

Redundancy Gain in Semantic
Categorisation

Peter Vivian Shepherdson

a thesis submitted for the degree of
Doctor of Philosophy
at the University of Otago, Dunedin,
New Zealand.

March 5, 2013

Abstract

Redundancy gain refers to the common finding in experimental psychology that the presentation of multiple, redundant stimuli tends to evoke responses more quickly and accurately when compared to presentation of a single stimulus. Two types of account of such findings are generally offered. First, *race models* suggest that redundancy gain is a result of statistical facilitation. Second, *coactivation models* suggest that redundancy gain results from each stimulus making some contribution to the eventual response.

Though redundancy gain has primarily been shown in relatively simple tasks (e.g., detection, perceptual discrimination), there have also been efforts to demonstrate comparable phenomena in tasks involving higher-order cognition. One example of this is in the work of Mohr and Pulvermüller in lexical decision tasks (LDTs) using redundant stimuli (e.g., Mohr, Pulvermüller, & Zaidel, 1994; Mohr, Pulvermüller, Rayman, & Zaidel, 1994; Mohr, Pulvermüller, Mittelstädt, & Rayman, 1996; Mohr, Endrass, Hauk, & Pulvermüller, 2007). Those authors explained redundancy gain in LDT on the basis of a cell assembly model of lexical representation (e.g., Pulvermüller & Mohr, 1996; Pulvermüller, 1999). According to this explanation, activity from redundant stimuli sums in the network of cells where the word is neurally represented, leading it to “ignite” more rapidly and effectively — a form of coactivation — which in turn leads to faster and more accurate responses.

I sought to determine whether a similar phenomenon would occur in a semantic categorisation task, and whether the same basic model could be used to account for such findings. To investigate this issue, I conducted a series of experiments based on the LDTs used by Mohr and Pulvermüller. In my experiments participants were asked to classify visually-presented lexical stimuli as members or non-members of a pre-specified target category, and make the appropriate (“target-present”/“target-absent”) response.

Experiments 1–3 showed that redundancy gain can be demonstrated in a semantic categorisation task, and that this can occur with both lateralised and non-lateralised stimulus presentation. The pattern of results from these experiments was strikingly similar to the results of Mohr et al. (1996): improved performance in redundant trials, and an advantage for stimuli presented in the right visual field over those presented to the left visual field when display was lateralised. I also found redundancy gains and visual field effects for “target-present” but not “target-absent” responses, analogous to findings in LDTs for “word” and “non-word” responses, respectively.

Experiment 4 showed that performance in LVF trials does not improve substantially when participants are allowed longer to respond, suggesting that the high error rates in that condition in preceding experiments are likely a result of data-rather than resource-limited processing. Experiment 5 showed that visual field effects in target-present trials are absent when stimuli are presented vertically rather than horizontally. Under the assumption that vertical presentation should disrupt lexical processing but not access of semantic representations, this implies that visual field effects in earlier experiments were not due to such representations being cerebrally asymmetrical.

Experiment 6 showed that when the task involves a decision between two target categories, rather than between targets and non-targets, redundancy gains and visual field effects are undiminished for both categories. Finally, Experiments 7 and 8 showed that redundancy gain does not decrease when multiple target categories are used and redundant trials involve the presentation of two stimuli from different categories (versus experiments with single target categories or redundant trials with two stimuli from the same category). This provides evidence against a cell assembly coactivation account, as this account would predict greater coactivation with same-category than different-category redundant targets.

Based on the results of Experiments 1–8, it is apparent that redundancy gain is not limited to simple tasks, but is rather a more generalisable phenomenon. In addition, as the cell assembly coactivation account appears inappropriate to explain the results of Experiments 7 and 8, other accounts (e.g., race models, response-level coactivation) are preferable. These accounts are explored in the General Discussion.

Acknowledgements

Given the horror stories one sometimes hears about completing a PhD thesis, I found conducting and writing up the following experiments to be a surprisingly enjoyable experience. A major factor in this has undoubtedly been the excellent supervision I have received from Jeff Miller, whose assistance, feedback, and general attitude toward my project have all been immensely helpful. I am also thankful to the Department of Psychology at the University of Otago for providing an ideal working environment, and to the Government and taxpayers of New Zealand for indirectly funding my study by way of the money received by the university itself. I hope I have not wasted the resources which have been invested in me over the past couple of years. Finally, thank you also to all those who have directly assisted my work in one way or another, be it through discussion, proof-reading, commenting, or some other unspecified means. Particularly, the involvement of the following individuals has been appreciated: Quinlan Buchlak, Tim Ko, Rachel Orzech, Ann Reynolds, Wolf Schwarz, Bill Shepherdson, David Shepherdson, Viv Shepherdson, and Judy Trevena.

CONTENTS

1. <i>Introduction</i>	1
1.1 Semantic memory and categorisation	2
1.2 Redundant target experiments and redundancy gain	8
1.3 Redundancy gain in tasks involving higher-level processing	13
1.4 Redundancy gain in semantic categorisation	22
1.5 Conclusion	30
2. <i>Experiments 1–3</i>	31
2.1 Experiments 1 and 2	31
Method	33
Results	37
Discussion	49
2.2 Experiment 3	51
Method	52
Results	52
Discussion	56
2.3 General Discussion	56
Possible explanatory models	57
RVF/HI advantages	59
Conclusion	60
3. <i>Experiment 4</i>	61
Method	62
Results	63
Discussion	69
Conclusion	72
4. <i>Experiments 5a and 5b</i>	73
Method	74
Results	76
Discussion	79
Conclusion	83
5. <i>Experiment 6</i>	84
Method	85
Results	86
Discussion	94
Conclusion	96

6. <i>Experiment 7</i>	97
Method	99
Results	100
Discussion	108
Conclusion	111
7. <i>Experiment 8</i>	112
Method	115
Results	116
Discussion	125
Conclusion	128
8. <i>General Discussion</i>	129
8.1 Does a semantic categorisation task show any redundancy gain at all?	129
8.2 How can the redundancy gain in semantic categorisation best be explained?	131
Assessment of the cell assembly model	132
A tentative race model	134
Can a serial model account for redundancy gain in semantic categorisation?	137
8.3 Locus of coactivation	141
8.4 Implications for cell assembly accounts of redundancy gain in lexical decision	143
8.5 Other issues	147
8.6 Conclusion	149
<i>References</i>	150
 <i>Appendix</i>	 160
A. <i>Category Norms</i>	161
Method	162
Results	163
Discussion	163
Conclusion	164

LIST OF TABLES

2.1	Sensitivity and bias for the three presentation conditions in Experiment 1.	44
2.2	Sensitivity and bias for the three presentation conditions and three categories in Experiment 2.	48
2.3	Sensitivity and bias for the three presentation conditions in Experiment 3.	55
3.1	Sensitivity and bias for the three presentation conditions in Experiment 4.	66
4.1	Sensitivity and bias for the three presentation conditions in Experiment 5.	79
5.1	Sensitivity and bias for the three presentation conditions in Experiment 6.	90
6.1	Sensitivity and bias for single-target trials in Experiment 7.	105
6.2	Sensitivity and bias for redundant trials in Experiment 7.	106
7.1	Sensitivity and bias for single-target trials in Experiment 8.	120
7.2	Sensitivity and bias for redundant trials in Experiment 8.	122
8.1	Key findings of Experiments 1–8	132
8.2	RT standard deviations for “target-present” and “target-absent” responses in Experiments 1–5	135
8.3	Observed and modelled redundant trial accuracy data	138
A.1	Categories used.	166
A.2	Responses for “A precious stone”.	167
A.3	Responses for “A metal”.	168
A.4	Responses for “A military title”.	169
A.5	Responses for “A four-footed animal”.	170
A.6	Responses for “A kitchen utensil”.	172
A.7	Responses for “A part of the human body”.	174
A.8	Responses for “An elected office”.	176
A.9	Responses for “A religious official”.	178
A.10	Responses for “A part of a building”.	179
A.11	Responses for “A musical instrument”.	181
A.12	Responses for “A flower”.	183
A.13	Responses for “A tree”.	185
A.14	Responses for “A fish”.	187
A.15	Responses for “A weapon”.	189
A.16	Responses for “A bird”.	191
A.17	Responses for “A vegetable”.	193

