

**The Evolution of Cognitive Models:
From Neuropsychology to Neuroimaging and back**

Cathy J. Price,

Wellcome Centre for Human Neuroimaging,
Institute of Neurology,
University College London, UK

Acknowledgements

With thanks to Glyn Humphreys for many years of inspiring and fascinating lessons and discussions on what we can learn from neuropsychology and neuroimaging. I am currently funded by the Wellcome [097720/Z/11/Z], [205103/Z/16/Z] and [203147/Z/16/Z]

Abstract

This paper provides a historical and future perspective on how neuropsychology and neuroimaging can be used to develop cognitive models of human brain functions. Section 1 focuses on the emergence of cognitive modelling from neuropsychology, why lesion location was considered to be unimportant and the challenges faced when mapping symptoms to impaired cognitive processes. Section 2 describes how established cognitive models based on behavioural data alone cannot explain the complex patterns of distributed brain activity that are observed in functional neuroimaging studies. This has led to proposals for new cognitive processes, new cognitive strategies and new functional ontologies for cognition. Section 3 considers how the integration of data from lesion, behavioural and functional neuroimaging studies of large cohorts of brain damaged patients can be used to determine whether inter-patient variability in behaviour is due to differences in the premorbid function of each brain region, lesion site or cognitive strategy. This combination of neuroimaging and neuropsychology is providing a deeper understanding of how cognitive functions can be lost and re-learned after brain damage – an understanding that will transform our ability to generate and validate cognitive models that are both physiologically plausible and clinically useful.

Introduction

The motivation for this paper was to describe a journey of thoughts and theories about cognitive models of human brain function that were initiated by conducting neuropsychological and neuroimaging studies with Glyn Humphreys. Previous discussions of how neuroimaging has contributed to cognitive models were the focus of a special issue of *Cortex* more than 10 years ago. The lead article (Coltheart, 2006a) argued in line with others previously (e.g. Marr and Poggio, 1977; Colby, 1978; Uttal, 2001; Harvey, 2004) that knowing about neural implementation of cognitive processing had not to date (2006) informed or changed our cognitive models. The debate centred on whether there was any evidence that neuroimaging had provided new insights that adjudicated between two alternative cognitive models. Although several examples were offered (Jack et al., 2006; Henson et al., 2006; Jonides et al., 2006; Seron and Fias, 2006; Vallar, 2006), Coltheart (2006b) and others (Page, 2006; Schutter et al., 2006) argued that none of them had contributed any more information than could have been gained from behavioural studies alone. More recently, in a special issue of *Perspectives in Psychological Science* (Mather et al., 2013a), Coltheart (2013) further emphasized that the contribution of neuroimaging data to a cognitive theory should not be based on the consistency of neuroimaging data with predictions from cognitive theory. It should be based on falsifying the predictions of a particular theory.

In the current paper, I take a different perspective and focus on how neuroimaging has changed the way we think about the functional computations (types of cognitive processing) that underlie behaviour. I start by introducing the rationale, fascination and limitations of neuropsychology. The bottom line is that we do not know how cognitive functions are implemented in the brain. We can only speculate and approximate on what the underlying computations are and how they are instantiated. I then discuss what neuroimaging has told us about the general principals of neuronal implementation and how the nature of the neuronal implementation constrains the nature of the computations and algorithms that are being performed. Therefore, this paper is not about the functions of different brain regions (i.e. the functional anatomy). It illustrates how learning about the anatomy can shed new light on what the computations underlying cognition might be.

The discussion of neuroimaging findings also highlights the fact that we don't know what is being coded and we do not yet have a formal terminology to assign functional labels to brain regions. For example, most cognitive models of reading and spelling refer to "orthographic processing". This simply means processing related to written text but it doesn't specify the nature of the processing or the degree to which this processing is shared by non-orthographic visual stimuli. I consider why current psychological nomenclature is

insufficient to describe the function of brain areas and how neuroimaging is motivating new terminology, new brain functions and new cognitive models.

In the final section, I highlight the benefits of integrating data from neuroimaging and neuropsychology. In brief, I show how neuroimaging can be used to distinguish between 3 different types of inter-patient variability: differences in (i) lesion site, (ii) the brain structures that compute a given function, and (iii) the cognitive strategy used for a given task even when the structure-function mapping is consistent at the individual process level. This helps to provide a deeper understanding of computational functions, processing pathways, co-occurring impairments and how the same functional impairment (and lesion site) can lead to different symptoms.

Section 1: Using Neuropsychology to inform cognitive models

Neuropsychology involves the study of behaviour in patients with neurological disorders. By indicating how brain damage impacts on behaviour, neuropsychological studies can test and infer models of the computations that underlie specific cognitive functions (e.g. language, memory, perception) in the neurologically normal brain. The most famous examples of neurological studies date back to the 19th Century when Paul Broca reported that patients with left posterior inferior frontal damage had more difficulty with speech production than speech comprehension; and conversely, Karl Wernicke noted that patients with damage to the left posterior superior temporal cortex had more difficulty with speech comprehension than production. This “double dissociation” in cognitive function (across different patients) indicated that speech production and comprehension are functionally independent of one another.

Bringing Broca’s and Wernicke’s findings together, Ludwig Lichtheim developed a simple processing model of language that linked auditory representations of speech (in Wernicke’s area) to motor representations of speech (in Broca’s area) via anatomical connections through the arcuate fasciculus. Jules Dejerine added to the model (1891) by including visual images of speech in the left angular gyrus/supramarginal gyrus where damage could result in a selective reading difficulty that dissociated from relatively preserved spoken language and writing abilities. Dejerine therefore coined the term “pure alexia” to describe a very specific deficit confined to the impaired processing of orthographic code rather than a more general perceptual disturbance (see Bub et al., 1993 for a full description).

Figure 1 illustrates the 19th Century neurological model of language and reading. Other 19th Century neurological investigations reported double dissociations in other cognitive functions leading to a deeper understanding of hand movement control and its breakdown in different types of apraxia (Liepmann, 1990) and object recognition and its break down in different types of agnosia (Lissauer, 1890).

Figure 1 about here.

After the early 19th Century attempts to localise mental functions to brain structures, most neuropsychologists in the 20th Century divorced themselves from the anatomy and focused on “functional architectures”, building box and arrow diagrams of the computational processes and representations that were needed to support complex cognitive functions. Information about the brain was considered misleading and largely irrelevant. The main concern was that the mapping between lesion site and cognitive deficit was inconsistent across patients, either because of premorbid differences in the computational functions of specific brain regions or because of differences in the degree to which the brain was able to reorganise itself when recovering from a lost function. Another reason to focus on computational rather than anatomical architectures was that, until the 1990s, neuroimaging of the lesion site (e.g. with CT scanning) was rarely precise enough to dissociate the brain regions that support different types of processing. In the last 2 decades, magnetic resonance imaging (MRI) has allowed us to match lesion sites across patients more precisely and demonstrated the very consistent effects that some lesion sites have on behaviour (Leff et al., 2006). It has also shown that the same lesion site can have dissimilar effects in different patients. For example, Figure 2 shows magnetic resonance images of the brains of two patients who both incurred damage to the left angular gyrus. One patient has difficulty reading and writing; the other has a slightly larger lesion that did not cause a reading impairment. Factors other than lesion site (e.g. premorbid reading experience, co-morbidity or differences in functional anatomy) must therefore explain the differences in reading ability after damage to the angular gyrus.

Figure 2 about here.

Throughout the 20th Century, cognitive models based on neuropsychological findings became increasingly sophisticated. For example, Marshall and Newcombe (1973) generated a dual route model of reading to account for why some patients have more difficulty reading words with irregular spellings (e.g. YACHT) than novel words (e.g. SHUP) and others exhibit the opposite pattern (more difficulty with novel words than familiar words with irregular

spelling). Patients with more difficulty reading irregularly spelled words were hypothesized to have a damaged lexical reading route with relative sparing of a sublexical reading route; whereas patients with more difficulty reading novel words than irregularly spelled words were hypothesized to have a damaged sublexical route with relative sparing of the lexical route (see Figure 3).

As computing hardware and software developed, the validity of neuropsychologically motivated information processing models could be tested by building computational models that formalise and implement the hypothesized processing steps. In the case of reading, this initially resulted in a more parsimonious model composed of three components (orthography, phonology and semantics) and two different routes between orthography and phonology: a direct route and an indirect route via semantics (e.g. Plaut et al. 1996), see Figure 3. Other more complex computational models have also been developed (Coltheart et al., 2001).

Figure 3 about here

These and many other studies of memory, object recognition, language and hand movements resulted in a resurgence of neuropsychology in the 1980s, including the launch of a journal dedicated to publishing neuropsychological findings (*Cognitive Neuropsychology*, 1984) and text books introducing neuropsychological methods to psychology students (Coltheart, Patterson and Marshall, 1980; Shallice, 1988; McCarthy and Warrington, 1990). I was one of these students and, in 1984, I volunteered as a research assistant to Glyn Humphreys and Jane Riddoch to help them screen large numbers of stroke patients in several London hospitals. Their goals were very clear. They were looking for patients with selective deficits in cognitive functions, conducting detailed case studies to determine which types of processing (computations) were impaired or preserved, gaining deeper insight into the complexity of cognition and generating information processing models of cognitive functions and dysfunctions. The importance of their work was also clear because, by re-thinking neuropsychological syndromes in information processing terms, they could offer more precise ways of assessing cognitive functions, more meaningful diagnoses and potentially clues to indicate how impaired processing could be re-learned to facilitate the most optimal recovery. Their findings resulted in numerous seminal papers that spanned many types of cognitive functions including unilateral visual attention (Riddoch and Humphreys, 1983), routes to object constancy (Humphreys and Riddoch, 1984), routes to action (Riddoch, Humphreys, Price, 1989), routes to reading (Humphreys and Evett, 1985), different types of agnosia (Riddoch and Humphreys, 1987; Humphreys et al., 1992;

Humphreys and Price, 1994) and category specific semantic impairments (Humphreys and Riddoch, 1987).

