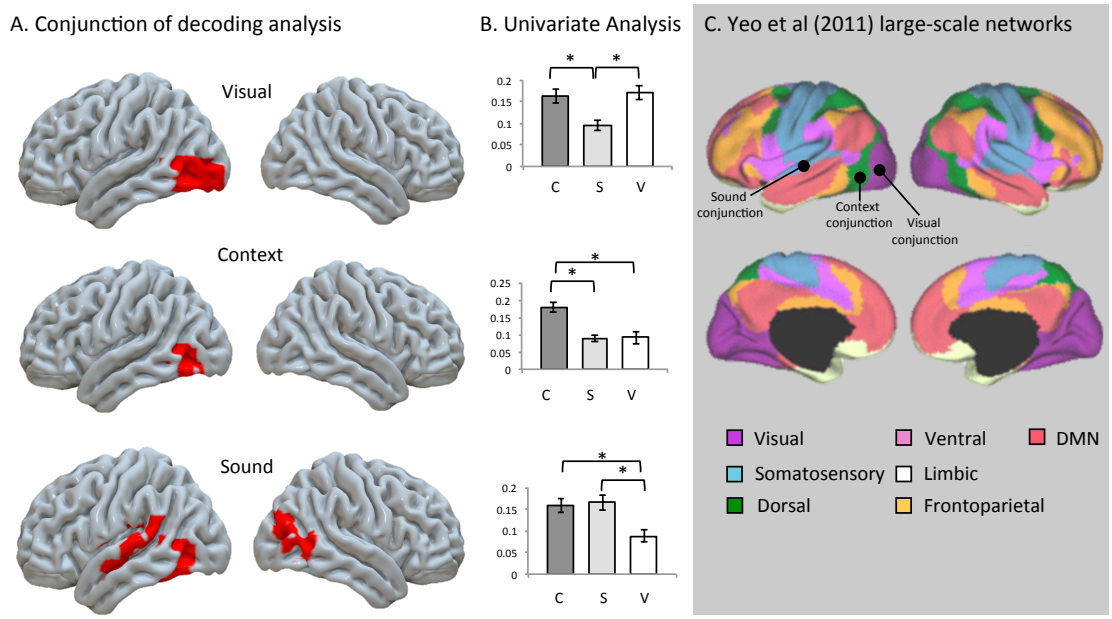

Figure 4.3. A. Represents brain regions where patterns of activity consistently informed the classifier for each of our three tasks (visual, context and sound). For visual patterns we looked at the conjunction of the two searchlight maps that decoded visual properties (sound vs. visual *and* visual vs. context). For context patterns we looked at the conjunction of the two searchlight maps that decoded context properties (visual vs. context *and* sound vs. context). For sound patterns we looked at the conjunction of the two searchlight maps that decoded sound properties (sound vs. visual *and* sound vs. context). B. Shows the univariate percent signal change for each of our three conditions taken from a 6mm sphere centered on the peak conjunction point (visual [-48 -70 2], context [-48 -60 0], sound [-52 -8 -10]). (C = context, S = sound, V = visual). * Indicates a significant different between conditions ($p < .05$). The unthresholded maps for each condition have been uploaded to the Neurovault database and can be found here http://neurovault.org/collections/2671/. C. Grey panel illustrates the 7 core intrinsic networks identified by Yeo et al (2011); Dark purple = visual network, light blue = somatosensory network, dark green = dorsal network, light pink = ventral network, white = limbin network, yellow/orange = frontoparietal network (FPN) and red = default mode network (DMN). The black circles highlight where our peak conjunction sites fall with respect to these network. Our peak visual conjunction fell within the Visual network, peak context conjunction fell within the dorsal network and peak sound conjunction site within the somatosensory network.

The conjunction of the MVPA searchlight maps revealed regions of sensory cortex that could decode different types of imagery (Figure 4.3A). As an additional complementary analysis, the percentage signal change was extracted for each condition from each of the three conjunction sites by placing a 6mm sphere around the peak (Figure 4.3B). A 3 (conjunction site; visual, sound, conceptually-complex context) by 3 (imagery type: visual, sound, conceptually-complex context) repeated-measures ANOVA revealed no significant main effect of conjunction site ($F(2,36) = 0.48$, $p = .622$) or imagery type ($F(2,36) = 2.30$, $p = .114$;, however there was a significant interaction between site and imagery type ($F(4,72) = 4.38$, $p = .003$). Planned comparisons in the form of repeated-measure t-tests revealed that our visual cluster showed significantly more activity for visual imagery than auditory imagery ($t(18) = 4.99$, $p < .001$) and for the context condition vs. auditory imagery ($t(18) = 4.61$, $p < .001$), but there was no significant difference between the visual and context conditions ($t(18) = .94$, $p = .36$). Our auditory cluster showed significantly more activity for auditory imagery than visual imagery ($t(18) = 4.64$, $p < .001$) and for the context condition vs. visual imagery ($t(18) = 5.602$, $p < .001$), but no significant difference between auditory and context imagery ($t(18) = -1.17$, $p = .25$). Finally, our context cluster revealed significantly more activity for the context condition compared to both visual ($t(18) = 5.56$, $p < .001$) and auditory imagery ($t(18) = 5.31$, $p < .001$), but no significant difference between visual and auditory imagery conditions ($t(18) = -.03$, $p = .97$).

These univariate analyses demonstrate that regions that were able to classify particular aspects of internally-driven conceptual retrieval also showed a stronger BOLD response to these conditions – i.e., greater activation to visual or auditory imagery in 'visual' and 'auditory' classifier areas, and more activation to complex conceptual contexts in areas that could reliably classify this context condition. Regions that could decode visual and auditory imagery also responded to the context condition, consistent with the view that there is a multi-sensory response to complex conceptual contexts. Moreover, the context classifier region showed a response across both visual and auditory conditions, suggesting this region is heteromodal; however, it also showed an increased response to imagery involving contexts, supporting the view

that this region responds most strongly to the unique demands of the construction process imposed by this condition. Finally, to determine which distributed networks our conjunction findings fall within, we compared our results with seven large-scale networks as defined by Yeo et al (2011) (Figure 4.3C). Both visual and sound conjunction clusters fell predominantly within unimodal sensory networks (visual and somatosensory respectively), while our context conjunction site was located within the dorsal attentional network.

Given our priori predictions regarding heteromodal cortex (e.g., ATL), we interrogated candidate heteromodal regions within the auditory vs. visual classifier map. The brain regions labelled on Figure 4.4 are the peaks representing the highest decoding accuracy taken from Table 4.2, with the exclusion of peaks in unimodal cortex (determined by the conjunction results). This analysis included a distributed network of putative transmodal regions, including supramarginal gyrus extending into pMTG, ventrolateral ATL (aMTG and aITG), thalamus, anterior parahippocampal gyrus and anterior cingulate cortex (aCC) (Figure 4.4A). As before, the percent signal change was extracted from each of these regions by placing a 6mm sphere around each peak; SMG [-60 -42 16], aMTG [-56 -6 -18], aCC [-4 34-2], thalamus [-12 26 2] and aPG [-36 -18 -18]. A 5 (location; SMG, aMTG, aCC, thalamus, aPG) by 3 (imagery type: visual, sound, conceptually-complex context) repeated-measures ANOVA revealed no significant main effect of conjunction site ($F(4,72) = 0.34$, $p = .71$) or imagery type ($F(4,72) = 2.02$, $p = .131$; nor was there a significant interaction between site and imagery type ($F(8,144) = 2.65$, $p = .102$). This equivalency across conditions is consistent with the characterization of these regions as transmodal. Finally, to quantify which intrinsic networks our clusters fall within we compared our results with seven large-scale networks as defined by Yeo et al (2011) (Figure 4.4B). The majority of clusters fell within transmodal cortices, including the default mode network and limbic system.

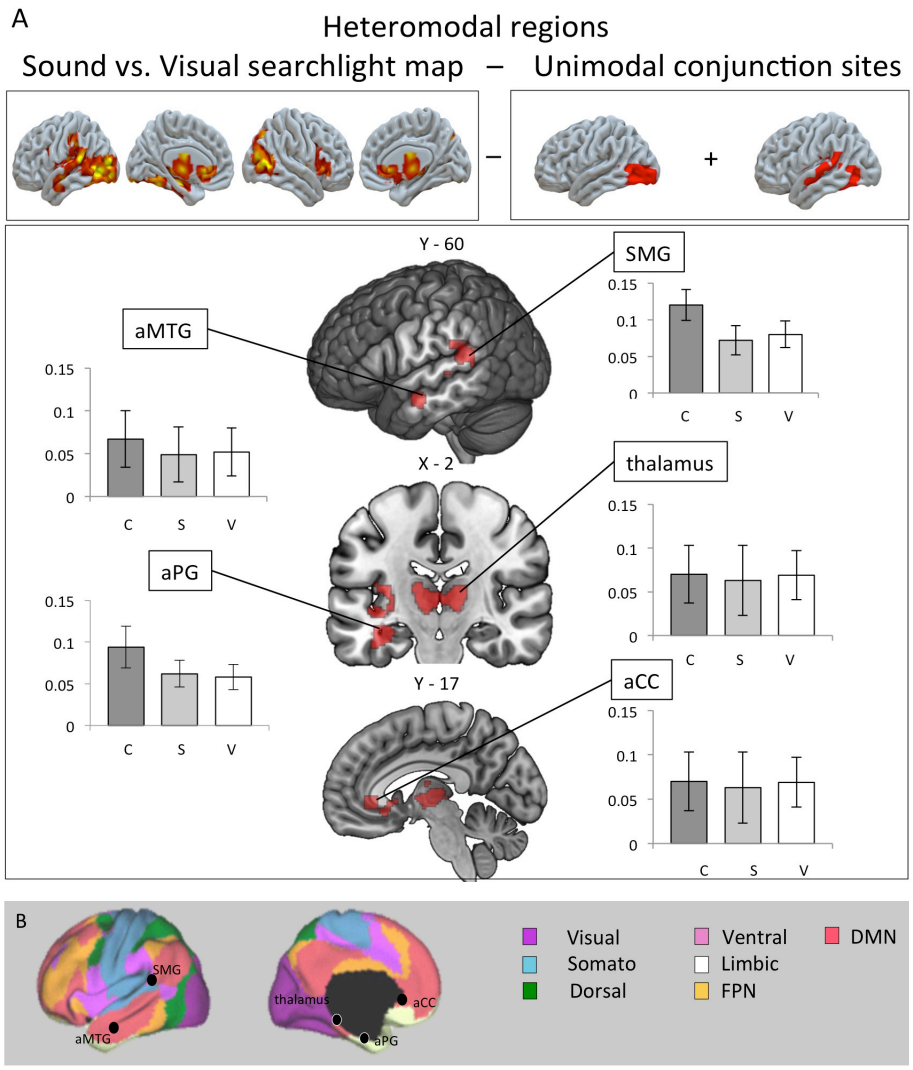

Figure 4.4. Heteromodal brain regions taken from the auditory vs. visual classifier map (Figure 4.2A). (A) Labelled regions highlight the peaks of decoding accuracy from Table 4.2 (excluding those peaks in unimodal cortex highlighted in our conjunction analysis for sound and visual imagination); SMG = supramarginal gyrus [-60 -42 16], aMTG = anterior middle temporal gyrus [-56 -6 -18], aCC = anterior cingulate cortex [-4 34-2], thalamus [-12 26 2], aPG= anterior parahippocampal gyrus [-36 -18 -18]. The bar graph shows the univariate percent signal change for each of our three conditions (C = context, S = sound, V = visual) extracted from a 6mm sphere centered on each labeled peak. There were no significant different between conditions across any of our ROIs ($p > .05$). The unthresholded map for can be found here http://neurovault.org/collections/2671/. (B) Grey panel illustrates the 7 core intrinsic networks identified by Yeo et al (2011); Dark purple = visual network, light blue = somatosensory network, dark green = dorsal network, light pink = ventral network, white = limbic network, yellow/orange = frontoparietal network (FPN) and red = default mode network

(DMN). The black circles highlight where our peak sites fall with respect to these network. SMG falls between ventral stream and somatomotor, aMTG, ACC fall within the default mode network, aPG falls within the limbic system. Subcortical regions (e.g., the thalamus) are not shown on the Yeo et al (2011) networks.

4.4.3. Intrinsic Connectivity

To provide a better understanding of the neural architecture that supported imagination in each condition, we explored the intrinsic connectivity of our unimodal conjunction sites (Figure 4.3) and transmodal sites (Figure 4.4) identified through MVPA, in resting-state fMRI. The results of the unimodal connectivity analysis are presented in Supplementary Table A.2.2 (Figure 4.5). For the visual and auditory conjunction sites, which peaked within visual and auditory cortex respectively, there was coupling beyond the sensory areas surrounding the seed regions, to include areas of transmodal cortex, including ATL, particularly the left medial surface, posterior middle temporal gyrus and precuneus. To quantify the interpretation of the functional connectivity of the visual, context and sound connectivity maps, we performed a decoding analysis using automated fMRI meta-analytic software NeuroSynth (right panel of Figure 4.5). Meta-analytic decoding of these spatial maps revealed domain specific networks and their associated function. The visual connectivity map correlated with terms related to visual processing (e.g., visual, objects), likewise our sound connectivity map correlated with terms related to auditory processing (e.g., speech, sound). The context connectivity map included both visual (e.g., objects) and higher-order terms (e.g., attention).

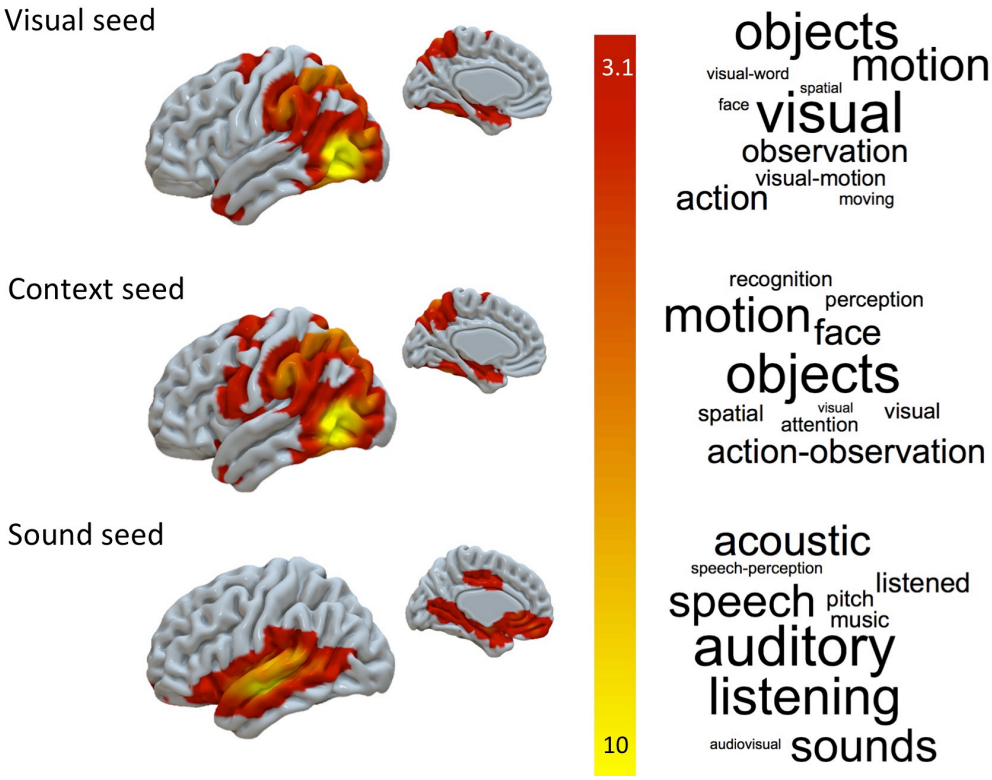

Figure 4.5. Resting state connectivity maps of unimodal regions projected on rendered brain, displaying left hemisphere and left medial view. Maps thresholded at z = 3.1, cluster corrected $p < .01$. Visual maps seeded from left inferior lateral occipital cortex [-48 70 -2]. Context maps seeded from left inferior lateral occipital cortex [-48 -60 0]. Sound maps seeded from left superior temporal gyrus [-52 -8 -10]. Word clouds represent the decoded function of each connectivity map using automated fMRI meta-analyses software (NeuroSynth, Yarkoni et al. 2011). This software computed the spatial correlation between each unthresholded zstat mask and every other meta-analytic map (n = 11406) for each term/concept stored in the database. The 10 meta-analytic maps exhibiting highest positive correlation for each sub-system was extracted, and the term corresponding to each of these meta-analyses is shown on the right. The font size reflects the size of the correlation. This allows us to quantify the most likely reverse inferences that would be drawn from these functional maps by the larger neuroimaging community.

The results of the heteromodal connectivity analysis are presented in Supplementary Table A.2.2 (Figure 4.6). Both our thalamus and SMG seed coupled extensively with sensorimotor regions and core portions of the DMN (thalamus = angular gyrus and posterior cingulate cortex; SMG = middle temporal gyrus and ATL). The three other seeds (aMTG, anterior parahippocampal gyrus and anterior cingulate cortex) all coupled with core transmodal networks (DMN and limbic system). To aid the interpretation of these connectivity maps, we performed a decoding analysis using automated fMRI meta-analytic software NeuroSynth (right panel of Figure 4.6). The thalamus connectivity map correlated with terms related to task demands and multisensory properties (e.g., anticipation, motivation, somatosensory), likewise our SMG connectivity map correlated with terms related to sensory processing (e.g., speech, sound), while in contrast aMTG, aPG and aCC connectivity maps all correlated with terms related to memory retrieval (e.g., semantic, memory, encoding, DMN).

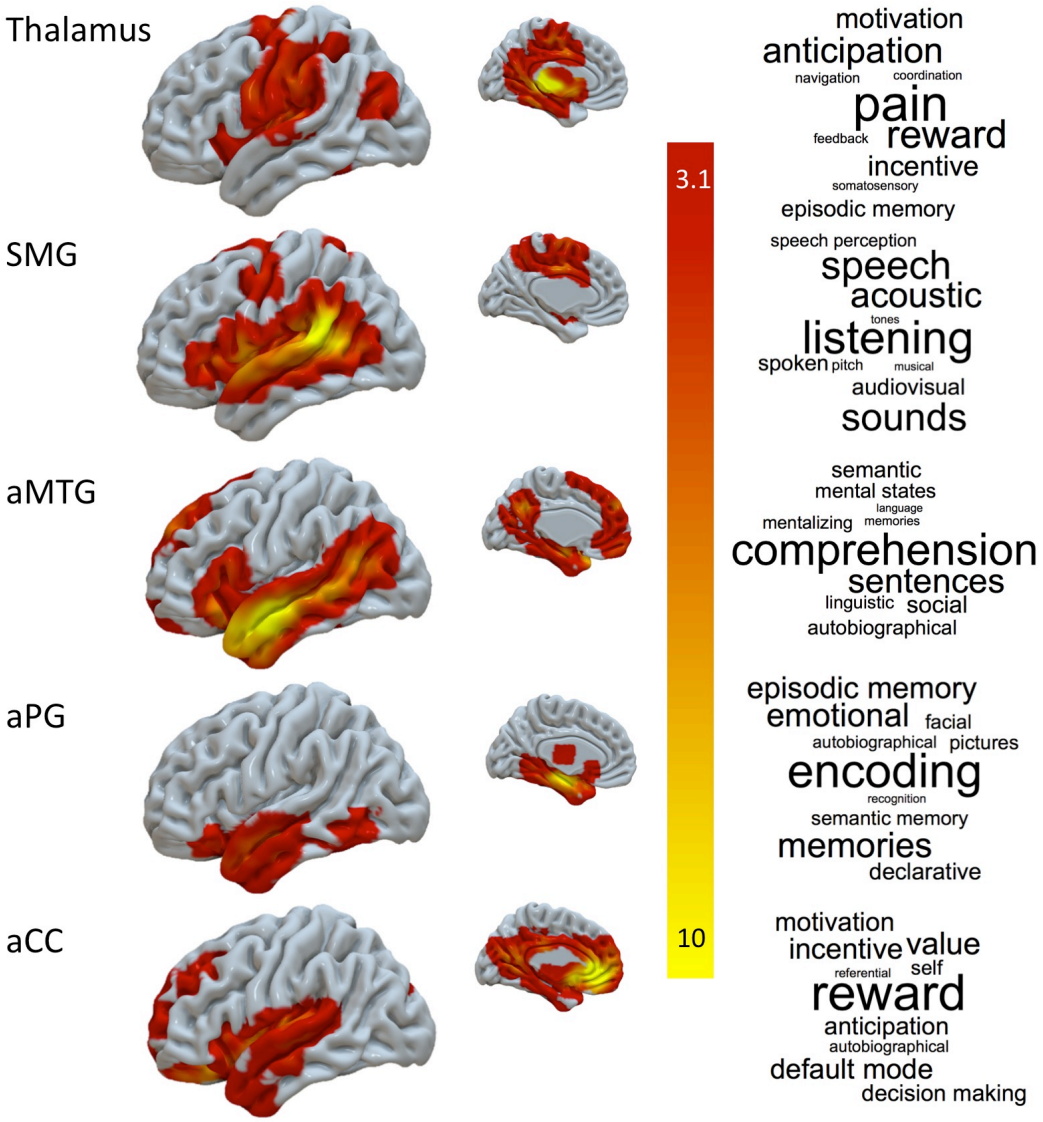

Figure 4.6. Resting state connectivity maps of heteromodal regions projected on rendered brain, displaying left hemisphere and left medial view. Maps thresholded at z = 3.1, cluster corrected p < .01. Thalamus maps seeded from [-48 -60 0]. Supramarginal gyrus (SMG) map seeded from [-48 -70 -2]. Anterior middle temporal gyrus (aMTG) seeded from [-56 -6 -18]. Anterior parahippocampal gyrus (aPG) seeded from [-36 -18 -18]. Anterior cingulate cortex (aCC) seeded from [-4 34-2]. Word clouds represent the decoded function of each connectivity map using automated fMRI meta-analyses software (NeuroSynth, Yarkoni et al. 2011). This software computed the spatial correlation between each unthresholded zstat mask and every other meta-analytic map (n = 11406) for each term/concept stored in the database. The 10 meta-analytic maps exhibiting highest positive correlation for each sub-system was extracted, and the term corresponding to each of these meta-analyses is shown on the right. The font size reflects the size of the correlation. This allows us to quantify the most likely reverse

inferences that would be drawn from these functional maps by the larger neuroimaging community.

Next, to localise whether there was a common network implicated across different forms of imagery, we performed a formal conjunction of the functional connectivity maps obtained from the visual, auditory and context MVPA conjunction peaks (see Figure 4.7A). This revealed a distributed network of clusters in bilateral parietal operculum cortex, supramarginal gyrus, posterior middle temporal gyrus, insular cortex and left anterior parahippocampal gyrus extending to hippocampus and temporal pole. Meta-analytic decoding of this conjunction map revealed strong correlations with terms related to visual processing (e.g., visual, objects) as well as higher-order cognitive functions such as attention, imagination and meaning. Finally, to explore whether our connectivity conjunction shared spatial properties with our "transmodal" MVPA connectivity maps (SMG, aMTG, anterior parahippocampal gyrus and anterior cingulate cortex), we calculated the number of overlapping voxels across the four maps. As the thalamus connectivity map deviated substantially from the connectivity profiles of the remaining four "transmodal seeds", it was omitted. This revealed a distributed network of regions encompassing lateral temporal cortex and anterior parahippocampal gyrus (see Figure 4.7B). The similarity of these two conjunction maps is evident by an overlap within anterior parahippocampal gyrus.

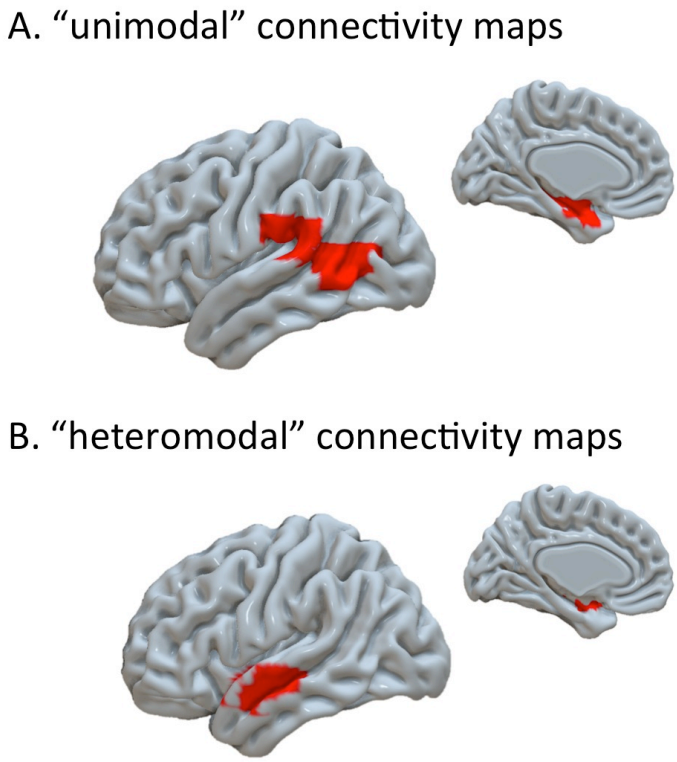

Figure 4.7. Overlap of resting-state connectivity maps. (A) The overlap between our “unimodal” connectivity maps depicted in Figure 4.5. (B) The overlap between our “heteromodal” connectivity maps depicted in Figure 4.6, with the exception of our thalamus seed as this connected to a substantial network of sensorimotor and thalamic regions. Overlap between these two maps is seen in anterior parahippocampal gyrus.

4.5. Discussion

Our study examined common and distinct components supporting conceptually-driven visual and auditory imagery. Multivariate whole-brain decoding identified aspects of secondary visual and auditory cortex (inferior lateral occipital cortex and superior temporal gyrus) in which the pattern of activation across voxels related to the modality of what was imagined. Using functional connectivity, we established that at rest these regions showed a pattern of differential connectivity with auditory or visual cortex, indicating that they reflected domain-specific aspects of imagination. We also

identified several heteromodal regions (including ventrolateral ATL, anterior parahippocampal gyrus and anterior cingulate cortex) that were also able to decode the difference between thinking about what a concept looked like and what it sounded like. Finally, a region within the dorsal attention network (inferior lateral occipital cortex) was predominantly recruited for contextually-complex imagery and reliably able to decode between all of our experimental conditions. Complementary investigation of the intrinsic connectivity of these regions confirmed their role in unimodal and heteromodal processing. These findings are consistent with the view that imagination emerges from a combined response within unimodal and transmodal regions.

The current fMRI study is one of only a few (e.g., Vetter et al., 2014) to identify patterns of activity in both visual and auditory association cortices that can reliably decode between different modalities of imagination (e.g., thinking about what a dog sounds like and what it looks like) within the same subjects. Our study is the first, to our knowledge, to investigate this issue whilst equating the visual and auditory input across our conditions. Typically neuroimaging studies of visual imagery have required participants to stare at a fixation cross while imagining an object, ensuring a consistent and simple visual input into the system (e.g., Albers et al. 2013; Dijkstra et al. 2017; Ishai et al. 2000; Lee, Kravitz & Baker, 2012; Reddy et al. 2010). In contrast, studies of auditory imagery typically require participants to imagine the sound of an object or piece of music in the presence of auditory input created by the scanner noise (e.g., Kraemer et al. 2005; Lima et al. 2015; 2016; Zattore & Halpern, 2005). In this study, we presented both visual and auditory random noise, providing more comparable visual and auditory baselines. This methodological advance allows a purer test of common and distinct neural contributions to imagination in different modalities than has been possible in prior studies.

Domain specific contributions to imagination

Our study provided evidence that neural recruitment occurs in primary sensory regions in order to support modality-specific imagery. However, the highest decoding accuracy and the location of our imagination conjunctions fell within secondary sensory regions

(superior temporal gyrus and inferior lateral occipital cortex respectively; Figure 4.3). Our functional connectivity analyses confirmed that although these regions fall outside of these systems as defined by Yeo and colleagues (2011), at rest these regions are functionally coupled to primary visual and auditory cortex respectively. These findings are in line with prior decoding and fMRI studies that have highlighted the relationship between imagery and secondary sensory regions (Albers et al. 2013; de Borst & de Gelder, 2016; Coutanche & Thompson-Schill, 2014; Chen et al. 1998; Daselaar et al. 2010; Halpern et al. 2004; Ishai et al. 2000; Lee et al. 2012; Reddy et al. 2010; Stokes et al. 2009; Vetter et al. 2014; Zvyagintsev et al. 2013). Interestingly, our results are consistent with the 'anterior shift' noted by Thompson-Schill (2003). She found that areas activated by semantic processing are not isomorphic to those used in direct experience, but rather are shifted anterior to those areas (for a wider review see Chatterjee, 2010; Binder & Desai, 2011; McNorgan et al. 2011; Meteyard et al. 2012).