LIST OF FIGURES

1.1	Possible models of redundancy gain in the lexical decision task	20
1.2	Possible models of redundancy gain in the semantic categorisation task	25
2.1	Example stimulus displays from Experiment 1.	35
2.2	Latency data for correct responses in Experiment 1.	38
2.3	Latency data for correct responses in Experiment 2.	40
2.4	Target \times Presentation interaction in RT data, Experiment 2.	41
2.5	Category \times Target interaction in RT data, Experiment 2.	41
2.6	Latency data for error responses in Experiment 2.	42
2.7	Response accuracy data from Experiment 1.	44
2.8	Response accuracy data from Experiment 2.	46
2.9	Category \times Target interaction for % response accuracy data, Experiment 2.	47
2.10	Example of a stimulus display from Experiment 3, showing a redundant target-present trial.	53
2.11	Latency data for correct responses in Experiment 3.	54
2.12	Response accuracy data for Experiment 3.	55
3.1	Latency data for correct responses in Experiment 4.	64
3.2	Response accuracy data from Experiment 4.	65
3.3	Mean peak force from correct trials in Experiment 4.	67
3.4	Latency and accuracy data for Experiments 1, 2, and 4.	68
3.5	Peak force data, summed across Experiments 1, 2, and 4.	70
4.1	Example of a stimulus display from Experiment 5b, showing a BVF-Present trial.	75
4.2	Latency data for correct responses in Experiments 5a and 5b (combined).	76
4.3	Response accuracy data from Experiments 5a and 5b (combined).	78
5.1	Latency data for correct responses in Experiment 6.	87
5.2	Latency data for error responses in Experiment 6.	88
5.3	Response accuracy data from Experiment 6.	89
5.4	Latency data for Body trials from Experiments 2 and 6.	91
5.5	Latency data for Building trials from Experiments 2 and 6.	91
5.6	Accuracy data for Body trials from Experiments 2 and 6.	92
5.7	Accuracy data for Building trials from Experiments 2 and 6.	93
5.8	Response force data for Building trials from Experiments 2 and 6.	93
6.1	Single-target trial latency data for correct responses in Experiment 7.	101
6.2	Redundant trial latency data for correct responses in Experiment 7.	103
6.3	Single-target trial response accuracy data from Experiment 7.	105
6.4	Redundant trial response accuracy data from Experiment 7.	106

7.1 Single-target trial latency data for correct responses in Experiment 8. 117
7.2 Redundant trial latency data for correct responses in Experiment 8 118
7.3 Single-target trial response accuracy data from Experiment 8. 120
7.4 Redundant trial response accuracy data from Experiment 8 121
7.5 CDFs for the Instrument/Animal condition in Experiment 8 123
7.6 CDFs for the four redundant trial types, along with independent race model CDFs
derived from single-target trial data 126

1. INTRODUCTION

Complex decision-making tends to require the integration of multiple pieces of information. Be it a professional footballer having to choose which teammate is best placed to receive the ball, a clinical psychologist determining which treatment method to use with a patient, or a prospective graduate student attempting to select an appropriate topic for their doctoral research, scenarios abound in which an agent aims to discern the optimal course of action, given the many sources of evidence at their disposal. Indeed, this is a characteristic shared — consciously or otherwise — by all organisms in possession of a nervous system, and thus might be considered an almost universal aspect of behaviour. Given the ubiquity of this integration as a component of the process of choice and decision, if one seeks an understanding of how these choices and decisions are made then the investigation of informational integration can provide one strategy for its achievement.

Many studies have explored the manner in which information is integrated in the context of decisions about simple perceptual features (e.g., Fischer & Miller, 2008; Gondan, Götze, & Greenlee, 2010; Miller, 1982, 2004; Mordkoff & Yantis, 1991, 1993). This line of research continues to provide valuable information about basic information processing, and about the ways in which such integration can occur. However, there have been fewer attempts to investigate comparable higher-level processing, such as that involving semantic information.

My aim in this thesis is to introduce, describe, and outline the theoretical implications of, a number of experiments I conducted with this end — the investigation of semantic informational integration — in mind. These experiments involved the use of a task where participants were asked to make speeded categorisation decisions (member vs. non-member of a target category or categories) on the basis of the semantic identities of visually-presented lexical stimuli. Each trial involved the simultaneous presentation of two stimuli; sometimes,

these were both informative, whereas at others, only one was (with the other being a distractor). My primary interest was in how the presence of multiple, redundant sources of semantic information would affect participants' ability to perform the task.

As the experiments I conducted drew together threads from divergent areas in the experimental and theoretical literature, in the following pages of this introduction I shall attempt to provide sufficient background on each of these areas that the experiments and discussions to follow should be comprehensible. First, I will introduce some theories of the structure of semantic memory and the processes thought to be involved in making semantic categorisation decisions, and outline what I see as the common elements from two purportedly disparate theoretical traditions. Second, I will summarise some of the key findings and models relating to redundant target experiments and redundancy gain, both of which are integral to the experiments I carried out. Third, I will discuss experiments in which a redundant targets design was used in tasks which might have involved higher-level processing (e.g., lexical decision, fame judgements). And finally, I will discuss the possibilities relating to redundancy gain in semantic categorisation, and outline the key questions I aimed to investigate through conducting semantic categorisation experiments with redundant targets.

1.1 Semantic memory and categorisation

Varying theories regarding the structure and organisation of semantic memory have been offered over the years, and discussions of categorisation have a long history in Western thought (e.g., Plato, 1892). More recently though, such theories have followed two distinct courses. On the one hand, some have conceived of semantic memory as an entity founded on abstract representations gleaned from experience with concepts and categories, where the features and properties of these concepts are learnt and stored. Such perspectives are sometimes referred to as “amodal” theories of knowledge, because they posit that sensory experience is used in constructing, but is not directly part of, conceptual representations. On the other hand, in recent decades there has been a proliferation of theories suggesting that knowledge is entirely grounded in sensory experience, and that its storage and manipulation is inextricably linked to sensory and perceptual functions and structures. I will refer to these as “embodied” theories of

knowledge. Both strands of thought are based on a considerable amount of theoretical and empirical work, a review of which is far beyond the scope of this brief summary. Instead, what follows is an attempt to describe an example from each of these schools, to illustrate what sorts of ideas they offer and how they deal with the process of semantic categorisation (which is integral to the experiments described in this thesis).

First, to a so-called “amodal” theory. One particularly well-known model of semantic processing which fits within this realm is the spreading activation model of Collins and Loftus (1975). Taking inspiration from earlier work by Quillian (e.g., Collins & Quillian, 1969; Quillian, 1962, 1967, 1969), Collins and Loftus’ model described semantic memory as a network of types and tokens, with connections between these nodes according to the strength of the semantic relation between the entities they represent. This theory suggests that items which share a close semantic relationship are also close together in the hypothetical semantic network. For instance, to use an example Collins and Loftus themselves offer, the node for “Fire Engine” is presumed to be close to that for “Ambulance”, slightly more distant from that for “Car”, and more distant still from that for “House”.

The key principle of Collins and Loftus’ (1975) theory is that access of particular memory nodes requires some form of “activation” to reach the appropriate part of the hypothesised semantic network. This activation then disperses through the nearby nodes and continues on throughout the semantic network, thus providing the “spreading activation” component of the theory’s name. With increasing distance from the initial node on which activity was centred, the level of activation declines. The implication of this is that when a node is activated, those nodes which are nearby in semantic space receive strong activation relative to nodes representing less closely related concepts.

In constructing their theory, Collins and Loftus (1975) aimed to account for the results from a variety of experiments involving access of semantic information. These included production experiments (where participants produce examples of a particular concept), sentence verification experiments (where participants determine whether a presented sentence is true or false), and semantic categorisation (where participants decide whether or not some item x is a member of some category y). Though it was important to Collins and Loftus that their theory account for production and sentence verification results, I shall focus here on their

explanation for how semantic categorisation could take place within the spreading activation framework, because this process is most relevant for the semantic categorisation experiments reported in this thesis.