Inspired by Glyn and Jane, I was a cognitive neuropsychology enthusiast. Three of the lessons I learnt greatly influenced our subsequent neuroimaging work. The first was the lack of “pure specificity”. Although patients might be more impaired using one function than another, cognitive impairments after brain damage are rarely confined to one function. I learnt this from conducting a review of 100 years of literature on “pure alexia” from boxes of papers that Max Coltheart gave me when he moved from Birkbeck College London to Macquarie University in Sydney. What I read over and over again was that reading impairments in patients diagnosed with “pure alexia” were typically reported to co-occur with object naming and/or colour naming difficulties. Therefore the evidence for a reading specific module was not compelling when patients’ perceptual skills were thoroughly tested. An alternative hypothesis I found appealing was that deficits that appeared to be selective for reading could be the consequence of a perceptual impairment in rapidly processing multiple features (particularly letters) in parallel (Kinsbourne and Warrington, 1962).

The second lesson was that the same symptom could arise from a breakdown at one or more processing levels. Specifically, by conducting multiple case studies of patients who were letter by letter readers, we illustrated how the same symptom (letter-by-letter reading) could result from difficulty in visual attention or processing visual features in parallel (Price and Humphreys, 1992). Likewise, Caramazza et al., (1990) reported that semantic errors during naming and reading were not necessarily caused by impairments in the same type of processing.

The third lesson was that, even with extensive assessments of all possible processing abilities, it is often difficult to describe which deficits are causing the symptoms. Co-occurring deficits can lead to complex sets of symptoms, and two patients with the same underlying deficit can adopt very different compensatory strategies resulting in very different patterns of behaviour. For example, we showed how reading words could be worse than letter naming when the patient had a mild anomia as well as difficulty switching attention across the letters (Price and Humphreys, 1994).

In Section 3, I will discuss how these neuropsychological observations can be investigated with neuroimaging.

Section 2: Using Neuroimaging to inform cognitive models

The opportunity to pursue neuropsychological models with neuroimaging in neurologically normal individuals was very exciting. This was expected to introduce an independent source of physiological validity to conclusions derived from patients with brain damage. The basic rationale involves 6 different steps. In Step 1, we take an information processing model of a given cognitive task with a known stimulus and response (e.g. Figure 3). In Step 2, we use functional neuroimaging to identify the brain areas that are activated during the task. In Step 3, we compare the brain responses to those observed during other tasks that differ in distinct ways (e.g. reading the word rabbit versus naming a picture of a rabbit). In Step 4, we infer the type of processing that a brain region supports by examining how the brain region responds over a range of different conditions, using factorial designs, interactions and conjunctions (Price, Moore, Friston, 1997). In Step 5, we compare the processing associated with each brain region in Step 4, with processing specified in the models used in Step 1; and finally in Step 6, we update the cognitive model to be consistent with all sources of data.

How can neuroimaging help adjudicate between two different cognitive models? The example I will give returns to the nature of orthographic processing. Are the brain responses involved in orthographic processing specific to reading or does orthographic processing rely on visual processes (e.g. parallel feature recognition) that is common to orthographic and non-orthographic visual stimuli? To investigate this question, we can use neuroimaging to evaluate whether there are any brain responses that are specific to orthographic inputs (e.g. activated during reading aloud but not during object naming). Such regions could be defined as orthographic processing modules. The alternative result would be that all the brain areas that respond to orthographic stimuli also respond to non-orthographic stimuli. In this case, it is not accurate to describe the brain regions as orthographic processing modules but we can specify which areas are activated when orthographic stimuli are presented, and how the degree and timing of activation in these areas changes with the type of stimulus, task or person being tested.

In the 25 years that I have been conducting neuroimaging experiments, including more than 50 experiments on the neural basis of reading, I have never observed a brain area where the response was consistent with what would be expected if it was indeed dedicated to orthographic processing. This is in line with the conclusions of neuropsychological studies that have shown how reading impairments, in the absence of aphasia, are typically accompanied by difficulties processing non-orthographic stimuli when perception and attention are thoroughly tested, and response times are measured as well as

accuracy (see Farah and Wallace, 1990; Behrmann et al., 1998; Starrfelt et al., 2009; Starrfelt and Behrmann., 2011; Roberts et al., 2013).

The trouble with reaching convincing conclusions about cognitive functions from neuroimaging data is multi-faceted. Some of the problems encountered have led to suggestions that neuroimaging has not told us anything useful about cognitive models so far (Coltheart, 2006a,b, 2013). The counter arguments are that neuroimaging provides richer data sets for contrasting cognitive models (Turner et al., 2013) and allows us to distinguish whether neural responses are selective to a specific function (or stimulus) or shared by multiple functions/stimuli (Mather et al., 2013b). While such debates are important for drawing attention to the type of inferences that can and cannot be made from neuroimaging data, there are multiple examples of how neuroimaging results have thrown into question the usefulness of cognitive models based on traditional neuropsychological data. The key point is that, if our traditional cognitive models are correct, we should be able to (i) map known cognitive functions/computations to brain responses; and (ii) use this knowledge to predict whether this type of processing was engaged in a new task on the basis of the activated brain areas (Rubin et al., 2017). In one sentence: Knowing what a region does, should indicate how this process contributes to a range of different cognitive tasks.

Below, I highlight five specific challenges involved in linking cognitive models to neuroimaging data. For consistency, the neural systems that support reading and object naming will be used as an example. I then highlight more generic principles learnt from neuroimaging that motivate a very thorough re-thinking of traditional cognitive models.

Five of the challenges involved in linking cognitive models to neuroimaging data

Defining regional specificity. When a region (or neuronal population) is selective for one type of stimulus over another, we cannot immediately rule out the possibility that there might be other types of stimuli or tasks that can activate the same region. For example, if we found an area that was activated for semantic decisions on written words but not for semantic decisions on picture of objects, we might conclude that this area was an “orthographic processing module” (Glezer et al., 2009). However, another experiment might reveal that the so called “orthographic processing module” responded more during object naming than reading aloud (Price et al., 2006). In this case, we cannot conclude that the region is dedicated to orthographic processing. Instead, we might hypothesise that the region played a role in linking familiar visual stimuli to their names, which occurs automatically during presentation of written words (irrespective of task), but only for pictures of objects when the task requires name retrieval (Glaser and Glaser, 1989). Pursuing strong evidence for “specificity” therefore requires many time consuming and expensive neuroimaging experiments to test all possible alternative hypotheses.

Specificity in functional connectivity: The absence of a brain region that is dedicated to orthographic processing does not exclude the possibility that there are other types of neural responses that distinguish orthographic from all other types of non-orthographic processing. For example, specificity for orthography might be observed in the combination of regions that are activated (even if each region is involved in many other functions), or it might be in the way that different regions communicate with one another (i.e. specificity in the functional connectivity). To test these hypotheses, we need to report (i) the relative degree of activation in distinct regions for different stimuli and tasks, (ii) how these regions connect to each other during different conditions, and (iii) the relative timing of activation in each region. Delivering the relevant data to test these hypotheses has been much more challenging than it might sound because it requires neural measurements and analyses that combine results from different techniques (with high spatial and temporal resolution) and across multiple different brain regions. This may take decades given the time, expertise and expense involved.

The spatial scale of specificity. When different types of stimuli and task activate the same brain regions, it could be argued that the spatial resolution of the neuroimaging technique was not sufficient to distinguish processing specific neurons. We have considered this possibility by using neuroimaging data with very high spatial resolution (Wright et al., 2008; Figure 4). Although we cannot rule out the possibility that there are orthographic specific neurons within commonly activated brain regions, the spatial resolution of neuroimaging data is much higher than that of lesions. A null result in neuroimaging (common activation for different stimuli) should therefore also result in a null result in neuropsychological data (impairment in all tasks that rely on the damaged functions/computations).

Figure 4 about here

Specificity in neuropsychological data that is not observed in neuroimaging data.

If neuroimaging shows that a cognitive function of interest activates brain regions that are also strongly activated by other functions, how can brain damage disproportionately affect one function relative to the other? For example, neuroimaging studies have shown that a region in the left ventral occipital cortex – often referred to as the visual word form area – is more activated by object naming than by reading (Price et al., 2006). In contrast, neuropsychological studies have shown that damage to this region can cause more severe difficulty recognising written words than objects (Starrfelt et al., 2009). This could occur if the undamaged (preserved) brain regions are able to support the recovery of object recognition but not that of reading. For example, if the damage the patient has incurred

preserves brain structures that provide partial visual clues (global shape/distinguishing features), this processing might be sufficient for accurate recognition of objects with distinctive shapes but not sufficient for accurate recognition of words which have very similar visual structures and therefore require parallel processing of multiple visual features. As noted above, however, a relative difference in accuracy between object and word recognition does not imply that object recognition is normal when response times are taken into account (Behrmann et al., 1998; Starrfelt et al., 2009; Starrfelt and Behrmann., 2011; Roberts et al., 2013).