Our whole-brain searchlight analysis also revealed patterns of activity supporting modality-specific imagination that extended beyond sensory cortex into semantic regions, including ATL (MTG, ventral and medial portions) and anterior cingulate cortex (see Figure 4.4). Functional connectivity analysis indicated that the majority of these regions showed extensive connectivity to other temporal lobe regions, encompassing both medial and lateral sites. Three of these regions also showed pre-frontal connectivity, primarily with connections to regions of the default mode network (anterior IFG and ventral and dorsal medial prefrontal cortex). Together this pattern of functional connectivity, suggests that these regions form a common network in the temporal lobe, and at least some of these regions are closely allied at rest with regions within the default mode network.

Domain general contributions to imagination

We found a cluster in left inferior lateral occipital cortex (LOC) that showed stronger activation in the context condition. This region was able to classify the distinction between all three conditions. Left lateral occipital cortex is traditionally thought to be a 'visual' region. However, this region predominantly falls within the dorsal attention network, as opposed to the visual network (Yeo et al. 2011). While this "task-positive"

network usually responds to demanding, externally-presented decisions (for review see Corbetta & Shulman, 2002), in this study we see engagement in a task in which imagery is being generated internally from memory. This pattern of results demonstrates that imagery not only recruits transmodal regions associated with memory but also sites implicated in attention, when the features that are being retrieved have to be shaped to suit the context, and/or when complex patterns of retrieval are required. One caveat is that our current experimental paradigm does not allow us to establish if this response in LOC is driven by the need to generate rich heteromodal content (i.e., 'dog races' can envision the sound of a crowd cheering *and* the visual properties of a race track), or the requirement to steer retrieval away from dominant features to currently-relevant information (since the fact that dogs go for walks is not pertinent to 'dog races', and might need to be suppressed to allow contextually-relevant information to come to the fore). Nevertheless, the findings do suggest that this specific region plays a greater role in supporting imagery of complex multimodal contexts as opposed to single features.

Seeding from our "heteromodal" MVPA sites highlighted extensive functional coupling with core transmodal networks including DMN and limbic systems (see Figure 4.6; Margulies et al. 2016; Mesulam, 1989; Yeo et al. 2011). Meta-analytic decomposition of these maps returned terms related to memory retrieval (e.g., semantic, memory, encoding, DMN). In addition, to transmodal networks, two of these sites (thalamus and SMG) also coupled to somatosensory and attentional networks. Thalamic influence has been previously reported during multisensory interplay (Driver & Noesselt, 2008), its role in multimodal processing may therefore explain why this region could decode between visual and auditory forms of imagination. Moreover, it has recently been suggested that SMG is crucial in the construction of mental representations (Benedek et al. 2017). As this region is connected to both attention and sensory networks, our findings converge with previous evidence suggesting that SMG integrates memory content in new ways and supports executively demanding mental simulations (Benedek et al. 2014; 2017; Fink et al. 2010). Finally, we identified a common intrinsic network across the different seed regions that extended from auditory and visual cortex into several transmodal regions, including bilateral temporal-parietal cortex and left anterior parahippocampal gyrus (see Figure 4.7).

These sites terminated in several regions identified in our “heteromodal” connectivity maps and thus provide a quantifiable explanation for how unimodal and transmodal regions communicate in order to produce dynamic retrieval of knowledge from memory.

4.6. Conclusion

In this investigation of semantic retrieval in the absence of meaningful stimuli in the external environment, we found extensive recruitment of sensory cortex, which was modulated by the modality of imagination required by the task. We also observed a role for transmodal brain regions in supporting internally-generated conceptual retrieval. These findings are consistent with the view that different types of imaginative thought depend upon patterns of common distinct neural recruitment that reflect the respective contributions of modality specific and modality invariant neural representations.

Chapter 5 - Isolated from input: Evidence of default mode network support for perceptually-decoupled and conceptually-guided cognition

This chapter is adapted from: Murphy, C., Jefferies, E., Rueschemeyer, S. A., Sormaz, M., Wang, H. T., Margulies, D., & Smallwood, J. (*in review*). Isolated from input: Transmodal cortex in the default mode network supports perceptually-decoupled and conceptually-guided cognition. bioRxiv, 150466. [4]

5.1. Abstract

The default mode network supports a variety of mental operations such as semantic processing, episodic memory retrieval, mental time travel and mind-wandering, yet the commonalities between these functions remains unclear. One possibility is that this system supports cognition that is independent of the immediate environment; alternatively or additionally, it might support higher-order conceptual representations that draw together multiple features. We tested these accounts using a novel paradigm that separately manipulated the availability of perceptual information to guide decision-making and the representational complexity of this information. Using task based imaging we established regions that respond when cognition combines both stimulus independence with multi-modal information. These included left and right angular gyri and the left middle temporal gyrus. Although these sites were within the default mode network, they showed a stronger response to demanding memory judgements than to an easier perceptual task, contrary to the view that they support

[4] The author, Charlotte Murphy, designed the experiment, analysed the results and wrote the article under the supervision of Prof. Beth Jefferies, Dr. Shirley-Ann Rueschemeyer and Dr. Jonathan Smallwood. Dr. Mladen Sormaz and Hao-Ting Wang contributed to the collection of the data, some of which occurred outside of the PhD – funded by Dr. Smallwood's ERC grant (Wandering Minds – 303701). Dr.Daniel Margulies provided the principal gradient maps.

automatic aspects of cognition. In a subsequent analysis, we showed that these regions were located at the extreme end of a macroscale gradient, which describes gradual transitions from sensorimotor to transmodal cortex. This shift in the focus of neural activity towards transmodal, default mode, regions might reflect a process of isolation from specific sensory inputs that enables conceptually rich cognitive states to be generated in the absence of input.

5.2. Introduction

Although initial studies characterized the default mode network (DMN) as "task negative", this network actively supports aspects of cognition (Spreng, 2012), including semantic processing (Binder, Desai, Graves, & Conant, 2009; Krieger-Redwood et al. 2016), episodic recollection (Rugg & Vilberg, 2013), working memory (Konishi, McLaren, Engen, & Smallwood, 2015; Spreng et al. 2014; Vatansever, Menon, Manktelow, Sahakian, & Stamatakis, 2015), autobiographical planning (Spreng, Gerlach, Turner, & Schacter, 2015; Spreng, Stevens, Chamberlain, Gilmore, & Schacter, 2010), self-generation of emotion (Engen, Kanske, & Singer, 2017) and imagining the future or the past (Schacter & Addis, 2007). Although we lack an over-arching account of the functions of the DMN, many of these situations involve memory retrieval – i.e., a requirement to focus cognition on previously-encoded knowledge, as opposed to information in the external environment. In line with this account, many regions within or allied to the DMN are considered to be heteromodal 'hubs' for memory-related processes, including the posterior cingulate cortex (Leech, Braga, & Sharp, 2012; Leech & Sharp, 2014), angular gyrus (Bonnici, Richter, Yazar, & Simons, 2016; Seghier, 2013), hippocampus (Moscovitch, Cabeza, Winocur, & Nadel, 2016) and anterior temporal lobes (Visser, Jefferies, & Lambon Ralph, 2010). In addition, cognitive states that activate the DMN tend to involve meaningful content that has personal relevance (Gusnard, Akbudak, Shulman, & Raichle, 2001).

The current study was motivated by the hypothesis that there might be common neurocognitive processes underpinning perceptually-decoupled and conceptually-guided cognition in the DMN. During states of episodic recollection, we

recreate past experiences that involve places, objects and people not currently present in the environment. Consequently, memory retrieval might necessitate a process of decoupling from sensorimotor systems, allowing cognition to be generated internally in a way that diverges from what is going on around us (Smallwood, 2013). These perceptually-decoupled states might be easier in brain regions whose neural computations are functionally independent, or distant, from systems important for perceiving and acting. This is consistent with the observation that the distributed regions of the DMN are maximally distant from primary visual and motor cortex, both in terms of their distinct patterns of functional connectivity and their positions along the cortical surface (Margulies et al. 2016).

In addition, DMN regions might support higher-order representations with predictive value across multiple situations and modalities, which integrate features from diverse sensorimotor regions. Contemporary accounts of semantic representation (Lambon Ralph, Jefferies, Patterson, & Rogers, 2017) envisage an interaction between unimodal brain regions that support knowledge about specific features (e.g., knowledge that BANANAS are YELLOW and CURVED in visual cortex) and heteromodal regions within or allied to the DMN, which extract deeper similarity structures across these domains (i.e., allow us to understand that BANANA and KIWI are conceptually related, despite salient differences in colour, shape etc.). This view is also consistent with the observation that DMN lies at the extreme end of a gradient from heteromodal to unimodal cortex (Margulies et al. 2016), since increasingly abstract and complex representations might be formed at greater distances along the gradient, where the influence of specific features linked to stimuli in the immediate environment is reduced (Buckner & Krienen, 2013; Margulies et al. 2016; Mesulam, 1998). Within the DMN, angular gyrus (Binder & Desai, 2011; Bonner et al. 2013) and anterior temporal cortex (Lambon Ralph et al. 2017; Patterson et al. 2007) are both implicated in heteromodal semantic processing. However, their roles remain controversial since other regions such as left inferior frontal gyrus and posterior aspects of the temporal lobe frequently show stronger task-induced activation in fMRI. Angular gyrus, in particular, typically shows a pattern of task-induced deactivation, which is greater for harder judgements in both semantic and non-semantic tasks (Humphreys et al. 2015; Humphreys & Lambon Ralph, 2015). In addition, despite

commonalities in the intrinsic connectivity of these regions, differences in semantic content have been proposed although not broadly accepted (Jackson, Hoffman, Pobric & Lambon Ralph, 2016): the anterior temporal lobes might support object identification, while angular gyrus is potentially more sensitive to thematic associations (Davey et al. 2015; Schwartz et al. 2011).

We developed a novel paradigm to identify brain regions important for stimulus independence, more complex memory representations and the combination of both features in cognition. Our experiment builds on prior work by Konishi and colleagues (Konishi et al. 2015). In their study, participants kept track of the location of pairs of simple shapes (triangles, squares and circles) presented either side of fixation. When probed with one shape from a prior trial and asked which side of the screen it was presented on, activity increased in regions including those within the DMN. The current study extended this paradigm by varying the complexity of the information to be encoded and retrieved. In one condition participants keep track of the location of pairs of stimuli that vary on a single feature (coloured patches), in a second they keep track of stimuli that vary in a more complex manner (pairs of familiar real world objects such as dogs or cars). Objects place greater demands on memory than do colours because they are distinguished based on a greater number of features. This allowed us to contrast higher and lower levels of representational complexity in the perceptual representations and memories that would be probed. We also manipulated whether these decisions were made when the relevant information is on the screen (0 – back) or when only the identity of the target upon which the decision is based is present (1 – back). In the latter case the relevant spatial information must be retrieved from memory, a manipulation that allowed us to explore the property of stimulus independence in cognition. This paradigm is presented schematically in Figure 5.1.

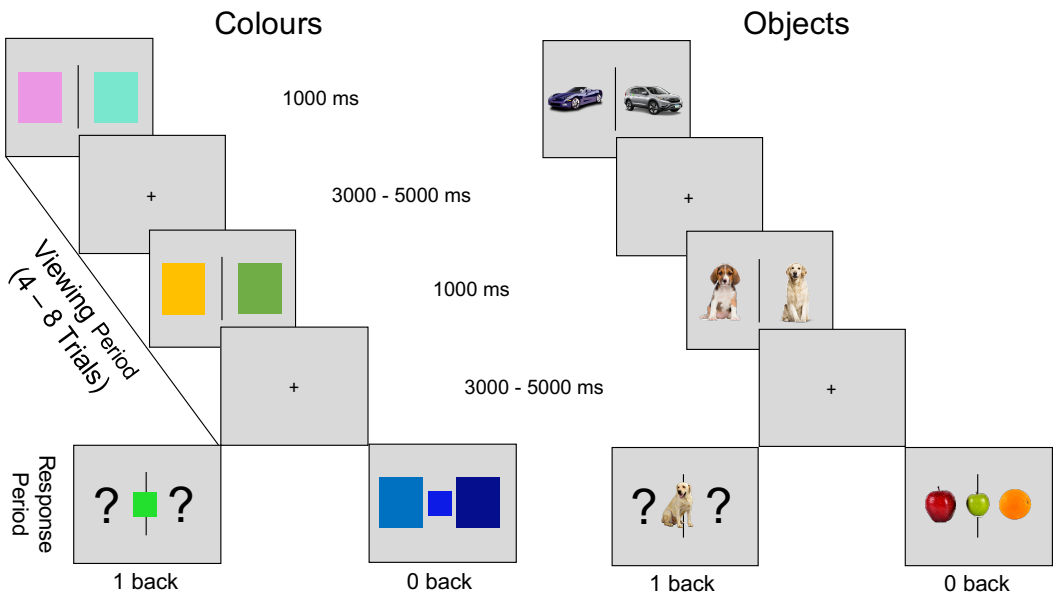

Figure 5.1. Experimental design. The four different judgments that participants made in this experiment.

We used this paradigm in a functional magnetic resonance imaging (fMRI) experiment to localize the brain regions that support the properties of stimulus independence and representational complexity. Our aim was to establish whether regions sensitive to perceptual decoupling and conceptual retrieval fall within the DMN, and whether these effects were located in overlapping or distinct regions. We identified regions of cortex that respond: (i) to stimuli with a rich multi-modal structure by comparing the response of objects to colours, (ii) when decision making has a higher reliance on memory by comparing decisions that are made in the 1-back condition with those made in the 0-back conditions, and (iii) to conditions that require a combination of both elements of cognition. We tested these hypotheses using both standard whole brain univariate analyses, as well as characterizing neural activity in each condition in terms of its position on the macroscale gradient from unimodal to heteromodal cortex described by Margulies et al. (2016).

If regions in the DMN are activated when spatial decisions are guided by information from memory, this would support a role in stimulus independent decision making. Alternatively, if DMN regions respond to decisions made regarding objects rather than colours, this would reflect a role in the processing of information with a high degree of representational complexity. Finally, if DMN regions show the strongest

response when spatial decisions are made based on objects from memory, these regions would support more complex stimulus independent representational states. This latter pattern would be consistent with the hypothesis that decoupling from perceptual input enables cognition to represent information that diverges from what is going on around us (Smallwood, 2013). Moreover, if this processing emerges in regions located towards the transmodal end of the principle gradient, this would support the hypothesis that the "distance" from systems important for perceiving and acting, provides a cortical mechanism that underpins the processing of complex representations derived from memory (Lambon Ralph et al. 2016; Margulies et al. 2016).

5.3. Material and Methods

5.3.1. Participants

Thirty right-handed native British-speaking participants with normal or corrected-to-normal vision were recruited from the University of York (16 female; mean age 22.68, range 18-34 years). One participant's data was excluded due to excessive motion artefacts, leaving twenty-nine subjects in the final analysis for (15 female; mean age 22.57, range 18-24 years). In a subsequent analysis we used a set of 60 participants' resting state data, including the same individuals who performed the task (34 female; mean age 20.32, range 18-29 years). Both studies were approved by the York Neuroimaging Centre (YNIC) Ethics Committee. Participant's provided informed consent prior to the start of the experimental session.

5.3.2. Stimuli

The task paradigm had four conditions: (A) Object 0-back, (B) Object 1-back, (C) Colour 0-back and (D) Colour 1-back using a block design. In all conditions, pairs of items were presented separated by a central line. In the colour conditions, these were different coloured squares, while in the object conditions, these were familiar and meaningful objects, taken from the same semantic category (i.e., different types of cars, fruit, dogs; see Figure 5.1). Items were presented once with no repetition. The contrast

between object and colour conditions allowed us to investigate regions that are important for the retrieval of conceptual information. The colour patches only varied on one feature (their colour), while the objects were meaningful multi-featural concepts. In addition, the contrast of 0-back and 1-back conditions allowed us to investigate the effect of stimulus-independent processing (1 back > 0 back).

5.3.3. Procedure

In the scanner, participants completed a total of four functional runs (average run time 8 min 32 s). Within each run, there were two blocks related to each of the 4 conditions (Object 1-back; Object 0-back; Colour 1-back; Colour 0-back). Each block began with written instructions stating the task type (0-back or 1-back). Blocks consisted of observing pairs of items (1000 ms); each pair was separated by a jittered inter-stimulus interval (ISI; 3000-5000 ms) in which a fixation cross was presented. At random intervals (4-8 trials), a third item was presented in the centre of the screen and participants were asked to indicate the location of one of the pair (left or right) that was most similar to this probe (see Figure 5.1). This paradigm also required participants to match items that were present and compared this with items in memory. In the 0-back catch-trials participants had to decide which stimulus (left or right of the screen) was most similar to this centrally-presented probe (i.e., all items were present on the screen). In the 1-back catch-trials, participants had to decide which stimulus (left or right of the screen) had been most similar to this centrally-presented probe on the previous trial (i.e., the critical stimulus was absent). Blocks consisted of 5 probes in total and lasted on average 64 s.

5.3.4. MRI Acquisition

Data for both experiments were acquired using a GE 3 T HD Excite MRI scanner at the YNIC. A Magnex head-dedicated gradient insert coil was used in conjunction with a birdcage, radio-frequency insert coil tuned to 127.4 MHz. A gradient-echo EPI sequence was used to collect data from 38 bottom-up axial slices aligned with the temporal lobe (TR = 2s, TE = 18ms, FOV = 192x192mm, matrix size = 64x64, slice thickness = 3mm, slice-gap = 1mm, flip-angle = 90°). Voxel size was 3x3x3mm.

Functional images were co-registered onto a T1-weighted anatomical image from each participant (TR = 7.8s, TE = 3ms, FOV = 290x290mm, matrix size = 256x256mm, voxel size = 1.13x1.13x1mm) using linear registration.

5.3.5. Pre-processing

All imaging data were pre-processed using a standard pipeline and analysed via FMRIB Software Library (FSL Version 6.0). Images were skull-stripped using a brain extraction tool [BET, (Smith, 2002)]. The first five volumes (10s) of each scan were removed to minimize the effects of magnetic saturation, and slice-timing correction with Fourier space time-series phase-shifting was applied. Motion correction (MCFLIRT, (Jenkinson, Bannister, Brady, & Smith, 2002)) was followed by temporal high-pass filtering (cut-off = 0.01Hz). Individual participant data was registered to their high-resolution T1-anatomical image, and then into a standard spare (Montreal Neurological Institute); this process included tri-linear interpolation of voxel sizes to 2x2x2 mm.

The resting state functional data used were pre-processed and analysed using the FMRI Expert Analysis Tool (FEAT). The individual subject analysis involved: motion correction using MCFLIRT; slice-timing correction using Fourier space time-series phase-shifting; spatial smoothing using a Gaussian kernel of FWHM 6mm; grand-mean intensity normalisation of the entire 4D dataset by a single multiplicative factor; high-pass temporal filtering (Gaussian-weighted least-squares straight line fitting, with sigma = 100 s); Gaussian low-pass temporal filtering, with sigma = 2.8s

5.3.6. Task-based fMRI

For our task-based analysis, the time points of interest were the probe trials where participants had to make a decision about something present (0-back) or absent (1-back) from the screen. We therefore used a box-car regressor to model the probe trial for each condition and another one to model the entire block. Modelling the entire block ensured any effect detected from our analysis can be attributed to the probe itself and not the general effect of the block. Box-car regressors for each probe/block, for each condition, for each run, were convolved with a double gamma hemodynamic response function. Regressors of no interest were included to account for head

motion. We computed four contrasts: (1) 0-back > 1-back, (2) 1-back > 0-back, (3) Object > Colour and (4) Colour > Object. A fixed effect design (FLAME, http://www.fmrib.ox.ac.uk/fsl) was conducted to average the four runs, within each individual. Individual participant data were then entered into a higher-level group analysis using a mixed effects design (FLAME, http://www.fmrib.ox.ac.uk/fsl) whole-brain analysis. Finally, our analysis focused on a conjunction of 1-back > 0-back and Object > Colour to identify regions engaged in both stimulus independent processing and conceptually abstract representations.

5.3.7. Resting-state fMRI

We extracted the time series from regions identified by univariate analysis and used these as explanatory variables in a connectivity analyses at the single subject level. In each analysis, we entered 11 nuisance regressors; the top five principal components extracted from white matter (WM) and cerebrospinal fluid (CSF) masks based on the CompCor method (Behzadi, Restom, Liau, & Liu, 2007), six head motion parameters and spatial smoothing (Gaussian) was applied at 6mm (FWHM). WM and CSF masks were generated from each individual's structural image (Zhang, Brady, & Smith, 2001). No global signal regression was performed, following the method implemented in Murphy, Birn, Handwerker, Jones, & Bandettini (2009).

All whole brain analyses were cluster corrected using a z-statistic threshold of 3.1 to define contiguous clusters. Multiple comparisons were controlled using Gaussian Random Field Theory at a threshold of $p < .05$ [34]. All statistical maps produced in these analyses are freely available at Neurosynth at the following URL: http://neurovault.org/collections/2296/.

5.4. Results

Table 5.1 presents behavioural performance, in the form of response efficiency (RT/ACC), for each of the four conditions of our task. Response efficiency controls for speed-accuracy trade-offs. These data were compared using a 2 (task; 0-back vs. 1-back) by 2 (condition; object vs. colour) repeated-measures analysis of variance (ANOVA). There was no significant differences between stimulus type ($F(1,28) = 2.55$, p

= .116) but a significant main effect of task ($F(1,28) = 15.38$, $p < .001$). There was no significant interaction ($p > .05$). These analyses demonstrate that performance on the 1-back task was less efficient than for the 0-back task but that object and colour conditions were well matched in terms of overall task difficulty.

Table 5.1. Behavioural results.

Condition	**Response Efficiency**	
	Mean	**SE**
Colour 1-back	1028	206
Colour 0-back	829	287
Object 1-back	1041	229
Object 0-back	841	295

Footnote: SE = standard error. Response efficiency = reaction time in milliseconds / percent accuracy.

We next generated statistical maps describing patterns of neural activity at the moments when participants responded in each of our four conditions. These maps were compared at the group level using a GLM. The contrast of 0-back > 1-back decisions captures perceptually-guided decision-making, revealing increased activity in the bilateral ventral visual stream, from occipital pole through to posterior fusiform cortex (presented in cool colours in the upper panel of Figure 5.2). These regions have a well-documented role in online visual processing. The contrast of 1-back > 0-back reflects stimulus independence in decision-making. This comparison revealed greater activation in bilateral angular gyrus and anterior temporal lobes, as well as medial structures in the posterior cingulate cortex and medial prefrontal cortex (these are presented in warm colours in the middle panel of Figure 5.2). Many of these regions fall within the DMN (58.44% of voxels fell within the DMN as defined by Yeo et al. 2011) and are spatially similar to the 'general recollection network' proposed by Rugg and Vilburg (2013). The comparison of Objects > Colours identifies brain areas that support the processing of multi-featural conceptual representations. These are

presented in warm colours in the lower panel in Figure 5.2. This contrast revealed a similar set of regions to the stimulus independence contrast (medial pre frontal cortex, left and right angular gyrus and anterior temporal lobe) with the addition of the right dorsolateral cortex (52.49 % of voxels fell within the DMN as defined by Yeo et al. 2011). The contrast of Colours > Objects yielded no significant whole-brain corrected results. To allow comparison with previous research, the spatial maps for the contrast of 1-back > 0-back from Konishi and colleagues are also displayed: similarities can be seen in posterior cingulate cortex, medial prefrontal cortex, right angular gyrus and dorsolateral cortex.

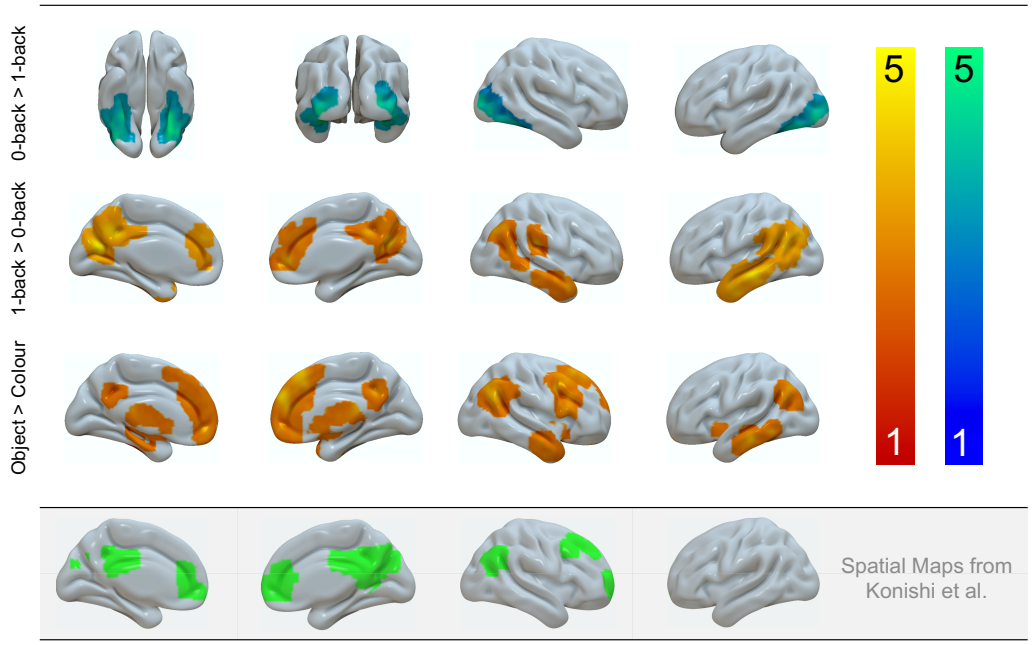

Figure 5.2. Neural activity produced when making decisions based on meaningful objects and when decisions are made from memory. (a) Activity elicited when decisions were made based using information from perception (b) Activity when decisions were made on the basis of information from memory and (c) when information from memory was more complex. Spatial maps were cluster corrected at Z = 3.1 FWE.

Our next analysis formally identifies regions that show a response to both stimulus independence and memory complexity. Figure 3 shows the results of a formal

conjunction of the contrasts of Object > Colour and 1-back > 0-back, revealing three regions – bilateral angular gyrus and lateral medial temporal gyrus in the left hemisphere. The left hand panel of Figure 5.3 summarizes the parameter estimates from each of these regions in each condition of our task. In every case the strongest response was when decisions were made in the Object 1-back condition. Importantly, although these regions fell within the DMN (88.07% of voxels within the conjunction mask fell within the DMN as defined by Yeo et al. 2011), their response profile indicated greater responding during a demanding condition (i.e. Object 1-back) ruling out a task-negative interpretation of these results.

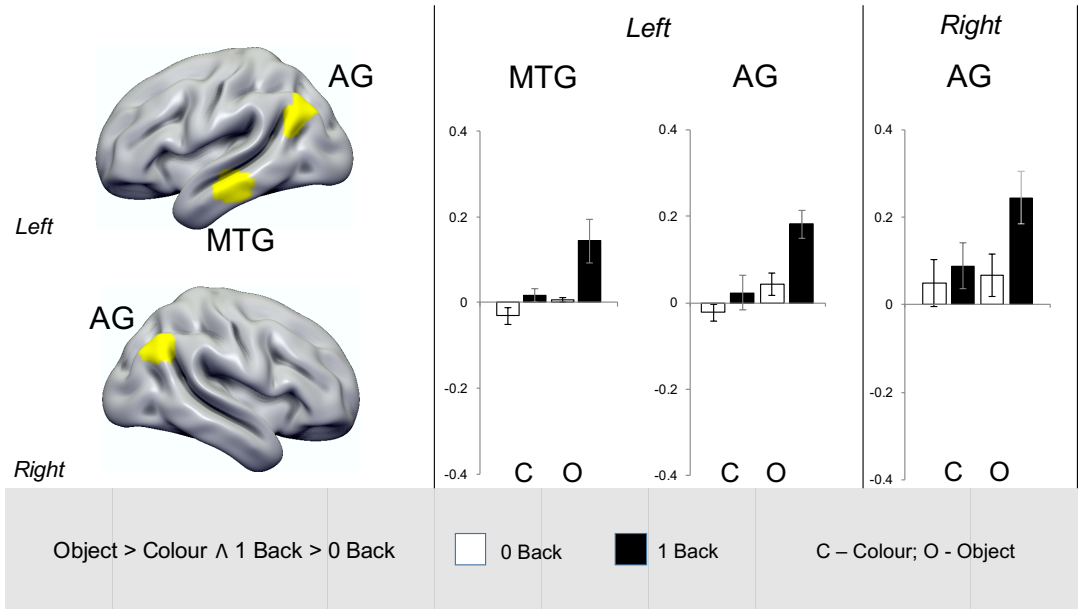

Figure 5.3. Locating peak activity during stimulus independent decisions regarding complex objects. (a) A conjunction of the neural activity when making decisions based on meaningful categories and when decisions are made in the absence of perceptual input revealed three regions: bilateral angular gyrus and in the left middle temporal gyrus. (b) Percent signal extracted from these regions confirmed an additive effect (i.e. these regions responded significantly more to the object condition when information was not present on the screen compared to all other conditions). The conjunction analysis was based on whole-brain cluster corrected spatial maps from Figure 5.2. Error bars indicated 95% confidence intervals.