Collins and Loftus (1975) described categorisation decisions — responses to questions such as, “Is x a y ?” (e.g., is a dog an animal?) — as a form of semantic matching. They proposed that to determine whether two concepts x and y match, activity spreads from the nodes representing both, and intersections of this activity are assessed. These intersections can provide both positive evidence (i.e., that supporting the “ x is a member of category y ” response), and negative evidence (i.e., that supporting the “ x is not a member of category y ” response), and this evidence is summed to reach an overall decision. Decisions are made when either a positive or negative criterion is reached. If neither criterion is met after an exhaustive search of the evidence, a “don’t know” response results.

Collins and Loftus (1975) also offered suggestions regarding some forms this evidence might take. First, they posited that superordinate connections between nodes — that is, links specifying that the concept represented by one node is or is not a subset of the concept represented by another — could lead to a decision without the aid of other evidence. According to the theory, these connections need not be direct, but can operate across multiple levels of the network. For instance, if y has a superordinate connection to z , and z has a superordinate connection to x , the relationship between x and y can be inferred. Second, they suggested that matches and mismatches between the properties possessed by two concepts could be used as a form of evidence. For instance, if both x and y share properties u , v , and w , this would provide positive evidence of a match; on the other hand, if y possesses a property s which x lacks, this would provide evidence of a mismatch (particularly where s is an especially integral property of y).

Third, in what Collins and Loftus (1975) referred to as “the Wittgenstein strategy”, they suggested that comparisons between the concept to be categorised and other concepts known to be examples of the target category could provide both positive and negative evidence of category membership (where properties match and mismatch, respectively). For example, if x shares properties v and w with z , and z is known to be a member of y , this would be evidence that x is also a member of y . Fourth, Collins and Loftus suggested that negative evidence could

be obtained through connections between type nodes indicating that their subordinate tokens are mutually exclusive. Collins and Loftus (1975, p. 415) illustrated this through the example of bird species: if such connections exist between ducks and eagles, then evidence that a mallard is a duck is evidence against it being an eagle. The final form of evidence Collins and Loftus suggested, again operating only as negative evidence, was the existence of counterexamples. In a categorisation decision, these would be used in instances where some examples of x are not members of category y . For instance, if asked whether writing implements are pens, items such as chalk or crayons would serve as counterexamples.

The above is a summary of how Collins and Loftus (1975) suggested responses to questions about category membership might take place within their theoretical framework. As is evident, though the basic principles of the theory were simple enough, its ability to account for contemporary experimental findings (e.g., Juola & Atkinson, 1971; Loftus & Cole, 1974; Rips, Shoben, & Smith, 1973) required a number of extra assumptions. Not included in these assumptions were any regarding how it might be implemented on a neurophysiological level, though the idea of a network constructed of nodes whose connections vary in strength seems to hint at some of the suggestions of Hebbian theory (e.g., Hebb, 1949). More actively appropriating Hebbian principles, however, is Barsalou's (1999) perceptual symbol systems theory, aspects of which I shall be summarising next to exemplify an embodied theory of semantic memory, and another account of the process of semantic categorisation.

Barsalou's theory was predicated on the idea that semantic representations are reflective of the perceptual nature of the information from which they are composed. He proposed that both "bottom-up" and "top-down" processes are involved in the construction and manipulation of conceptual knowledge: by way of selective attention (top-down), sensory information (bottom-up) is stored in associational areas of cortex, and then elements of this stored information are reactivated (top-down) when access to specific concepts is required. Through ongoing reiteration of these processes, specification of and differentiation between concepts increases, and understanding of these concepts is developed.

The role of selective attention is important in the construction of conceptual knowledge according to Barsalou's (1999) theory. Barsalou's suggestion was not that all sensory experience is stored and retrieved in whole. Rather, he proposed that components of perceptual

experiences are added to long-term memory, supplementing and refining existing representations. This, he states, is what distinguishes a perceptual *recording system* from a *conceptual system*: whereas the former “captures a physical system by creating attenuated (not exact) copies of it. . . [but] does not interpret what each part of a recording contains”, the latter “binds specific tokens in perception. . . to knowledge for general types of things in memory” (Barsalou, 1999, p. 581). Basically, through repeated attention to and extraction of features or items in perceptual experience, representations of these features or items are built in long-term memory. Such representations, rather than reflecting a specific individual item, reflect the sum of many experiences with items of the same general type.

Key to Barsalou’s (1999) proposal, and particularly relevant for the process of semantic categorisation, was the idea of what he termed “simulators” and “simulations”. Simulators essentially refer to the processes involved in top-down activation of concepts which have been developed through perceptual experience. For example, each time one has an experience involving trees, the sensory information derived therefrom adds to one’s representation (“frame”, using Barsalou’s terminology) of the concept *tree*. Additionally, it increases one’s ability to produce novel instances of the concept *tree* using this representation. The latter are the simulations to which Barsalou referred. As they are based on the representations formed through experience, each time a simulator is used, this usage changes it. That is to say, rather than becoming static, Barsalou’s perceptual symbols are dynamic entities.

Barsalou (1999) suggested that the process of simulating is essential to categorisation decisions. According to perceptual symbol systems theory, if attempting to determine whether x is a member of category y , one uses the simulator for y and attempts to construct a simulation of x . When this succeeds, one concludes that x is a y ; if it fails, one concludes otherwise. This process is slightly different where one has prior experience with x in comparison to instances where x is novel. In the former case, the perceived x is matched to an existing simulation of that entity; the latter case requires construction of a novel simulation to match the entity.

One implication of information being used in such a way is that simulators for superordinate categories rely on knowledge obtained through experiences of items which have been identified as members of those categories. Again, Barsalou (1999) did not posit this to be

a purely associational process, whereby all sensory information received is weighted equally in the formation of concepts. Rather, he theorised that selective attention is involved in determining which information is stored and how it is allocated as the purview of one or more simulators. Nonetheless, this would imply that — similarly to the spreading activation theory of Collins and Loftus (1975) — an understanding of higher-order concepts is founded on content provided by subordinate instances.

Like Collins and Loftus (1975), Barsalou (1999) did not directly specify how his perceptual model was likely to be implemented in the brain. However, he did make reference to Pulvermüller's (1999) suggestions that lexical knowledge is represented as Hebbian cell assemblies (that is, collections of neurons the repeated and concerted stimulation of which has led them to form a collective unit), suggesting similar representations might underlie the perceptual symbols on which he built his theory. Indeed, cell assembly theory would seem to lend itself to Barsalou's perception-based theory of knowledge, given suggestions that different areas of the brain store information from different sensory modalities (e.g., Binder, Desai, Graves, & Conant, 2009; Pulvermüller, Shtyrov, & Hauk, 2009), and from different semantic categories (e.g., Martin & Chao, 2001).

As should be apparent, the models of Collins and Loftus (1975) and Barsalou (1999) come from distinct theoretical traditions, and incorporate considerably different terminology and underlying metaphors. Yet, despite this, numerous similarities between the two can be drawn, including the manner in which they deal with semantic categorisation. First, as hinted earlier, both theories seem to suggest that representations of superordinate categories are reliant on the properties (be they formal, or perceptual) of their subordinate instances. In spreading activation theory this results from the hierarchical network relationships between types and tokens, whereas in perceptual symbol systems theory it is inherent in the construction of frames through repeated perceptual experience with similar entities. Second, both theories conceptualise categorisation as a matching process, either between a perceived entity and information from the existing memory store, or between items in that store. In spreading activation theory, this is accomplished by assessing the connections between nodes in the semantic network; whereas in perceptual symbol systems theory, it occurs by an attempt to simulate the item being categorised through the use of the stored perceptual information to

which the relevant simulator has access. Third, both theories seem to imply that conceptual space can be mapped based on the similarities items share. In Collins and Loftus' theory this is an inbuilt feature of the semantic network, whereas for Barsalou it is a result of similar concepts being constructed from similar perceptual information.