Other complexities and inconsistencies: Perhaps the greatest challenge in relating neuroimaging data to cognitive models is that specific cognitive functions (e.g. orthographic processing) are typically associated with activation in multiple brain regions. This is known as “distributed processing”. At a superficial level, distributed processing is not necessarily challenging for cognitive models. It just adds complexity to the description of the functional anatomy. For example, we can describe the functional anatomy of semantic processing as a distributed set of brain regions that includes anterior and posterior inferior temporal cortex and the angular gyri (Binder et al., 2011). However, when different experiments report different sets of regions for the same type of processing, we need theoretical accounts to explain why the brain activations are changing across studies: Does each region within a system support computational functions that are not yet specified in the cognitive model? Or does variation from one study to another reflect inter-subject variability in cognitive strategy or the brain regions that support a particular type of processing? These questions need to be addressed by re-considering the computational mechanisms that are shared by different tasks (Patterson and Lambon Ralph, 1999; Price and Friston, 2005; Humphreys & Lambon Ralph, 2015) and conducting more experiments with large samples of participants so that inter-subject variability can be investigated.

Implications of neuroimaging findings for interpreting neuropsychological data and re-thinking cognitive models

Despite the challenges faced, neuroimaging has already transformed our understanding of cognitive models. This has happened indirectly by gradually illustrating that well-established cognitive models (i.e. those taught in psychology text books) are not sufficiently detailed to predict patterns of brain activation during tasks that are known to weight different types of processing. One might argue that it is not important for cognitive models to predict brain responses. However, as I will describe below, the validity of current cognitive models is directly challenged by observations that the mapping between known

cognitive functions/computations and brain structures does not indicate a one-to-one relationship. Many regions can be associated with the same type of processing, and conversely, the same brain region can be assigned multiple different functions depending on the cognitive model that the investigator is using. Together, this results in a many-to-many mapping between cognitive function and brain structure which is highly relevant for understanding the basic computations of the human mind, how they break down after brain damage, and how the brain might compensate for the lost functions. Below, I provide four examples of how this many-to-many mapping is informing cognitive models.

Introducing new cognitive functions. As described above, neuroanatomical descriptions that associate multiple brain regions with a single cognitive function become unmanageably complex in the context of observing that the set of regions that comprise the distributed neural system for one type of processing do not always activate together and may differentially contribute to other neural systems involved in other types of processing. In this case, a neuroanatomist might ask: What does each region do? How do we predict when a region will be activated or not? How do different computations combine to generate increasingly complex cognitive functions? If successful, the neuroanatomist will have created their own cognitive model with their own processing components. These biologically informed cognitive models can be compared to those based on behavioural data alone – and will have the added advantage that they predict brain responses from behaviour, and conversely should be able to predict behaviour from brain damage.

A recent example of a type of processing that was not predicted from cognitive models comes from an investigation into the response properties of different parts of the supramarginal gyrus (Oberhuber et al., 2017), see Figure 5. In this study, we compared activation for 8 types of stimuli (familiar written words, familiar heard words, unfamiliar written pseudowords, unfamiliar heard pseudowords, pictures of objects, sounds of objects, meaningless colour patterns and meaningless humming in male and female voices), each presented during two different tasks (speech production and one back matching). Within the supramarginal gyrus, we found a ventral posterior part that responded when written or spoken speech was being processed (consistent with some form of phonological processing) and a more anterior ventral part that responded to the demands on articulatory planning (consistent with output from phonology). However, it was difficult to identify the type of processing that was supported by the other two regions. For example, a posterior part of the dorsal posterior supramarginal gyrus (blue in Figure 5) was more activated for reading and repeating words than all other conditions which collectively controlled for semantic processing, lexical phonological retrieval, the mapping of orthography to phonology and

speech output. Previous neuroimaging studies have shown that the same dorsal posterior supramarginal region has also been associated with the acquisition of vocabulary (Richardson et al., 2010) and is anatomically linked to semantic processing regions in the angular gyrus and speech processing regions in more anterior supramarginal gyrus areas (Lee et al., 2007). We therefore hypothesized that it was actively involved in integrating lexical and sublexical phonological inputs which is more important when reading or repeating words than any other condition (Oberhuber et al., 2017). Critically, this type of processing is not specified in any of the boxes or arrows of traditional cognitive models of reading (see Figure 3).

Figure 5

Re-defining old cognitive functions. As just described, we can sometimes forge new cognitive labels to explain the response properties of a region. One of the consequences of this is that different researchers will produce different labels for the same region, depending on their interests. For example, a region in the left ventral occipito-temporal cortex is referred to as a visual word form area” by those who are interested in reading (Dehaene and Cohen, 2011), “the ventral object recognition system” by those who are interested in object recognition (Kravitz et al. 2013), and a region that integrates visuospatial features abstracted from sensory inputs with higher level phonological and semantic representations (Price and Devlin, 2011). Again, one might argue that it is irrelevant whether a computation of interest shares a biological substrate with a computation from another cognitive model. However, observations that there are common components for different cognitive tasks provides unique insights into how we could ultimately generate a single cognitive framework that includes generalised sensory and motor functions that support multiple cognitive tasks. Figure 6 provides an example of a physiologically constrained cognitive model of heard and seen speech and non-speech processing.

FIGURE 6 about here

Although generalised models are inevitably more complex than those focusing on individual functions (e.g. Figure 6 vs. Figure 3), simplicity is gained when different task models are integrated into an internally consistent, scalable framework. We have referred to this type of modelling as “functional ontologies for cognition” (Price and Friston, 2005). The goal is to provide a framework of all the computations that are shared or distinct for a range of different cognitive tasks (e.g., reading, repetition, object and colour naming), with each computation corresponding to the type of processing that is implemented in functionally distinct brain regions. Obviously, achieving such a framework will be challenging because it

requires a standardized definition for cognitive processes across diverse communities of scientists, harmonization of conflicting results from different techniques and an enormous number of new studies and data to test the validity of the model (Hastings et al., 2014). This should surely be our goal.

Multiple routes for the same task. The third useful insight that “distributed processing” has offered cognitive models is that, when there is a wide network of multiple regions involved in the same task, these regions can inter-connect with one another in different ways. This provides a rich set of alternative neural pathways for translating the same sensory input into the same response output. We describe alternative neural pathways for the same task in terms of “degeneracy” (Price and Friston, 2002) which is defined (in Wikipedia) as “the ability of elements that are structurally different to perform the same function or yield the same output”. It is evident throughout biological and physical systems, for example, in genetic codes or body function. For example, most people can write their names with a pen in their right hand or their left hand, even if they are better using one than the other (due to inherent preferences and practice).

By testing the functional connectivity between different brain regions, neuroimaging has confirmed that there are indeed multiple ways that the same stimulus (e.g. a written word) can be converted into the same output (e.g. its spoken name). In neuropsychological terms, these alternative neural pathways can be equated to different processing routes or cognitive strategies (Binder et al., 2005; Mechelli et al., 2005; Graves et al., 2014; Hoffman et al., 2015) but at present there is not a clear correspondence between the number of neural pathways that are being identified (see Figure 7) and the number of routes included in traditional models (Figure 3).

Inter-subject variability. When there are multiple ways of doing the same thing, neurologically normal individuals differ in which cognitive strategy/neural pathway they prefer to use (Seghier et al., 2008; Kherif et al., 2009; Miller et al., 2012; Hoffman et al., 2015). Neuroimaging has provided rich evidence for inter-subject variability in the degree to which neurologically normal individuals use different pathways (Figure 7). The next step is to understand whether there are any behavioural or demographic markers that indicate which neural pathway an individual is likely to be using. This would allow us to link cognitive strategies to neural pathways and examine how these pathways/strategies are learnt or relearnt (see Section 3 for a more detailed discussion).

Figure 7 about here.

Section 3: Combining neuropsychology with neuroimaging

This section briefly summarises how neuroimaging can be used to inform neuropsychological studies and how the integration of results from both types of data provides the most clinically useful, physiologically plausible, models of cognition.

Understanding co-occurring functional impairments: Using behavioural data alone, a neuropsychologist can identify which type of processing a patient has impaired or preserved, as well as the severity of each impairment. The goal is to show that distinct computations can be independently impaired in different patients (see Section 1). It is also possible to (i) show how the same impairment can affect multiple disparate task domains (Patterson et al., 2006), (ii) explore how individual differences in specific processing impairments affect performance on tasks of interest (Wooliams et al. 2007) and (c) use principal component analyses (PCA) on larger scale behavioural data to identify, in a data driven manner, patterns of co-occurrence in neuropsychological data (Butler et al., 2014; Halai et al., 2016).

Nevertheless, it can be challenging to interpret why two different functions (defined on the basis of the cognitive model being tested) are always observed to be impaired together (e.g. word and object recognition). In this case, behavioural data alone cannot distinguish between two alternative explanations: (1) that co-occurring impairments are the consequence of the two functions being co-located (i.e. in close proximity in the brain) and are therefore commonly affected by brain damage; or (2) that both cognitive functions rely on another undefined lower level function that explains both co-occurring functional impairments but is not part of the cognitive model being tested.

With functional neuroimaging studies of neurologically normal individuals, we can test if the two different types of processing are co-located in the brain. With structural neuroimaging of the patients, we can test whether the lesion site in the patient includes the areas that are normally activated by the lost computations; or whether there is damage to the white matter pathways that connect different cortical and subcortical regions. By integrating all the available data with a good prior knowledge of the function of different brain regions, we can make informed hypotheses about which types of processing are likely to be impaired.

Same functional impairment results in different symptoms. Despite the challenges (e.g. of co-occurring deficits), some patients have selective deficits that fit with meaningful functional impairments. Take the case of “anomia” as an example. Patients with anomia have good object recognition, semantic memory and auditory repetition skills but struggle

when trying to retrieve the names of familiar objects. Their impairment can therefore be described at the level of “word retrieval”. Nevertheless, patients with supposedly common functional impairments may vary in the severity and duration of their anomia, the type of errors they make and also in their ability to perform other functions. Using behavioural data alone, we do not know if this inter-patient variability is due to differences in: (i) the type of processing that is supported by specific brain regions, (ii) lesion site or (ii) the set of computations (and regions) used to complete a specific task (i.e. the cognitive strategy). With functional neuroimaging of neurologically normal individuals, we can investigate the degree of inter-subject variability in functional anatomy (how consistently does a region respond to a specific type of processing). With structural neuroimaging, we can stratify patients in neuropsychological studies on the basis of their lesion site.