We also explored the intrinsic architecture of conjunction regions responding to Object > Colour and 1-back > 0-back in an independent resting-state data set. The results of this analysis are presented in Figure 5.4 and reveal coupling beyond the seed regions to the posterior cingulate cortex, dorsolateral prefrontal cortex and pre-supplementary motor area bilaterally. Some of these regions fall outside the DMN, as defined by Yeo and colleagues, and instead are members of the frontoparietal network linked with cognitive control (Yeo et al. 2011).

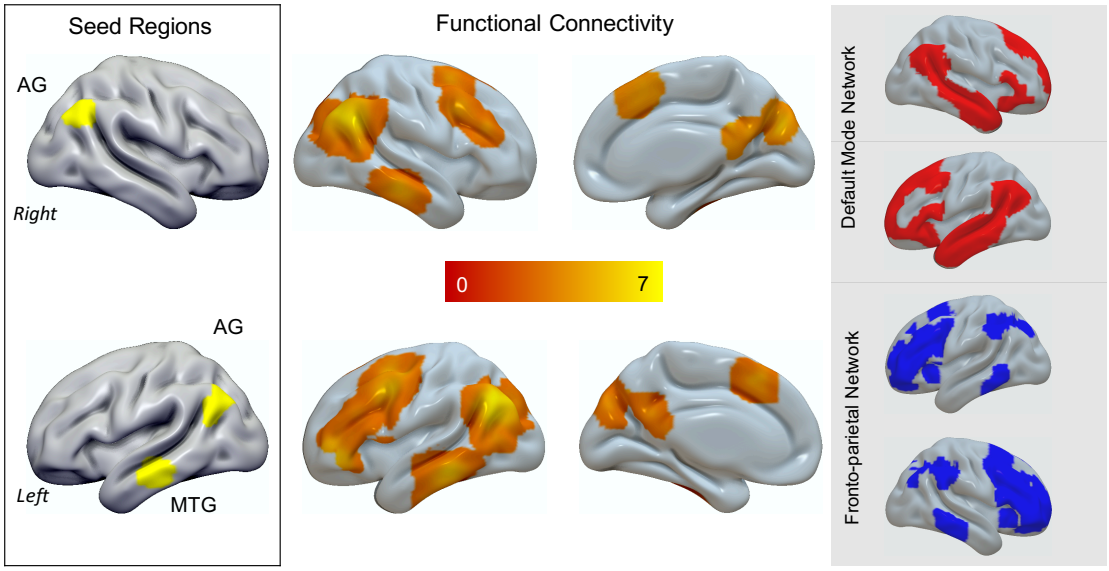

Figure 5.4. Peak areas during stimulus independent decision regarding complex stimuli involve both regions of the default mode network (DMN) and the fronto-parietal network (FPN). These regions show functional connectivity at rest with both the pre-supplementary cortex and the dorsolateral pre-frontal cortex. Although the regions identified in our conjunction analysis fall within the DMN they show functional communication with regions in the FPN, including the right dorsolateral prefrontal cortex. The spatial networks in the grey panel are from the decomposition of Yeo and colleagues. The conjunction analysis was based on whole-brain cluster corrected spatial maps from Figure 5.1. For the connectivity analyses spatial maps were cluster corrected at Z = 3.1 FWE.

We also conducted a supplementary analysis contrasting Object and Colour decisions separately in the 1-back and 0-back conditions to confirm regions important

for stimulus-independent decisions (see Supplementary Figure A.1.1). This analysis showed that 1-back trials involving meaningful objects activated regions including angular gyrus, middle temporal gyrus and right dorsolateral prefrontal regions more than colours. In contrast, the comparison of Objects > Colours in the 0-back condition only revealed greater activity in fusiform cortex.

Together these analyses highlight a network of regions that are important when spatial decisions are made in the absence of external sensory support, and when they involve multi-feature concepts (Figure 5.5). Common regions responding to the two task contrasts (1-back > 0-back; Object > Colour) include angular gyrus and middle temporal gyrus. In the right hemisphere, two of the three maps also include the right dorsolateral cortex. All of these right hemisphere regions responded to a similar 1-back > 0-back contrast involving abstract shapes (circle, triangle, square) in the study by Konishi and colleagues (2015).

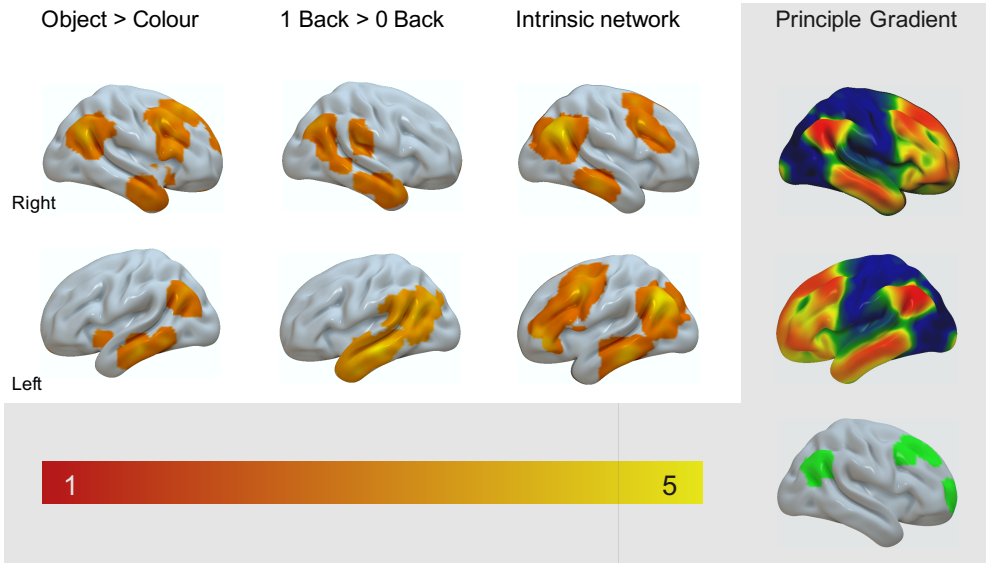

Figure 5.5. Regions linked to during stimulus independent decisions regarding complex stimuli form localized clusters in transmodal cortex. Regions in the main panel reflect data generated in this experiment. The grey sub panel presents the spatial distribution of the principle gradient from Margulies and colleagues (2016) coloured blue-red, and the cluster corrected map from Konishi and colleagues (2015) coloured green.

In Figure 5.5, we summarise the spatial maps produced by this experiment and present these alongside the principal gradient from Margulies and colleagues (2016), which describes a functional spectrum of intrinsic connectivity across the cortical surface, extending from primary sensorimotor systems to regions of the DMN at the other extreme. More similar colours on this gradient reflect greater similarity in connectivity. It can be seen that common regions implicated in stimulus-independent and conceptual processing are all localized towards the transmodal end of the principal gradient.

Our final analysis characterizes the similarity between the neural patterns captured by our task and the spatial distribution of the principle gradient from Margulies et al. (2016) in a more formal manner. Following Margulies et al. we divided the principle gradient into 20 equally sized bins. Next for each participant we calculated the average signal in each bin for each condition of our task. The left hand panel in Figure 5.6 presents these data plotted across the principle gradient separately for each condition; the shaded bars represent the 95% confidence intervals. It can be see that the conditions are most distinct towards the transmodal end, with the highest values when participants made judgments about objects from memory. To quantify these patterns, we compared their distribution using a 2 (stimulus independence) X (stimulus complexity) X 20 (Gradient Bin) ANOVA. This revealed a significant 3-way interaction [$F(19, 532) = 5.136$, $p < .001$]. To follow up this interaction, we performed a principle components analysis (PCA) on the condition level data, describing the dynamics captured in the left hand panel of Figure 5.6. The results revealed two components with eigenvalues greater than 1 accounting for over 86% of the variance (component 1 = 70.49%; component 2 = 15.73%) across the principal gradient bins. The first two components are presented in the right hand panel of Figure 5.6. The second component describes a gradual transition showing increasing levels of BOLD activity from the unimodal end of the gradient towards the transmodal end. Projecting the values from component 2 back onto the task conditions, and averaging them at the group-level, revealed that this pattern of variance loaded almost exclusively on the 'object' 1-back condition. There was a significant positive fit between the spatial map

of the principle gradient and recruitment in the Object 1-Back task, but not other conditions.

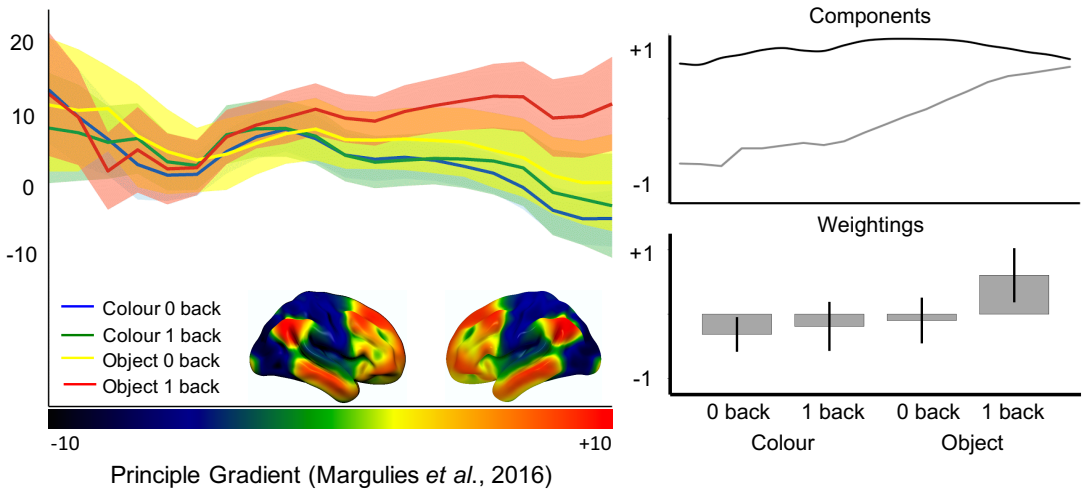

Figure 5.6. Stimulus independent decisions regarding meaningful objects leads to a whole-brain shift towards the transmodal-end of the gradient. (a) A regions-of-interest analysis using bins of the principal gradient revealed that the decisions that are made on objects rather than colours when these stimuli are not available to perception led to higher activity towards the transmodal-end of the principal gradient. (b) Decomposition using PCA revealed that this difference was related to a gradual shift in the locus of neural activity away from regions on the principal gradient related to perception and action and towards transmodal regions of cortex. Error bars indicated 95% confidence interval.

5.5. Discussion

Our experiment contrasted judgments when stimuli to be decided upon were present in the current trial (stimulus-dependent 0-back decisions) with identical decisions where the information was from the previous trial (stimulus-independent 1-back decisions, see Figure 5.1). We also varied whether decisions were made on uni-dimensional stimuli (colour patches) or more complex multi-dimensional stimuli (objects in conceptual categories). In each task, the items to be matched were not perceptually identical: participants selected the closest hue or the closest concept from similar distractors. This allowed us to identify regions that capture cognitive

processes important for representing (i) information that is decoupled from stimulus input, (ii) representations that are complex and multi-dimensional in nature and (iii) a combination of both.

Using conventional whole brain analyses we identified overlapping regions in the DMN that are sensitive to both perceptual decoupling (i.e., the requirement to make decisions based on memory, as opposed to the immediate environment) and when these decisions regarded more complex conceptual categories of stimuli (i.e., decisions based on objects rather than colours). We also used a novel analytic approach to demonstrate that these isolated clusters of activity can be seen as part of a whole brain shift in the locus of neural activity towards the extreme end of a gradient from unimodal to heteromodal cortex (Margulies et al. 2016). These findings have broad implications for the role of DMN in cognition, and also contribute to our understanding of specific DMN regions, particularly angular gyrus and lateral temporal lobe. We first consider the results in terms of their implications for functional accounts of these regions. Secondly, we consider the macroscale organisation of the cortex, focusing on approaches which can explain the functional similarity of these distributed clusters and their relative position on the cortical surface.

Functional implications for the angular gyri: There were stronger responses in left and right angular gyri, as well as in left middle temporal gyrus, when conceptual decisions were based on information that was no longer present in the environment. These findings are inconsistent with several existing accounts of the contribution of angular gyrus to memory and semantic cognition. First, they do not easily align with the proposal that specific aspects of meaning are represented in angular gyrus – namely thematic associations, but not item identity (Davey et al. 2015; Schwartz et al. 2011). Our conceptual task involved matching items on the basis of their identity, yet it robustly activated this region. Secondly, the findings are at odds with the proposal that the angular gyri only activate during contrasts of easier versus harder tasks, and for "automatic" and not "controlled" patterns of retrieval (Humphreys et al. 2015; Humphreys & Lambon Ralph, 2017). The 1-back condition was harder than the 0-back condition and still elicited a greater response.

Instead, our findings are consistent with suggestions that the role of the angular gyri role is to allocate attention to complex memory representations. The

angular gyri show a stronger response to a range of memory retrieval situations in which the retrieved representations are detailed, specific or precise (Binder et al. 2005; Price, et al. 2015; Bonnici et al. 2016; Davey at al. 2015). In our study, the 1-back trials required attention to switch from an encoding mode, to a retrieval mode when task relevant information is represented internally. This pattern of responding in the angular gyrus is consistent with the purported role of inferior parietal cortex in focusing attention on memory (Cabeza et al. 2011). In our study, the angular gyri did not activate to the same extent when stimuli were matched on colour, suggesting this region is especially important when heteromodal representations from long-term memory are retrieved.

Functional implications for temporal cortex: Angular gyrus shows strong intrinsic connectivity with ventral anterior temporal cortex (Davey et al. 2016; Jackson et al. 2016), which is proposed to support the integration of multiple features and modalities to capture 'deep' conceptual similarities between items with diverse 'surface' features (e.g., items such as PINEAPPLE and KIWI that have different colours, sizes, shapes, phonology etc.; for a review see Lambon (Ralph et al. 2017). Semantic dementia patients with atrophy focussed on this region show highly consistent degradation of conceptual knowledge across tasks (Bozeat et al. 2000; Jefferies & Lambon Ralph, 2006), while neuroimaging studies of healthy participants localise the response during heteromodal conceptual processing to ventral anterior temporal lobes and anterior middle temporal gyrus (Murphy et al. 2017; Visser et al. 2011). Word meaning can be decoded within anterior middle and inferior temporal gyri, while patterns of activation in superior temporal gyrus instead reflect the presentation format (Murphy et al. 2017).

The ventrolateral anterior temporal lobes are thought to provide a "graded hub" in which different unimodal features are gradually integrated to form heteromodal concepts, with visual information reaching this region along the ventral visual pathway (fusiform cortex), auditory and motor information arriving from superior temporal gyrus and frontal cortex, and social/emotional information merging from the temporal pole (Lambon Ralph et al. 2017). Nevertheless, the peak response in the anterior temporal lobes in the current study was in lateral MTG, and not in the site of the putative hub in ventrolateral anterior temporal cortex (Murphy et al. 2017).

Visser et al. (2012) observed evidence compatible with two gradients of information convergence in the temporal lobes: first, there is a posterior-to-anterior axis: posterior temporal lobe regions proximal to visual and auditory cortex show largely unimodal responses, while more anterior regions integrate across these types of input to support heteromodal conceptual processing. Secondly, there may be integration from superior and inferior regions, implicated in auditory and visual processing respectively: towards middle temporal gyrus response become more heteromodal response along the length of the temporal cortex. The site we observed in the conjunction of semantic and perceptually-decoupled decisions in the current study corresponds to the extreme heteromodal end of both of these temporal lobe gradients.

Implications for the default mode network: We replicated prior demonstrations that transmodal regions in the DMN are engaged when participants make decisions that rely on information from memory rather than input from perception, even though the 1-back task was more difficult than the 0-back task (Konishi et al. 2015). This pattern of task-positive behaviour adds to a growing body of evidence that the DMN contributes in an active manner to demanding external cognitive tasks (Konishi et al. 2015; Krieger-Redwood et al. 2016; Spreng et al. 2014; Spreng et al. 2015; Spreng et al. 2010; Vatansever et al. 2015). The contribution of DMN to controlled cognitive states appears to reflect situations in which DMN regions work in tandem with the frontoparietal network. Prior work has established the combination of these networks is important for tasks including controlled semantic retrieval (Krieger-Redwood et al. 2016), working memory (Vatansever et al. 2015), autobiographical planning (Spreng et al. 2014; Spreng et al. 2015), retrieving memories of close personal friends (de Caso et al. 2017) and the control of spontaneous thoughts in a deliberate manner (Golchert et al. 2017). Our study shows that right angular gyrus, within the DMN, and right dorsolateral prefrontal cortex, a member of the frontoparietal network, activate together when participants make judgments about meaningful objects from memory rather than colours (see Supplementary Figure A.1.1). Our functional connectivity analysis also demonstrates that these regions are correlated at rest, suggesting they form an intrinsic network. The right dorsolateral cluster replicates the spatial distribution observed from the prior study by Konishi et al. (Konishi et al. 2015) and overlaps with a region of greater grey matter associated with more deliberate mind-

wandering (Golchert et al. 2017). Both 1-back retrieval in our paradigm, and more deliberate spontaneous thought, require memory retrieval to be shaped in a goal-directed fashion. It is possible that a range of states requiring the goal-directed control of memory depend on co-operation between these two large-scale networks.

At the most general level our study supports the hypothesis that the capacity for complex memory representations to influence cognition emerges from the topographical arrangement of neural processes across the cortex. Prior work highlighted regions of transmodal cortex, such as the default mode network, as having the greatest distance from uni-modal sensorimotor cortex in both functional and structural space (Margulies et al. 2016). Our finding builds on this observation by showing an association between ongoing neural activity and this dimension of connectivity under situations when the demands placed on cognition require a combination of memory complexity and stimulus independency. Using both standard and novel methods of analysis, we demonstrated that the neural activity associated with this type of activity is prevalent in transmodal regions (Figure 5.5) and can be represented as a whole brain shift in the balance of neural activity, away from sensorimotor regions cortex and towards the transmodal end of the gradient (Figure 5.6). This topographical shift in the distribution of neural processing is consistent with theoretical accounts that assume that more abstract cortical functions are facilitated through functional isolation from incoming input (Buckner & Krienen, 2013; Margulies et al. 2016; Mesulam, 1998; Smallwood, 2013). Consistent with this interpretation of DMN function, a recent study found that strong connectivity within the DMN (including an overlapping region of left temporal cortex) was linked to poor performance on tasks which depend on encoding information from the environment but not for those that depended on retrieving information from memory (Poerio et al. 2017). The findings of Poerio and colleagues, in combination with those from the current study, provide converging evidence that regions of the DMN support a state where cognition is guided by memory rather than input, regardless of whether it is beneficial to the task or not.

There are a number of limitations that should be borne in mind when considering the results of this study. First, our comparison of semantic and colour decisions allowed us to demonstrate a neural pattern associating conceptual

processing with stimulus independency. This comparison is too crude a manipulation to determine which aspects of the semantic judgements gave rise to this response in the DMN; for example, is it the richness of concepts such as Labrador or apple, their heteromodal nature or the fact that they are acquired over a lifetime, which dissociates them from colours? Future studies could probe different features of retrieval, such as whether the target is a concrete or abstract concept, whether it has to be identified at a specific or superordinate level, and whether there are differences according to the modality of the representation being probed. Second, the nature of our design precludes the ability to separate different aspects of memory retrieval engaged during 1-back decisions. In our paradigm, these decisions require both the integration of appropriate information from memory, as well as the inhibition of the memory representation for the non-probed item. Interestingly, studies have implicated dorsolateral prefrontal cortex in the suppression of memories (Anderson et al. 2004) whereas the angular gyrus has been linked to the integration of appropriate semantic features (Wagner et al. 2015) and the retrieval of specific information (Davey et al. 2015). It is possible that the angular gyrus and dorsolateral prefrontal region are performing distinct roles in integration of relevant associations and suppression of irrelevant information during retrieval in our paradigm. Future work could address this question by manipulating the level of featural overlap between target and probe during retrieval in a similar paradigm as in this experiment.

Chapter 6 - Thesis Summary and Discussion

6.1. Summary of Research Questions

This thesis sought to understand the contributions of unimodal sensory and transmodal cortex to the retrieval of conceptual knowledge. The empirical work in this thesis employed paradigms that varied the extent to which semantic access was directly driven by sensory inputs or generated internally from memory. This is an interesting avenue for investigation since semantic retrieval extends beyond the here-and-now, to draw on abstract knowledge that has been extracted across multiple experiences; for instance, we can easily bring to mind what a dog looks and sounds like, regardless of whether or not there is a dog present in our immediate environment. However, a clear understanding of the neural substrates that support patterns of semantic retrieval that are not immediately driven by stimuli in the environment is lacking. For instance, while previous research has resulted in a number of theories regarding the broad organisation of semantic cognition in the brain (e.g., Barsalou et al. 2003; Binder & Deasi, 2011; Damasio, 2008; Koenig & Grossman, 2007; Lambon Ralph et al. 2017; Meteyard et al. 2012; Patterson et al. 2007; Pulvermüller, 2013; Tranel, Damasio & Damasio, 1997), most of this work has considered processes such as word comprehension or picture recognition; less is known about patterns of semantic retrieval that are perceptually-decoupled (i.e., not immediately driven by stimuli in the environment). Furthermore, many theories of semantic representation propose that concepts are computed through the interaction of sensory-motor features in unimodal cortex with abstract or transmodal representations which integrate these features (e.g., Lambon Ralph et al. 2017; Margulies et al. 2016; Patterson et al. 2007): however, the way in which these components are recruited might differ depending on whether retrieval is tightly coupled to the external world or internally-focused.

Notably, much of the empirical work in this thesis takes inspiration from modern accounts of transmodal brain regions (e.g., Lambon Ralph et al. 2017; Margulies et al. 2016). These have helped to form hypotheses about the role of the ATL as a hub that

integrates modality-specific information from the spokes, to form abstract transmodal conceptual representations. As well as the default mode networks role in perceptually-decoupled and conceptually-guided cognition. Previous literature indicates the DMN is optimal for the abstraction of conceptual representations from multiple experiences and modalities, and the generation of mental content that is not directly mirrored in the external world. These functions occur by virtue of the distance from DMN to primary sensory-motor regions on the cortical surface and in intrinsic connectivity (Bullmore & Sporns, 2009; Hagmann et al. 2008; Margulies et al. 2016). This functional separation is characterised by a whole-brain principal gradient of connectivity, described by Margulies and colleagues (2016). Increasing distance from sensory-motor regions might also allow the integration of multiple feature types, supporting the emergence of complex, high-dimensional representations, as envisaged by contemporary models of semantic processing (e.g., hub-and-spoke model; Patterson et al. 2007; the graded-hub account; Lambon Ralph et al. 2017). This process of integration might be detectable as "echoes of integration" in connectivity decompositions (Braga et al. 2013; Leech et al. 2012). These explanatory frameworks are used throughout the thesis.

The thesis employed a combination of task-based fMRI and machine learning approaches plus analyses of intrinsic connectivity to investigate the retrieval of conceptual knowledge. MVPA is used to localise functions of unimodal and transmodal brain regions; this permits the exploration of regions that can classify different aspects of knowledge. Univariate analysis is employed to identify the circumstances in which semantic regions, including those in the DMN, show activation and deactivation, while functional connectivity is used to express measures of intrinsic organisation. The specific aims of this thesis were:

- To explore the role of unimodal sensory-motor cortex in semantic representations. Using MVPA, chapters 3 and 4 will explore whether the patterns of activity in unimodal cortex represent modality of presentations (Spoken vs. written; Chapter 3), modality of word meaning (loud vs. bright; Chapter 3) and modality of retrieved memory (thinking about what a dog sounds like vs. what it looks like; Chapter 4).

- To localise where within the ATL abstract heteromodal semantic representations are supported. Using a searchlight MVPA, Chapter 3 investigates where, within the ATLs, patterns of activity for a concept activated through the visual domain map on to the same concept activated through the auditory domain. This cross-classification will permit the localisation of a 'hub' region that captures abstract meaning irrespective of presentation modality. Moreover, analysis of the intrinsic connectivity of this region will provide evidence of whether this region is embedded within a network that facilitates abstract transmodal processing (e.g., default mode network).
- To investigate perceptually-decoupled semantic retrieval states such as imagination (Chapter 4) and judgements from memory (Chapter 5). Chapter 4 will examine perceptually-decoupled forms of semantic retrieval and establish which brain regions can decode sensory features in imagination, in the absence of input. While Chapter 5 assessed the conjunction of cognitive processes that required (i) multi-featural abstract concepts and (ii) decoupling. Notably, this later study combines the two dominant features of chapter 3 (abstract conceptual representations) and perceptual-decoupling (retrieval of knowledge from memory in the absence of input). Both of these experiments will interrogate whether unimodal and/or transmodal regions are necessitated by perceptually-decoupled semantic retrieval and measure the intrinsic connectivity of these brain regions.
- To identify circumstances in which regions of the transmodal DMN shows above baseline activation during semantic tasks. This is addressed in Chapter 5 by comparing semantic conditions (multi-featural concepts) with perceptual conditions (simple colour patches) in a univariate fMRI analysis.

6.2. Main Findings

6.2.1. Chapter 3

This study established the functional organisation of both sensory cortex and ATL, to modality-invariant conceptual retrieval. In an fMRI experiment, participants listened to

or viewed words that referred to either visual or auditory features (e.g., bright or loud). Using a linear classification algorithm, this study found a functional dissociation between superior and ventrolateral ATL, such that superior portions hold patterns of activity regarding sensory modality, and the ventrolateral portions hold patterns of activity about abstract modality-invariant semantic representations. Moreover, the activity within this ventrolateral region was equated from both spoken and written format, suggesting that ventrolateral ATL is transmodal in nature, as it does not show preference for spoken or written input. An additional functional connectivity analysis revealed different patterns of connectivity within superior and ventrolateral portions of the ATL, with superior ATL showing stronger connectivity to language, auditory and motor regions, while ventrolateral ATL showed connectivity to other transmodal semantic regions, such as angular gyrus, posterior cingulate and hippocampus within the default mode network and limbic system. Furthermore, primary auditory and visual cortex held patterns of activity for both the sensory modality and the semantic category. However, these patterns were not overlapping suggesting that sensory experience and meaning are not equated in sensory cortex.

These findings therefore reveal a region in ventrolateral ATL that captures meaning irrespective of presentation modality, as proposed by the graded-hub account (Lambon Ralph et al. 2017). Moreover, the finding that ventrolateral ATL 'hub', but not aSTG, is functionally coupled to the default mode network is consistent with the view that the integration of increasingly abstract representations are facilitated by regions at the farthest end of the gradient (Margulies et al. 2016)

6.2.2. Chapter 4

Having found that presentation format could be separated from semantic concepts in both unimodal and transmodal brain regions in Chapter 3, the next study examined perceptually-decoupled forms of semantic retrieval, to establish which brain regions could decode sensory features in imagination, in the absence of input. Notably, this design permitted the investigation of transmodal brain regions, because this network has been previously implicated in cognition that is independent of the immediate environment. However, only a handful of studies directly compare different forms of

memory retrieval within this network (e.g., Daselaar et al. 2010). Using a novel paradigm, where participants were presented with a constant source of visual and auditory noise, to control for sensory input across conditions, participants were prompted to generate imagery, in the mind's ear and/or eye, relating to the experimental conditions (i.e., focus on the sound of a concept, its visual features or a conceptually complex context). MVPA was then used to identify brain regions that held patterns of activity informative for the different forms of imagination while functional connectivity was performed to show the intrinsic architecture within which these results were embedded. This combination of methods permitted the investigation of whether the content of retrieved memories recruits common or distinct brain regions. The key findings were that:

a. In line with previous imagery literature, there were consistent patterns of activity regarding retrieved memories in secondary sensory cortex (Albers et al. 2013; Daselaar et al. 2010; de Borst & de Gelder, 2016; Halpern & Zatorre, 1999; Ishai et al. 2000; Reddy et al. 2010; Vetter et al. 2014; Zvyagintsev et al. 2013). Specifically, STG activated to sound and context imagery and held patterns of activity that could decode between sound and other forms of imagery, whereas inferior lateral occipital cortex, predominantly activated to visual and context imagery and held patterns of activity informative for decoding visual conditions. These results are consistent with the 'anterior shift' noted by Thompson-Schill (2003), where areas activated by semantic processing are not isomorphic to those used in direct experience in primary sensory regions, but rather are shifted anterior to those areas.
b. Several regions allied to or within the extended default mode network (including anterior middle temporal gyrus, anterior parahippocampal gyrus and anterior cingulate gyrus) could decode between sound and visual imagery. Although the MVPA analysis revealed differences in the patterns of activity within default mode regions, the univariate analysis could not find any statistical differences between the imagery types. This finding is interesting as, although the default mode is considered transmodal (and the univariate activity is comparable across all forms of imagery), the *patterns*

of activity within these regions are informative to decoding unimodal memory retrieval (thinking of a dog barking vs. what a dog looks like), and thus must contain traces of sensory information. These results therefore align with previous literature suggesting that regions capable of performing integration, such as default mode and heteromodal cortices, might be expected to contain traces or "echoes" of the neural signals from multiple networks (Braga et al. 2013; Leech, Braga & Sharp, 2012).

c. Finally, context based imagery could only be decoded in regions of cortex that fell within the dorsal attention network (inferior lateral occipital cortex; LOC). More broadly, this pattern of results demonstrates that imagery not only recruits transmodal regions associated with memory representation and retrieval (i.e., areas allied to or within the extended DMN) but also sites implicated in executive control when the features that are being retrieved have to be shaped to suit the context, and/or when complex patterns of retrieval are required (Corbetta & Shulman, 2002; Dosenbach et al. 2009).