As will become apparent, each of these features shared between spreading activation and perceptual symbol systems theory is relevant to understanding the redundant semantic categorisation task outlined later, and its use in the experiments described in the chapters to come. Key to the use of this task though, was the phenomenon of redundancy gain. Therefore, what follows next is a review of the literature revolving around this phenomenon.

1.2 *Redundant target experiments and redundancy gain*

Redundancy gain is a term used to describe the common finding in experimental psychology that when someone is presented with two pieces of information prompting the same response, that response tends to occur more frequently and rapidly than it would were either of these two pieces presented separately. Investigation of redundancy gain dates back at least a century. The starting point for such investigation is generally considered to be the work of Todd (1912), who found that when his participants were simultaneously presented with any combination of visual and auditory stimuli and mild electric shocks, the latency for a simple response decreased when compared with their responses to any of these stimuli in isolation.¹

An explanation Todd forwarded for this finding was that “simultaneous stimuli summate in excitatory effect and send their discharge down one common tract to the reacting finger” (1912, p. 63). Todd's description can be used as an example of one of the two major types of explanation which have since been forwarded for findings of redundancy gain: coactivation models. According to coactivation models, evidence from redundant stimuli is combined in some way, and this summed evidence results in enhanced responding — be it measured by shorter latencies (e.g., Miller, 1982), higher accuracy (e.g., Mohr & Pulvermüller, 2002), more forceful responses (e.g., Giray & Ulrich, 1993), or some other relevant variable. To illustrate by way of analogy, according to coactivation models, stimuli are akin to lumberjacks felling a tree:

¹ Though Todd was not the first to show faster responses to dual stimuli (see, e.g., Dunlap & Wells, 1910).

the more axes at work, the quicker the tree falls. Similarly, the more stimuli which provide evidence that a particular response is appropriate, the more rapidly that response will be made. Coactivation models are frequently applied in situations where the summation of evidence is assumed to take the form of increased neural activity (e.g., Mohr et al., 1996); however, this assumption is not an inherent part of these models.

The main alternatives to coactivation models are so-called “race models”. Race models suggest that redundant stimuli are processed in separate channels; the first of the stimuli for which processing is completed triggers the appropriate response, this being the origin of the “race” aspect of the name. According to race models, redundancy gain is a form of statistical facilitation. Again, this can be illustrated by analogy. Consider a series of races between two (or more) competitors. Assuming every race is not won by the same racer, the average winning time will be faster than the average time for each individual competitor.

As an explanation for redundancy gain, statistical facilitation was initially proposed by Raab (1962), in an attempt to explain the results of Todd (1912) and those of Hershenson (1962).² Hershenson had conducted experiments involving simple reactions to visual and auditory stimuli, based on those of Todd (1912), but with redundant stimuli presented at varying stimulus onset asynchronies (SOAs). Although statistical facilitation provided a simple explanation for most of the redundancy gains found, Raab (1962) acknowledged that the data occasionally showed more facilitation than would be predicted by a race model, and left open the possibility that some summation might occur. In subsequent years, however, race models have come to act as somewhat of a null hypothesis when it comes to redundancy gain.

The primary method of falsifying this hypothesis in recent decades has been by way of Miller’s (1982) race model inequality (RMI). The inequality provides a boundary for the redundancy gains which could occur by way of statistical facilitation. This is achieved via the assumption that however extensive statistical facilitation is, the proportion of responses to redundant signals occurring by time t should not be greater than the sum of the proportions of responses having been made by the same t to each of the single stimuli — that is, statistical facilitation should not make the “racers” themselves faster. Formally, the RMI is expressed

² See also work by Pirene (1943) and Mulligan and Shaw (1980) for the application of similar models in the context of detection probability.

thus:

$$F_R(t) \leq F_1(t) + F_2(t) \tag{1.1}$$

where $F_1(t)$ and $F_2(t)$ are the cumulative distribution functions (CDFs) of reaction time (RT) in the two single-target conditions and $F_R(t)$ is the CDF in the redundant condition. Using data from single-target RTs, this boundary can be estimated, and the distribution of redundant RT data tested against it. If the inequality is violated, then a race model is presumed insufficient to account for the extent of the redundancy gain. In such cases, coactivation models of some sort tend to be accepted as the most effective explanation.³

Aside from race and coactivation models, certain classes of serial models are also capable of accounting for redundancy gains. Serial models can be defined on the basis of whether they propose exhaustive (where all stimuli are processed, irrespective of their status as targets or non-targets) or self-terminating (where a response is made once a target has been identified) search. For instance, Townsend and Nozawa (1997) outlined an exhaustive serial model of performance in tasks with redundant targets which could account for not only redundancy gain, but RMI violations as well. This model requires the assumption that the processing of non-targets is slower than the processing of targets. In this case, the absence of a target in the second channel on single-target trials slows RTs because a slow non-target must be processed instead, and this leads to the gain in redundant target trials. Like the race and coactivation models, this can be illustrated analogically: imagine a sniffer dog tasked with determining whether two suitcases contain illicit drugs. The dog will stop searching a suitcase sooner when the drugs are present, than when they are not; thus, his or her search of both suitcases will be faster when both contain drugs than when only one does, or when neither do. The same principle applies in Townsend and Nozawa's exhaustive serial model. Self-terminating serial models also predict redundancy gains in many cases; however, where target number and display size are confounded (e.g., both single-target and redundant trials contain only targets), redundancy gain can sometimes be used as an argument against such models (see, e.g., Townsend, 1990).

Though the work of Todd (1912), Raab (1962), and Hershenson (1962) consisted entirely

³ Though for another type of non-coactivation model capable of producing RMI violations, see Mordkoff and Yantis (1991).

of tasks involving simple detection, redundancy gain has also been investigated in more complex arenas. Indeed, not only redundancy gain, but violations of the RMI were demonstrated by Miller (1982) in go/no-go and choice-RT tasks, in addition to simple detection. Findings of redundancy gain have now been presented in areas as diverse as lexical decision (e.g., Mohr, Pulvermüller, & Zaidel, 1994; Mohr et al., 1996; Yoshizaki, 2001), letter classification (e.g., Grice, Canham, & Boroughs, 1984; Grice & Reed, 1992), flavour perception (e.g., Dalton, Doolittle, Nagata, & Breslin, 2000; Veldhuizen, Shepard, Wang, & Marks, 2010), emotion recognition (e.g., Collignon et al., 2008, 2010; Tamietto, Adenzato, Geminiani, & de Gelder, 2007), rhyme judgement (Banich & Karol, 1992), intersensory facilitation (e.g., Barutchu, Crewther, & Crewther, 2009; Barutchu et al., 2010), and face perception (Mohr, Landgrebe, & Schweinberger, 2002), to name but a few.

In the absence of RMI violations, redundancy gain is frequently explained by recourse to race models (e.g., Mordkoff, Miller, & Roch, 1996); and in their presence, coactivation (e.g., Miller, 1986) is often the most plausible explanation.⁴ However, no clear pattern is evident when it comes to the task-related factors responsible for one or the other account providing the most suitable explanation for findings in a particular experiment. One suggestion has been that the “perceptual modules” — that is, the processes operating on information about a specific perceptual feature, such as colour, shape, or orientation — within which stimuli are processed are important; that stimuli race within such modules, and coactivation occurs between them (Mordkoff & Yantis, 1993). However, the applicability of such a suggestion to tasks where perceptual features are irrelevant to responding (e.g., lexical decision) is obviously limited.

When RMI violations are found, and redundancy gains are concluded to result from some form of coactivation, attempts are frequently made to determine the hypothetical level of processing (e.g., perceptual, decisional, motor) at which this coactivation has occurred.