After stratifying patients on the basis of their lesion site, we can investigate how matching lesion sites can have inconsistent effects in different patients. This requires an understanding of inter-subject variability in cognitive strategy before and after brain damage. We can investigate the neural pathways (i.e. sets of brain regions and their functional connectivity) that an individual uses to perform a task using functional neuroimaging. By comparing the identified pathways across neurologically normal controls and patients with specific lesion sites, we can estimate how many pathways there might be and how frequently each pathway is adopted. This may provide vital clues as to how cognitive functions can be recovered after brain damage. For example, if we establish from neurologically normal individuals that a task is typically performed by one of two possible pathways, the effect of damage to only one of these pathways will depend on whether the patient pre-dominantly used the damaged or undamaged pathway prior to their stroke. Let's refer to the pathway a patient uses most as their “dominant pathway” and the pathway they use less as their “non-dominant pathway”. If the non-dominant pathway is damaged, the ability to perform the task should not be severely affected because the patient can still use their dominant pathway. On the other hand, if the dominant pathway is damaged, then the patient needs to use another pathway (i.e. functionally re-organise) which might happen spontaneously or require practice and/or intervention.

By conducting an iterative combination of lesion, behavioural and functional neuroimaging studies in large cohorts of patients with diverse and carefully matched lesion sites, we can document which neural pathways are used for a given task and the conditions that determine when a pathway will be adopted. This will involve searching for any markers (behavioural, demographic or neuroimaging) that are associated with a given pathway. In this way, neuroimaging and the concept of degeneracy provide a framework for investigating (i) how a cognitive system can survive damage; (ii) how functional reorganisation can be

supported, (iii) how the influence of behavioural and demographic factors (e.g. vision, sight, general health) depends on lesion site and (iv) how this information can be used to predict the speed of recovery that a new patient could make.

Summary and Conclusions

In Section 1, I raised three types of observations that hamper the interpretation of neuropsychological data. The first is the lack of “pure specificity”. Although patients might be more impaired using one type of processing than another, cognitive impairments after brain damage are rarely confined to one type of processing. The second and third lessons were that the same symptom can arise from many different types of impairments and, conversely, the same underlying impairment can result in very different symptoms. In Section 2, I discussed the challenges faced when trying to test cognitive models with neuroimaging. I then proposed that the neuroimaging perspective that provides the greatest insight into cognitive models involves a many-to-many mapping between brain structure and known cognitive functions/computations/types of processing. This has led to proposals for new computational processes, new cognitive strategies and new ontologies for cognition that predict structure from function and function from structure (Rubin et al., 2017). It also provides an organised framework for understanding how cognitive functions are learnt or relearnt after damage and how learning can be influenced by training.

Finally, in Section 3, I considered how neuroimaging can inform neuropsychological studies by distinguishing between three different types of inter-patient variability: differences in premorbid functional anatomy, lesion site or cognitive strategy. By unveiling all possible neural pathways for a task, in normal and damaged brains, we can reconsider the underlying computational units and processing pathways. Moreover, by combining all sources of data, we can generate cognitive models that will be most informative for predicting, explaining and improving cognitive function after brain damage.

Figure legends

Figure 1: The Neurological model of Language.

An illustration of the anatomical and functional processing pathways that were hypothesized on the basis of post mortem studies conducted in the late 19th Century.

Figure 2: The “same” lesion site can have different effects in different patients.

MRI images from two different patients who both have damage to the left angular gyrus. This caused reading and writing difficulties in Patient 1 but not in Patient 2.

Figure 3: Mapping cognitive processes to brain activation.

Left: Dual route model of reading that dissociates the processing of lexical and sublexical orthographic inputs. Middle: A simpler model of reading that explains the same symptoms without having separate pathways for lexical and sublexical processing (Seidenberg and McClelland, 1989; Plaut et al., 1996). Black text is used for inputs, processing type and outputs. White text, in grey boxes, indicates the type of stimulus that would be impaired when a specific pathway was damaged. Right: Brain activations for reading segregated into visual/orthographic (red), semantic (yellow) and speech production (green) processing. The brain activations shown were identified by re-analysis of data from the experiment described in Seghier and Price (2012).

Figure 4: Common activation in the left ventral occipito-temporal “reading area” for reading aloud and object naming. Top: a section from a T1 image of the whole brain highlighting the area (in white dashed box) that we focus on in our high resolution functional neuroimaging data below. Middle: Red shows activation during reading aloud highlighting the left ventral occipito-temporal reading area in a white circle. Bottom: Yellow shows activation in exactly the same place for object naming. Data from Wright et al., (2008).

Figure 5: Four different reading responses in the left supramarginal gyrus are not predicted by cognitive models. The blue region was more activated for reading and repeating words than naming objects (from pictures or sounds), reading/repeating pseudowords and colour/gender naming. Activation is therefore highest when phonology can be generated from both lexical and sublexical phonological information. This was not expected from the cognitive models shown in Figure 3. The yellow region was more activated by reading pseudowords and naming objects than reading words or any of the other conditions. This cannot be explained by any the cognitive processes in Figure 3 (see Oberhuber et al., 2017). The green area was more activated for reading words and

pseudowords than naming objects or colours. This would be consistent with “the mapping of orthography to phonology”, except that the same region was more activated by all auditory stimuli than all visual stimuli, even when stimuli and tasks did not involve speech processing. The brown area was activated during speech production compared to one-back matching on the same stimuli. Activation was also higher for producing different object names on every trial compared to naming a limited number of colours repeatedly. Its response was therefore consistent with overt speech articulation. The white area within the black border shows other parts of the supramarginal gyrus as defined anatomically according to the IBASPM software in SPM 12. Data and explanations are from Oberhuber et al., (2017).

Figure 6: A physiologically constrained model of word processing

The model describes processing that is required for speech recognition but not specific to speech recognition. Incoming visual or auditory stimuli (e.g. a written or spoken word) are first processed in the primary sensory areas of the brain. By integrating these sensory features with prior knowledge, we form a visual or auditory mental image of the presented stimulus (that the subject may or may not be aware of). Auditory images of speech are equivalent to phonological (input) representations but the model uses generic terms to emphasize that the same brain regions are also involved in auditory images of non-speech sounds (Dick et al., 2007; Leech et al., 2009; Price et al., 2005; Saygin et al., 2003). If the sensory inputs carry semantic cues (e.g. familiar words, pictures of familiar objects or sounds of familiar objects), semantic associations can be retrieved and linked to the articulatory patterns associated with the word or object name (word retrieval stage). If there are no semantic cues available, articulatory plans can only be retrieved from the non-semantic parts of speech stimuli, e.g. the sublexical parts of an unfamiliar pseudoword (a pronounceable nonword). Finally, the articulatory plans are used to drive motor activity in the face, mouth and larynx when the task involves a speech response. This generates auditory and somatosensory processing (i.e. we hear and feel the movement in the speech articulators). This model was adapted from that in Price (2012) and updated with Philipp Ludersdorfer and Marion Oberhuber.

Figure 7: Neural pathways for reading outnumber predictions from cognitive models.

(A) Brain regions and functional connections for reading that have been shown to dissociate for different types of word stimuli (Mechelli et al., 2005) and in different subjects reading the same words (Seghier et al., 2010; 2014; Richardson et al., 2011). The dissociation of these pathways can be demonstrated by showing that as use of one pathway increases, use of the other pathway decreases. (B) shows inter-subject variability in the engagement of the putamen reading pathway. Group 1 used the putamen (put) reading pathway more than other pathways. Group 2 did not use the putamen pathway more or less than other pathways. Group 3 used the putamen pathway less than the other pathways. Data from Seghier et al., (2010).

REFERENCES

Behrmann, M., Nelson, J., Sekuler, E.B. (1998). Visual complexity in letter-by-letter reading: "pure" alexia is not pure. *Neuropsychologia*. 36(11):1115-32.

Binder, J.R., Desai, R.H. (2011). The neurobiology of semantic memory. *Trends in Cognitive Sciences* 15(11),527-36. DOI:10.1016/j.tics.2011.10.001

Bub, D.N (1993). Jules Dejerine and His Interpretation of Pure Alexia. *Brain and Language* 45(4), 531-59. DOI:10.1006/brln.1993.1059

Butler, R.A., Lambon Ralph, M.A., Woollams, A.M. (2014). Capturing multidimensionality in stroke aphasia: mapping principal behavioural components to neural structures. *Brain*, 137(Pt 12):3248-66. doi: 10.1093/brain/awu286

Cohen, L., Martinaud, O., Lemer, C., Lehericy, S., Samson, Y., Obadia, M., Slachevsky, A., Dehaene, S.. (2003). Visual word recognition in the left and right hemispheres: anatomical and functional correlates of peripheral alexias. *Cereb Cortex*. 13(12):1313-33.

Colby, K.M. (1978). Mind models: an overview of current work *Mathematical Biosciences* 39(3-4), 159-185. [https://doi.org/10.1016/0025-5564\(78\)90052-4](https://doi.org/10.1016/0025-5564(78)90052-4).

Coltheart, M. (2006a). What has functional neuroimaging told us about the mind (so far)? *Cortex* 42(3), 323-31. [https://doi.org/10.1016/S0010-9452\(08\)70358-7](https://doi.org/10.1016/S0010-9452(08)70358-7).