This chapter corroborates the assertion in Chapter 3 that both unimodal and heteromodal brain regions are critical for semantic retrieval in the absence of relevant sensory input. This study revealed extensive recruitment of sensory cortex, which was modulated by the modality of imagination required by the task. In addition, there was also an observed role for transmodal brain regions in supporting internally-generated conceptual retrieval. These findings emphasise the bi-directional connections between hub and spoke regions to permit the complete conceptualization of knowledge (Lambon Ralph et al. 2010; Pobric, Jefferies & Lambon Ralph, 2010; Reilly, Peelle, Garcia, & Crutch, 2016). Moreover, this data is highly coherent with recent perspectives on information-integration within transmodal regions of the brain that suggest graded information convergence (e.g., Buckner et al. 2009; Bullmore & Sporns, 2009; Hagmann et al. 2008; Lambon Ralph et al. 2017; Margulies et al. 2016).

6.2.3. Chapter 5

As both Chapters 3 and 4 revealed similar portions of unimodal and heteromodal brain regions for both abstract processing and perceptual decoupling, the concluding fMRI study assessed the conjunction of cognitive processes that required (i) multi-featural concepts as opposed to unidimensional perceptual stimuli and (ii) perceptual-decoupling. This fMRI experiment therefore used a matching paradigm and manipulated whether decisions were semantic or perceptual in nature (based on single-feature colour patches, or images of multi-featural concepts such as types of dogs or cars) and the availability of information in the external world (present vs. absent – i.e., the decision was based on memory). This chapter also characterised the whole-brain activity for each of our conditions in terms of their position on the macro-scale gradient from unimodal to heteromodal cortex described by Margulies et al (2016). This allowed us to test the hypothesis that isolation from input is a critical feature shared by cognitive states activating the DMN. Notably, this experiment benefitted from an increased sample size compared to the previous two empirical chapters. The findings revealed:

a. Perceptually guided decision-making (present > absent) revealed increased activity in bilateral ventral visual stream, regions which have a well-documented role in online visual processing.
b. In contrast, both decoupled states (absent > present) and more conceptually complex states (multi-featural concepts > uni-featural concepts) activated bilateral angular gyrus, anterior temporal lobes, posterior cingulate cortex and medial prefrontal cortex; regions that fall predominantly within the default mode network, or in networks allied to the DMN.
c. A formal conjunction of these contrasts (absent > present AND multi-featural > uni-featural) revealed consistent activation in bilateral angular gyrus and left middle temporal gyrus – showing that core regions of the default mode network are critical for making conceptual decisions based on information that was no longer present in the environment.
d. Finally, the intrinsic architecture of conjunction regions using functional connectivity was explored. The results of this analysis reveal coupling to core parts of the default mode network as well as portions of frontoparietal

network. One explanation for why these regions recovered both default mode and frontoparietal networks, might relate to the fact that the 1-back condition was more demanding than 0-back; suggesting that integration between executive and default mode networks is necessary for demanding tasks.

These data highlight that the conjunction of conceptual and memory demands recruits bilateral angular gyrus and left lateral middle temporal gyrus. These three regions fall at the highest end of the principal gradient, even further than ventrolateral ATL, supporting the notion that the most abstract forms of cognition are processed at the highest end of the representational hierarchy (Margulies et al. 2016).

6.3 Linking Data to Theory

6.3.1. Sensory Cortex

Embodied accounts of semantic processing postulate that neural regions generally used for perception and action, are also recruited during semantic processing (Barsalou, 1999; 2008; Humphreys & Forde, 2001; Martin, 2007; Patterson et al. 2007; Pulvermüller, 2005). These regions are considered modality-specific as they represent modality-specific attributes and are located in areas proximal to and reciprocally linked to primary sensory and motor regions. For instance, knowledge about the sound a dog makes would be represented in auditory cortex. It is suggested that retrieving a concept will engage neural pathways that encode items distinct colour, shape, sound, motor properties and so on (e.g., Patterson et al. 2007; Pulvermüller, 2005). Indeed, a plethora of functional neuroimaging studies have provided compelling evidence that these neural pathways are, to some extent, shared with perception and action systems (Martin et al. 1995; Goldberg, Perfetti Charles & Schneider Walter, 2006; Rueschemeyer et al. 2014). For instance, words denoting actions (e.g., kick) and manipulable objects (e.g., hammer) elicit activation in the brain's motor system (Hauk, Johnsrude & Pulvermüller, 2004) whereas words associated with specific smells (e.g., cinnamon) elicit activation in olfactory cortex (Gonzalez et al. 2006). Taken together,

these findings inform cognitive accounts of perception and knowledge representation. Clearly, however, perception and internally-driven cognition, are not the same phenomena, and many questions remain regarding the relationship between sensory experience and meaning. For example, despite sensory and motor cortices *activating* to both perception and states of imagination or semantic retrieval to words, it remains unclear whether the *representations* elicited by these states are overlapping or whether these situations involve a common neural region, but are coded in a unique format. Recent evidence in the motor domain suggests that sensory experience and meaning are coded in unique formats (e.g., Rueschemeyer et al. 2010; 2014), however it is not clear whether this is true of other modalities of testing, such as in the visual or auditory domain.

The data presented in this thesis suggest that sensory cortex is recruited during semantic retrieval for both word meaning (Chapter 3) and perceptually decoupled states (Chapter 4). The overlap in sensory recruitment across the two studies, taken from the MVPA analysis, is depicted in Figure 6.1. Both auditory and visual cortices held patterns of activity not only regarding sensory input (spoken, written) but they were also sensitive to modality-specific meaning. In Chapter 3 this meaning was related to the sensory-feature of a word (i.e., the word 'loud' refers to a concept perceived through the auditory modality whereas 'bright' refers to a concept perceived through the visual modality; Figure 6.1A), while Chapter 4 showed that sensory regions were recruited when imagining the sound or visual properties of a concept (i.e., thinking about a dog barking versus thinking about what it looks like; Figure 6.1A-B).

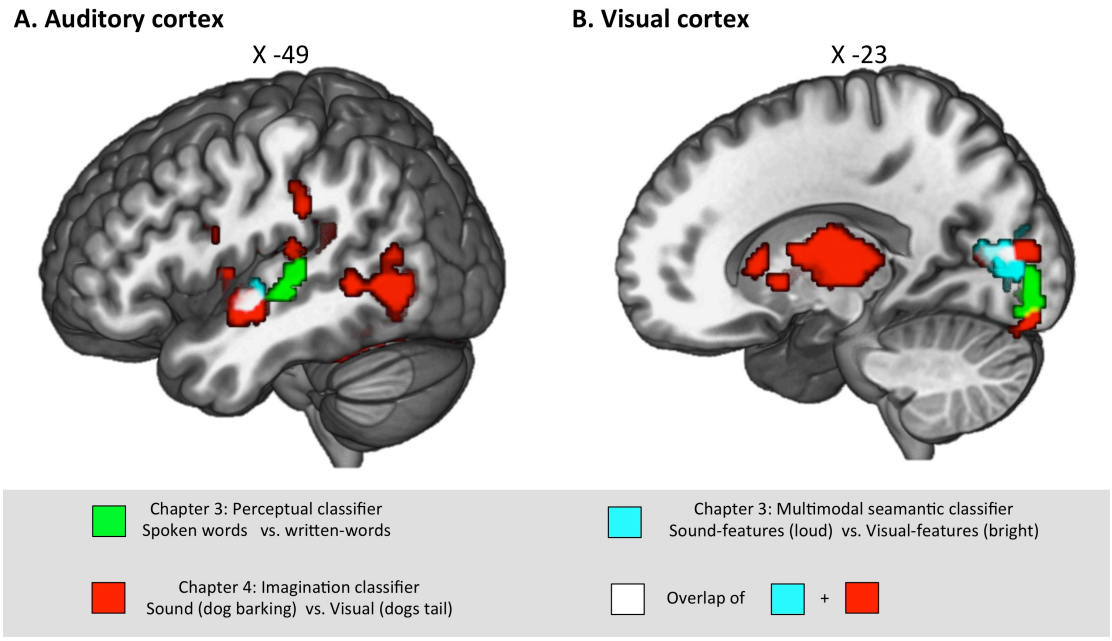

Figure 6.1. MVPA results from chapter 3 and 4 highlighting similarity of results in primary auditory and visual cortex. (A) Results within auditory cortex, the overlap between semantic classifier (sound vs. bright) and imagination classifier (dog barking vs. dog tail) projected in white in left auditory cortex (Planum polare). (B) Results within visual cortex, the overlap between semantic classifier (sound vs. bright) and imagination classifier (dog barking vs. dog tail) projected in white (intracalcarine cortex and lingual gyrus).

Interestingly, the results in Chapter 3 highlight that, although primary auditory and visual cortex held patterns of activity for both the sensory modality (green clusters in Figure 6.1) and the semantic category (cyan clusters in Figure 6.1), these patterns were predominantly non-overlapping, suggesting that sensory experience and meaning are not equated in primary sensory cortex. Moreover, in Chapter 3 it was not possible to cross-classify between modality of meaning and modality of perception. Meaning and sensory experience did not share a common neural pattern, consistent with a growing body of literature that suggests meaning is not a direct reflection of sensory experience in sensorimotor cortex (e.g., Rueschemeyer et al. 2010; 2014). For instance, a recent MVPA study conducted by Coutanche and Thompson-Schill (2014) interrogated whether the pattern of activity for an imagined concept (e.g., orange) was comparable to the activity when perceiving an orange. They found that only left ATL was able to successfully perform this cross-classification, suggesting that meaning and experience are not equated in unimodal portions of cortex but instead are

represented in heteromodal convergence zones, arguably because such regions code conceptual knowledge in an abstract form that is comparable across memory and perceptual experience.

Furthermore, both Chapter 3 and 4 show that overlapping regions anterior to perception hold patterns of activity regarding semantic content (see white clusters in Figure 6.1). In Chapter 3 the clusters that could decode word meaning (loud versus bright) were anterior to those clusters that could decode input (spoken versus written), while Chapter 4 identified clusters in anterior secondary sensory cortices that could decode between different forms of imagined features (dog barking versus dog tail). One possible explanation for the recruitment of sensory association cortices is the 'anterior shift' noted by Thompson-Schill (2003). She found that areas activated by semantic processing are not isomorphic to those used in direct experience, but rather are shifted anterior to those areas (for a wider review see Chatterjee, 2010; Binder & Desai, 2011; McNorgan et al. 2011; Meteyard et al. 2012). This anterior shift therefore suggests that information in modality-specific regions is abstracted from direct experience during retrieval of semantic concepts from memory. Such accounts are therefore consistent with a gradient of 'abstraction', where, as one moves way from primary sensory and motor cortex, more complex conjunctions are captured (Bajada et al. 2017; Buckner & Krienen, 2013; Lambon Ralph et al. 2017; Margulies et al. 2016; Mesulam, 1998; Meteyard et al. 2012; Patterson et al. 2007; Plaut, 2002; Visser et al. 2010).

Taken together these findings therefore do not align with strong embodied accounts, which propose that neural and phenomenological processes that occur during sensory perception and semantic processing are similar in function and structure due to shared underlying neural mechanisms (for a review, see Meteyard et al. 2012; Pulvermueller, 2005; Glenberg & Kaschak, 2002). Instead, the findings align with integrative embodied theories that do not postulate meaning and experience are equated in sensory cortex, but instead, emphasise the contribution of heteromodal brain regions to assist semantic processing (for review see Meteyard, et al. 2012; Caramazza et al. 1990; Damasio, 1989; Lambon Ralph, Sage, Jones & Mayberry, 2010; Martin, 2007; Patterson et al. 2007; Pulvermueller, 2013; Riddoch et al. 1988; Rogers et al. 2004). This interpretation is further supported by the findings that modality-

specific information (e.g., the meaning of the word 'loud' – Chapter 3; or imagining what a dog sounds like – Chapter 4) activates brain regions that extend beyond sensory cortex in to heteromodal structures such as anterior temporal lobe, and portions of the default mode network.

One plausible explanation of why we discovered both unimodal and heteromodal regions in this thesis relates to top-down activation hypotheses (Friston, 2012; Mechelli et al. 2004; Miller & D'Esposito, 2005; Naselaris et al. 2015; Shulman et al. 1997). Such theories suggest that over time heteromodal brain regions, such as anterior temporal lobe, extract relevant featural information from sensory cortex during perception in order to form abstract generalizable concepts (the concept 'dog' is formed in heteromodal brain regions, due to the integration of what a dog sounds like, looks like, how it moves etc.) (e.g., Berkes et al. 2011; Patterson et al. 2007; Lambon Ralph et al. 2017). This in turn creates heteromodal representations of concepts that are no longer tied to sensory input. In line with predictive coding theories, higher order regions, such as anterior temporal lobe, maintain these heteromodal internal representations, which provide predictions about the environment to the early sensory regions through top-down activation (Friston, 2012; Mechelli et al. 2004; Miller & D'Esposito, 2005; Naselaris et al. 2015; Shulman et al. 1997). For example, during visual imagery transmodal brain regions mediate activity in early visual areas through feedback connections (de Borst et al. 2012; Mechelli et al. 2004). It has been argued that these feedback connections permit the primary sensory cortices to represent features of objects during imagery, which may facilitate interpretation of sensory input. Furthermore, the involvement of default mode regions has been shown to facilitate the combination and modification of stored perceptual information in novel ways (Kosslyn, Ganis & William, 2001). This, for instance, allows us to imagine scenarios that we have never experienced before, such as imaging the dog races, despite never having attended this event.

Take home message: Collectively, the findings from this thesis align with integrative embodied accounts (e.g., hub-and-spoke model; Patterson et al. 2007) which highlight that modality-specific areas reflect general representations of content-features (e.g., auditory cortex reflects the sound a dog makes, whereas visual cortex reflects what a dog looks likes); however the spatial characteristics of modality-specific

activations differ for semantic meaning and sensory experience. Moreover, the results of this thesis suggest that in addition to sensory portions of the cortex, transmodal brain regions – such as anterior temporal lobe and the default mode network – are recruited to facilitate semantic retrieval, arguably through top-down activation of the relevant conceptual-feature in sensory cortices.

6.3.2. Anterior Temporal lobe

Theoretical accounts of semantic cognition have suggested that in addition to sensory portions of the cortex - outlined above - cross-modal interactions for all modality-specific information are mediated, at least in part, by convergence zones (Barsalou et al. 2003; Binder & Desai, 2011; Damasio, 1989; 2008; Mesulam, 2000; Meteyard et al. 2012; Patterson et al. 2007; Pulvermüller, 2013; Simmons & Martin, 2009; Tranel, Damasio & Damasio, 1997). One prominent theory suggests this convergence zone or hub is situated bilaterally in the ATLs (e.g., hub-and-spoke model; Patterson et al. 2007). Notably, however, the ATLs are not one homogenous region and the debate surrounding the heterogeneity of the ATLs has been fuelled by compelling evidence highlighting both structural (Brodmann, 1909) and functional differences across superior-to-ventral ATL (Baylis, Rolls & Leonard, 1987; Binney, Parker & Lambon Ralph, 2012; Geranmayeh, Leech & Wise, 2015; Jackson et al. 2016; Rice, Hoffman & Lambon Ralph, 2015; Skipper, Ross & Olson, 2011; Visser & Lambon Ralph, 2011; Visser et al. 2012; Yeo et al. 2011). As a result of such findings, contemporary accounts of semantic cognition have extended the predictions of the hub-and-spoke model and postulate a broader graded function across the ATL, with superior and ventral portions showing sensitivity to sensory input and ventrolateral ATL showing the properties of an integrative multi-modal hub (graded-hub account; Lambon Ralph et al. 2017). Despite both neuropsychological and neuroimaging literature converging on ventrolateral ATL as the site of a transmodal integrative hub, few studies have investigated whether patterns within this area reflect truly transmodal representations (Chapter 3) or whether this region is recruited during perceptually decoupled semantic retrieval (Chapter 4 & 5). Figure 6.2 illustrates the graded function of the ATL, as outlined by the

graded-hub account (adapted from Lambon Ralph et al. 2017; Figure 6.2A) and summarizes the ATL findings across the three experimental chapters (Figure 6.2C).

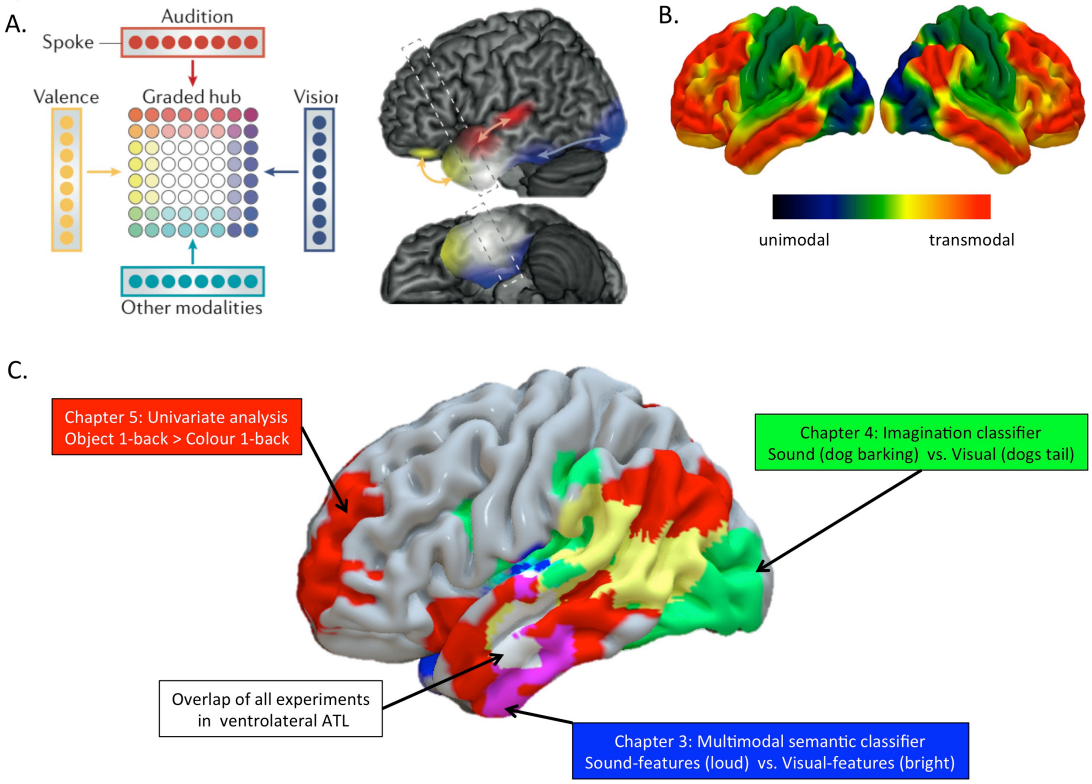

Figure 6.2. Comparison of our ATL findings across the three experimental chapter. (A) Modified version of the graded hub account (Lambon Ralph et al. 2017) highlighting ventrolateral ATL as the site of the putative 'hub' in white. (B) Modified version of the principal gradient account (Margulies et al. 2016) highlighting the convergence of unimodal regions into transmodal cortices. (C) MVPA results from Chapter 3 - multimodal semantic classifier (loud vs. bright) project in blue, MVPA results from chapter 4 - imagination classifier (imagining what a dog sounds like versus what a dog looks like) projected in green and univariate results from chapter 5 – conceptual retrieval from memory (object 1-back > colour 1-back) projected in red. Overlap of (i) 3 and 4 in cyan, (ii) Chapter 3 and 5 in magenta (iii) Chapter 4 and 5 in yellow and (iv) all three experiments (Chapter 3, 4 and 5) in ventrolateral ATL projected in white.

The MVPA findings from Chapter 3 and Chapter 4 reveal a similar cluster within ventrolateral ATL that could decode semantic content in a modality-invariant manner, that is the patterns for decoding between loud vs. bright in the spoken modality where comparable to the patterns for decoding loud vs. bright in the written modality

(Chapter 3). This region also held patterns of activity that could decode between imaging the auditory properties of a concept and imagining the visual properties of a concept (Chapter 4). Notably, both of these studies show transmodal representations about both sound (e.g., dog barking) and visual concepts (e.g., dogs fluffy tail) are represented within ventrolateral ATL. This is an interesting discovery for several reasons. First, previous literature has attempted to utilise MVPA to address similar questions regarding representational content in the ATL (e.g., Peelen & Caramazza, 2012; Correia et al. 2014). However, these studies typically report very different portions of ATL as the putative 'hub' site. For instance, Peelen and Caramazza (2012) found that bilateral ventral ATL encodes information about the abstract conceptual properties of objects, whereas Correia et al (2014) identified superior ATL in a crossmodal study investigating semantic representations in Dutch-English bilinguals. One plausible explanation for these conflicting results is the notion that presentation format influences the location of heteromodal processing with the ATLs. Peelen and Caramazza used picture stimuli which have been shown to recruit ventral ATL regions in line with a graded effect of modality-input (Plaut, 2002; Visser et al. 2010) and typically depend heavily on visual-feature knowledge. While Correia and colleagues utilized spoken concrete nouns, in both English and Dutch, which depend heavily on orthographic and phonological knowledge. These previous results are therefore consistent with graded effect of input modality (Lambon Ralph et al.2017; Margulies et al. 2016; Plaut, 2002; Visser et al. 2010), with language and auditory semantic processing recruiting superior portions of the ATL and visual and picture processing recruiting ventral portions of the ATL; however they do not provide compelling evidence for a 'transmodal' hub as these effects seem to be, at least in part, driven by the modality of presentation. The current data therefore adds to this literature by showing transmodal representations, not driven by either auditory or visual input, are represented in ventrolateral ATL.

Second, most studies of ATL function have focused on the representation of object concepts, which heavily rely on visual feature knowledge (Binney et al. 2012; Bright, Moss & Tyler, 2004; Correia et al.2014; Countanche & Thompson-Schill, 2014; Lambon Ralph et al.2009; Peelen & Caramazza, 2012; Visser et al. 2010; 2012). As pointed out in a review by Bonner & Price (2013), what is less known about the ATL is

whether it also encodes categories that have weak visual-feature associations, such as auditory concepts like “loud” and “thunder”. The data presented in this thesis lends tentative support for the notion that visual-feature and auditory-feature concepts are represented, or at least decodable, within ventrolateral ATL.

Furthermore, the findings demonstrate that recruitment of ventrolateral ATL, but not superior ATL, is necessitated by perceptually decoupled memory retrieval, in the form of imagination (Chapters 4 & 5). Specifically, this site held unique patterns of activity that coded different forms of imagination (thinking about what a dog sounds like versus what it looks like), and responded more to the retrieval of concepts from memory compared to simple colour patches. Interestingly, Margulies and colleagues (2016) situate this ventrolateral ATL site higher along the principal gradient than more superior or ventral portions of ATL, such as fusiform gyrus. The current findings therefore complement this account as transmodal regions in ventrolateral ATL which are further along the principal gradient, are implicated in cortical functions by virtue of its distance from both auditory (projecting along superior temporal gyrus) and visual (projecting along the ventral visual stream through the fusiform) input. Therefore, the findings of Margulies et al (2016), in combination with those from the current experimental chapters, provide converging evidence that regions of transmodal cortex, namely ventrolateral ATL, support a state where cognition is guided by memory rather than input.

Collectively, Chapters 3 and 4 demonstrate that ventrolateral ATL recruitment is required for a diverse range of functions, including transmodal conceptual representations and perceptually decoupled memory retrieval. However, the conjunction of multi-featural conceptual knowledge and perceptually decoupled retrieval, as investigated by Chapter 5, revealed a site posterior to this ventrolateral cluster, in lateral middle temporal gyrus (MTG). This suggests that ventrolateral ATL may be important for both heteromodal conceptual representations and perceptual decoupling, but the two processes converge in neighbouring lateral MTG. Notably, this lateral MTG site has been identified as a candidate ‘hub’ in several theories of semantic memory (Binder & Desai, 2011; Bonnici et al. 2016; Mesulam, 1985). We therefore find compelling evidence that ventrolateral ATL and lateral MTG both show properties consistent with heteromodal hubs, that integrate increasingly abstract and

complex representations, however the precise function differs; ventrolateral ATL captured transmodal representations while lateral MTG is necessitated by conceptually complex perceptually-decoupled states.

Interestingly, although the two temporal sites (ventrolateral ATL and lateral MTG) differ in their peak activity (see Figure 6.3 clusters c-e), it is worth emphasizing that both sites are not only intrinsically linked to one another but are also embedded within the DMN (Yeo et al. 2011), with the latter being situated at the highest point within the principal gradient (Marguiles et al. 2016). This findings also fit with the two gradient account of Visser et al. (2011), who showed (i) a first gradient of convergence from superior and ventral portions into ventrolateral ATL and (ii) a second gradient of processing that extends longitudinally (caudal-rostral) along the middle temporal gyrus, converging in regions of transmodal cortex to facilitate abstract conceptual processing (Buckner et al. 2009; Bullmore & Sporns, 2009; Hagmann et al. 2008; Lambon Ralph et al. 2017; Marguiles et al. 2016; Mesulam, 1998; Plaut, 2002; Visser et al. 2011). Specifically, the current data builds on research that suggest a gradient from the most rostral and caudal temporal regions into lateral MTG (i.e., Temporal pole → lateral MTG ← posterior MTG) to facilitate transmodal conceptual processing (Binney et al. 2012; Binder et al. 2009; Lambon Ralph et al. 2017; Visser et al. 2012). The most rostral temporal regions receive inputs from limbic systems regarding valence and social concepts while the posterior regions receive modality-specific information from sensory cortices (Jackson et al. 2016; 2017; Pascual et al. 2013; Hurley et al. 2015; Saur et al. 2010).

Finally, it is worth noting that, although ventrolateral ATL was consistently activated across all fMRI experiments in this thesis, the paradigms varied extensively across the three experimental chapters. Both Chapters 3 and 4 employed MVPA, which measures generalizable patterns of activity across multiple voxels, whereas Chapter 5 utilized a univariate analysis, which makes use of the overall level of activity within individual voxels. To understand whether the neighbouring brain regions were a manifestation of the methods used, future studies could optimize the experiment from Chapter 5 for MVPA analysis. However, notwithstanding experimental differences, all studies proposed in this thesis share a commonality in that they investigated semantic retrieval that was not explicitly tied to sensory input from the external world (either

abstracted away from input in Chapter 3, or perceptually decoupled in Chapters 4 and 5). Moreover, despite a lack of direct overlap between the lateral temporal sites (ventrolateral ATL and lateral MTG), the functional connectivity profiles extracted across the three chapters show a common intrinsic network that reflects transmodal brain regions including both the semantic cognition network and DMN. One could therefore speculate that using multiple methodologies (univariate analysis, MVPA and functional connectivity) is advantageous as instead of relying on cluster-based univariate methods which are intrinsically tied to the arbitrary statistical thresholds applied (Woo, Krishnan & Wager, 2014), they provide converging support for a shared distributed network in which the cluster-based results lie within. Therefore, even if the peak results differ in their coordinates, their affiliation with a common neural network (e.g., semantic cognition network and DMN) provides an explanation for their cognitive function; this network is implicated in a host of cognitive tasks that focus cognition on previously encoded knowledge, as opposed to information in the external environment.