Arguments have been made for coactivation at the perceptual (e.g., Krummenacher, Muller, & Heller, 2002; Miniussi, Girelli, & Marzi, 1998; Murray, Foxe, Higgins, Javitt, & Schroeder, 2001; Savazzi & Marzi, 2008; Schwarz & Ischebeck, 1994), motor (e.g., Cavina-Pratesi, Bricolo, Prior, & Marzi, 2001; Giray & Ulrich, 1993), and decisional (e.g., Miller, 1982; Mordkoff et al., 1996)

⁴ Though, as mentioned, channel interaction (e.g., Mordkoff & Yantis, 1991) and exhaustive serial processing (e.g., Townsend & Nozawa, 1997) can also explain race model violations in some situations.

levels. There is no clear consensus on this issue, and it might vary with changes in the task used. More complex tasks also open up the possibility of coactivation in higher-level processing, such as the lexical coactivation posited by Mohr, Pulvermüller, and Zaidel (1994), which is considered at length in the following section.

Being able to infer the locus of coactivation is important if one wishes to use redundant target experiments to investigate informational integration, as it has the potential to shed light on how this integration is taking place and by extension in what other circumstances similar integration might occur. On the other hand, where redundancy gains are best explained by race models, the facilitation resulting from redundant targets is an entirely statistical phenomenon. As such, determining the stage(s) of processing at which this “race” is occurring seems generally to be less studied. However, this is not to say that such information would not be informative; under certain plausible assumptions, races that vary in extent as a proportion of the overall processing involved in task completion could also be expected to have differing levels of distributional variance, and thus lead to different amounts of redundancy gain (for the application of a similar principle, see, e.g., Miller, 1986).

Though redundancy gain has been demonstrated in some tasks which may have involved higher-order processing, the majority of research has focused on more basic tasks, such as simple detection and perceptual discrimination. These tasks provide an ideal means of investigating decisional processes in an environment relatively free from potentially complicating and confounding factors. The drawback of this, though, is that the extent to which findings of redundancy gain are generalisable to tasks involving — or indeed, relying on — such complicating factors is limited. Are coactivation, or “races” between stimuli, phenomena which reflect general principles of cognitive processing? Or are they limited to basic decisions about simple, uncomplicated stimuli, occurring in regulated experimental contexts? This question remains very much open. In the following section, I will review some of the evidence from experiments where redundancy gains have been found in tasks which may have required higher-order processing. In doing so, though not scratching the surface of an answer to that question, I hope to highlight some of the proposals which have been forwarded by those attempting to explain such gains, as well as some difficulties inherent in attempting to investigate redundancy gain in complex tasks. These issues are particularly relevant in the

context of the semantic categorisation experiments I conducted.

1.3 Redundancy gain in tasks involving higher-level processing

As mentioned, redundant target experiments have primarily been undertaken with what on the surface appear to be far simpler tasks than something like semantic categorisation — simple detection, perceptual discrimination, et cetera. One concern in attempting to extrapolate findings from these tasks to others requiring more intuitively complex processing is that the models used to explain results in simpler contexts might not be applicable. In this section I review evidence from studies in which the basic idea of redundant target experiments has been adapted to tasks potentially involving higher-order processing, discuss the explanations offered for the derived findings, and highlight potential issues with some of these explanations.

One area in which this adaptation has been undertaken is that of the processing of emotions. Two groups of researchers have investigated this in slightly different ways. In an effort to assess hemispheric contributions to emotional processing, Tamietto et al. (Tamietto, Latini Corazzini, de Gelder, & Geminiani, 2006; Tamietto et al., 2007) conducted experiments involving lateralised presentation of photos of actors making emotional expressions. Taking a different path, Collignon et al. (2008, 2010) presented their participants with video and auditory stimuli representing different emotions, to investigate how the information from these sensory modalities would interact. Both sets of studies involved participants either being asked to detect the presence of a “target” emotion, or to discriminate between two possible emotions. Redundancy gains were found in all cases, such that responses were faster when emotional stimuli were presented bilaterally (Tamietto et al., 2006, 2007) or bimodally (Collignon et al., 2008, 2010), to the extent that violations of Miller’s (1982) race model inequality occurred. Tamietto et al. (2007) suggested that their findings were indicative of neural summation from the dual stimulation of transcortical cell assemblies — an account of redundancy gain in higher level processing I shall explain in more detail later in this section. Collignon et al. (2008), however, were non-committal regarding the specific process leading to improved performance in audio-visual trials, suggesting merely that some form of integration of multi-sensory emotional information must be responsible.

One aspect of Tamietto et al.'s (2006) findings worth noting was that redundancy gains persisted when redundant stimuli were photos of different actors conveying the same expression. According to Tamietto et al., this implies that their results are due to redundancy in the emotional expressions, rather than the physical properties of the stimuli — though responses were faster when both stimuli were also physically identical. Intuitively though, this claim seems questionable: emotional facial expressions are defined by their physical features, making complete dissociation of their visual and emotional components problematic.

Collignon et al.'s (2008) claim that their findings are a result of the integration of high-level emotional information can also be questioned, on the basis of highly repetitive stimulus presentation in their experiments. Participants completed 480 trials, involving the presentation of two emotional expressions (fear and disgust) from six actors. This could have led to the formation of associations between low-level stimulus features and responses, and a situation where more complex processing of the stimuli was unnecessary. If so, it is difficult to be sure that any informational integration occurred in emotional processing, rather than at a lower level (e.g., the response level).

Object recognition is another form of higher-level processing which has been investigated using redundant target experiments, in studies by Molholm, Ritter, Javitt, and Foxe (2004) and Suied, Bonneel, and Viaud-Delmon (2009). In Molholm et al.'s study, participants responded when either visual, auditory, or joint audio-visual representations of one of eight animal species were presented. Redundancy gains were found such that responses were faster to audio-visual target trials than to single-stimulus trials (or incongruent audio-visual trials, with the target animal in one modality and a non-target animal in the other) to an extent that Molholm et al. suggested violated the race model inequality, leading them to conclude that some form of “neural interaction of visual and auditory object information” (p. 457) was responsible (though for reasons described below, this argument is questionable). Suied et al. (2009) also found RMI violations for bimodal target trials in a similar experiment where participants had to respond to the image or sound of a telephone, whether presented together, separately, or in concert with the image or sound of a distractor (a frog). As a control experiment where redundant targets were semantically incongruent (the sound of a frog and the image of a telephone) failed to show similar gains, they concluded that facilitation was dependent on semantic congruence.

However, these conclusions of high-level redundancy gain too are potentially problematic, especially with regard to the idea of neural coactivation. Both Molholm et al. (2004) and Suied et al.'s (2009) studies again involved the repetitive presentation of a small set of stimuli, meaning that lower-level processes might have been responsible for redundancy gains. Though Suied et al. attempted to discount this possibility by way of their control experiment, their use of the same two items as both targets and non-targets (that is, the sound of a frog and image of a telephone were targets, whilst the image of a frog and sound of a telephone were non-targets) makes the task substantially more difficult (as evidenced by greatly increased RTs), and thus the suitability of this control questionable.

Molholm et al. based their suggestions of neural interaction on the fact that redundant RTs were faster than those a race model would predict, yet the race model they tested was unusual in two respects. First, rather than using single-stimulus trials to model the separate "racers", they used incongruent bimodal trials. This could lead to interference between the incongruent stimuli, slowing responses and making the resultant RTs inappropriate for use in modelling the race which hypothetically occurs in redundant trials. Second, the race model they composed from these incongruent trials had an additional assumption of stochastic independence of channel processing, making it less conservative than Miller's (1982) race model inequality. In other words, their results represented a violation of the predictions of one particular race model, rather than those of all possible race models. Again, then, the redundancy gains found by both Molholm et al. (2004) and Suied et al. (2009) could easily have been a result of lower-level processes.