Coltheart, M. (2006b). Perhaps functional Neuroimaging has not told us anything about the mind (so far). *Cortex* 42(3):422-427. [https://doi.org/10.1016/S0010-9452\(08\)70374-5](https://doi.org/10.1016/S0010-9452(08)70374-5)

Coltheart, M. (2013) How Can Functional Neuroimaging Inform Cognitive Theories? *Perspect Psychol Sci*. 8(1):98-103. doi: 10.1177/1745691612469208.

Coltheart, M., Patterson, K., Marshall, J.C. Editors. (1980). *Deep dyslexia*. ([1st publ.] ed.). London: Routledge & Kegan Paul. <https://www.isbns.net/isbn/9780710004567/> ISBN 0-7100-0456-7.

Coltheart M, Rastle K, Perry C, Langdon R, Ziegler J. (2001).DRC: a dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*;108(1):204-56.

Caramazza, A., Hillis, A.E.(1990) Where Do Semantic Errors Come From? *Cortex*, 26 (1), 95–122.

Dehaene, S., Cohen, L. (2011). The unique role of the visual word form area in reading. *Trends in Cognitive Sciences* 15(6), 254-62. DOI: 10.1016/j.tics.2011.04.003.

Dick, F., Saygin, A.P., Galati, G., Pitzalis, S., Benvivuto, S., D'Amico, S., Wilson, S., Bates, E., Pizzamiglio, L. (2007) What is involved and what is necessary for complex linguistic and nonlinguistic auditory processing: evidence from functional magnetic resonance imaging and lesion data. *Journal of Cognitive Neuroscience*, 19:799-816.

Farah, M.J. Wallace, M.A. (1991). Pure Alexia as a Visual Impairment: A Reconsideration *Cognitive Neuropsychology* 8(3-4), 313-334. <http://dx.doi.org/10.1080/02643299108253376>

Glaser, W. R., Glaser, M. O. (1989). Context effects in Stroop-like word and picture processing. *Journal of Experimental Psychology: General* 118, 13–42.

Glezer, L. S., Jiang, X., & Riesenhuber, M. (2009). Evidence for highly selective neuronal tuning to whole words in the “visual word form area”. *Neuron*, 62(2), 199-204.

Halai, A.D., Woollams, A.M., Lambon Ralph, M.A.. (2017). Using principal component analysis to capture individual differences within a unified neuropsychological model of chronic post-stroke aphasia: Revealing the unique neural correlates of speech fluency, phonology and semantics. *Cortex*. 86:275-289. doi: 10.1016/j.cortex.2016.04.016.

Hastings, J., Frishkoff, G.A., Smith, B., Jensen, M., Poldrack, R.A., Lomax, J., Bandrowski, A., Imam, F., Turner, J.A., Martone, M.E. (2014). Interdisciplinary perspectives on the development, integration, and application of cognitive ontologies. *Frontiers in Neuroinformatics* 8, 62. DOI: 10.3389/fninf.2014.00062.

Graves, W.W., Binder, J.R., Desai, R.H., Humphries, C., Stengel, B.C., Seidenberg, M.S. (2014). Anatomy is strategy: skilled reading differences associated with structural connectivity differences in the reading network. *Brain Lang.* 133:1-13. doi: 10.1016/j.bandl.2014.03.005.

Harley, T.A. (2004) Promises, promises. *Cogn Neuropsychol.* 21(1):51-6. doi: 10.1080/02643290342000212.

Henson, R. (2006). What has (neuro)psychology told us about the mind (so far)? *Cortex* 42, 387-392. [https://doi.org/10.1016/S0010-9452\(08\)70365-4](https://doi.org/10.1016/S0010-9452(08)70365-4).

Hoffman, P., Lambon Ralph, M.A., Woollams, A.M. (2015). Triangulation of the neurocomputational architecture underpinning reading aloud. *Proc Natl Acad Sci U S A.* 112(28):E3719-28. doi: 10.1073/pnas.1502032112.

Humphreys, G.F., Lambon Ralph, M.A. (2013). Fusion and Fission of Cognitive Functions in the Human Parietal Cortex. *Cereb Cortex.* 2015 Oct;25(10):3547-60. doi: 10.1093/cercor/bhu198.

Humphreys, G.W. Evett, L.J. (1985). Are there independent lexical and nonlexical routes in word processing? An evaluation of the dual-route theory of reading. *Behavioral and Brain Science* 8(4), 689-705.

Humphreys, G.W., Price, C.J., Riddoch, M.J. (1999). From objects to names: a cognitive neuroscience approach. *Psychological Research* 62(2-3), 118-30. DOI: 10.1007/s004260050046

Humphreys, G.W., Riddoch, M.J. (1984). Routes to object constancy: implications from neurological impairments of object constancy. *The Quarterly Journal of Experimental Psychology Section A.* 36(3), 385-415. <http://dx.doi.org/10.1080/14640748408402169>

Humphreys, G.W., Riddoch, M.J. (1987). On telling your fruit from your vegetables: a consideration of category-specific deficits after brain damage. *Trends in Neurosciences* 10(4), 145-148. [https://doi.org/10.1016/0166-2236\(87\)90040-3](https://doi.org/10.1016/0166-2236(87)90040-3)

Humphreys, G.W., Riddoch, M.J., Price, C.J. (1997). Top-down processes in object identification: evidence from experimental psychology, neuropsychology and functional anatomy. *Philosophical Transactions of the Royal Society of London B Biological Sciences* 352(1358),1275-82. DOI: 10.1098/rstb.1997.0110

Humphreys, G.W., Riddoch, M.J., Quinlan, P.T., Price, C.J., Donnelly, N. (1992). Parallel pattern processing and visual agnosia. *Canadian Journal of Psychology* 46(3), 377-416.

Jack, I.A., Sylvester, C.M., Corbetta, M. (2006). Losing our brainless minds: How neuroimaging informs cognition. *Cortex*, 42, 418-421. [https://doi.org/10.1016/S0010-9452\(08\)70373-3](https://doi.org/10.1016/S0010-9452(08)70373-3)

Jonides, J., Nee, D.E., Berman, M.G. (2006). What has functional neuroimaging told us about the mind? So many examples, so little space. *Cortex* 42(3) 414-7. [https://doi.org/10.1016/S0010-9452\(08\)70372-1](https://doi.org/10.1016/S0010-9452(08)70372-1).

Kherif, F., Josse, G., Seghier, M.L, Price, C.J. (2009). The main sources of intersubject variability in neuronal activation for reading aloud. *J Cogn Neurosci*. 21(4):654-68. doi: 10.1162/jocn.2009.21084.

Kinsbourne, M. and Warrington, E.K. (1962). A disorder of simultaneous form perception. *Brain*. 85:461-86.

Kravitz, D.J., Saleem, K.S., Baker, C.I., Ungerleider, L.G., Mishkin, M. (2013). The ventral visual pathway: an expanded neural framework for the processing of object quality. *Trends in Cognitive Sciences* 17(1), 26-49. DOI: 10.1016/j.tics.2012.10.011.

Lee, H., Devlin, J.T., Shakeshaft, C., Stewart, L.H., Brennan, A., Glensman, J., Pitcher, K., Crinion, J., Mechelli, A., Frackowiak, R.S., Green, D.W., Price, C.J. (2007). Anatomical traces of vocabulary acquisition in the adolescent brain. *Journal of Neuroscience* 27(5), 1184-9. DOI:10.1523/JNEUROSCI.4442-06.2007

Leech, R., Holt, L.L., Devlin, J.T., Dick, F. (2009) Expertise with artificial nonspeech sounds recruits speech-sensitive cortical regions. *Journal of Neuroscience*, 29:5234-5239.

Leff, A.P., Spitsyna, G., Plant, G.T., Wise, R.J. (2006). Structural anatomy of pure and hemianopic alexia. *Journal of Neurology, Neurosurgery and Psychiatry* 77(9), 1004-7. <http://dx.doi.org/10.1136/jnnp.2005.086983>

- Liepmann, H. (1900). Das Krankheitsbild der apraxie (motorischen asymbolie). *Monatsschrift für Psychiatrie und Neurologie* 8, 15-44.
- Marr, D., Poggio, T., (1977). "From understanding computation to understanding neural circuitry", in *Neuronal Mechanisms in Visual Perception*. Eds E., Poppel, R, Held, J.E. Dowling. *Neurosciences Research Program Bulletin* 15(3), 470–488.
- Marshall, J.C., Newcombe, F. (1973). Patterns of paralexia: a psycholinguistic approach. *Journal of Psycholinguistic Research* 2(3),175-99. DOI: 10.1007/BF01067101
- Mather, M., Cacioppo, J.T., Kanwisher, N. (2013a) Introduction to the Special Section: 20 Years of fMRI-What Has It Done for Understanding Cognition? *Perspect Psychol Sci.* 8(1):41-3. doi: 10.1177/1745691612469036.
- Mather, M., Cacioppo, J.T., Kanwisher, N. (2013b) How fMRI Can Inform Cognitive Theories. *Perspect Psychol Sci.* 8(1):108-13. doi: 10.1177/1745691612469037.
- McCarthy, R.A, Warrington, E.K. (1990). *Cognitive Neuropsychology: A Clinical Introduction* (Cognitive neuropsychology: a clinical introduction ed.). San Diego: Academic Press. <https://www.abebooks.co.uk/book-search/isbn/0124818463/> ISBN 0124818463.
- Mechelli, A., Crinion, J.T., Long, S., Friston, K.J., Lambon, Ralph, M.A., Patterson, K., McClelland, J.L., Price, C.J. (2005). Dissociating reading processes on the basis of neuronal interactions. *Journal of Cognitive Neuroscience* 17(11), 1753-65. <http://www.mitpressjournals.org/doi/pdf/10.1162/089892905774589190>.
- Miller, M.B., Donovan, C.L., Bennett, C.M., Aminoff, E.M., Mayer, R.E. (2012). Individual differences in cognitive style and strategy predict similarities in the patterns of brain activity between individuals. *Neuroimage.* 59(1):83-93. doi: 10.1016/j.neuroimage.2011.05.060.
- Patterson, K., Lambon Ralph, M. A. (1999). Selective disorders of reading? *Current Opinion in Neurobiology.* 9:235–239.
- Patterson, K., Lambon Ralph, M.A., Jefferies, E., Woollams, A., Jones, R., Hodges, J.R., Rogers, T.T. (2006). "Presemantic" cognition in semantic dementia: six deficits in search of an explanation. *J Cogn Neurosci.* 18(2):169-83.