***Take home message*:** Results across the three fMRI experiments converge within ventrolateral ATL (white cluster in Figure 6.2B), a site consistent with the putative 'hub' outlined in the graded-hub account (see Figure 6.2A; Lambon Ralph et al. 2017). This ventrolateral cluster was revealed during (i) modality-invariant representations of modality-specific concepts (loud vs. bright) (Chapter 3), (ii) perceptually-decoupled processing of semantic-features (sound of a concept vs. visual-features of a concept) (Chapter 4) and (iii) during the retrieval of multi-featural concepts from memory (Chapter 5) compared to the retrieval of unimodal colour features from memory. In addition, the findings of Chapter 5 revealed a neighbouring cluster in lateral MTG was recruited for the conjunction of (i) conceptually complex semantic processing and (ii) perceptually decoupled retrieval from memory. Collectively, the findings provide tentative support for the notion that there are two principal directions of information convergence in the temporal lobes: (i) laterally (STG → MTG/ITG ← FG) and (ii) longitudinally (ATL → MTG ← pMTG). Moreover, both the ventrolateral ATL and lateral MTG sites are allied with the DMN, a network implicated in a host of cognitive tasks that focus cognition on previously encoded knowledge, as opposed to information in the external environment. Consequently, there may be

common neurocognitive processes shared between states that activate the multimodal semantic network outlined in the graded hub account and those that recruit the wider DMN.

6.3.3. Default Mode Network

To-date an established body of work has routinely shown that the DMN actively supports several aspects of cognition (Spreng, 2012), including semantic processing (Binder et al. 2009; Humphreys et al. 2015; Irish & Piguet, 2013; Krieger-Redwood et al. 2016), autobiographical and episodic recollection (Andrews-Hanna, 2012; Buckner et al. 2008; Rugg & Vilberg, 2013), working memory (Konishi et al. 2015; Spreng et al. 2014; Vatansever et al. 2015), mental imagery (Hassabis et al. 2007), self-generation of emotion (Engen, Kanske & Singer, 2017) and imagining the future or recalling the past (Huijbers et al. 2009; Schacter et al. 2007; Spreng et al. 2009; Svoboda et al. 2006; Szpunar et al. 2007). This evidence therefore goes against historical accounts of the DMN as "task-negative" by showing that the DMN does activate under a variety of task conditions. Notably, many of these situations involve memory retrieval – i.e., a requirement to focus cognition on previously-encoded knowledge as opposed to information in the external environment. The current thesis was therefore motivated by the hypothesis that there might be common neurocognitive processes underpinning perceptually-decoupled and conceptually-guided cognition in the DMN. During states of episodic recollection, we recreate past experiences that involve places, objects and people not currently present in the environment. Consequently, memory retrieval might necessitate a process of decoupling from sensorimotor systems, allowing cognition to be generated internally in a way that diverges from what is going on around us (Smallwood, 2013). These perceptually-decoupled states might be easier in brain regions whose neural computations are functionally independent, or distant, from systems important for perceiving and acting.

Notably, recent macroscale decompositions of brain connectivity have helped characterize the neural regions that are likely to be important for abstract memory representations in a more formal manner. Margulies and colleagues (2016) described a principal gradient of connectivity with unimodal sensory regions at one end and

transmodal regions including posterior cingulate cortex, medial prefrontal cortex and angular gyrus at the other – regions that are collectively are known as the DMN (Raichle et al. 2001). DMN regions are maximally distant in functional and structural space from primary landmarks of unimodal function such as the calcarine sulcus or the central sulcus (Margulies et al. 2016). This topographic architecture suggests the DMN can represent information that differs from the current state of the external world, which is reflected in sensory-motor systems. Moreover, the DMN is situated at the top of a representational hierarchy allowing these regions to integrate information across systems (Buckner et al. 2009; Bullmore & Sporns, 2009; Hagmann et al. 2008; Margulies et al. 2016; Mesulam, 1998). However, the specific role of the DMN in semantic cognition is still unclear; for instance, is it perceptually-decoupled states, conceptually complex or the combination of these two cognitive states that the DMN preferentially responds to, moreover there is little clarity regarding which regions of the transmodal DMN shows above baseline activation during semantic tasks.

The results of this thesis go some way to helping to elucidate the role of this network in semantic retrieval. For example, not only has the current thesis confirmed the finding that the DMN is engaged during perceptually decoupled retrieval of information from memory (e.g., Konishi et al. 2015), but the results show that this recruitment was enhanced when the content being retrieved was conceptual (i.e., Dalmatian) as opposed to perceptual (i.e., colour patches) (Chapter 5); suggesting that this network has a specific role to play in semantic cognition that may lie beyond perceptual decoupling. If the DMN were only involved in perceptually decoupled retrieval, there should have been no difference between the perceptually-decoupled retrieval of uni-featural colour patches or multi-featural concepts. However, an almost identical network was retrieved when participants were thinking about multi-featural concepts in both perceptually decoupled retrieval from memory and perceptually guided retrieval. Furthermore, a conjunction of the two manipulations (conceptually-guided cognition and perceptually-decoupled retrieval) revealed core sites within the DMN; bilateral AG and lateral MTG. Taken together, the results of this thesis speak to prior suggestions that there might be common neurocognitive processes underpinning perceptually-decoupled and conceptually-guided cognition in the DMN (Buckner & Krienen, 2013; Margulies et al. 2016; Mesulam, 1998; Smallwood, 2013). These regions

are maximally distant in functional and structural space from primary landmarks of unimodal function such as the calcarine sulcus or the central sulcus (Margulies et al. 2016). This topographic architecture suggests regions, such as lateral temporal cortex and angular gyrus, are situated at the top of a representational hierarchy allowing these regions to integrate information across systems. Such regions are therefore able to support higher-order representations with predictive value across multiple situations and modalities, which integrate features from diverse sensorimotor regions. In other words, increasingly abstract and complex representations might be formed at greater distances along the gradient, where the influence of specific features linked to stimuli in the immediate environment is reduced (Buckner & Krienen, 2013; Margulies et al. 2016; Mesulam, 1998; Plaut, 2002).

Speculatively, the current findings are also broadly consistent with the account that there are multiple levels of integration (see Figure 6.3). Across all experimental chapters perceptually-guided semantic retrieval engaged unimodal portions of sensory cortex. Next, following the principal gradient account, heteromodal properties of semantic retrieval began to emerge in ventrolateral ATL, a region allied with the core DMN; such activity was common to both perceptually guided heteromodal representations (Chapter 3) and perceptually decoupled semantic retrieval (Chapters 4 & 5). Finally portions at the highest end of the principal gradient – bilateral AG and lateral MTG – were recruited for the most abstracted and complex representations from memory. Interestingly, the lack of angular gyrus findings in Chapter 3 and 4, discussed previously, fits within this 'multiple levels of integration' narrative. The current data indicated that angular gyrus did not show heteromodal properties of semantic retrieval as it was not able to cross-classify between unimodal forms of conceptual knowledge, however previous literature has suggested both angular gyrus and lateral temporal cortex play a specific role in multimodal as opposed to unimodal representations (Bonnici et al. 2016). Given the functional profile and anatomical location of the findings, the results lend support to the notion that the macroscale organisation of the cortex directly relates to its cognitive function (e.g., Buckner & Krienen, 2013; Mesulam, 1998; Margulies et al. 2016); the gradient encapsulates multiple levels of integration from unimodal portions of cortex through to the most

transmodal regions of the DMN, with the later supporting abstract states (e.g., transmodal semantic concepts from memory) that are not tied to the external world.

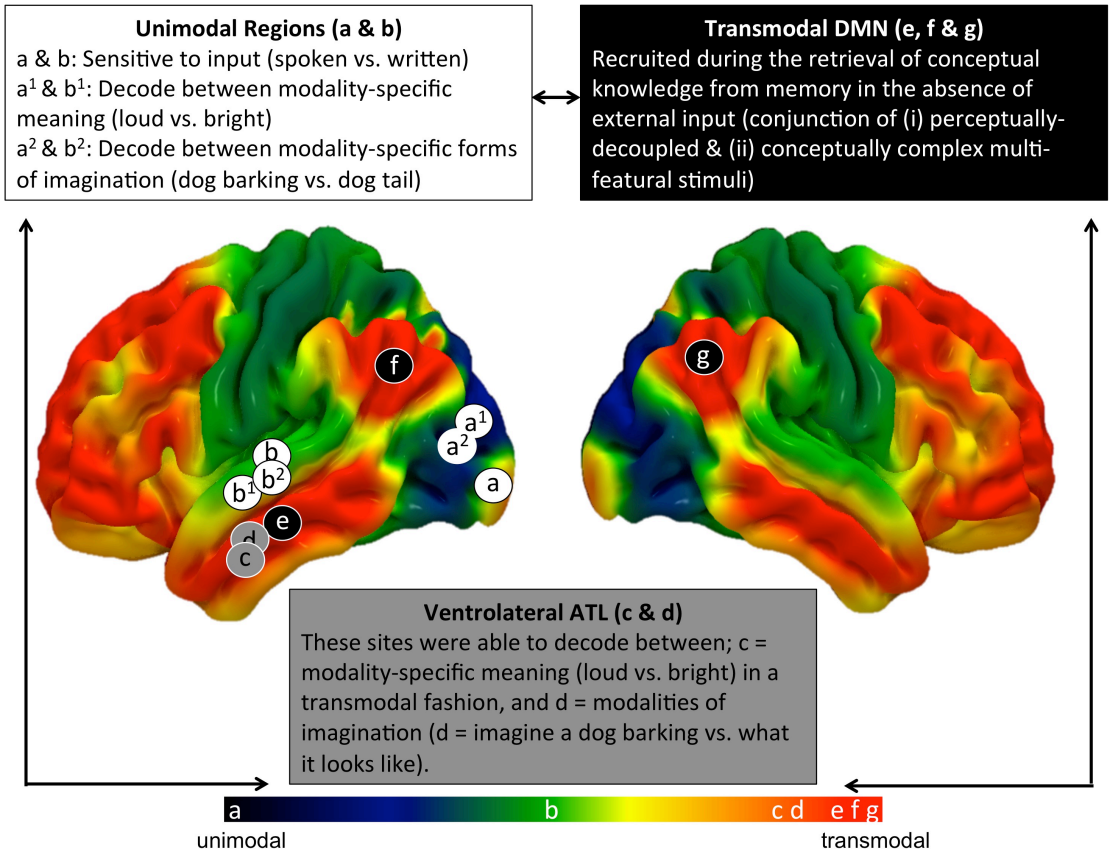

Figure 6.3. Schematic illustration of key findings projected on to the principal gradient (Margulies et al. 2016).

Finally, in line with the notion of multiple levels of integration, the results outlined in both Chapters 3 and 4 align with other studies that suggest transmodal brain regions, including the DMN and heteromodal cortices, contain traces or "echoes" of neural signals represented in other intrinsic connectivity networks, such as primary sensory networks (Braga et al. 2013; Leech, Braga & Sharp, 2012). For instance, this thesis shows that core regions of the DMN, including posterior and anterior portions of the temporal lobe, hippocampus and anterior cingulate cortex, were recruited equally by both auditory and visual imagery in the univariate analysis (Chapter 4), however the MVPA analysis could decode between imaging what a dog sounds like versus imagining what a dog looks like, therefore different patterns must be generated depending on the sensory features that need to be brought to mind. Notably, although portions of

the DMN were able to decode between auditory and visual meaning, these regions did not capture the difference between these conditions and complex conceptual contexts – perhaps because the context condition involved both auditory and visual features (e.g., thinking about the dog races may envisage the visual properties of a greyhound and a race track as well as the acoustic properties of crowds cheering and dogs barking). Compatible with the "echoes" of integration proposal (Braga et al. 2013; Leech et al. 2012) one possible explanation is that the difference between imagining a dog's sounds and visual features might be decodable within DMN because these conditions have very different "echoes" from sensory cortex; however the context condition might involve overlapping sensory inputs with both the visual and auditory imagery conditions. Moreover, this interpretation is consistent with the findings from Chapter 3, where ventrolateral ATL contained transmodal patterns of activity relating to sensory semantic concepts (auditory words, such a *loud*, could be distinguished from visual word concepts, such as *bright*). Taken together, these findings align with the view that transmodal regions receive both types of information (i.e., about sound *and* visual features) in order to permit integration of information effectively and allow for memory retrieval in the absence of sensory input.

***Take home message*:** This thesis presents a complex network of brain areas that are often both modality-invariant (Chapter 3) and sensitive to perceptually decoupled states (Chapter 4 and 5). Combining the data in this thesis with the literature suggests that the DMN shows specialization for the most abstract forms of cognition. Regions allied with the DMN, such as ventrolateral ATL, are necessary for processing heteromodal semantic concepts, whereas core portions of the DMN, such as lateral MTG and angular gyrus, are critical for retrieving conceptually complex representations from memory. Furthermore, this thesis provides tentative support for multiple levels of integration. Following Margulies and colleagues' (2016) principal gradient account, the data highlight that both perceptually-coupled (spoken and written words) and perceptually-decoupled (imagine what something looks like and what it sounds like) semantics engage extensive portions of unimodal cortex, whilst heteromodal representations were identified in regions further up the gradient, such as ventrolateral ATL, and finally regions at the furthest end of the gradient, such as lateral MTG and angular gyrus, were implicated in the most abstract cognitive

processes (i.e., retrieving complex semantic concepts from memory). The DMN therefore plays a critical role in semantic processing as it is situated at the top of a representational hierarchy allowing these regions to integrate information across systems: this also explains why the neural activity within the DMN contains echoes of the information represented in other regions of cortex (Braga et al. 2013; Leech, Braga & Sharp, 2012).

6.4. Limitations and Future Directions

This thesis provides evidence for a distributed network of brain regions contributing to both modality-invariant and perceptually decoupled semantic retrieval. Naturally, there are questions that remain to be answered in view of the findings presented here. An important first step might be to replicate the findings in Chapter 3, using a wider variety of modalities. For example, future studies could interrogate whether modality-invariant representations exist in ventrolateral ATL by comparing concepts across multiple input types (e.g., pictures, sounds, words, actions, smells), to test whether representations in ventrolateral ATL are invariant to all modalities of testing. Several recent studies have adopted such paradigms to investigate whether different input types necessitate common representations (e.g., Kumar, Federmeier, Fei-Fei & Beck, 2017; Man, Damasio, Meyer & Kaplan, 2015; Simanova et al. 2014), although few have reported ventrolateral ATL as a site for transmodal representations. One explanation for the lack of consistency in the results and recent literature is that some studies opt for a regions-of-interest analysis that neglects anterior temporal lobe (e.g., Man et al. 2015). Moreover, failure to find significant cross-modal decoding in the ATL may reflect the fact that this region is vulnerable to magnetic susceptibility artefacts causing signal loss and distortion. For instance, unlike the experiments outlined in the current thesis, Kumar et al. (2017) acknowledge that they did not optimize their scanning protocol for the ATLs. This resulted in lower tSNR within ATL regions (~40 compared to > 70 in the current thesis), and thus it is difficult to draw conclusions about failures to find decoding. Future research could therefore interrogate whether

modality-invariant representations exist in ventrolateral ATL by comparing concepts across multiple input types (e.g., pictures, sounds, words, actions, smells), whilst also optimizing their scanning protocol for ATL, such as aligning their slices with the temporal lobe, using a short echo time or adopting a dual-echo sequence (Halai et al. 2014). Also, the data presented in this thesis go some way in clarifying the role of ventrolateral ATL in processing concrete concepts in a transmodal fashion. However, concrete concepts are definable by their combination of features (e.g., a dog has four legs, fur and barks) and so the results presented here are not generalizable to abstracts concepts such as 'truth', which are not definable by their combination of features. Future studies may therefore wish to clarify the degree to which this region represents abstract concepts.

There are also a number of limitations that should be borne in mind when considering the results of this thesis in regards to sensory cortex. First, the experimental designs have not permitted us to directly investigate directionality between unimodal sensory and transmodal brain regions, and therefore cannot provide direct evidence for top-down activation hypotheses. Future research could employ dynamic causal modelling to investigate directionality (Stephan et al. 2010). Although this method has been questioned regarding whether robust statements of directionality can be made (see Daunizaeu, David & Stephan, 2011). Alternatively, similar research questions could be investigated using magnetoencephalography (MEG) to resolve the timing of recruitment of brain regions, which can help infer directionality (e.g., van Diessen et al. 2015). Therefore, this thesis can only conclude that the findings are consistent with a top-down activation account. Second, the nature of the designs in both Chapter 3 and 4 preludes the ability to directly compare perception and meaning of the *same* concept. In Chapter 3 participants retrieved the meaning of modality-specific words, such as 'loud', but they did not perceive something loud. Moreover, in Chapter 4 participants imagined modality-specific features of concepts, such as what a dog looks likes, but they were not presented with an image of a dog. In light of the findings, it can only be concluded that the general mechanism for thinking about modality-specific meaning (thinking about sound-features) and perception (hearing something) is not equitable in sensory cortex. A wealth of literature has investigated the relationship between imagination and

perception in sensory cortex (Albers et al. 2013; Lee, Kravitz & Baker, 2012; Reddy et al. 2010; Vetter et al. 2014), however the simultaneous investigation of multiple forms of imagination and perception (e.g., visual, auditory, audiovisual, emotional) in *both* unimodal and transmodal brain regions is still lacking.

Finally, further elucidating the role of the DMN in semantic cognition would be a beneficial avenue for investigation, given that this network has been implicated in a variety of tasks that focus cognition on previously encoded knowledge, as opposed to information in the external environment, yet a unifying account for the role of the DMN is still lacking. For example, the current data has shown that (i) the DMN is recruited during perceptually-decoupled retrieval and thus the findings are consistent with previous work suggesting that this decoupling process is necessary for the generation of precise representations from memory (e.g., Davey, 2015) and (ii) is activated during both perceptually-driven and perceptually-decoupled retrieval of semantic concepts (Chapter 5); but it remains unclear whether the same network is recruited when information available in the environment is coherent with the representations retrieved from memory. Future research could therefore investigate whether this network is involved in retrieval of semantic concepts from memory when coherent representations are available in the external world. For instance, in addition to 0-back and 1-back conditions outlined in Chapter 5, future work could include a recognition condition where during probe trials participants are shown two concepts (similar to 0-back probes) and are required to retrieve the concepts they were shown on the previous trial (similar to the 1-back probes) and decide whether the concepts on screen match their retrieved representation from memory. Investigations along this line would be important as it would clarify whether the DMN is engaged during semantic retrieval in both perceptually decoupled and coherent states, an issue of contention in the current literature (e.g., Huijbers, Pennartz, Cabeza & Daselaar, 2011).

Furthermore, one potential limitation of the experimental investigation of the DMN is the comparison of semantic and colour decisions in Chapter 5. This contrast allowed us to demonstrate a neural pattern associating conceptual processing with stimulus independency, however it is too crude a manipulation to determine which aspects of the semantic judgments gave rise to this response in the DMN; for example, is it the richness of concepts such as Labrador, their heteromodal nature or the fact

that they are acquired over a lifetime, which dissociates them from colours? Future studies could probe different features of retrieval, such as whether the target is a concrete or abstract concept, whether it has to be identified at a specific or superordinate level, and whether there are differences according to the modality of the representation being probed. Moreover, the imagination task in Chapter 4 only investigated task-driven mental imagery. As the majority of literature regarding the DMN focuses on spontaneous mind wandering states (Addis et al. 2007; Addis et al. 2009; Binder, Desai, Graves & Conant, 2009; Buckner & Carroll, 2007; Christoff, Gordon, Smallwood, Smith & Schooler, 2009; Hassabis & Maguire, 2007; Mason et al. 2007; Rugg & Vilberg, 2013; Schacter & Addis, 2007), it is unclear whether this network is recruited in a similar or distinctive manner during spontaneously generated semantic retrieval.

6.5. Conclusions

This thesis sought to investigate the neural basis of semantic retrieval within unimodal and heteromodal brain networks, whilst manipulating the availability of information in the environment, by using convergent methods of task-based fMRI, machine learning and functional connectivity analyses. Much of the empirical work in this thesis takes inspiration from modern accounts of transmodal brain regions (e.g., Lambon Ralph et al. 2017; Marguiles et al. 2016), which suggest the ATL and default mode network support both abstraction and decoupling, through the particular interactions they have with unimodal cortices. The data suggests a gradient of processing from superior to ventral ATL, such that superior portions preferentially process information about sensory modality, and ventrolateral portions process abstract modality-invariant semantic representations. Moreover, ventrolateral ATL was also necessary for decoupled semantic processing during imagination. In addition, this thesis found comparable networks recruited for both conceptual processing and perceptually-decoupled retrieval that corresponded to the broader DMN. Further interrogation of these sites, confirmed lateral MTG and bilateral angular gyrus were pivotal in the combination of conceptual retrieval from memory. Taken together this data suggests

that there may be multiple levels of integration occurring across heteromodal brain regions; such that those situated farthest from sensory input systems in both functional and connectivity space are required for processing the most abstract forms of cognition.

Appendices

A.1. Supplementary Figures

A.1.1. Chapter 5

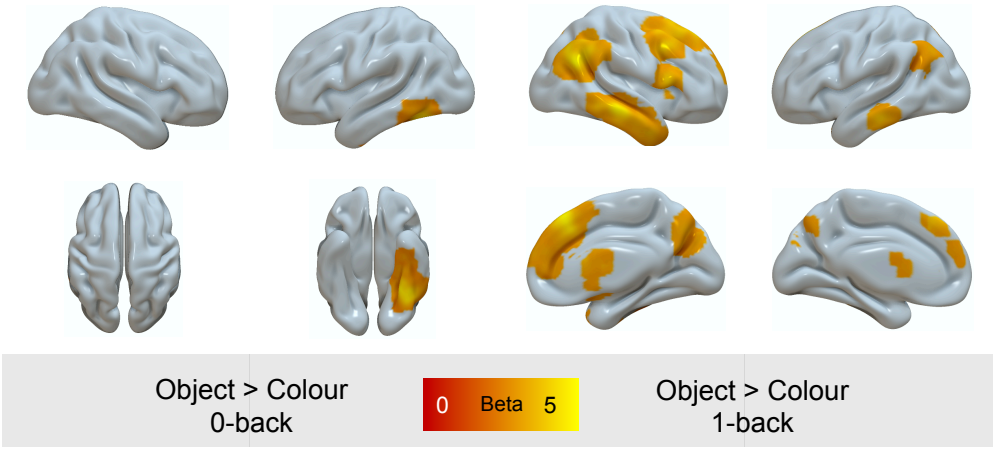

Figure A.1.1. *Comparison of complex memory representation in the presence or absence of relevant perceptual input.* Spatial maps were cluster corrected at Z = 3.1 FWE.

A.2. Supplementary Tables

A.2.1 Chapter 4

Table A.2.1. List of stimuli

	Sound	Visual	Context
Dog	"Sound Dog"	Visual Dog"	"Fast Dog"
			"Dangerous Dog"
			"Old Dog"
			"New Dog"
			"Muddy Dog"
			"Clean Dog"
			"Abandoned Dog"
			"Family Dog"
Car	"Sound Car"	Visual Car"	"Fast Car"
			"Dangerous Car"
			"Old Car"
			"New Car"
			"Muddy Car"
			"Clean Car"
			"Abandoned Car"
			"Family Car"

Footnote: Prompt for each experimental conditions depicted in " ".

Table A.2.2. Coordinates of peak clusters in the resting-state connectivity analyses.

Seed Region	Cluster	Cluster Extent	Z-score	x	Y	z
Context seed	*Increased Correlation*					
	L. Lateral occipital cortex, inferior division	15566	16.4	-50	-64	0
	L. Superior frontal gyrus	566	8.18	-22	-8	54
	R. Planum polare	256	5.45	42	-10	-8
	Reduced Correlation					
	R. Lingual gyrus	6653	7.27	4	-88	14
	R. Anterior cingulate gyrus	5584	7.14	6	26	30
	R. Insular Cortex	2324	6.46	38	14	-10
	L. Postcentral Gyrus	340	4.75	-60	-6	14
	L. Frontal Pole	296	4.43	-36	50	12
	R. Lateral occipital pole, superior division	265	4.8	48	-64	48
Visual seed	*Increased Correlation*					
	L. Lateral occipital cortex, inferior division	7797	15.3	-48	-68	0
	R. Lateral occipital cortex, inferior division	6793	10.9	50	-64	2
	L. Hippocampus	346	5.29	-20	-10	-20
	L. Superior Frontal gyrus	342	7.47	-22	-8	54
	Reduced Correlation					
	R. Lingual gyrus	6688	7.35	4	-70	-4
	R. Insular cortex	2463	6.31	40	12	-6
	R. Paracingulate gyrus	2369	6.85	10	22	34
	R. Frontal pole	2270	6.17	38	40	18
	L. Insular cortex	856	5.42	-36	4	2
	L. Frontal pole	388	5.25	-34	50	8
	R. Posterior cingulate gyrus	354	4,59	2	-32	26

Sound seed	*Increased Correlation*					
	L. Superior temporal gyrus	17702	15.8	-46	-10	-6
	R. Intracalcarine cortex	614	5.45	20	-62	10
	L. Lingual gyrus	564	5.27	-16	-51	0
	R. Anterior cingulate gyrus	511	4.54	6	-14	42
	Reduced Correlation					
	R. Thalamus	1961	6.28	16	-14	10
	L. Superior frontal gyrus	1685	4.98	-20	10	62
	R. Cerebellum	1445	5.58	36	-58	-48
	L. Cerebellum	1187	5.6	-36	-50	-48
	L. Lateral occipital cortex, superior division	1110	5.16	-26	-72	30
	R. Lateral occipital cortex, superior division	670	5.86	26	-78	34
	R. Superior frontal gyrus	571	5.68	26	-4	52
	L. Lateral occipital cortex, inferior division	364	4.38	-48	-78	-12
	L. Frontal pole	291	5.07	-26	54	2
Thalamus	*Increased Correlation*					
	L. Thalamus	22269	17.1	-12	-26	2
	L. Lateral occipital cortex, superior division	270	5.17	-42	-72	24
	Reduced Correlation					
	L. Cerebellum	19581	7.66	-40	-74	-32
	L. Frontal pole	786	5.58	-26	54	20
	L. Planum polare	258	6.29	-44	-10	-12
SMG	*Increased Correlation*					
	L. Supramarginal gyrus, posterior division	9745	15.1	-60	-42	16
	R. Planum temporale	7485	8.64	52	-32	18
	L. Cingulate gyrus, anterior division	4128	7.26	-6	-12	36

	L. Precentral gyrus	330	4.96	-46	-8	44
	R. Cerebellum	289	5.43	26	-72	-56
	Reduced Correlation					
	R. Lateral occipital cortex, superior division	6353	7.45	26	-66	52
	L. Lateral occipital cortex, superior division	3955	6.63	-28	-60	46
	R. Middle frontal gyrus	1529	6.19	38	8	60
	L. Superior frontal gyrus	552	5.95	-26	18	58
	L. Cerebellum	245	5.23	-44	-68	-46
aMTG	*Increased Correlation*					
	L. Middle temporal gyrus, anterior division	10430	15.3	-56	-6	-18
	R. Middle temporal gyrus, posterior division	7048	10.2	50	-12	-16
	L. Posterior cingulate gyrus	2696	7.34	-8	-54	32
	L. Superior frontal gyrus	1606	7.69	-8	52	32
	L. Frontal pole	821	5.94	-6	56	-14
	Reduced Correlation					
	R. Frontal pole	3034	6.85	46	46	12
	L. Frontal pole	1397	6.56	-46	42	16
	R. Angular gyrus	1178	6.25	42	-52	50
	L. Supramarginal gyrus, posterior division	1158	6.90	-50	-42	44
	L. Cerebellum	1108	6.15	-32	-70	-34
	R. Paracingulate gyrus	781	6.91	4	20	42
	R. Superior frontal gyrus	734	5.47	20	16	56
	R. Cerebellum	648	5.69	40	-56	-54
	L. Superior frontal gyrus	490	5.38	-24	2	56
	L. Lingual gyrus	337	4.54	-2	-82	-24
	Thalamus	246	4.64	0	-4	2

	L. Precuneous	210	4.87	-14	-74	42
aPG	*Increased Correlation*					
	L. Parahippocampal gyrus, anterior division/temporal fusiform cortex	15370	15.6	-36	-16	-18
	L. Thalamus	207	4.88	-2	-14	6
	Reduced Correlation					
	R. Middle frontal gyrus	7768	7.09	34	16	50
	R. Lateral occipital cortex, superior division	2232	6.83	46	-62	30
	Intracalcarine cortex	2115	4.71	12	-82	4
	L. Middle frontal gyrus	1893	5.90	-34	2	50
	L. Angular gyrus	1016	5.50	-54	-58	36
	L. Thalamus	659	5.79	-8	-14	-2
ACC	*Increased Correlation*					
	L. Cingulate gyrus, anterior division	28384	15.4	-4	34	-2
	R. Lateral occipital cortex, superior division	315	5.56	52	-68	20
	L. Middle frontal gyrus	272	6.21	-24	32	34
	Reduced Correlation					
	R. Cerebellum	7277	7.76	12	-80	-34
	L. Inferior frontal gyrus, pars opercularis	3364	6.43	-54	14	20
	R. Inferior frontal gyrus, pars opercularis	2065	5.47	52	16	18
	L. Lateral occipital cortex, superior division	1782	6.82	-30	-64	40
	R. Lateral occipital cortex, superior division	750	4.93	36	-66	46
	L. Paracingulate gyrus	468	4.61	-4	28	44

Footnote: The table shows peak clusters in the resting-state connectivity analysis from eight seed regions. Three “unimodal” regions; context seed [-48 60 0], visual seed [-48 -70 -2] and sound seed [52 -8 -10]. Results are thresholded at p<.01 (cluster corrected). Five “heteromodal” regions; Thalamus seed [-48 -60 0], supramarginal gyrus (SMG) seed [-48 -70 -2], anterior middle temporal gyrus (aMTG) seed [-56 -6 -18], anterior

parahippocampal gyrus (aPG) seed [-36 -18 -18] and anterior cingulate cortex (aCC) seed [-4 34-2]. L=left, R=right.