Multisensory interactions in the processing of what could be considered higher-level concepts in a redundant targets context were also the focus of a study by Laurienti, Kraft, Maldjian, Burdette, and Wallace (2004). They presented coloured circles, and/or auditory or visually-presented colour names (red, blue, or green), and asked participants to press one button when a red stimulus was presented and another for a blue stimulus. Responses were faster when auditory names and coloured circles were congruent than when either stimulus was presented alone or matched with a non-target colour (green), but there was no similar gain when redundant stimuli were circles and written colour names. As gains in the former condition were also purported to violate race model predictions, this led Laurienti et al. to

suggest that integration of multi-sensory information is reliant on both semantic congruence, and cross-modal presentation. However, like Molholm et al. (2004), the race model against which Laurienti et al. tested their data was one where stochastic independence was assumed, meaning that their findings may be accounted for by race models lacking this assumption. In addition, Laurienti et al.'s claims for the importance of semantic congruence are undermined by the fact that there were no conditions in their experiment where incongruent dual stimuli led to the same response; rather, the colour green (which was included in all incongruent trials, as red and blue were never presented together) was not associated with any response by itself. That is to say, since their design did not allow the possibility of testing incongruent redundant stimuli (i.e., different colours assigned to the same response), they cannot rule out the possibility that these would produce an equally large redundancy gain. Additionally, these experiments once again involved high levels of stimulus repetition, which is problematic for reasons described in relation to studies mentioned earlier (e.g., Collignon et al., 2008, 2010; Molholm et al., 2004; Tamietto et al., 2006, 2007).

A final study involving multisensory stimuli in what could be conceived of as a higher-level context was that of Dijkstra, Frauenfelder, and Schreuder (1993), who used a redundant target design to investigate the processing of phonemic and graphemic representations of vowels. They instructed participants to make the target-present response if either of two target vowels was present (e.g., “a” and “u”), and found larger audio-visual redundancy gains for congruent than incongruent redundant targets (e.g., auditory and visual representations of the letter “a”, compared to an auditory “a” combined with a visual “u”). Because of this congruency effect, they argued that the representation of a given vowel can be activated cross-modally. Unlike many of the studies described previously, Dijkstra et al.'s results cannot be attributed to repetitive stimulus presentation, because though stimuli *were* heavily repeated the key comparison was between congruent and incongruent redundant trials. As such, though repetition might have contributed to the gains found for both stimulus types, it cannot be responsible for the congruency effects.

Another task in which the redundant targets paradigm has been used to investigate higher level processing is that of fame judgements for faces. Mohr et al. (2002) presented participants with pictures of faces presented to the LVF or RVF, or identical pictures presented

bilaterally, and asked them to determine whether the pictures were of famous people. The participants showed a bilateral redundancy gain when identifying familiar (famous) faces, but not unfamiliar faces, which Mohr et al. suggested was due to participants having neural representations of the famous faces, allowing them to benefit from dual stimulation where the non-famous faces could not. This explanation was endorsed by Schweinberger, Baird, Blümner, Kaufmann, and Mohr (2003) and Baird and Burton (2008), who replicated the finding in similar experiments with identical (two copies of the same photo; Baird & Burton, 2008; Schweinberger et al., 2003) and non-identical (two photos of the same person; Baird & Burton, 2008) redundant targets.

Lexical decision is another area where redundant target experiments have been conducted, and probably offers the closest comparison with the semantic categorisation experiments reported in this thesis. In fact, lexical decision task (LDT) research provided both the inspiration for, and the account used to explain findings of, many of the studies of redundancy gain mentioned previously (e.g., Baird & Burton, 2008; Mohr et al., 2002; Schweinberger et al., 2003; Tamietto et al., 2006, 2007). This account explains redundancy gain in higher level processing through a model suggesting that this results from neural summation in Hebbian-style transcortical cell assemblies. The cell assembly model has its origin in the work of Mohr and Pulvermüller (e.g., Mohr, Pulvermüller, & Zaidel, 1994; Mohr, Pulvermüller, Rayman, & Zaidel, 1994; Mohr et al., 1996; Mohr, Pulvermüller, Cohen, & Rockstroh, 2000; Mohr & Pulvermüller, 2002; Mohr et al., 2007), who initially adapted the redundant targets paradigm to the realm of lexical decision. These experiments have generally taken a common form, in which words and non-words are presented visually either singly to the left or right of fixation, or redundantly to both sides, and participants are asked to decide whether or not the stimuli are words, making one response if either stimulus is a word and another if either is a non-word (words and non-words are never presented together).

This mode of presentation was designed to take advantage of the fact that information from the left visual field (LVF) is initially processed in the right hemisphere, and information from the right visual field (RVF) is initially processed in the left hemisphere. Prior to its use in tasks with redundant targets, lateralised presentation of lexical stimuli had a long history in studies of linguistic processing (e.g., Bradley & Garrett, 1983; Broerse & Zwaan, 1966;

Chiarello & Nuding, 1987; Heron, 1957; Liederman, Merola, & Martinez, 1985; Mishkin & Forgays, 1952; Young & Ellis, 1985), and it was this tradition which Mohr and Pulvermüller's early work using redundant targets in lexical decision (e.g., Mohr, Pulvermüller, Rayman, & Zaidel, 1994; Mohr, Pulvermüller, & Zaidel, 1994; Mohr et al., 1996) followed. They sought to test a model of language processing based on Hebbian principles (e.g., Hebb, 1949), according to which the neural representations of words take the form of cell assemblies distributed across both neural hemispheres.

Mohr, Pulvermüller, and Zaidel (1994) argued that comparing lexical decision performance when words and non-words were presented in either the LVF or the RVF, to performance when the same stimuli were presented to both visual fields (BVF), should allow a test of this model. They outlined three potential outcomes of BVF presentation: first, interhemispheric inhibition, leading to worse performance in the redundant than the single-target conditions; second, monohemispheric control, where the BVF performance mirrors that in one of the unilateral conditions due to one of the two hemispheres dominating processing; and third, interhemispheric cooperation, leading to superior BVF performance compared to the unilateral conditions. The cell assembly model they forwarded would predict the latter outcome, and this prediction was borne out in their results. This model has since been elaborated at considerable length (e.g., Pulvermüller & Mohr, 1996; Pulvermüller, 1999, 2001) — though not with specific reference to redundancy gain.

At this point, it is worth noting that though Mohr, Pulvermüller, and Zaidel's (1994) findings were *consistent* with their suggestion of lexical cell assemblies, this is not the only model able to account for a redundancy gain in lexical decision. As mentioned in the section on redundancy gain, there are two main “species” of model generally forwarded to explain such gains. The first of these is coactivation models, where the information from both redundant targets is summed to produce a response. Mohr and Pulvermüller's cell assembly-based explanation involved redundant stimuli providing extra activation to a word's neural representation, which can be considered a form of coactivation (neural summation at the level of the lexical cell assembly). The second “species” is race models, in which redundant stimuli race to activate a response. Mohr, Pulvermüller, and Zaidel did not discuss the possibility of a race model account of their results. However, later work by the same group presented some

arguments for preferring a cell assembly-based coactivation account, as described next.

Mohr et al. (1996) replicated the earlier findings of a bilateral redundancy gain in lexical decision, and further showed that gains were also evident when two identical stimuli were presented to a single visual field (be it left or right), and that even larger gains occurred with the presentation of four copies of the same word. Again, the cell assembly model was forwarded as the best explanation for these results; however, the validity of the justifications for this preference over the race model is questionable. Outlines of the two models Mohr et al. considered are displayed in Figure 1.1 — with the race model in 1.1a, and the cell assembly coactivation model in 1.1b (1.1c illustrates a model not considered by Mohr et al.; this model will be discussed later). There were three main findings Mohr et al. suggested were sympathetic with their cell assembly model. First, redundancy gains had been evident only in data from trials where words were presented (that is, not in data from non-word trials). The cell assembly model provided a simple explanation for this: words possess existing neural representations which can benefit from dual stimulation, whereas non-words do not. However, certain race models could offer an equally simple explanation: If the “race” between stimuli only ends where either a word is identified, or processing reaches some form of exhaustive state in which all stimuli have been processed — analogous to Ratcliff’s (1978) theory of memory retrieval — this finding would be expected. This is because a “race” would only speed “word” responses; “non-word” responses would not be able to benefit in such a manner.