Price, C.J., McCrory, E., Noppeney, U., Mechelli, A., Moore, C.J., Biggio, N., Devlin, J.T. How reading differs from object naming at the neuronal level. *Neuroimage*. 2006 Jan 15;29(2):643-8.

Oberhuber, M., Hope, T.M., Seghier, M.L., Parker Jones, O., Prejawa, S., Green, D.W., Price, C.J. (2016). Four Functionally Distinct Regions in the Left Supramarginal Gyrus Support Word Processing. *Cerebral Cortex* 26 (11), 4212-4226.
DOI:<https://doi.org/10.1093/cercor/bhw251>

Page, M.P.A. (2006). What can't functional neuroimaging tell the cognitive psychologist? *Cortex*, 42, 428-443. [https://doi.org/10.1016/S0010-9452\(08\)70375-7](https://doi.org/10.1016/S0010-9452(08)70375-7).

Pearce, J.M.S. (2009). Hugo Karl Liepmann and apraxia. *Clinical Medicine (London)* 9(5), 466-70. DOI:10.7861/clinmedicine.9-5-466.
<http://www.clinmed.rcpjournal.org/content/9/5/466.long>

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

Poldrack, R.A. (2010). Mapping Mental Function to Brain Structure: How Can Cognitive Neuroimaging Succeed? *Perspective on Psychological Science* 5(6), 753-61. DOI: 10.1177/1745691610388777.

Price, C.J. (2012) A review and synthesis of the first 20 years of PET and fMRI studies of heard speech, spoken language and reading. *Neuroimage*. 15;62(2):816-47. doi: 10.1016/j.neuroimage.2012.04.062.

Price, C.J., Friston K.J. (2002). Degeneracy and cognitive anatomy. *Trends in Cognitive Sciences* 6(10), 416-421. [https://doi.org/10.1016/S1364-6613\(02\)01976-9](https://doi.org/10.1016/S1364-6613(02)01976-9)

Price, C.J., Friston, K.J. (2005). Functional ontologies for cognition: The systematic definition of structure and function. *Cognitive Neuropsychology* 22(3), 262-75.
DOI: 10.1080/02643290442000095.

Price, C.J. Humphreys, G.W. (1992). Letter by letter reading? Functional deficits and compensatory strategies. *Cognitive Neuropsychology* 9(5), 427-457.
<http://dx.doi.org/10.1080/02643299208252067>

Price, C.J. Humphreys, G.W. (1994). Attentional dyslexia, the effect of co-occurring deficits. *Cognitive Neuropsychology* 10(6), 569-592.
<http://dx.doi.org/10.1080/02643299308253474>.

Price, C.J., Moore, C.J., Humphreys, G.W., Frackowiak, R.S., Friston, K.J. (1996). The neural regions sustaining object recognition and naming. *Proceedings of the Royal Society B Biological Sciences* 263(1376), 1501-7. DOI:10.1098/rspb.1996.0219

Price, C.J., Moore, C.J., Friston, K.J. (1997). Subtractions, conjunctions, and interactions in experimental design of activation studies. *Hum Brain Mapp* 5, 264-272.
DOI:10.1002/(SICI)1097-0193.

Price, C.J., Moore, C.J., Humphreys, G.W., Wise, R.J. (1997). Segregating Semantic from Phonological Processes during Reading. *Journal of Cognitive Neuroscience* 9(6), 727-33.
DOI: 10.1162/jocn.1997.9.6.727.

Price, C., Thierry, G., Griffiths, T (2005) Speech-specific auditory processing: where is it? *Trends in Cognitive Science*, 9:271-276.

Richardson, F.M., Thomas, M.S., Filippi, R., Harth, H., Price, C.J. (2010). Contrasting effects of vocabulary knowledge on temporal and parietal brain structure across lifespan. *Journal of Cognitive Neuroscience* 22(5), 943-54. DOI: 10.1162/jocn.2009.21238.

Richardson, F.M., Seghier, M.L., Leff, A.P., Thomas, M.S., Price, C.J. (2011) Multiple routes from occipital to temporal cortices during reading. *Journal of Neuroscience*, 31(22):8239-47. doi: 10.1523/JNEUROSCI.6519-10.2011.

Riddoch, M.J., Humphreys, G.W. (1983). The effect of cueing on unilateral neglect *Neuropsychologia*, 21(6), 589-599.
<http://www.sciencedirect.com/science/article/pii/0028393283900568>

Riddoch, M.J., Humphreys, G.W. (1987). A case of integrative visual agnosia. *Brain* 110 (Pt 6), 1431-62. <https://doi.org/10.1093/brain/110.6.1431>

Riddoch, M.J., Humphreys, G.W. (1987). Visual Object Processing in Optic Aphasia: A Case of Semantic Access Agnosia. *Cognitive Neuropsychology* 4(2), 131-185. <http://dx.doi.org/10.1080/02643298708252038>

Riddoch, M.J., Humphreys, G.W., Price, C.J. (1989). Routes to action: Evidence from apraxia. *Cognitive Neuropsychology* 6 (5), 437-454. <http://dx.doi.org/10.1080/02643298908253424>

Roberts, D.J., Woollams, A.M., Kim, E., Beeson, P.M., Rapcsak, S.Z., Lambon Ralph, M.A. (2013). Efficient visual object and word recognition relies on high spatial frequency coding in the left posterior fusiform gyrus: evidence from a case-series of patients with ventral occipito-temporal cortex damage. *Cerebral Cortex*. 23(11):2568-80. doi:10.1093/cercor/bhs224.

Rubin, T.N., Koyejo, O., Gorgolewski, K.J., Jones, M.N., Poldrack, R.A., Yarkoni, T. (2017). Decoding brain activity using a large-scale probabilistic functional-anatomical atlas of human cognition. *PLoS Comput Biol*. 13(10):e1005649. doi: 10.1371/journal.pcbi.1005649.

Saygin, A.P., Dick, F., Wilson, S.M., Dronkers, N.F., Bates, E. (2003) Neural resources for processing language and environmental sounds: evidence from aphasia. *Brain* 126:928-945.

Seghier M.L., Lee, H.L., Schofield, T., Ellis, C.L., Price, C.J.(2008). Inter-subject variability in the use of two different neuronal networks for reading aloud familiar words. *Neuroimage*. 42(3):1226-36. doi: 10.1016/j.neuroimage.2008.05.029.

Seghier, M.L., Bagdasaryan J., Jung D.E., Price C.J. (2014) The importance of premotor cortex for supporting speech production after left capsular-putaminal damage. *J Neurosci.*, 34(43):14338-48. doi: 10.1523/JNEUROSCI.1954-14.2014.

Seghier, M.L., Price, C.J. (2010) Reading aloud boosts connectivity through the putamen. *Cerebral Cortex*. 2010;20:570–582. doi: 10.1093/cercor/bhp123

Seghier, M.L., Price, C.J. (2012). Functional Heterogeneity within the Default Network during Semantic Processing and Speech Production. *Front Psychol*. 2012 Aug 13;3:281. doi: 10.3389/fpsyg.2012.00281.

Seidenberg, M.S., McClelland, J.L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review* 96(4), 523-68.

<http://dx.doi.org/10.1037/0033-295X.96.4.523>

Seron, X., Fias, W. (2006). How images of the brain can constrain cognitive theory: The case of numerical cognition. *Cortex* 42, 406-410. [https://doi.org/10.1016/S0010-9452\(08\)70370-8](https://doi.org/10.1016/S0010-9452(08)70370-8)

Shallice, T. (1988). *From neuropsychology to mental structure* (Reprint. ed.). Cambridge [England]: Cambridge University Press.

p. 112. <http://www.cambridge.org/catalogue/catalogue.asp?isbn=0521308747>

ISBN 0521308747

Schutter, D.J.L.G., de Haan, E.H.F., van Honk, J. (2006). Scaling problems in the brain-mind conundrum. *Cortex* 42, 411-413. [https://doi.org/10.1016?20010-9452\(08\)70371-X](https://doi.org/10.1016?20010-9452(08)70371-X)

Starrfelt, R., Habekost, T., Leff, A.P. (2009). Too little, too late: reduced visual span and speed characterize pure alexia. *Cerebral Cortex*.19(12):2880-90. doi: 10.1093/cercor/bhp059.

Starrfelt, R., Behrmann, M. (2011). Number reading in pure alexia--a review. *Neuropsychologia*. 49(9):2283-98. doi: 10.1016/j.neuropsychologia.2011.04.028.

Uttal, W.R. (2001). *The New Phrenology. The Limits of Localizing Cognitive Processes in the Brain* Cambridge, MA. MIT Press. ISBN: 9780262210171.

Vallar, G. (2006). Mind, brain, and functional neuroimaging. *Cortex* 42, 402-405. [https://doi.org/10/1016/S0010-9452\(08\)70369-1](https://doi.org/10/1016/S0010-9452(08)70369-1).