References

Aertsen, A. M., Gerstein, G. L., Habib, M. K., & Palm, G. (1989). Dynamics of neuronal firing correlation: modulation of" effective connectivity". *Journal of neurophysiology*, *61*(5), 900-917.

Abraham, A., & Bubic, A. (2015). Semantic memory as the root of imagination. Frontiers in psychology, 6.

Addis, D. R., Pan, L., Vu, M. A., Laiser, N., & Schacter, D. L. (2009). Constructive episodic simulation of the future and the past: Distinct subsystems of a core brain network mediate imagining and remembering. Neuropsychologia, 47(11), 2222-2238.

Addis, D. R., Wong, A. T., & Schacter, D. L. (2007). Remembering the past and imagining the future: common and distinct neural substrates during event construction and elaboration. Neuropsychologia, 45(7), 1363-1377.

Albers, A. M., Kok, P., Toni, I., Dijkerman, H. C., & de Lange, F. P. (2013). Shared representations for working memory and mental imagery in early visual cortex. Current Biology, 23(15), 1427-1431.

Alderson-Day, B., & Fernyhough, C. (2015). Inner speech: development, cognitive functions, phenomenology, and neurobiology. Psychological bulletin, 141(5), 931.

Allport, D. A. (1985). Distributed memory, modular subsystems and dysphasia. Current perspectives in dysphasia, 32-60.

Amedi, A., Malach, R., & Pascual-Leone, A. (2005). Negative BOLD differentiates visual imagery and perception. Neuron, 48(5), 859-872.

Amunts, K., Malikovic, A., Mholberg, H., Schormann, T., & Zilles, K. (2000). Brodmann's Area 17 and 18 brought into stereostaxic space – where and how variable? *NeuroImage, 11,* 68-84.

Anderson, J. R. (1978). Arguments concerning representations for mental imagery. Psychological Review, 85(4), 249.

Anderson, M. C., Ochsner, K. N., Kuhl, B., Cooper, J., Robertson, E., Gabrieli, S. W., . . . Gabrieli, J. D. (2004). Neural systems underlying the suppression of unwanted memories. Science, 303(5655), 232-235. doi:10.1126/science.1089504

Andrews-Hanna, J. R. (2012). The brain's default network and its adaptive role in internal mentation. The Neuroscientist, 18(3), 251-270.

Andrews-Hanna, J. R., Reidler, J. S., Sepulcre, J., Poulin, R., & Buckner, R. L. (2010). Functional-anatomic fractionation of the brain's default network. Neuron, 65(4), 550-562.

Andrews-Hanna, J. R., Smallwood, J., & Spreng, R. N. (2014). The default network and self-generated thought: component processes, dynamic control, and clinical relevance. *Annals of the New York Academy of Sciences, 316,* 29-52.

Andrews-Hanna, J. R., Snyder, A. Z., Vincent, J. L., Lustig, C., Head, D., Raichle, M. E., & Buckner, R. L. (2007). Disruption of large-scale brain systems in advanced aging. *Neuron* 56, 924–935.

Antrobus, J. S., Singer, J. L., & Greenberg, S. (1966). Studies in the stream of consciousness: experimental enhancement and suppression of spontaneous cognitive processes. Perceptual and Motor Skills, 23, 399- 417.

Badre, D., Poldrack, R. A., Paré-Blagoev, E. J., Insler, R. Z., & Wagner, A. D. (2005). Dissociable controlled retrieval and generalized selection mechanisms in ventrolateral prefrontal cortex. Neuron, 47(6), 907-918.

Bajada, C. J., Haroon, H. A., Azadbakht, H., Parker, G. J., Ralph, M. A. L., & Cloutman, L. L. (2016). The tract terminations in the temporal lobe: Their location and associated functions. *Cortex*.

Bajada, C. J., Jackson, R. L., Haroon, H. A., Azadbakht, H., Parker, G. J., Lambon Ralph, M. A., & Cloutman, L. L. (2017). A graded tractographic parcellation of the temporal lobe. NeuroImage.

Baker, D. H., Karapanagiotidis, T., Coggan, D. D., Wailes-Newson, K., & Smallwood, J. (2015). Brain networks underlying bistable perception. *NeuroImage*, *119*, 229-234.

Baron, S. G., & Osherson, D. (2011). Evidence for conceptual combination in the left anterior temporal lobe. Neuroimage, 55(4), 1847-1852.

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioural and Brain Sciences, 22*, 577-660.

Barsalou, L. W. (2008). Grounded Cognition. *The Annual Review of Psychology, 59,* 617-645.

Barsalou, L. W., Simmons, W. K., Barbey, A. K., & Wilson, C. D. (2003). Grounding conceptual knowledge in modality-specific systems. Trends in cognitive sciences, 7(2), 84-91.

Baylis, G. C., Rolls, E. T., & Leonard, C. M. (1987). Functional subdivisions of the temporal lobe neocortex. Journal of Neuroscience, 7(2), 330-342.

Beckmann, C. F., DeLuca, M., Devlin, J. T., & Smith, S. M. (2005). Investigations into resting-state connectivity using independent component analysis. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *360*(1457), 1001-1013.

Behzadi, Y., Restom, K., Liau, J., & Liu, T. T. (2007). A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. *Neuroimage*, *37*(1), 90-101.

Benedek, M., Jauk, E., Fink, A., Koschutnig, K., Reishofer, G., Ebner, F., & Neubauer, A. C. (2014). To create or to recall? Neural mechanisms underlying the generation of creative new ideas. NeuroImage, 88, 125-133.

Benedek, M., Schues, T., Beaty, R. E., Jauk, E., Koschutnig, K., Fink, A., & Neubauer, A. (2017). To create or to recall original ideas: Brain processes associated with the imagination of novel object uses. *Cortex.*

Bhattacharyya, P. K., & Lowe, M. J. (2004). Cardiac-induced physiologic noise in tissue is a direct observation of cardiac-induced fluctuations. *Magnetic resonance imaging*, *22*(1), 9-13.

Bianciardi, M., Fukunaga, M., van Gelderen, P., Horovitz, S. G., de Zwart, J. A., & Duyn, J. H. (2009). Modulation of spontaneous fMRI activity in human visual cortex by behavioral state. *Neuroimage*, *45*(1), 160-168.

Binder, J. R., & Desai, R. H. (2011). The neurobiology of semantic memory. Trends in cognitive sciences, 15(11), 527-536.

Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral Cortex*, *19*(12), 2767-2796.

Binder, J. R., Gross, W. L., Allendorfer, J. B., Bonilha, L., Chapin, J., Edwards, J. C., ... & Koenig, K. (2011). Mapping anterior temporal lobe language areas with fMRI: a multicenter normative study. Neuroimage, 54(2), 1465-1475.

Binder, J. R., Gross, W. L., Allendorfer, J. B., Bonilha, L., Chapin, J., Edwards, J C., Grabowski, T. J., Langfitt, J. T., Loring, D. W., Lowe, M. J., Koenig, K., Morgan, P. S., Ojemann, J. G., Rorden, C., Szaflarski, J. P., Tivarus, M. E., & Weaver, K. E. (2011). Mapping anterior temporal lobe language areas with fMRI: A multicenter normative study. NeuroImage, 54, 1465-1475.

Binney, R. J., Embleton, K, V., Jefferies, E., Parker, G. J. M., & Lambon Ralph, M, A. (2010). The ventral and inferolateral aspects of the anterior temporal lobe are crucial in semantic memory: Evidence from a novel direct comparison of distortion-corrected fMRI, rTMS and semantic dementia. *Cerebral Cortex, 20,* 2728-2738.

Binney, R. J., Parker, G. J., & Ralph, M. A. L. (2012). Convergent connectivity and graded specialization in the rostral human temporal lobe as revealed by diffusion-weighted imaging probabilistic tractography. Journal of cognitive neuroscience, 24(10), 1998-2014.

Birn, R. M., Diamond, J. B., Smith, M. A., & Bandettini, P. A. (2006). Separating respiratory-variation-related fluctuations from neuronal-activity-related fluctuations in fMRI. *Neuroimage*, *31*(4), 1536-1548.

Birn, R. M., Murphy, K., & Bandettini, P. A. (2008). The effect of respiration variations on independent component analysis results of resting state functional connectivity. *Hum. Brain Mapp.* 29, 740–750.

Bishop, C. M. (2001). Bishop Pattern Recognition and Machine Learning. Springer, New York.

Biswal, B., Yetkin, F. Z., Haughton, V. M., & Hyde, J. S. (1995). Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. *Magn. Reson. Med.* 34, 537–541.

Blaizot, X., Mansilla, F., Insausti, A. M., Constans, J. M., Salinas-Alaman, A., Pro-Sistiaga, P., ... & Insausti, R. (2010). The human parahippocampal region: I. Temporal pole cytoarchitectonic and MRI correlation. Cerebral Cortex, 20(9), 2198-2212.

Bonner, M. F., & Price, A. R. (2013). Where is the anterior temporal lobe and what does it do?. Journal of Neuroscience, 33(10), 4213-4215.

Bonnici, H. M., Richter, F. R., Yazar, Y., & Simons, J. S. (2016). Multimodal Feature Integration in the Angular Gyrus during Episodic and Semantic Retrieval. J

Neurosci, 36(20), 5462-5471. doi:10.1523/JNEUROSCI.4310-15.2016

Bookheimer, S. (2002). Functional MRI of language: new approaches to understanding the cortical organization of semantic processing. Annual review of neuroscience, 25(1), 151-188.

Booth, J. R., Burman, D. D., Meyer, J. R., Gitelman, D. R., Parrish, T. P., & Mesulam, M. M. (2002). Modality indepdence of word comprehension. Human Brain Mapping, 16(4), 251-261.

Bozeat, S., Ralph, M. A. L., Patterson, K., Garrard, P., & Hodges, J. R. (2000). Non-verbal semantic impairment in semantic dementia. Neuropsychologia, 38(9), 1207-1215.

Brodmann, K. (1909). Vergleichende Lokalisationslehre der Grosshirnrinde in ihren Prinzipien dargestellt auf Grund des Zellenbaues. Barth.

Braga, R. M., Sharp, D. J., Leeson, C., Wise, R. J., & Leech, R. (2013). Echoes of the brain within default mode, association, and heteromodal cortices. Journal of Neuroscience, 33(35), 14031-14039.

Brett, M., Johnsrude, I. S., & Owen, A. M. (2002). The problem of functional localization in the human brain. *Nature reviews neuroscience*, *3*(3), 243-249.

Bright, P., Moss, H., & Tyler, L. K. (2004). Unitary vs multiple semantics: PET studies of word and picture processing. *Brain and language*, *89*(3), 417-432.

Bright, M. G., & Murphy, K. (2013). Removing motion and physiological artifacts from intrinsic BOLD fluctuations using short echo data. *Neuroimage*, *64*, 526-537.

Brosch, J. R., Talavage, T. M., Ulmer, J. L., & Nyenhuis, J. A. (2002). Simulation of human respiration in fMRI with a mechanical model. *IEEE Transactions on Biomedical Engineering*, *49*(7), 700-707.

Bruffaerts, R., Dupont, P., Peeters, R., De Deyne, S., Storms, G., & Vandenberghe, R. (2013). Similarity of fMRI activity patterns in left perirhinal cortex reflects semantic similarity between words. Journal of Neuroscience, 33(47), 18597-18607.

Buckner, R. L., Andrews-Hanna, J. R., & Schacter, D. L. (2008). The brain's default network. *Annals of the New York Academy of Sciences*, *1124*(1), 1-38.

Buckner, R. L., & Carroll, D. C. (2007). Self-projection and the brain. Trends in cognitive sciences, 11(2), 49-57.

Buckner, R. L., & Krienen, F. M. (2013). The evolution of distributed association networks in the human brain. Trends in Cognitive Sciences, 17(12), 648- 665.

Buckner, R. L., Sepulcre, J., Talukdar, T., Krienen, F. M., Liu, H., Hedden, T., ... & Johnson, K. A. (2009). Cortical hubs revealed by intrinsic functional connectivity: mapping, assessment of stability, and relation to Alzheimer's disease. Journal of Neuroscience, 29(6), 1860-1873.

Buckner, R. L., & Vincent, J. L. (2007). Unrest at rest: default activity and spontaneous network correlations. *Neuroimage*, *37*(4), 1091-1096.

Bullmore, E., & Sporns, O. (2009). Complex brain networks: graph theoretical analysis of structural and functional systems. Nature reviews. Neuroscience, 10(3), 186.

Bunzeck, N., Wuestenberg, T., Lutz, K., Heinze, H. J., & Jancke, L. (2005). Scanning silence: mental imagery of complex sounds. Neuroimage, 26(4), 1119-1127.

Butler, C. R., Brambati, S. M., Miller, B. L., & Gorno-Tempini, M. L. (2009). The neural correlates of verbal and non-verbal semantic processing deficits in neurodegenerative disease. Cognitive and behavioral neurology: official journal of the Society for Behavioral and Cognitive Neurology, 22(2), 73.

Cabeza, R., Mazuz, Y. S., Stokes, J., Kragel, J. E., Woldorff, M. G., Ciaramelli, E., . . . Moscovitch, M. (2011). Overlapping parietal activity in memory and perception: evidence for the attention to memory model. J Cogn Neurosci, 23(11), 3209-3217. doi:10.1162/jocn_a_00065.

Caramazza, A., Hillis, A. E., Rapp, B. C., & Romani, C. (1990). The multiple semantics hypothesis: Multiple confusions?. Cognitive neuropsychology, 7(3), 161-189.

Cerf-ducastel, B., & Murphy, C. (2004). Validation of a stimulation protocol suited to the investigation of odor-taste interactions with fMRI. Physiology & Behaviour, 81(3), 389.396.

Chan, A. M., Halgren, E., Marinkovic, K., & Cash, S. S. (2011). Decoding word and category-specific spatiotemporal representations from MEG and EEG. *Neuroimage*, *54*(4), 3028-3039.

Chang, C. C., & Lin, C. J. (2011). LIBSVM: a library for support vector machines. *ACM Transactions on Intelligent Systems and Technology (TIST)*, *2*(3), 27.

Chao, L. L., & Martin, A. (1999). Cortical regions associated with perceiving, naming, and knowing about colors. Journal of Cognitive Neuroscience, 11(1), 25-35.

Chao, L. L., & Martin, A. (2000). Representation of manipulable man-made objects in the dorsal stream. Neuroimage, 12(4), 478-484.

Chatterjee, A. (2010). Disembodying cognition. Language and cognition, 2(1), 79- 116.

Chen, W., Kato, T., Zhu, X. H., Ogawa, S., Tank, D. W., & Ugurbil, K. (1998). Human primary visual cortex and lateral geniculate nucleus activation during visual imagery. Neuroreport, 9(16), 3669-3674.

Chen, Y., Shimotake, A., Matsumoto, R., Kunieda, T., Kikuchi, T., Miyamoto, S., ... & Ralph, M. L. (2016). The 'when'and 'where'of semantic coding in the anterior temporal lobe: Temporal representational similarity analysis of electrocorticogram data. Cortex, 79, 1-13.

Christoff, K., Gordon, A. M., Smallwood, J., Smith, R., & Schooler, J. W. (2009). Experience sampling during fMRI reveals default network and executive system contributions to mind wandering. Proceedings of the National Academy of Sciences, 106(21), 8719-8724.

Coccia, M., Bartolini, M., Luzzi, S., Provinciali, L., & Lambon Ralph, M. A. (2004). Semantic memory is an amodal, dynamic system: Evidence from the interaction of naming and object use in semantic dementia. Cognitive Neuropsychology, 21(5), 513-527.

Coifman, R. R., Lafon, S., Lee, A. B., Maggioni, M., Nadler, B., Warner, F., & Zucker, S. W. (2005). Geometric diffusions as a tool for harmonic analysis and structure definition of data: Diffusion maps. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(21), 7426-7431.

Cohen, L., Jobert, A., Le Bihan, D., & Dehaene, S. (2004). Distinct unimodal and multimodal regions for word processing in the left temporal cortex. NeuroImage, 23(4), 1256-1270.

Cole, D. M., Smith, S. M., & Beckmann, C. F. (2010). Advances and pitfalls in the analysis and interpretation of resting-state FMRI data. *Frontiers in systems neuroscience*, *4*, 8.

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. Nature reviews. Neuroscience, 3(3), 201.

Cordes, D., Haughton, V. M., Arfanakis, K., Carew, J. D., Turski, P. A., Moritz, C. H., ... & Meyerand, M. E. (2001). Frequencies contributing to functional connectivity

in the cerebral cortex in “resting-state” data. *American Journal of Neuroradiology*, *22*(7), 1326-1333.

Correia, J., Formisano, E., Valente, G., Hausfeld, L., Jansma, B., & Bonte, M. (2014). Brain-based translation: fMRI decoding of spoken words in bilinguals reveals language-independent semantic representations in anterior temporal lobe. *The Journal of Neuroscience*, *34*(1), 332-338.

Coutanche, M. N. (2013). Distinguishing multi-voxel patterns and mean activation: why, how, and what does it tell us?. *Cognitive, Affective, & Behavioral Neuroscience*, *13*(3), 667-673.

Coutanche, M. N., & Thompson-Schill, S. L. (2014). Creating concepts from converging features in human cortex. *Cerebral Cortex*, *25*(9), 2584-2593.

Cox, D. D., & Savoy, R. L. (2003). Functional magnetic resonance imaging (fMRI)“brain reading”: detecting and classifying distributed patterns of fMRI activity in human visual cortex. *Neuroimage*, *19*(2), 261-270.

Cree, G. S., & McRae, K. (2003). Analyzing the factors underlying the structure and computation of the meaning of chipmunk, cherry, chisel, cheese, and cello (and many other such concrete nouns). Journal of Experimental Psychology: General, 132(2), 163.

Dagli, M. S., Ingeholm, J. E., & Haxby, J. V. (1999). Localization of cardiac-induced signal change in fMRI. *Neuroimage*, *9*(4), 407-415.

Damoiseaux, J. S., & Greicius, M. D. (2009). Greater than the sum of its parts: a review of studies combining structural connectivity and resting-state functional connectivity. *Brain Structure and Function*, *213*(6), 525-533.

Damoiseaux, J. S., Rombouts, S. A. R. B., Barkhof, F., Scheltens, P., Stam, C. J., Smith, S. M., & Beckmann, C. F. (2006). Consistent resting-state networks across healthy subjects. *Proceedings of the national academy of sciences*, *103*(37), 13848-13853.

Damasio, A. R. (1989). The Brain Binds Entities and Events by Multiregional Activation from Convergence Zones. Neural Computation, 1(1), 123-132.

Damasio, A. R. (2008). The brain binds entities and events by multiregional activation from convergence zones. Brain, 1(1).

Daselaar, S. M., Porat, Y., Huijbers, W., & Pennartz, C. M. (2010). Modality-specific and modality-independent components of the human imagery system. Neuroimage, 52(2), 677-685.

Daunizeau, J., David, O., & Stephan, K. E. (2011). Dynamic causal modelling: a critical review of the biophysical and statistical foundations. Neuroimage, 58(2), 312-322.

Davey, J., Thompson, H. E., Hallam, G., Karapanagiotidis, T., Murphy, C., De Caso, I., ... & Jefferies, E. (2016). Exploring the role of the posterior middle temporal gyrus in semantic cognition: Integration of anterior temporal lobe with executive processes. NeuroImage.

Davey, J., Cornelissen, P. L., Thompson, H. E., Sonkusare, S., Hallam, G., Smallwood, J., & Jefferies, E. (2015). Automatic and Controlled Semantic Retrieval: TMS Reveals Distinct Contributions of Posterior Middle Temporal Gyrus and Angular Gyrus. J Neurosci, 35(46), 15230-15239. doi:10.1523/JNEUROSCI.4705-14.2015

Davis, T., LaRocque, K. F., Mumford, J. A., Norman, K. A., Wagner, A. D., & Poldrack, R. A. (2014). What do differences between multi-voxel and univariate analysis mean? How subject-, voxel-, and trial-level variance impact fMRI analysis. *NeuroImage*, *97*, 271-283.

Deichmann, R., Gottfried, J. A., Hutton, C., & Turner, R. (2003). Optimized EPI for fMRI studies of the orbitofrontal cortex. *Neuroimage*, *19*(2), 430-441.

Dentico, D., Cheung, B. L., Chang, J. Y., Guokas, J., Boly, M., Tononi, G., & Van Veen, B. (2014). Reversal of cortical information flow during visual imagery as compared to visual perception. Neuroimage, 100, 237-243.

Desjardins, A. E., Kiehl, K. A., & Liddle, P. F. (2001). Removal of confounding effects of global signal in functional MRI analyses. *Neuroimage*, *13*(4), 751-758.

Devereux, B. J., Clarke, A., Marouchos, A., & Tyler, L. K. (2013). Representational similarity analysis reveals commonalities and differences in the semantic processing of words and objects. The Journal of Neuroscience, 33(48), 18906-18916.

Devlin, J. T., Russell, R. P., Davis, M. H., Price, C. J., Wilson, J., Moss, H. E., ... & Tyler, L. K. (2000). Susceptibility-induced loss of signal: comparing PET and fMRI on a semantic task. *Neuroimage*, *11*(6), 589-600.

de Borst, A. W., & de Gelder, B. (2016). fMRI-based multivariate pattern analyses reveal imagery modality and imagery content specific representations in primary somatosensory, motor and auditory cortices. Cerebral Cortex, 1- 15.

de Caso, I., Karapanagiotidis, T., Aggius-Vella, E., Konishi, M., Margulies, D. S., Jefferies, E., & Smallwood, J. (2017). Knowing me, knowing you: Resting-state functional connectivity of ventromedial prefrontal cortex dissociates memory related to self from a familiar other. Brain Cogn, 113, 65-75. doi:10.1016/j.bandc.2017.01.004

de la Porte, J., Herbst, B. M., Hereman, W., & Van Der Walt, S. J. (2008). An introduction to diffusion maps. In *Proceedings of the 19th Symposium of the Pattern Recognition Association of South Africa (PRASA 2008), Cape Town, South Africa* (pp. 15-25).

de Luca, M., Beckmann, C. F., De Stefano, N., Matthews, P. M., and Smith, S. M. (2006). fMRI resting state networks define distinct modes of long-distance interactions in the human brain. *Neuroimage,* 29, 1359–1367.

de Panfilis, C., & Schwarzbauer, C. (2005). Positive or negative blips? The effect of phase encoding scheme on susceptibility-induced signal losses in EPI. *Neuroimage*, *25*(1), 112-121.

de Pasquale, F., Della Penna, S., Snyder, A. Z., Marzetti, L., Pizzella, V., Romani, G. L., & Corbetta, M. (2012). A cortical core for dynamic integration of functional networks in the resting human brain. Neuron, 74(4), 753-764.

Diedrichsen, J., Wiestler, T., & Ejaz, N. (2013). A multivariate method to determine the dimensionality of neural representation from population activity. *Neuroimage*, *76*, 225-235.

Dijkstra, N., Zeidman, P., Ondobaka, S., van Gerven, M. A. J., & Friston, K. (2017). Distinct top-down and bottom-up brain connectivity during visual perception and imagery. Scientific Reports, 7.

Dosenbach, N. U. F., & Petersen, S. E. (2009). Attentional networks.

Driver, J., & Noesselt, T. (2008). Multisensory interplay reveals crossmodal influences on 'sensory-specific' brain regions, neural responses, and judgments. Neuron, 57(1), 11-23.

Duda, R. O., Hart, P. E., & Stork, D. G. (2001). *Pattern Classification*. New York, NY:

John Wiley & Sons.

Duda, R. O., Hart, P. E., & Stork, D. G. (2012). *Pattern classification*. John Wiley & Sons.

Engen, H. G., Kanske, P., & Singer, T. (2017). The neural component-process architecture of endogenously generated emotion. Soc Cogn Affect Neurosci, 12(2), 197-211. doi:10.1093/scan/nsw108

Fair, D. A., Cohen, A. L., Power, J. D., Dosenbach, N. U., Church, J. A., Miezin, F. M., ... & Petersen, S. E. (2009). Functional brain networks develop from a "local to distributed" organization. PLoS computational biology, 5(5), e1000381.

Fairhall, S. L., & Caramazza, A. (2013). Brain regions that represent transmodal conceptual knowledge. *The Journal of Neuroscience*, *33*(25), 10552-10558.

Farah, M. J. (1984). The neurological basis of mental imagery: A componential analysis. Cognition, 18(1), 245-272.

Fernandino, L., Humphries, C. J., Conant, L. L., Seidenberg, M. S., & Binder, J. R. (2016). Heteromodal cortical areas encode sensory-motor features of word meaning. Journal of Neuroscience, 36(38), 9763-9769.

Fink, A., Grabner, R. H., Gebauer, D., Reishofer, G., Koschutnig, K., & Ebner, F. (2010). Enhancing creativity by means of cognitive stimulation: Evidence from an fMRI study. NeuroImage, 52(4), 1687-1695.

Friedman, L., Glover, G. H., & Fbirn Consortium. (2006). Reducing interscanner variability of activation in a multicenter fMRI study: controlling for signal- to-fluctuation-noise-ratio (SFNR) differences. Neuroimage, 33(2), 471-481.

Friston, K. J., Harrison, L., & Penny, W. (2003). Dynamic causal modelling. Neuroimage, 19(4), 1273-1302.

Friston, K., Moran, R., & Seth, A. K. (2013). Analysing connectivity with Granger causality and dynamic causal modelling. Current opinion in neurobiology, 23(2), 172-178.

Fox, M. D., & Raichle, M. E. (2007). Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. *Nature Reviews Neuroscience*, *8*(9), 700-711.

Fox, M. D., Snyder, A. Z., Vincent, J. L., Corbetta, M., Van Essen, D. C., & Raichle, M. E. (2005). The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proc. Natl. Acad. Sci. U.S.A.* 102, 9673–

9678.

Fox, M. D., Snyder, A. Z., Vincent, J. L., Corbetta, M., Van Essen, D. C., & Raichle, M. E. (2005). The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(27), 9673-9678.

Fox, M. D., Zhang, D., Snyder, A. Z., & Raichle, M. E. (2009). The global signal and observed anticorrelated resting state brain networks. *Journal of neurophysiology*, *101*(6), 3270-3283.

Fransson, P. (2005). Spontaneous low-frequency BOLD signal fluctuations: an fMRI investigation of the resting-state default mode of brain function hypothesis. *Hum. Brain Mapp.* 26, 15–29.

Friedman, L., Glover, G. H., & Fbirn Consortium. (2006). Reducing interscanner variability of activation in a multicenter fMRI study: controlling for signal-to-fluctuation-noise-ratio (SFNR) differences. Neuroimage, 33(2), 471-481.

Friston, K. J., Bastos, A., Litvak, V., Stephan, K. E., Fries, P., & Moran, R. J. (2012). DCM for complex-valued data: cross-spectra, coherence and phase-delays. Neuroimage, 59(1), 439-455.

Friston, K. J., Holmes, A. P., Worsley, K.J., Poline, J. P., Frith, C. D., & Frackowiak, R. S. (1994). Statistical parametric maps in functional imaging: a general linear approach. *Human brain mapping. 2*(4):189–210.

Gabrieli, J. D., Brewer, J. B., Desmond, J. E., & Glover, G. H. (1997). Separate neural bases of two fundamental memory processes in the human medial temporal lobe. Science, 276(5310), 264-266.

Gainotti, G. (2011). The organization and dissolution of semantic-conceptual knowledge: is the 'amodal hub'the only plausible model?. Brain and cognition, 75(3), 299-309.

Galton, C. J., Patterson, K., Graham, K., Lambon-Ralph, M. A., Williams, G., Antoun, N., ... & Hodges, J. R. (2001). Differing patterns of temporal atrophy in Alzheimer's disease and semantic dementia. Neurology, 57(2), 216-225.