Second, though redundancy gains had been found in experiments with non-clinical samples, in an experiment with a split-brain participant (Mohr, Pulvermüller, Rayman, & Zaidel, 1994) no bilateral redundancy gain was evident. Mohr et al. (1996) argued that this was evidence against a race model because such a model would predict the two stimuli, processed by separate hemispheres, to race irrespective of whether or not the corpus callosum is intact. By contrast, the cell assembly model would require intact transcallosal connections for the activation caused by bilateral stimulation to sum; thus, its presence in participants with intact corpora callosa, and its absence in a split-brain participant, fit the cell assembly model’s predictions. This finding is equally open to alternative explanation from a race model. Data from the split-brain study (Mohr, Pulvermüller, Rayman, & Zaidel, 1994) show that the patient’s mean RTs for LVF and RVF stimuli differed by approximately 200 ms. A race model

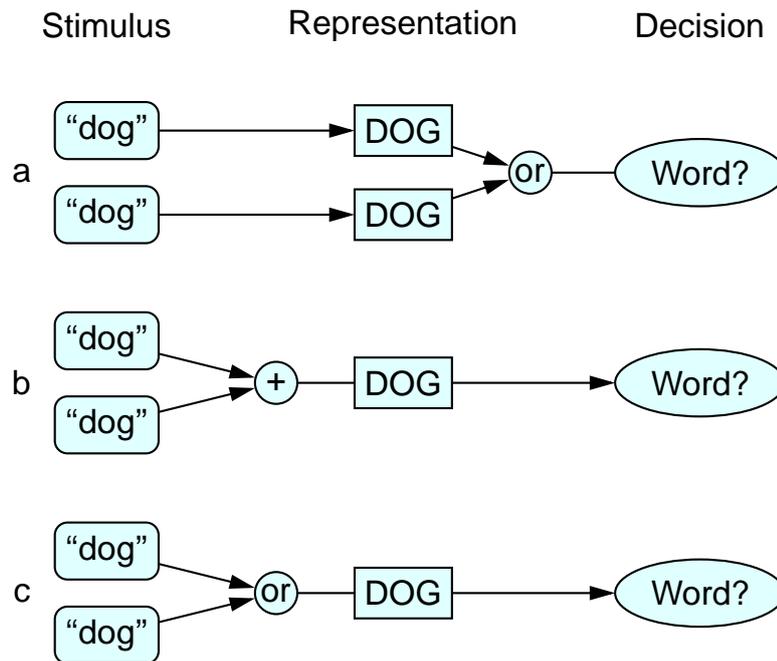


Fig. 1.1: Different schematic models of redundancy gain in the lexical decision task. “a” shows the hemispheric race model considered by Mohr et al. (1996); “b” the cell assembly coactivation model those authors proposed; and “c” an alternative race model, in which stimuli race to activate a single neural representation. The stages in the models (L to R) are stimulus presentation, activation of lexical representation, and “word or non-word” decision.

will only lead to redundancy gain where there is overlap between the distributions of the single racers, and the greater the overlap, the greater the expected gain (Raab, 1962). Here, the extent of that overlap is questionable because of the large difference between LVF and RVF processing times, meaning that the non-significant redundancy gain found for the split-brain patient is not inconsistent with race model explanations.

Mohr et al. (1996) also offered a third argument in support of their cell assembly explanation, based on the finding that further redundancy gains occurred with four than with two stimuli. This involved the suggestion that to account for such a result, a race model would require people to possess not only separate lexicons (that is, stores of word forms) in the left and right hemispheres, but *multiple* representations of words in each hemisphere. This

reasoning seems to imply that Mohr et al. considered it necessary for any “race” which was occurring to necessitate separate neural representations of each stimulus “racing”. However, this is a narrow conceptualisation of race models, and access to separate representations need not be required for a race to occur. For example, if processing of each of the four stimuli involved sensory or perceptual access of a single representation, there seems no obvious reason why this would preclude a “race” between these stimuli for such access. Such a model is illustrated in Figure 1.1c.

The implication of all this is that though results from LDTs clearly show that redundancy gain can occur in tasks requiring higher-order processing, the cell assembly-based explanation offered to explain these might not be appropriate. Each finding Mohr et al. (1996) suggested as evidence against a race model can be accounted for by a less restricted version of such a model than those authors had envisaged. This leaves open the question of how redundancy gain in lexical decision might occur.

Nonetheless, it would appear that the only experiments which can unequivocally be said to have demonstrated redundancy gain in higher-level processing are the lexical decision experiments of Mohr and Pulvermüller (e.g., Mohr, Pulvermüller, & Zaidel, 1994; Mohr, Pulvermüller, Rayman, & Zaidel, 1994; Mohr et al., 1996), and similarly-designed fame judgement tasks (Baird & Burton, 2008; Mohr et al., 2002; Schweinberger et al., 2003); and possibly Dijkstra et al.’s (1993) vowel identification study (depending on the extent to which vowel identification can be considered “higher-level” processing). Though claims of cell assembly-based coactivation offered by Mohr and Pulvermüller might be premature, such an account does at least offer a framework within which further investigation can be undertaken, and would also seem to be generally consistent with the findings of Dijkstra et al. (1993). In addition, with some minor alterations, a similar task to that used by Mohr and Pulvermüller could provide a method of conducting such investigation in other areas than lexical decision or fame judgement — both of which could be argued to require judgements of familiarity rather than any more complex processing.

One such area is that of semantic categorisation. Though a number of studies have investigated the effect of multiple sources of information on the process of identifying or categorising items (e.g., Chen & Spence, 2010; Kuipers & La Heij, 2008), there does not appear

to have been any exploration of this in a redundant targets context, where multiple stimuli are directly relevant to the decisions being made (rather than as task-irrelevant flankers, primes, or similar). This implies that using a semantic categorisation task in a redundant targets paradigm could provide information relevant to understanding both semantic information processing, and redundancy gain in the context of higher-level processing.

This is precisely what I did in the experiments described in the following chapters of this thesis. Next, I shall detail the manner in which I adapted Mohr and Pulvermüller's basic experimental design to the realm of semantic categorisation. In doing so, I aim to highlight both the issues I hoped to investigate using a semantic categorisation task with redundant stimuli, and the manner in which the task I used was designed to obviate the potential problems arising in many of the experiments I have just finished describing in this section.

1.4 Redundancy gain in semantic categorisation

In the previous three sections I have given brief summaries of two differing models of semantic memory and the categorisation process, and of the redundancy gain literature; this has been followed by a slightly more detailed review and critique of research which involved the application of redundant target experiments to tasks potentially involving higher level processing. Now, I shall provide a summary of the main issues I hoped to investigate in my study of redundancy gain in semantic categorisation. I shall conclude by outlining the factors which needed to be considered when adapting the redundant targets task used by Mohr and Pulvermüller (e.g., Mohr, Pulvermüller, & Zaidel, 1994) to the realm of semantic categorisation, and describing how I attempted to deal with these.

The question which was prominent above all others was whether a semantic categorisation task with redundant targets would show any redundancy gains at all. Though both race and coactivation models would generally predict gains, there are also situations where multiple target stimuli lead to a redundancy *loss* (e.g., Koppen, Alsius, & Spence, 2008; Molholm et al., 2004). In addition, on an intuitive level semantic categorisation seems a more complex task than many of those where redundancy gains have previously been shown, meaning that deleterious effects of resource-limited processing (e.g., Norman & Bobrow, 1975)

could occur. This has a precedent in, for instance, Jamet and Lebohec's (2007) finding that less information from a multimedia presentation was retained when the same text was presented simultaneously in written and verbal forms than when only presented in the visual modality. Jamet and Lebohec suggested this deterioration in retention (i.e., a redundancy loss) was due to an "overload" of information in bimodal presentation. Though multimedia presentation and reaction time tasks are very different, similar principles could apply given the complex processing involved in semantic categorisation. Consequently, I sought evidence for any redundancy gains whatsoever in this task.

Of course, as important as — if not more important than — the presence or absence of a redundancy gain is some understanding of how any gain has come about. Based on the general redundancy gain literature, and the specific suggestions of Mohr and Pulvermüller (e.g., Mohr et al., 1996), two accounts of redundancy gains are immediately apparent. First, there is the possibility of a race between redundant stimuli, leading to gains through statistical facilitation. Second, there might be of some form of coactivation, potentially of the type Mohr et al. (1996) outlined with reference to Pulvermüller's cell assembly model (Pulvermüller, 1999, 2001; Pulvermüller & Mohr, 1996).