Woollams, A.M., Lambon Ralph, M.A., Plaut, D.C., Patterson, K. (2010) SD-squared revisited: reply to Coltheart, Tree, and Saunders. *Psychol Rev.* 117(1):273-81; discussion 282-3. doi: 10.1037/a0017641. Review.

Wright, N.D., Mechelli, A., Noppeney, U., Veltman, D.J., Rombouts, S.A., Glensman, J., Haynes, J.D., Price, C.J. (2008). Selective activation around the left occipito-temporal sulcus for words relative to pictures: individual variability or false positives? *Hum Brain Mapp.* 29(8):986-1000. <https://doi.org/10.1002/hbm.20443>

Figure legends

Figure 1: The Neurological model of Language.

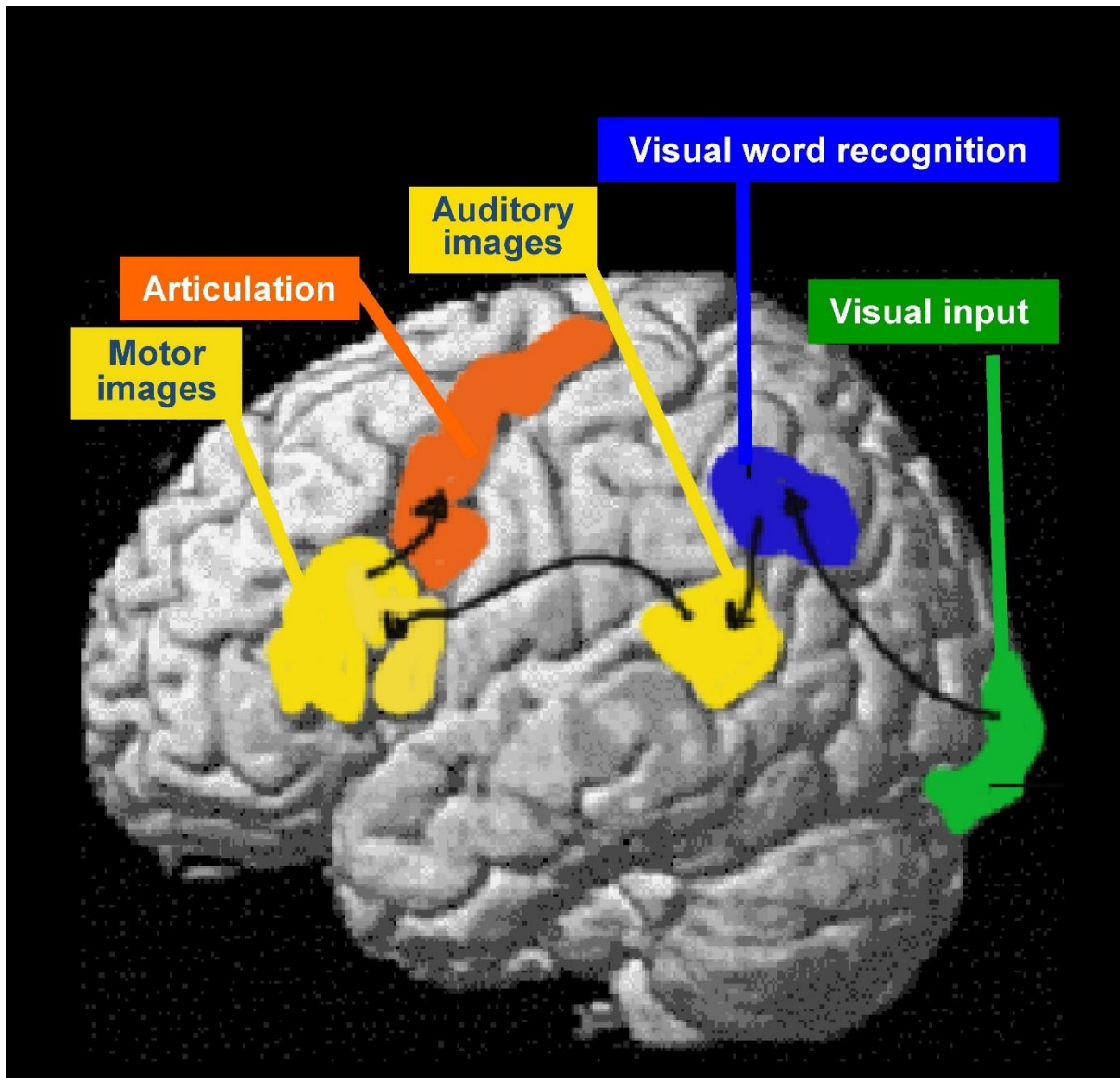


Figure 2: The “same” lesion site can have different effects in different patients.

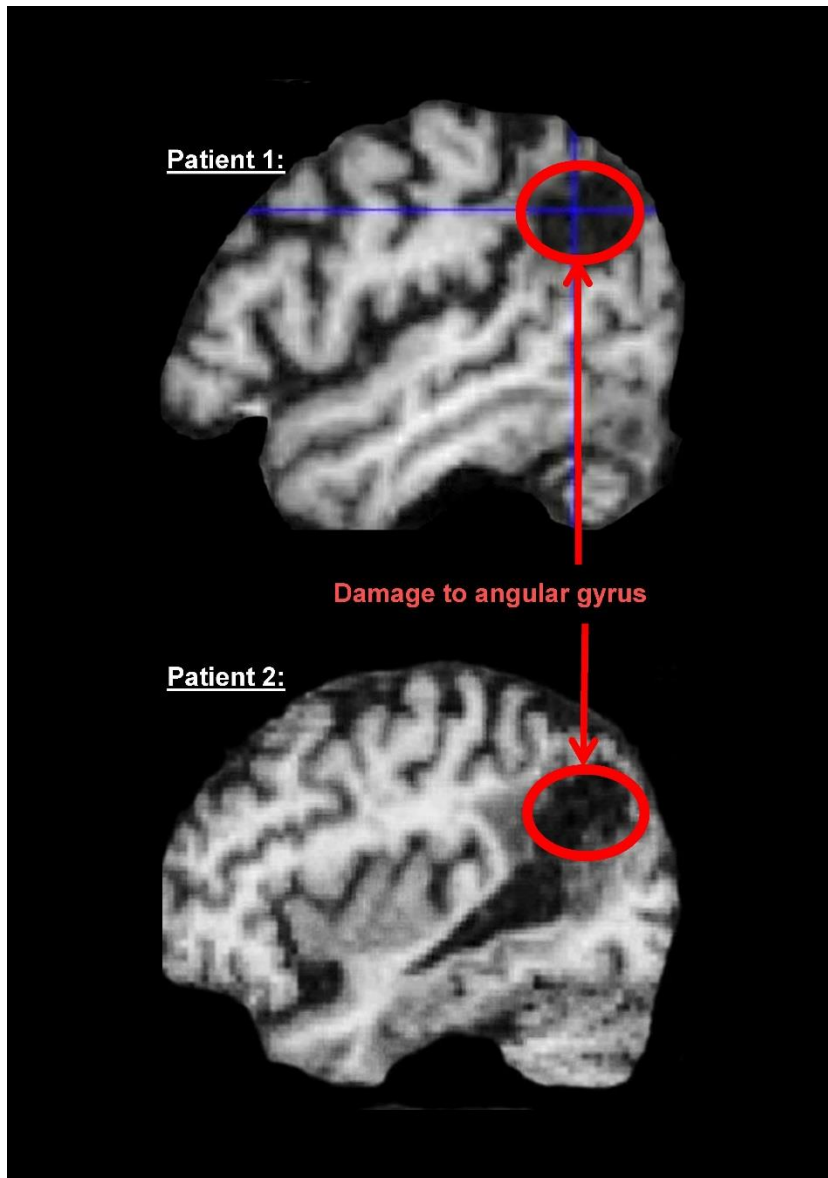


Figure 3: Mapping cognitive processes to brain activation.

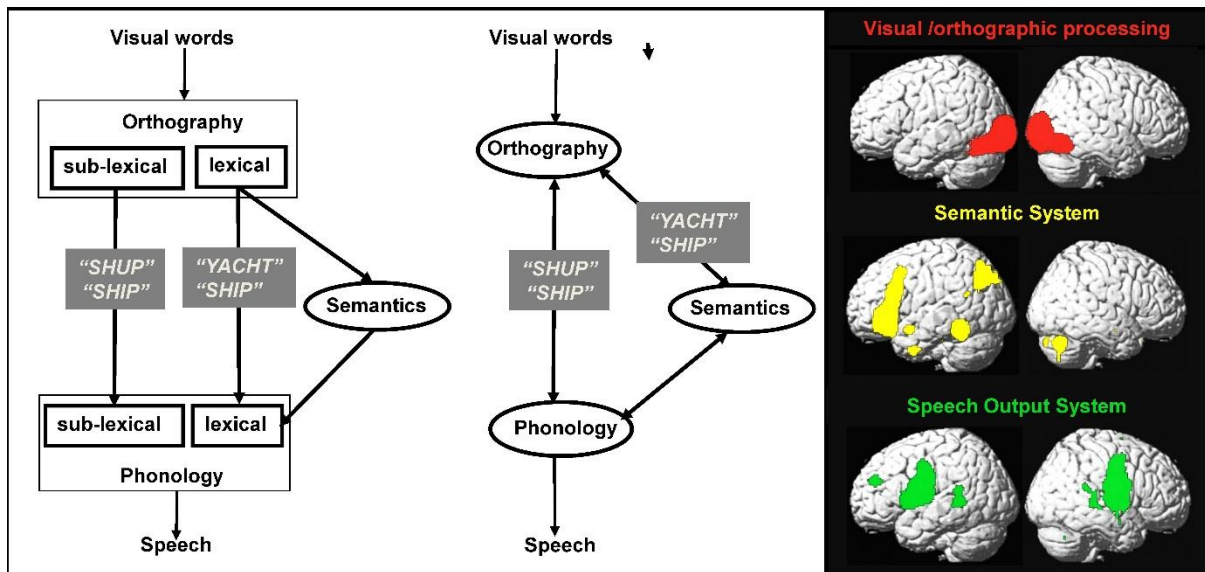


Figure 4: Common activation in the left ventral occipito-temporal “reading area” for reading aloud and object naming.