Ganis, G., & Schendan, H. E. (2008). Visual mental imagery and perception produce opposite adaptation effects on early brain potentials. Neuroimage, 42(4), 1714-1727.

Geranmayeh, F., Leech, R., & Wise, R. J. (2015). Semantic retrieval during overt picture description: left anterior temporal or the parietal lobe?. Neuropsychologia, 76, 125-135.

Geschwind, N. (1965). Disconnexion syndromes in animals and man. Brain, 88(3), 585-585.

Glenberg, A. M., & Kaschak, M. P. (2002). Grounding language in action. Psychonomic bulletin & review, 9(3), 558-565.

Gloor, P. (1997). The temporal lobe and limbic system. Oxford University Press, USA.

Golchert, J., Smallwood, J., Jefferies, E., Seli, P., Huntenburg, J. M., Liem, F., . . . Margulies, D. S. (2017). Individual variation in intentionality in the mind-wandering state is reflected in the integration of the default-mode, fronto-parietal, and limbic networks. Neuroimage, 146, 226-235. doi:10.1016/j.neuroimage.2016.11.025

Goldberg, R. F., Perfetti, C. A., & Schneider, W. (2006). Perceptual knowledge retrieval activates sensory brain regions. Journal of Neuroscience, 26(18), 4917-4921.

Gonzalez, J., Barros-Loscertales, A., Pulvermüller, F., Meseguer, V., Sanjuan, A., Belloch, V., & Avila, C. (2006). Reading cinnamon activates olfactory brain regions. NeuroImage, 32(2), 906-912.

Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in neurosciences*, *15*(1), 20-25.

Greicius, M. D., Krasnow, B., Reiss, A. L., & Menon, V. (2003). Functional connectivity in the resting brain: a network analysis of the default mode hypothesis. *Proc. Natl. Acad. Sci. U.S.A.* 100, 253–258.

Greicius, M. D., Srivastava, G., Reiss, A. L., & Menon, V. (2004). Default-mode network activity distinguishes Alzheimer's disease from healthy aging: evidence from functional MRI. Proceedings of the National Academy of Sciences of the United States of America, 101(13), 4637-4642.

Gusnard, D. A., Akbudak, E., Shulman, G. L., & Raichle, M. E. (2001). Medial prefrontal cortex and self-referential mental activity: relation to a default mode of brain function. Proc Natl Acad Sci U S A, 98(7), 4259-4264. doi:10.1073/pnas.071043098

Guyon, I., & Elisseeff, A. (2003). An introduction to variable and feature selection.

Journal of machine learning research, *3*(Mar), 1157-1182.

Hagmann, P., Cammoun, L., Gigandet, X., Meuli, R., Honey, C. J., Wedeen, V. J., & Sporns, O. (2008). Mapping the structural core of human cerebral cortex. PLoS biology, 6(7), e159.

Halai, A. D., Welbourne, S. R., Embleton, K., & Parkes, L. M. (2014). A comparison of dual gradient-echo and spin-echo fMRI of the inferior temporal lobe. *Human brain mapping*, *35*(8), 4118-4128.

Halpern, A. R. (2001). Cerebral substrates of musical imagery. Annals of the New York Academy of Sciences, 930(1), 179-192.

Halpern, A. R., & Zatorre, R. J. (1999). When that tune runs through your head: a PET investigation of auditory imagery for familiar melodies. Cerebral cortex, 9(7), 697-704.

Halpern, A. R., Zatorre, R. J., Bouffard, M., & Johnson, J. A. (2004). Behavioral and neural correlates of perceived and imagined musical timbre. Neuropsychologia, 42(9), 1281-1292.

Hampson, M., Peterson, B. S., Skudlarski, P., Gatenby, J. C., & Gore, J. C. (2002). Detection of functional connectivity using temporal correlations in MR images. *Human brain mapping*, *15*(4), 247-262.

Hanke, M., Halchenko, Y. O., Sederberg, P. B., Hanson, S. J., Haxby, J. V., & Pollmann, S. (2009). PyMVPA: a python toolbox for multivariate pattern analysis of fMRI data. *Neuroinformatics*, *7*(1), 37-53.

Hanson, S. J., Matsuka, T., & Haxby, J. V. (2004). Combinatorial codes in ventral temporal lobe for object recognition: Haxby (2001) revisited: is there a "face" area?. *Neuroimage*, *23*(1), 156-166.

Hart, J., Anand, R., Zoccoli, S., Maguire, M., Gamino, J., Tillman, G., ... & Kraut, M. A. (2007). Neural substrates of semantic memory. Journal of the International Neuropsychological Society, 13(5), 865-880.

Hassabis, D., & Maguire, E. A. (2007). Deconstructing episodic memory with construction. Trends in cognitive sciences, 11(7), 299-306.

Hauk, O., Johnsrude, I., & Pulvermüller, F. (2004). Somatotopic representation of action words in human motor and premotor cortex. Neuron, 41(2), 301-307.

Hauk, O., & Tschentscher, N. (2013). The body of evidence: what can neuroscience tell us about embodied semantics?. Frontiers in Psychology, 4.

Haynes, J. D., & Rees, G. (2006). Decoding mental states from brain activity in humans. *Nature Reviews Neuroscience*, *7*(7), 523-534.

Haxby, J. V., Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L., & Pietrini, P. (2001). Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science*, *293*(5539), 2425-2430.

Hertz, U., & Amedi, A. (2010). Disentangling unisensory and multisensory components in audiovisual integration using a novel multifrequency fMRI spectral analysis. Neuroimage, 52(2), 617-632.

Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews Neuroscience*, *8*(5), 393-402.

Hodges, J. R., Graham, N., & Patterson, K. (1995). Charting the progression in semantic dementia: Implications for the organisation of semantic memory. Memory, 3(3-4), 463-495.

Hodges, J. R., Patterson, K., Oxbury, S., & Funnell, E. (1992). Semantic dementia: progressive fluent aphasia with temporal lobe atrophy. Brain, 115(6), 1783-1806.

Hoffman, P., Binney, R. J., & Ralph, M. A. L. (2015). Differing contributions of inferior prefrontal and anterior temporal cortex to concrete and abstract conceptual knowledge. Cortex, 63, 250-266.

Holmes, A. P., & Friston, K. J. (1998). Generalisability, random E ects & population inference. Neuroimage, 7, S754.

Huettel, S. A., Song, A. W., & McCarthy, G. (2004). *Functional magnetic resonance imaging* (Vol. 1). Sunderland: Sinauer Associates.

Huettel, S. A., Song, A. W., & McCarthy, G. (2009). *Functional Magnetic Resonance Imaging, 2nd Edition* (Vol. 23). Massachusetts: Sinauer.

Huijbers, W., Pennartz, C. M., Cabeza, R., & Daselaar, S. M. (2009). When learning and remembering compete: a functional MRI study. PLoS biology, 7(1), e1000011.

Huijbers, W., Pennartz, C. M., Cabeza, R., & Daselaar, S. M. (2011). The hippocampus is coupled with the default network during memory retrieval but not during

memory encoding. PloS one, 6(4), e17463.

Humphreys, G.W., & Forde, E.M.E. (2001). Hierarchies, similarity, and interactivity in object recognition: "Category-specific" neuropsychological deficits. Behavioural & Brain Sciences, 24, 453-509.

Humphreys, G. F., Hoffman, P., Visser, M., Binney, R. J., & Ralph, M. A. L. (2015). Establishing task-and modality-dependent dissociations between the semantic and default mode networks. Proceedings of the National Academy of Sciences, 112(25), 7857-7862.

Hunter, M. D., Eickhoff, S. B., Miller, T. W. R., Farrow, T. F. D., Wilkinson, I. D., & Woodruff, P. W. R. (2006). Neural activity in speech-sensitive auditory cortex during silence. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(1), 189-194.

Hurley, R. S., Bonakdarpour, B., Wang, X., & Mesulam, M. M. (2015). Asymmetric Connectivity between the Anterior Temporal Lobe and the Language Network. Journal of cognitive neuroscience.

Hutton, C., Bork, A., Josephs, O., Deichmann, R., Ashburner, J., & Turner, R. (2002). Image distortion correction in fMRI: a quantitative evaluation. *Neuroimage*, *16*(1), 217-240.

Ikeda, M., Patterson, K., Graham, K. S., Ralph, M. L., & Hodges, J. R. (2006). A horse of a different colour: do patients with semantic dementia recognise different versions of the same object as the same?. Neuropsychologia, 44(4), 566-575.

Ishai, A., Ungerleider, L. G., & Haxby, J. V. (2000). Distributed neural systems for the generation of visual images. Neuron, 28(3), 979-990.

Irish, M. (2016). Semantic memory as the essential scaffold for future-oriented mental time travel. *Seeing the Future: Theoretical Perspectives on Future-Oriented Mental Time Travel*, 389-408.

Irish, M. (2016). Elucidating a Core Semantic Network in the Brain—Implications for Disorders of Semantic Cognition. Journal of Neuroscience, 36(23), 6144-6146.

Irish, M., & Piguet, O. (2013). The pivotal role of semantic memory in remembering the past and imagining the future. Frontiers in Behavioral Neuroscience, 7.

Jackson, R. L., Hoffman, P., Pobric, G., & Ralph, M. A. L. (2015). The nature and neural correlates of semantic association versus conceptual similarity. Cerebral

Cortex, 25(11), 4319-4333.

Jackson, R. L., Hoffman, P., Pobric, G., & Lambon Ralph, M. A. (2016). The Semantic Network at Work and Rest: Differential Connectivity of Anterior Temporal Lobe Subregions. The Journal of Neuroscience, 36(5), 1490-1501.

James, W. (1890). The principles of psychology. 2 vols. New York: Henry Holt and Company.

Jefferies, E. (2013). The neural basis of semantic cognition: converging evidence from neuropsychology, neuroimaging and TMS. Cortex, 49(3), 611-625.

Jefferies, E., & Lambon Ralph, M. A. (2006). Semantic impairment in stroke aphasia versus semantic dementia: a case-series comparison. Brain, 129(8), 2132-2147.

Jenkinson, M., Bannister, P., Brady, M., & Smith, S. (2002). Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage*, *17*(2), 825-841.

Jezzard, P., & Clare, S. (1999). Sources of distortion in functional MRI data. Human brain mapping, 8(2-3), 80-85.

Jiang, T., He, Y., Zang, Y., & Weng, X., (2004). Modulation of functional connectivity during the resting state and the motor task. *Hum. Brain Mapp. 22* (1), 63–71.

Jimura, K., & Poldrack, R. A. (2012). Analyses of regional-average activation and multivoxel pattern information tell complementary stories. *Neuropsychologia*, *50*(4), 544-552.

Jobard, G., Vigneau, M., Mazoyer, B., & Tzourio-Mazoyer, N. (2007). Impact of modality and linguistic complexity during reading and listening tasks. NeuroImage, 34(2), 784-800.

Jordan, A. (2002). On discriminative vs. generative classifiers: A comparison of logistic regression and naive bayes. *Advances in neural information processing systems*, *14*, 841.

Jung, J., Cloutman, L. L., Binney, R. J., & Ralph, M. A. L. (2016). The structural connectivity of higher order association cortices reflects human functional brain networks. Cortex.

Kalkstein, J., Checksfield, K., Bollinger, J., & Gazzaley, A. (2011). Diminished top-down control underlies a visual imagery deficit in normal aging. Journal of Neuroscience, 31(44), 15768-15774.

Kamitani, Y., & Tong, F. (2005). Decoding the visual and subjective contents of the human brain. *Nature neuroscience*, *8*(5), 679-685.

Kane, M. J., Brown, L. H., McVay, J. C., Silvia, P. J., Myin-Germeys, I., & Kwapil, T. R. (2007). For whom the mind wanders, and when: An experience-sampling study of working memory and executive control in daily life. Psychological science, 18(7), 614-621.

Kiefer, M., & Pulvermüller, F. (2012). Conceptual representations in mind and brain: Theoretical developments, current evidence and future directions.Cortex, 48(7), 805-825.

Killingsworth, M. A., & Gilbert, D. T. (2010). A wandering mind is an unhappy mind. Science, 330(6006), 932-932.

Kiviniemi, V., Kantola, J. H., Jauhiainen, J., Hyvarinen, A., & Tervonen, O. (2003). Independent component analysis of nondeterministic fMRI signal sources. *Neuroimage,* 19, 253–260.

Kiviniemi, V., Starck, T., Remes, J., Long, X., Nikkinen, J., Haapea, M., Veijola, J., Moilanen, I., Isohanni, M., Zang, Y. F., & Tervonen, O. (2009). Functional segmentation of the brain cortex using high model order group PICA. *Hum. Brain Mapp.* 30, 3865–3886.

Klein, S. B. (2013). Making the case that episodic recollection is attributable to operations occurring at retrieval rather than to content stored in a dedicated subsystem of long-term memory. Frontiers in Behavioral Neuroscience, 7.

Klinger, E., & Cox, W. M. (1987). Dimensions of thought flow in everyday life. Imagination, Cognition and Personality, 7(2), 105-128.

Knauff, M., Kassubek, J., Mulack, T., & Greenlee, M. W. (2000). Cortical activation evoked by visual mental imagery as measured by fMRI. Neuroreport, 11(18), 3957-3962.

Koenig, P., & Grossman, M. (2007). Process and content in semantic memory.

Kohavi, R., & John, G. H. (1997). Wrappers for feature subset selection. *Artificial intelligence*, *97*(1-2), 273-324.

Konishi, M., McLaren, D. G., Engen, H., & Smallwood, J. (2015). Shaped by the Past: The Default Mode Network Supports Cognition that Is Independent of Immediate

Perceptual Input. PLoS One, 10(6), e0132209. doi:10.1371/journal.pone.0132209.

Kosslyn, S. M. (2005). Mental images and the brain. Cognitive Neuropsychology, 22(3-4), 333-347.

Kosslyn, S. M., Ganis, G., & Thompson, W. L. (2001). Neural foundations of imagery. Nature reviews. Neuroscience, 2(9), 635.

Kosslyn, S. M., Pascual-Leone, A., Felician, O., Camposano, S., Keenan, J. P., Ganis, G., & Alpert, N. M. (1999). The role of area 17 in visual imagery: convergent evidence from PET and rTMS. Science, 284(5411), 167-170.

Kosslyn, S. M., & Thompson, W. L. (2003). When is early visual cortex activated during visual mental imagery?. Psychological bulletin, 129(5), 723.

Kosslyn, S. M., Thompson, W. L., & Alpert, N. M. (1997). Neural systems shared by visual imagery and visual perception: A positron emission tomography study. Neuroimage, 6(4), 320-334.

Kraemer, D. J., Macrae, C. N., Green, A. E., & Kelley, W. M. (2005). Musical imagery: sound of silence activates auditory cortex. Nature, 434(7030), 158-158.

Kragel, P. A., Carter, R. M., & Huettel, S. A. (2012). What makes a pattern? Matching decoding methods to data in multivariate pattern analysis. *Frontiers in neuroscience*, *6*.

Krieger-Redwood, K., Jefferies, E., Karapanagiotidis, T., Seymour, R., Nunes, A., Ang, J. W. A., ... & Smallwood, J. (2016). Down but not out in posterior cingulate cortex: Deactivation yet functional coupling with prefrontal cortex during demanding semantic cognition. Neuroimage, 141, 366-377.

Kriegeskorte, N., Goebel, R., & Bandettini, P. (2006). Information-based functional brain mapping. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(10), 3863-3868.

Kriegeskorte, N., Simmons, W. K., Bellgowan, P. S., & Baker, C. I. (2009). Circular analysis in systems neuroscience: the dangers of double dipping. *Nature neuroscience*, *12*(5), 535-540.

Krienen, F. M., & Buckner, R. L. (2009). Segregated fronto-cerebellar circuits revealed by intrinsic functional connectivity. *Cerebral cortex*, *19*(10), 2485-2497.

Krienen, F. M., & Sherwood, C. C. (2016). Gradients of Connectivity in the Cerebral Cortex. *Trends in Cognitive Sciences*.

Ku, S. P., Gretton, A., Macke, J., & Logothetis, N. K. (2008). Comparison of pattern recognition methods in classifying high-resolution BOLD signals obtained at high magnetic field in monkeys. *Magnetic resonance imaging*, *26*(7), 1007-1014.

Kuhl, B. A., Rissman, J., Chun, M. M., & Wagner, A. D. (2011). Fidelity of neural reactivation reveals competition between memories. Proceedings of the National Academy of Sciences, 108(14), 5903-5908.

Kumar, M., Federmeier, K. D., Fei-Fei, L., & Beck, D. M. (2017). Evidence for similar patterns of neural activity elicited by picture-and word-based representations of natural scenes. NeuroImage.

Kundu, P., Brenowitz, N. D., Voon, V., Worbe, Y., Vértes, P. E., Inati, S. J., ... & Bullmore, E. T. (2013). Integrated strategy for improving functional connectivity mapping using multiecho fMRI. *Proceedings of the National Academy of Sciences*, *110*(40), 16187-16192.

LaConte, S., Strother, S., Cherkassky, V., Anderson, J., & Hu, X. (2005). Support vector machines for temporal classification of block design fMRI data. *NeuroImage*, *26*(2), 317-329.

Lambon Ralph, M. A., Jefferies, E., Patterson, K., & Rogers, T. T. (2017). The neural and computational bases of semantic cognition. *Nature Reviews Neuroscience*.

Lambon Ralph, M. A., Lowe, C., & Rogers, T. T. (2007). Neural basis of category-specific semantic deficits for living things: evidence from semantic dementia, HSVE and a neural network model. Brain, 130(4), 1127-1137.

Lambon Ralph, M. A., Sage, K., Jones, R., & Mayberry, E. (2010). Coherent concepts are computed in the anterior temporal lobes. Proceedings of the National Academy of Sciences of the United States of America, 107, 2717-2722.

Lambon Ralph, M. A., & Patterson, K. (2008). Generalization and differentiation in semantic memory. Annals of the New York Academy of Sciences, 1124(1), 61-76.

Larson-Prior, L.J., Zempel, J.M., Nolan, T.S., Prior, F.W., Snyder, A.Z., & Raichle, M.E., (2009). Cortical network functional connectivity in the descent to sleep. Proc. *Natl. Acad. Sci. U. S. A.* 106 (11), 4489–4494.

Ledoit, O., Wolf, M. (2003). Improved estimation of the covariance matrix of stock returns with an application to portfolio selection. Journal of Empirical Finance, 10, 603–21.

Lee, S. H., Kravitz, D. J., & Baker, C. I. (2012). Disentangling visual imagery and perception of real-world objects. Neuroimage, 59(4), 4064-4073.

Leech, R., Braga, R., & Sharp, D. J. (2012). Echoes of the brain within the posterior cingulate cortex. Journal of Neuroscience, 32(1), 215-222.

Leech, R., & Sharp, D. J. (2014). The role of the posterior cingulate cortex in cognition and disease. Brain, 137(Pt 1), 12-32. doi:10.1093/brain/awt162Levenshtein, V. I. (1965). Binary codes with correction for deletions and insertions of the symbol 1. Problemy Peredachi Informatsii, 1(1), 12-25.

Lewis-Peacock, J. A., & Norman, K. A. (2013). Multi-Voxel Pattern Analysis of fMRI Data. In M. S. Gazzaniga (Ed.), The cognitive neurosciences (5th ed, pp. 911-920). Cambridge, MA: MIT Press.

Lima, C. F., Lavan, N., Evans, S., Agnew, Z., Halpern, A. R., Shanmugalingam, P., ... & Warren, J. E. (2015). Feel the noise: Relating individual differences in auditory imagery to the structure and function of sensorimotor systems. Cerebral cortex, 25(11), 4638-4650.

Logothetis, N. K., Pauls, J., Augath, M., Trinath, T., & Oeltermann, A. (2001). Neurophysiological investigation of the basis of the fMRI signal. *Nature*, *412*(6843), 150-157.

Lowe, M. J., Dzemidzic, M., Lurito, J. T., Mathews, V. P., & Phillips, M. D. (2000). Correlations in low-frequency BOLD fluctuations reflect cortico-cortical connections. *Neuroimage*, *12*(5), 582-587.

Lowe, M. J., B. J. Mock, and J. A. Sorenson. "Functional connectivity in single and multislice echoplanar imaging using resting-state fluctuations." *Neuroimage* 7.2 (1998): 119-132.

Lund, T. E., Nørgaard, M. D., Rostrup, E., Rowe, J. B., & Paulson, O. B. (2005). Motion or activity: their role in intra-and inter-subject variation in fMRI. *Neuroimage*, *26*(3), 960-964.

Macey, P. M., Macey, K. E., Kumar, R., & Harper, R. M. (2004). A method for removal of global effects from fMRI time series. *Neuroimage*, *22*(1), 360-366.

Maguire, E. A. (2001). Neuroimaging studies of autobiographical event memory. Philosophical Transactions of the Royal Society of London B: Biological Sciences, 356(1413), 1441-1451.

Mahmoudi, A., Takerkart, S., Regragui, F., Boussaoud, D., & Brovelli, A. (2012). Multivoxel pattern analysis for fMRI data: A review. *Computational and mathematical methods in medicine*, *2012.*

Man, K., Damasio, A., Meyer, K., & Kaplan, J. T. (2015). Convergent and invariant object representations for sight, sound, and touch. *Human brain mapping, 36(9),* 3629-3640.

Margulies, D. S., Ghosh, S. S., Goulas, A., Falkiewicz, M., Huntenburg, J. M., Langs, G., ... & Jefferies, E. (2016). Situating the default-mode network along a principal gradient of macroscale cortical organization. *Proceedings of the National Academy of Sciences*, *113*(44), 12574-12579.

Margulies, D. S., Kelly, A. M., Uddin, L. Q., Biswal, B. B., Castellanos, F. X., and Milham, M. P. (2007). Mapping the functional connectivity of anterior cingulate cortex*. Neuroimage, 37*, 579–588.

Margulies, D. S., Vincent, J. L., Kelly, C., Lohmann, G., Uddin, L. Q., Biswal, B. B., ... & Petrides, M. (2009). Precuneus shares intrinsic functional architecture in humans and monkeys. *Proceedings of the National Academy of Sciences,* 106(47), 20069-20074.

Marinkovic, K. *et al*. Spatiotemporal dynamics of modality-specific and suprtransmodal word processing. *Neuron* **38**, 487–497 (2003).

Martin, A. (2007). The representation of object concepts in the brain. *Annual Review of Psychology, 58, 25-45.*

Martin, A., & Chao, L. L. (2001). Semantic memory and the brain: structure and processes. *Current opinion in neurobiology, 11*(2), 194-201.

Martin, A., Haxby, J. V., Lalonde, F. M., Wiggs, C. L., & Ungerleider, L. G. (1995).

Discrete cortical regions associated with knowledge of color and knowledge of action. *Science, 270*(5233), 102.

Mason, M. F., Norton, M. I., Van Horn, J. D., Wegner, D. M., Grafton, S. T., & Macrae, C. N. (2007). Wandering minds: the default network and stimulus- independent thought. Science, 315(5810), 393-395.

McClelland, J. L., Botvinick, M. M., Noelle, D. C., Plaut, D. C., Rogers, T. T., Seidenberg, M. S., & Smith, L. B. (2010). Letting structure emerge: connectionist and dynamical systems approaches to cognition. Trends in cognitive sciences, 14(8), 348-356.

McClelland, J. L., & Rogers, T. T. (2003). The parallel distributed processing approach to semantic cognition. Nature reviews. Neuroscience, 4(4), 310.

McKiernan, K. A., D'angelo, B. R., Kaufman, J. N., & Binder, J. R. (2006). Interrupting the "stream of consciousness": an fMRI investigation. Neuroimage, 29(4), 1185-1191.

McNorgan, C., Reid, J., & McRae, K. (2011). Integrating conceptual knowledge within and across representational modalities. Cognition, 118(2), 211-233.

Mechelli, A., Price, C. J., Friston, K. J., & Ishai, A. (2004). Where bottom-up meets top-down: neuronal interactions during perception and imagery. Cerebral cortex, 14(11), 1256-1265.

Mesulam, M. M. (Ed.). (1985). Principles of behavioral neurology (No. 26). Oxford University Press, USA.

Mesulam, M. M. (1998). From sensation to cognition. Brain, 121 (Pt 6), 1013-1052.

Mesulam, M. (2012). The evolving landscape of human cortical connectivity: facts and inferences. Neuroimage, 62(4), 2182-2189.

Meteyard, L., Cuadrado, S. R., Bahrami, B., & Vigliocco, G. (2012). Coming of age: A review of embodiment and the neuroscience of semantics. *Cortex, 48(7), 788-804.*

Miller, B. T., & D'Esposito, M. (2005). Searching for "the top" in top-down control. Neuron, 48(4), 535-538.

Mion, M., Patterson, K., Acosta-Cabronero, J., Pengas, G., Izquierdo-Garcia, D., Hong, Y, T., et al. (2010). What the left and right anterior fusiform gyri tell us about semantic memory. *Brain, 133,* 3256-3268.

Misaki, M., Kim, Y., Bandettini, P. A., & Kriegeskorte, N. (2010). Comparison of multivariate classifiers and response normalizations for pattern-information fMRI. *Neuroimage*, *53*(1), 103-118.

Mitchell, T. M., Hutchinson, R., Niculescu, R. S., Pereira, F., Wang, X., Just, M., & Newman, S. (2004). Learning to decode cognitive states from brain images. *Machine Learning*, *57*(1-2), 145-175.

Monti, M. M. (2011). Statistical analysis of fMRI time-series: a critical review of the GLM approach. Frontiers in human neuroscience, 5(28).

Morosan, P., Rademacher, J., Schleicher, A., Amunts, K., Schormann, T., & Zilles, K. (2001). Human primary auditory cortex: Cytoarchitectonic subdivisions and mapping into a spatial reference system. *NeuroImage, 13*(4), 684-701.

Moscovitch, M. (1992). A neuropsychological model of memory and consciousness. Neuropsychology of memory, 2, 5-22.

Moscovitch, M., Cabeza, R., Winocur, G., & Nadel, L. (2016). Episodic Memory and Beyond: The Hippocampus and Neocortex in Transformation. Annu Rev Psychol, 67, 105-134. doi:10.1146/annurev-psych-113011-143733

Mummery, C. J., Patterson, K., Price, C. J., Ashburner, J., Frackowiak, R. S. J., & Hodges, J. R. (2000). A voxel-based morphometry study of semantic dementia: relationship between temporal lobe atrophy and semantic memory. Annals of neurology, 47(1), 36-45.

Mur, M., Bandettini, P. A., & Kriegeskorte, N. (2009). Revealing representational content with pattern-information fMRI—an introductory guide. *Social cognitive and affective neuroscience*, nsn044.

Murphy, K., Birn, R. M., Handwerker, D. A., Jones, T. B., & Bandettini, P. A. (2009). The impact of global signal regression on resting state correlations: are anti-correlated networks introduced?. *Neuroimage*, *44*(3), 893-905.

Murphy, K., Birn, R. M., & Bandettini, P. A. (2013). Resting-state fMRI confounds and cleanup. *Neuroimage*, 80, 349-359.

Murphy, B., Poesio, M., Bovolo, F., Bruzzone, L., Dalponte, M., & Lakany, H. (2011). EEG decoding of semantic category reveals distributed representations for single concepts. *Brain and language*, *117*(1), 12-22.

Murphy, C., Rueschemeyer, S. A., Watson, D., Karapanagiotidis, T., Smallwood, J., & Jefferies, E. (2017). Fractionating the anterior temporal lobe: MVPA reveals differential responses to input and conceptual modality. NeuroImage, 147, 19-31.

Muschelli, J., Nebel, M. B., Caffo, B. S., Barber, A. D., Pekar, J. J., & Mostofsky, S. H. (2014). Reduction of motion-related artifacts in resting state fMRI using aCompCor. *Neuroimage*, *96*, 22-35.

Naselaris, T., Olman, C. A., Stansbury, D. E., Ugurbil, K., & Gallant, J. L. (2015). A voxel-wise encoding model for early visual areas decodes mental images of remembered scenes. Neuroimage, 105, 215-228.

Ng, A. Y., & Jordan, M. I. (2002). On discriminative vs. generative classifiers: A comparison of logistic regression and naive bayes. *Advances in neural information processing systems*, *2*, 841-848.

Noonan, K. A., Jefferies, E., Corbett, F., & Ralph, M. A. L. (2010). Elucidating the nature of deregulated semantic cognition in semantic aphasia: evidence for the roles of prefrontal and temporo-parietal cortices. Journal of Cognitive Neuroscience, 22(7), 1597-1613.

Noonan, K. A., Jefferies, E., Visser, M., & Ralph, M. A. L. (2013). Going beyond inferior prefrontal involvement in semantic control: evidence for the additional contribution of dorsal angular gyrus and posterior middle temporal cortex. Journal of cognitive neuroscience, 25(11), 1824-1850.