Figure 5: Four different reading responses in the left supramarginal gyrus are not predicted by cognitive models.

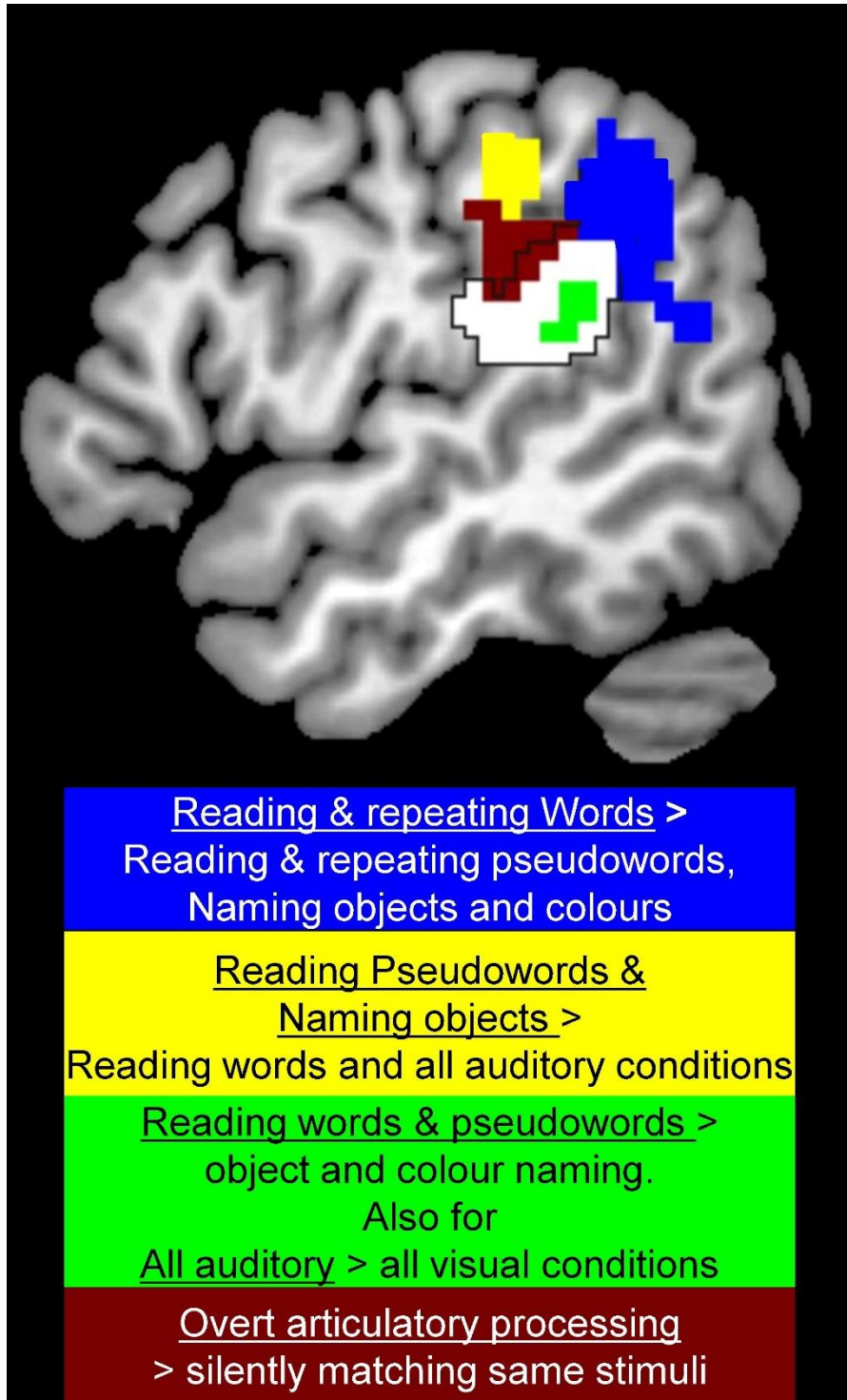


Figure 6: A physiologically constrained model of word processing

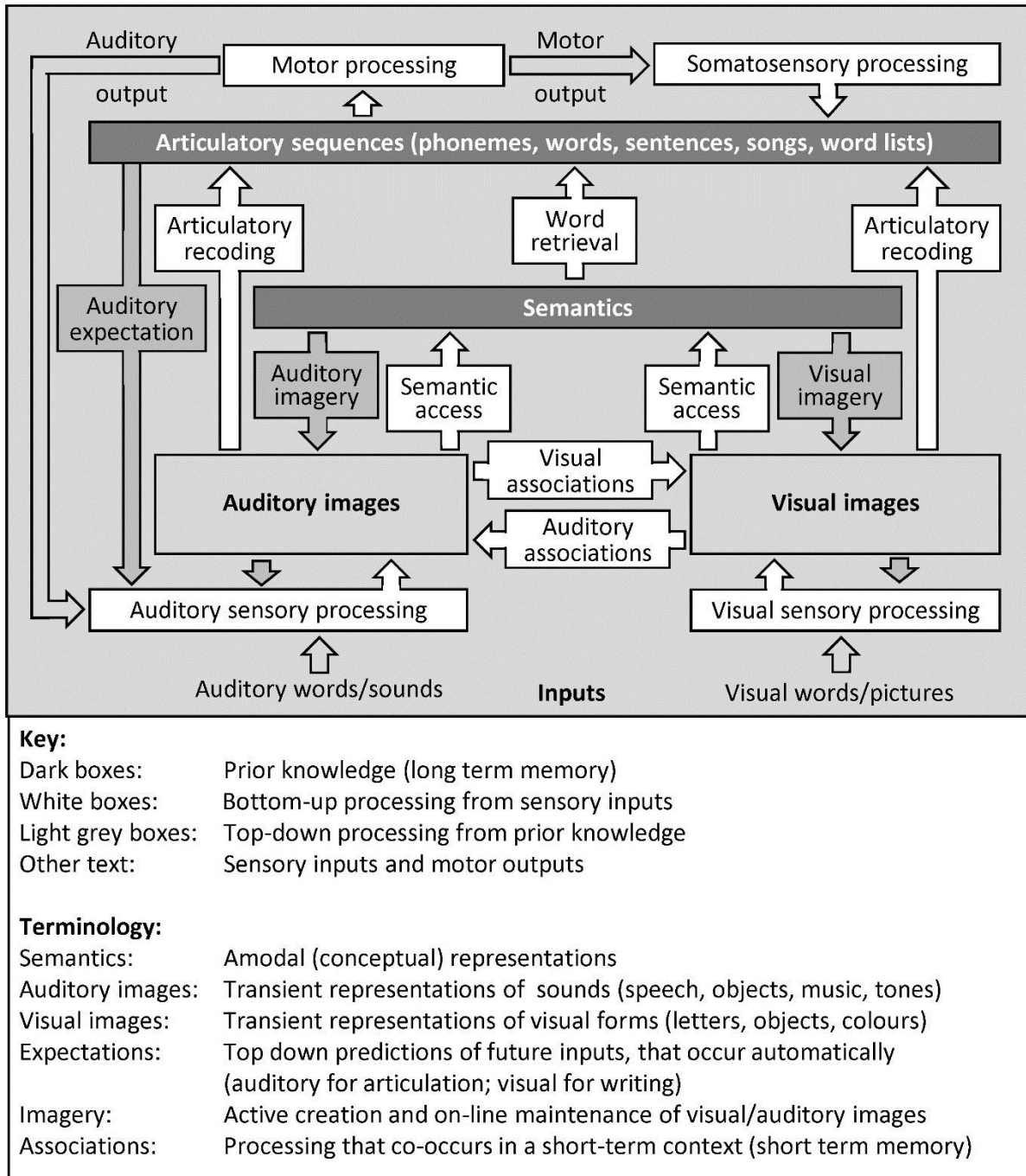
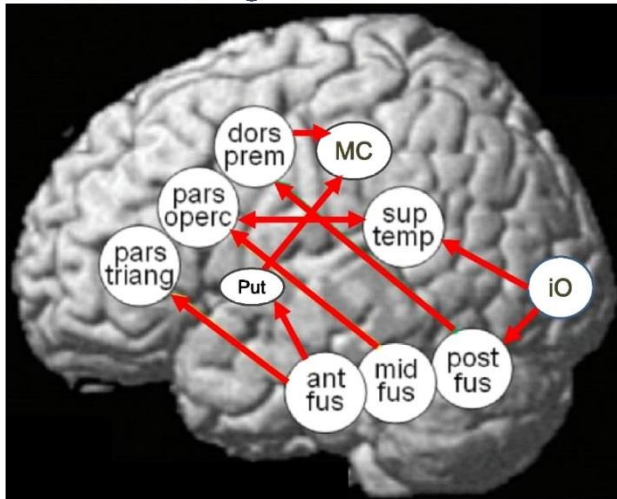


Figure 7: Neural pathways for reading outnumber predictions from cognitive models.

A) Dissociated reading routes



B) Inter-subject variability in processing route