Noppeney, U., Patterson, K., Tyler, L. K., Moss, H., Stamatakis, E. A., Bright, P., ... & Price, C. J. (2007). Temporal lobe lesions and semantic impairment: a comparison of herpes simplex virus encephalitis and semantic dementia. Brain, 130(4), 1138-1147.

Norman, K. A., Polyn, S. M., Detre, G. J., & Haxby, J. V. (2006). Beyond mind-reading: multi-voxel pattern analysis of fMRI data. *Trends in cognitive sciences*, *10*(9), 424-430.

Ogawa, S., Lee, T. M., Kay, A. R., & Tank, D. W. (1990). Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proceedings of the National Academy of Sciences*, *87*(24), 9868-9872.

O'reilly, J. X., Beckmann, C. F., Tomassini, V., Ramnani, N., & Johansen-Berg, H. (2010). Distinct and overlapping functional zones in the cerebellum defined by resting state functional connectivity. *Cerebral cortex*, *20*(4), 953-965.

Pascual, B., Masdeu, J. C., Hollenbeck, M., Makris, N., Insausti, R., Ding, S. L., & Dickerson, B. C. (2015). Large-scale brain networks of the human left temporal pole: a functional connectivity MRI study. *Cerebral Cortex*, *25*(3), 680-702.

Patterson, K., & Lambon Ralph, M. A. (1999). Is a picture worth a thousand words? Evidence from concept definitions by patients with semantic dementia. *Brain and Language, 70*(3), 309-335.

Patterson, K., Nestor, P. J., & 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.

Pearson, J., Rademaker, R. L., & Tong, F. (2011). Evaluating the mind's eye: the metacognition of visual imagery. Psychological Science, 22(12), 1535-1542.

Peelen, M. V., & Caramazza, A. (2012). Conceptual object representations in human anterior temporal cortex. *The Journal of Neuroscience*, *32*(45), 15728-15736.

Pereira, F., Mitchell, T., & Botvinick, M. (2009). Machine learning classifiers and fMRI: a tutorial overview. *Neuroimage*, *45*(1), S199-S209.

Plaut, D. C. (2002). Graded modality-specific specialisation in semantics: A computational account of optic aphasia. Cognitive Neuropsychology, 19(7), 603-639.

Pobric, G., Jefferies, E., & Lambon Ralph, M. A. (2007). Anterior temporal lobes mediate semantic representation: mimicking semantic dementia by using rTMS in normal participants. *Proceedings of the National Academy of Sciences*, *104*(50), 20137-20141.

Pobric, G., Ralph, M. A. L., & Jefferies, E. (2009). The role of the anterior temporal lobes in the comprehension of concrete and abstract words: rTMS evidence. *Cortex*, *45*(9), 1104-1110.

Pobric, G., Jefferies, E., & Lambon Ralph, M. A. (2010). Category-specific versus category-general semantic impairment induced by transcranial magnetic stimulation. *Current Biology*, *20*(10), 964-968.

Poerio, G. L., Totterdell, P., & Miles, E. (2013). Mind-wandering and negative mood: Does one thing really lead to another?. Consciousness and Cognition, 22(4), 1412-1421.

Poldrack, R. A. (2007). Region of interest analysis for fMRI. *Social cognitive and affective neuroscience*, *2*(1/4), 67.

Poldrack, R. A., Mumford, J. A., & Nichols, T. E. (2011). *Handbook of functional MRI data analysis*. Cambridge University Press.

Postle, N., McMahon, K. L., Ashton, R., Meredith, M., & Zubicaray, G, I, D. (2008). Action word meaning representations in cytoarchitectonically defined primary and premotor cortices. *NeuroImage, 43*, 634-644.

Power, J. D., Barnes, K. A., Snyder, A. Z., Schlaggar, B. L., & Petersen, S. E. (2012). Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *Neuroimage*, *59*(3), 2142-2154.

Power, J. D., Mitra, A., Laumann, T. O., Snyder, A. Z., Schlaggar, B. L., & Petersen, S. E. (2014). Methods to detect, characterize, and remove motion artifact in resting state fMRI. *Neuroimage*, *84*, 320-341.

Power, J. D., Schlaggar, B. L., & Petersen, S. E. (2015). Recent progress and outstanding issues in motion correction in resting state fMRI. *Neuroimage*, *105*, 536-551.

Price, A. R., Bonner, M. F., Peelle, J. E., & Grossman, M. (2015). Converging evidence for the neuroanatomic basis of combinatorial semantics in the angular gyrus. Journal of Neuroscience, 35(7), 3276-3284.

Pulvermüller, F. (2005). Brain mechanisms linking language and action, Nature Reviews Neuroscience, 6, 576-582.

Pulvermüller, F. (2012). Meaning and the brain: The neurosemantics of referential, interactive, and combinatorial knowledge. Journal of Neurolinguistics, 25(5), 423-459.

Pulvermüller, F. (2013). How neurons make meaning: brain mechanisms for embodied and abstract-symbolic semantics. Trends in cognitive sciences, 17(9), 458-470.

Raichle, M. E. (2010). Two views of brain function. Trends in cognitive sciences, 14(4), 180-190.

Raichle, M. E. (2011). The restless brain. Brain connectivity, 1(1), 3-12.

Raichle, M. E. (2015). The brain's default mode network. *Annual review of neuroscience*, (0).

Raichle, M. E., MacLeod, A. M., Snyder, A. Z., Powers, W. J., Gusnard, D. A., & Shulman, G. L. (2001). A default mode of brain function. *Proceedings of the National Academy of Sciences*, *98*(2), 676-682.

Raj, D., Anderson, A. W., & Gore, J. C. (2001). Respiratory effects in human functional magnetic resonance imaging due to bulk susceptibility changes. *Physics in medicine and biology*, *46*(12), 3331.

Raizada, R. D., & Lee, Y. S. (2013). Smoothness without smoothing: why Gaussian naive Bayes is not naive for multi-subject searchlight studies. *PloS one*, *8*(7), e69566.

Reddy, L., Tsuchiya, N., & Serre, T. (2010). Reading the mind's eye: decoding category information during mental imagery. NeuroImage, 50(2), 818-825.

Reilly, J., Garcia, A., & Binney, R. J. (2016). Does the sound of a barking dog activate its corresponding visual form? An fMRI investigation of modality-specific semantic access. Brain and language, 159, 45-59.

Reilly, J., Peelle, J. E., Garcia, A., & Crutch, S. J. (2016). Linking somatic and symbolic representation in semantic memory: the dynamic multilevel reactivation framework. Psychonomic bulletin & review, 23(4), 1002-1014.

Rice, G. E., Hoffman, P., Lambon Ralph, M. A., & Matthew, A. (2015). Graded specialization within and between the anterior temporal lobes. Annuals of the New York Academy of Sciences, 1359(1), 84-97.

Rice, G. E., Lambon Ralph, M. A. & Hoffman, P. The roles of left versus right anterior temporal lobes in conceptual knowledge: An ALE meta-analysis of 97 functional neuroimaging studies. *Cerebral Cortex, 26*(2), 1-18.

Riddoch, M. J., Humphreys, G. W., Coltheart, M., & Funnell, E. (1988). Semantic systems or system? Neuropsychological evidence re-examined. Cognitive Neuropsychology, 5(1), 3-25.

Rogers, T. T., Hocking, J., Noppeney, U., Mechelli, A., Gorno-Tempini, M. L., Patterson, K., & Price, C. J. (2006). Anterior temporal cortex and semantic

memory: reconciling findings from neuropsychology and functional imaging. *Cognitive, Affective, & Behavioral Neuroscience*, *6*(3), 201-213.

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

Rogers, T. T., Patterson, K., Jefferies, E., Lambon Ralph, M. A. (2015). Disorders of representation and control in semantic cognition: Effects of familiarity, typicality and specificity. *Neuropsychologia, 76,* 220-239.

Rosa, M. G. P. (2002). Visual maps in the adult primate cerebral cortex: some implications for brain development and evolution. *Brazilian Journal of Medical and Biological Research*, *35*(12), 1485-1498.

Rueschemeyer, S-A., Brass, M., Friederici, A, D. (2007). Comprehending prehending: Neural correlates of processing verbs with motor stems. *Journal of Cognitive Neuroscience, 19*(5), 855-865.

Rueschemeyer, S-A., Rooik, D. V., Lindemann, O., Willems, R. M., & Bekkering, H. (2010). The function of words: Distinct neural correlates for words denoting differently manipulable objects. *Journal of Cognitive Neuroscience, 22*(8), 1844-1851.

Rueschemeyer, S.-A., Eckmann, M., van Ackeren, & Kilner, J. (2014). Observing, performing and understanding actions: Revisiting the role of cortical motor areas in processing of action words. Journal of Cognitive Neuroscience, 26(8), 1644-1653.

Rugg, M. D., & Vilberg, K. L. (2013). Brain networks underlying episodic memory retrieval. Current opinion in neurobiology, 23(2), 255-260.

Salvador, R., Suckling, J., Schwarzbauer, C., & Bullmore, E. (2005). Undirected graphs of frequency-dependent functional connectivity in whole brain networks. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *360*(1457), 937-946.

Sanjuán, A., Hope, T. M., Jones, Ō. P., Prejawa, S., Oberhuber, M., Guerin, J., ... & Price, C. J. (2015). Dissociating the semantic function of two neighbouring

subregions in the left lateral anterior temporal lobe. *Neuropsychologia*, *76*, 153-162.

Satterthwaite, T. D., Wolf, D. H., Loughead, J., Ruparel, K., Elliott, M. A., Hakonarson, H., ... & Gur, R. E. (2012). Impact of in-scanner head motion on multiple measures of functional connectivity: relevance for studies of neurodevelopment in youth. *Neuroimage*, *60*(1), 623-632.

Saur, D., Schelter, B., Schnell, S., Kratochvil, D., Küpper, H., Kellmeyer, P., ... & Mader, W. (2010). Combining functional and anatomical connectivity reveals brain networks for auditory language comprehension. Neuroimage, 49(4), 3187-3197.

Schacter, D. L., & Addis, D. R. (2007). The cognitive neuroscience of constructive memory: remembering the past and imagining the future. Philosophical Transactions of the Royal Society B: Biological Sciences, 362(1481), 773-786.

Schacter, D. L., Addis, D. R., & Buckner, R. L. (2007). Remembering the past to imagine the future: the prospective brain. Nature reviews. Neuroscience, 8(9), 657.

Schapiro, A. C., Rogers, T. T., Cordova, N. I., Turk-Browne, N. B., & Botvinick, M. M. (2013). Neural representations of events arise from temporal community structure. Nature neuroscience, 16(4), 486-492.

Schölvinck, M. L., Maier, A., Frank, Q. Y., Duyn, J. H., & Leopold, D. A. (2010). Neural basis of global resting-state fMRI activity. *Proceedings of the National Academy of Sciences*, *107*(22), 10238-10243.

Schooler, J. W., Smallwood, J., Christoff, K., Handy, T. C., Reichle, E. D., & Sayette, M. A . (2011). Meta-awareness, perceptual decoupling and the wandering mind. Trends in cognitive sciences, 15(7), 319-326.

Scott, S. K., Blank, C. C., Rosen, S., & Wise, R. J. (2000). Identification of a pathway for intelligible speech in the left temporal lobe. *Brain*, *123*(12), 2400-2406.

Scott, S. K., & Johnsrude, I. S. (2003). The neuroanatomical and functional organization of speech perception. *Trends in neurosciences*, *26*(2), 100-107.

Scott, S. K., Leff, A. P., & Wise, R. J. (2003). Going beyond the information given: a neural system supporting semantic interpretation. *Neuroimage*, *19*(3), 870-876.

Seghier, M. L. (2013). The angular gyrus: multiple functions and multiple subdivisions. Neuroscientist, 19(1), 43-61. doi:10.1177/1073858412440596

Seghier, M. L., Fagan, E., & Price, C. J. (2010). Functional subdivisions in the left angular gyrus where the semantic system meets and diverges from the default network. Journal of Neuroscience, 30(50), 16809-16817.

Seghier, M. L., & Price, C. J. (2012). Functional heterogeneity within the default network during semantic processing and speech production. Frontiers in psychology, 3.

Shehzad, Z., Kelly, A. M., Reiss, P. T., Gee, D. G., Gotimer, K., Uddin, L. Q., Lee, S. H., Margulies, D. S., Roy, A. K., Biswal, B. B., Petkova, E., Castellanos, F. X., & Milham, M. P. (2009). The resting brain: unconstrained yet reliable. *Cereb. Cortex* 19, 2209–2229.

Shulman, G. L., Corbetta, M., Buckner, R. L., Raichle, M. E., Fiez, J. A., Miezin, F. M., & Petersen, S. E. (1997). Top-down modulation of early sensory cortex. Cerebral cortex (New York, NY: 1991), 7(3), 193-206.

Simanova, I., Van Gerven, M., Oostenveld, R., & Hagoort, P. (2010). Identifying object categories from event-related EEG: toward decoding of conceptual representations. *PloS one*, *5*(12), e14465.

Simanova, I., Van Gerven, M. A., Oostenveld, R., & Hagoort, P. (2014). Predicting the semantic category of internally generated words from neuromagnetic recordings. Journal of cognitive neuroscience.

Simmons, W. K., & Martin, A. (2009). The anterior temporal lobes and the functional architecture of semantic memory. Journal of the International Neuropsychological Society, 15(5), 645-649.

Simmons, W. K., Martin, A., & Barsalou, L. W. (2005). Pictures of appetizing foods activate gustatory cortices for taste and reward. Cerebral Cortex, 15(10), 1602-1608.

Simmons, W. K., Ramjee, V., Beauchamp, M. S., McRae, K., Martin, A., & Barsalou, L. W. (2007). A common neural substrate for perceiving and knowing about color. Neuropsychologia, 45(12), 2802-2810.

Singer, J. L. (1966). Daydreaming: An introduction to the experimental study of inner experience.

Singer, W. (1993). Synchronization of cortical activity and its putative role in information processing and learning. Annual review of physiology, 55(1), 349-374.

Sitnikova, T., West, W. C., Kuperberg, G. R., & Holcomb, P. J. (2006). The neural organization of semantic memory: Electrophysiological activity suggests feature-based segregation. Biological psychology, 71(3), 326-340.

Skipper, L. M., Ross, L. A., & Olson, I. R. (2011). Sensory and semantic category subdivisions within the anterior temporal lobes. Neuropsychologia, 49(12), 3419-3429.

Sladky, R., Friston, K. J., Tröstl, J., Cunnington, R., Moser, E., & Windischberger, C. (2011). Slice-timing effects and their correction in functional MRI. *Neuroimage*, *58*(2), 588-594.

Slotnick, S. D., Thompson, W. L., & Kosslyn, S. M. (2005). Visual mental imagery induces retinotopically organized activation of early visual areas. Cerebral cortex, 15(10), 1570-1583.

Smallwood, J. (2013). Distinguishing how from why the mind wanders: a process-occurrence framework for self-generated mental activity. Psychol Bull, 139(3), 519-535. doi:10.1037/a0030010

Smallwood, J., Karapanagiotidis, T., Ruby, F., Medea, B., de Caso, I., Konishi, M., ... & Jefferies, E. (2016). Representing Representation: Integration between the Temporal Lobe and the Posterior Cingulate Influences the Content and Form of Spontaneous Thought. *PloS one*, *11*(4), e0152272.

Smallwood, J., & Schooler, J. W. (2006). The restless mind. Psychological bulletin, 132(6), 946.

Smallwood, J., & Schooler, J. W. (2015). The science of mind wandering: empirically navigating the stream of consciousness. *Annual review of psychology*, *66*, 487-518.

Smith, S. M. (2002). Fast robust automated brain extraction. *Human Brain Mapping, 17*(3), 143-155.

Smith, S. M., Fox, P. T., Miller, K. L., Glahn, D. C., Fox, P. M., Mackay, C. E., Filippini, N., Watkins, K. E., Toro, R., Laird, A. R., & Beckmann, C. F. (2009).

Correspondence of the brain's functional architecture during activation and rest. *Proc. Natl. Acad. Sci. U.S.A.* 106, 13040–13045.

Snowden, J. S., Goulding, P. J., & Neary, D. (1989). Semantic dementia: a form of circumscribed cerebral atrophy. Behavioural Neurology.

Song, M., Zhou, Y., Li, J., Liu, Y., Tian, L., Yu, C., & Jiang, T., (2008). Brain spontaneous functional connectivity and intelligence. Neuroimage 41 (3), 1168–1176.

Sormaz, M., Jefferies, E., Bernhardt, B. C., Karapanagiotidis, T., Mollo, G., Bernasconi, N., ... & Smallwood, J. (2017). Knowing what from where: Hippocampal connectivity with temporoparietal cortex at rest is linked to individual differences in semantic and topographic memory. NeuroImage, 152, 400-410.

Spitsyna, G., Warren, J. E., Scott, S. K., Turkheimer, F. E., & Wise, R. J. (2006). Converging language streams in the human temporal lobe. *The Journal of Neuroscience*, *26*(28), 7328-7336.

Spreng, R. N. (2012). The fallacy of a "task-negative" network. Frontiers in psychology, 3.

Spreng, R. N., DuPre, E., Selarka, D., Garcia, J., Gojkovic, S., Mildner, J., . . . Turner, G. R. (2014). Goal-congruent default network activity facilitates cognitive control. J Neurosci, 34(42), 14108-14114. doi:10.1523/JNEUROSCI.2815-14.2014

Spreng, R. N., & Grady, C. L. (2010). Patterns of brain activity supporting autobiographical memory, prospection, and theory of mind, and their relationship to the default mode network. Journal of cognitive neuroscience, 22(6), 1112-1123.

Spreng, R. N., Gerlach, K. D., Turner, G. R., & Schacter, D. L. (2015). Autobiographical Planning and the Brain: Activation and Its Modulation by Qualitative Features. J Cogn Neurosci, 27(11), 2147-2157. doi:10.1162/jocn_a_00846

Spreng, R. N., Mar, R. A., & Kim, A. S. (2009). The common neural basis of autobiographical memory, prospection, navigation, theory of mind, and the default mode: a quantitative meta-analysis. *Journal of cognitive neuroscience*, *21*(3), 489-510.

Spreng, R. N., Stevens, W. D., Chamberlain, J. P., Gilmore, A. W., & Schacter, D. L. (2010). Default network activity, coupled with the frontoparietal control

network, supports goal-directed cognition. Neuroimage, 53(1), 303-317. doi:10.1016/j.neuroimage.2010.06.016

Stephan, K. E., Penny, W. D., Moran, R. J., den Ouden, H. E., Daunizeau, J., & Friston, K. J. (2010). Ten simple rules for dynamic causal modeling. Neuroimage, 49(4), 3099-3109.

Stokes, M., Thompson, R., Cusack, R., & Duncan, J. (2009). Top-down activation of shape- specific population codes in visual cortex during mental imagery. Journal of Neuroscience, 29(5), 1565-1572.

Suddendorf, T., & Corballis, M. C. (2007). The evolution of foresight: What is mental time travel, and is it unique to humans?. Behavioral and Brain Sciences, 30(3), 299-313.

Svoboda, E., McKinnon, M. C., & Levine, B. (2006). The functional neuroanatomy of autobiographical memory: a meta-analysis. Neuropsychologia, 44(12), 2189-2208.

Szpunar, K. K., Watson, J. M., & McDermott, K. B. (2007). Neural substrates of envisioning the future. Proceedings of the National Academy of Sciences, 104(2), 642-647.

Thompson-Schill, S. L. (2003). Neuroimaging studies of semantic memory: inferring “how” from “where”. Neuropsychologia, 41(3), 280-292.

Thompson-Schill, S. L., D’Esposito, M., Aguirre, G. K., & Farah, M. J. (1997). Role of left inferior prefrontal cortex in retrieval of semantic knowledge: a reevaluation. Proceedings of the National Academy of Sciences, 94(26), 14792-14797.

Tranel, D., Damasio, H., & Damasio, A. R. (1997). A neural basis for the retrieval of conceptual knowledge. Neuropsychologia, 35(10), 1319-1327.

Tulving, E. (1972). Episodic and semantic memory. Organization of memory, 1, 381-403.

Tusche, A., Smallwood, J., Bernhardt, B. C., & Singer, T. (2014). Classifying the wandering mind: revealing the affective content of thoughts during task-free rest periods. *Neuroimage*, *97*, 107-116.

Tyler, L. K., & Moss, H. E. (2001). Towards a distributed account of conceptual knowledge. Trends in cognitive sciences, 5(6), 244-252.

Tyler, L. K., Stamatakis, E. A., Bright, P., Acres, K., Abdallah, S., Rodd, J. M., & Moss,

H. E. (2004). Processing objects at different levels of specificity. *Journal of cognitive neuroscience*, *16*(3), 351-362.

Van Ackeren, M. & Rueschemeyer, S.-A. (2014). Theta power and beta coherence predict multimodal semantic integration at different cortical scales. PLOS ONE, 9(7), e101042.

Van Den Heuvel, M. P., & Pol, H. E. H. (2010). Exploring the brain network: a review on resting-state fMRI functional connectivity. *European Neuropsychopharmacology*, *20*(8), 519-534.

Van Diessen, E., Numan, T., Van Dellen, E., Van Der Kooi, A. W., Boersma, M., Hofman, D., ... & Stam, C. J. (2015). Opportunities and methodological challenges in EEG and MEG resting state functional brain network research. Clinical Neurophysiology, 126(8), 1468-1481.

Van Dijk, K. R., Hedden, T., Venkataraman, A., Evans, K. C., Lazar, S. W., & Buckner, R. L. (2010). Intrinsic functional connectivity as a tool for human connectomics: theory, properties, and optimization. *J. Neurophysiol.* 103, 297–321.

Van Dijk, K. R., Sabuncu, M. R., & Buckner, R. L. (2012). The influence of head motion on intrinsic functional connectivity MRI. *Neuroimage*, *59*(1), 431-438.

van Wijk, B. C., Litvak, V., Friston, K. J., & Daffertshofer, A. (2013). Nonlinear coupling between occipital and motor cortex during motor imagery: a dynamic causal modeling study. NeuroImage, 71, 104-113.

Vandenberghe, R., Nobre, A. C., & Price, C. J. (2002). The response of left temporal cortex to sentences. *Journal of Cognitive Neuroscience*, *14*(4), 550-560.

Vandenberghe, R., Price, C., Wise, R., Josephs, O. & Frackowiak, R. S. J. (1996). Functional-anatomy of a common semantic system for words and pictures. *Nature* **383**, 254–256 (1996).

Vatansever, D., Menon, D. K., Manktelow, A. E., Sahakian, B. J., & Stamatakis, E. A. (2015). Default mode network connectivity during task execution. Neuroimage, 122, 96-104. doi:10.1016/j.neuroimage.2015.07.053

Vetter, P., Smith, F. W., & Muckli, L. (2014). Decoding sound and imagery content in early visual cortex. Current Biology, 24(11), 1256-1262.

Vincent, J. L., Kahn, I., Snyder, A. Z., Raichle, M. E., & Buckner, R. L. (2008). Evidence

for a frontoparietal control system revealed by intrinsic functional connectivity. *Journal of neurophysiology*, *100*(6), 3328-3342.

Vincent, J. L., Patel, G. H., Fox, M. D., Snyder, A. Z., Baker, J. T., Van Essen, D. C., ... & Raichle, M. E. (2007). Intrinsic functional architecture in the anaesthetized monkey brain. *Nature*, *447*(7140), 83-86.

Visser, M., Jefferies, E., & Lambon Ralph, M. (2010). Semantic processing in the anterior temporal lobes: a meta-analysis of the functional neuroimaging literature. Journal of cognitive neuroscience, 22(6), 1083-1094.

Visser, M. & Lambon Ralph, M. A. Differential contributions of bilateral ventral anterior temporal lobe and left anterior superior temporal gyrus to semantic processes. *J. Cogn. Neurosci.* **23**, 3121–3131 (2011).

Visser, M., Embleton, K. V., Jefferies, E., Parker, G. J. M., & Lambon Ralph, M. A. (2010). The anterior temporal lobes and semantic memory clarified: Novel evidence from distortion-corrected spin-echo EPI fMRI. *Neuropsycholgia, 48,* 1689-1696.

Visser, M., Jefferies, E., Embleton, K. V., & Ralph, M. A. L. (2012). Both the middle temporal gyrus and the ventral anterior temporal area are crucial for multimodal semantic processing: distortion-corrected fMRI evidence for a double gradient of information convergence in the temporal lobes. *Journal of Cognitive Neuroscience*, *24*(8), 1766-1778.

Visser, M., Jefferies, E., & Ralph, M. L. (2010). Semantic processing in the anterior temporal lobes: a meta-analysis of the functional neuroimaging literature. *Journal of cognitive neuroscience*, *22*(6), 1083-1094.

Wagner, A. D., Paré-Blagoev, E. J., Clark, J., & Poldrack, R. A. (2001). Recovering meaning: left prefrontal cortex guides controlled semantic retrieval. Neuron, 31(2), 329-338.

Wagner, I. C., van Buuren, M., Kroes, M. C., Gutteling, T. P., van der Linden, M., Morris, R. G., & Fernandez, G. (2015). Schematic memory components converge within angular gyrus during retrieval. Elife, 4, e09668. doi:10.7554/eLife.09668

Wang, J., Conder, J. A., Blitzer, D. N., & Shinkareva, S. V. (2010). Neural representation of abstract and concrete concepts: A meta-analysis of neuroimaging studies. Human brain mapping, 31(10), 1459-1468.

Wang, T., Deng, J., & He, B. (2004). Classifying EEG-based motor imagery tasks by means of time–frequency synthesized spatial patterns. *Clinical Neurophysiology*, *115*(12), 2744-2753.

Warrington, E. K. (1975). The selective impairment of semantic memory. The Quarterly journal of experimental psychology, 27(4), 635-657.

Warrington, E. K., & Shallice, T. (1984). Category specific semantic impairments. Brain, 107(3), 829-853.

Weiskopf, N., Hutton, C., Josephs, O., Turner, R., & Deichmann, R. (2007). Optimized EPI for fMRI studies of the orbitofrontal cortex: compensation of susceptibility-induced gradients in the readout direction. *Magnetic Resonance Materials in Physics, Biology and Medicine*, *20*(1), 39.

Weissenbacher, A., Kasess, C., Gerstl, F., Lanzenberger, R., Moser, E., & Windischberger, C. (2009). Correlations and anticorrelations in resting-state functional connectivity MRI: a quantitative comparison of preprocessing strategies. *Neuroimage*, *47*(4), 1408-1416.

Whitney, C., Kirk, M., O'sullivan, J., Lambon Ralph, M. A., & Jefferies, E. (2010). The neural organization of semantic control: TMS evidence for a distributed network in left inferior frontal and posterior middle temporal gyrus. Cerebral Cortex, 21(5), 1066-1075.

Willems, R. M., Toni, I., Hagoort, P., & Casasanto, D. (2009). Body-specific motor imagery of hand actions: neural evidence from right-and left-handers. Frontiers in Human Neuroscience, 3.

Willems, R. M., Toni, I., Hagoort, P., & Casasanto, D. (2010). Neural dissociations between action verb understanding and motor imagery. Journal of Cognitive Neuroscience, 22(10), 2387-2400.

Wilson, M. (2002). Six views of embodied cognition. Psychonomic bulletin & review, 9(4), 625-636.

Wirth, M., Jann, K., Dierks, T., Federspiel, A., Wiest, R., & Horn, H. (2011). Semantic memory involvement in the default mode network: a functional neuroimaging study using independent component analysis. Neuroimage, 54(4), 3057-3066.

Worsley, K. J., Liao, C. H., Aston, J., Petre, V., Duncan, G. H., Morales, F., & Evans, A. C. (2002). A general statistical analysis for fMRI data. *Neuroimage*, *15*(1), 1-15.

Yarkoni, T. (2009). Big correlations in little studies: Inflated fMRI correlations reflect low statistical power—Commentary on Vul et al.(2009). *Perspectives on Psychological Science*, *4*(3), 294-298.

Yarkoni, T., Poldrack, R. A., Nichols, T. E., Van Essen, D. C., & Wager, T. D. (2011). Large-scale automated synthesis of human functional neuroimaging data. *Nature methods*, *8*(8), 665-670.

Yeo, B. T., Krienen, F. M., Sepulcre, J., Sabuncu, M. R., Lashkari, D., Hollinshead, M., ... & Fischl, B. (2011). The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *Journal of neurophysiology*, *106*(3), 1125-1165.

Zatorre, R. J., & Halpern, A. R. (2005). Mental concerts: musical imagery and auditory cortex. Neuron, 47(1), 9-12.

Zhang, Y., Brady, M., & Smith, S.(2001). Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. *IEEE Trans Med Imag, 20*(1), 45-57.

Zvyagintsev, M., Clemens, B., Chechko, N., Mathiak, K. A., Sack, A. T., & Mathiak, K. (2013). Brain networks underlying mental imagery of auditory and visual information. European Journal of Neuroscience, 37(9), 1421-1434.