Though I have previously raised issues questioning the certainty with which the cell assembly account was applied to explain results in Mohr and Pulvermüller's lexical decision experiments, even if it is appropriate in that context there could still be problems with extending it to semantic categorisation. Mohr et al.'s (1996) characterisation of the way coactivation of cell assemblies occurred with redundant stimulation was that the activity from each of the two stimuli summed in the shared assembly, and that this led participants to be more effective at making the "word" response. An analogous explanation was offered for experiments requiring a familiar/unfamiliar decision about face stimuli (Mohr et al., 2002). As mentioned earlier, in each of these cases, the task could be conceived of as requiring an assessment only of the familiarity of the stimuli presented, rather than any higher-order characteristics they possessed. Though this might be appropriate for lexical decision and face identification, according to the theories of both Collins and Loftus (1975) and Barsalou (1999) a semantic categorisation task necessitates the operation of a more complex matching process. As such, the familiarity of the stimuli does not offer the same diagnostic information. Given

suggestions that familiarity judgements can be made on the basis of the level of general (e.g., node non-specific) neural activity resulting from stimulus presentation (e.g., on a purely quantitative basis; Yonelinas, 2002), coactivation might lead to gains in tasks requiring familiarity assessments without this necessarily implying that similar principles will enhance performance in the categorisation task.

In addition, coactivation of lexical representations (that is, the representations of physical word forms themselves) might not extend to coactivation of semantic representations (that is, representations of their meanings). Many researchers suggest that lexical and semantic systems are at least partially distinct (e.g., Borowsky & Besner, 1993; Caramazza, 1997; Collins & Loftus, 1975; Endress & Potter, 2012; Stolz & Besner, 1996). Collins and Loftus (1975), for instance, refer to a dictionary-like lexical network which is separate from but connected to semantic memory. Studies with aphasic patients also show a dissociation between the two, such that reading of word forms can remain intact whilst comprehension of their meaning is impaired (e.g., Shallice, Warrington, & McCarthy, 1983), or vice versa (e.g., Warrington & Shallice, 1980). These include patients who have intact reading for irregular words, impaired non-word reading, and impaired understanding of word meanings (Funnell, 1983). This lexical/semantic dissociation is reinforced by Pulvermüller et al.'s (2009) suggestion that spoken and written word forms are represented in separate neural locations than the perceptual details of the items to which they correspond (e.g., the word “cat” has a representation which is distinct from, yet connected to, one’s understanding of what cats *are*). Given this distinction, the form of coactivation posited by Mohr and Pulvermüller may well be limited to lexical cell assemblies.

Of course, even if results from the semantic categorisation task do show evidence of coactivation, a remaining question will be the locus of this coactivation. As described previously, traditionally questions about the locus of coactivation have received one of three possible answers: that it occurs at the perceptual, decisional, or motor level. Obviously, in a task using word stimuli and requiring semantic processing the relevant levels of processing are slightly different from those generally considered important in redundancy gain experiments, and there at least two further processing levels to add: lexical and semantic. As is described in the following chapter, the task used was designed to exclude the possibility of perceptual and lexical coactivation. If a coactivation account of redundancy gain in semantic categorisation

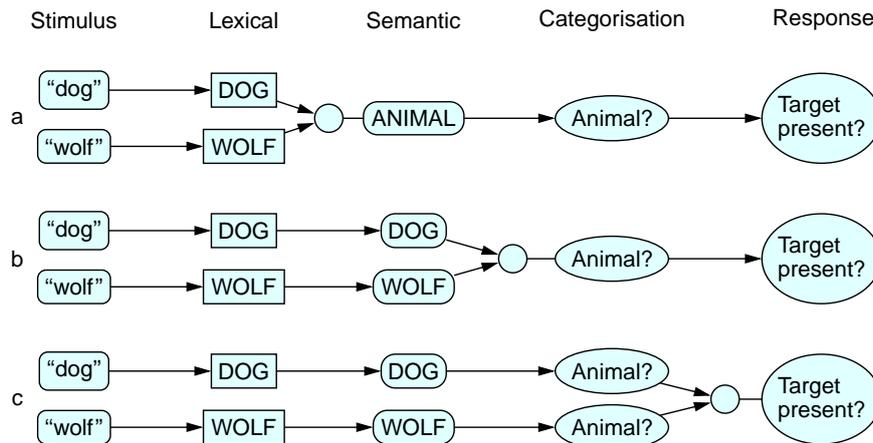


Fig. 1.2: Different schematic models of redundancy gain in the semantic categorisation task. Circles indicate points at which either coactivation or a race might occur; “a” shows this in the activation of a semantic representation; “b” shows this in categorisation; and “c” in response selection. The stages in the models (L to R) are stimulus presentation, activation of lexical representation, activation of semantic representation, categorisation, and response (target-present or target-absent). Other levels of processing at which coactivation might occur (e.g., perceptual and motor processing) are not included in this figure.

seems appropriate, it will be informative to know which of the remaining levels of processing is responsible.

In considering this issue the suggestions of both Collins and Loftus (1975) and Barsalou (1999) regarding how categorisation is accomplished — that is, by way way of some form of matching process — might be worth keeping in mind. What role could coactivation play in such a process? Figure 1.2 shows a basic schematic of some models of the way semantic categorisation with redundant targets might occur, using the target category “Animals” as an example. I describe these separately next. In all of the models in Figure 1.2, processing commences with the activation of lexical representations by the stimuli which have been presented (here, “dog” and “wolf”). In 1.2a, race and coactivation models diverge (illustrated by the empty circle, implying either a race or coactivation can occur here) at the point where lexical representations activate semantic representations. In the race version of the model, whichever lexical representation is quickest to activate the semantic representation of the

higher-order category to which they belong (“animal”) wins the race. In the coactivation version of the model, both the lexical representations provide activation directly to the “animal” representation. Given both spreading activation and perceptual symbol systems frameworks imply that semantic representations of superordinate categories are in some way composed of the items belonging to those categories, the presentation of redundant targets from within the same category should have the opportunity for coactivation of its semantic representation — in other words, semantic coactivation. This would be analogous to the form of coactivation proposed by Mohr and Pulvermüller (e.g., Mohr et al., 1996) in lexical decision, but obviously involving semantic rather than lexical (word-form) representations.

In Figure 1.2b, the stimuli “dog” and “wolf” activate separate lexical and semantic representations, and the divergence between race and coactivation models occurs subsequent to this. In the race version of the model, whichever semantic representation is activated most quickly is assessed for its category membership. In the coactivation version, their “animalness” is assessed in a unitary fashion after initial activation of item-level semantic representations (that is, unlike in 1.2a, the representation of the “animal” category to which they belong is not directly activated). That is to say, the “matching process” of categorisation is accomplished on a holistic basis, such that the semantic representational nodes activated by the two stimuli are considered as a whole. To illustrate this by example, recall some of the types of evidence which Collins and Loftus suggested were involved in making category membership decisions: for instance, property comparisons. If the matching processes (synonymous with assessments of “target-ness”) are not separate for each stimulus, the summed set of properties of redundant targets could be compared to those of the category to be matched. This form of coactivation involves evidential summation in the determination of category membership itself.

Finally, Figure 1.2c shows a divergence between race and coactivation models in response selection. The coactivation version involves separate assessments of category membership (here “animalness”) for each stimulus being followed by central summation of the results of these assessments. In other words, here coactivation does not occur in the matching process itself, but rather in a trial-level assessment of “target-ness” which follows. The race version involves separate processing for each stimulus at every stage throughout, such that the first of the stimuli to be categorised as a target (in this case, an “animal”) wins the race to activate the

“target-present” response.

Consequently, if any redundancy gains in semantic categorisation seem to result from coactivation, these levels of processing illustrated in Figure 1.2a–c — coactivation of semantic representation, in categorisation, and in response selection — would seem to provide the most likely sources. Note that the models displayed in Figure 1.2 do not represent an exhaustive set of all possible models. Rather, I have attempted to illustrate those which have some precedent in or draw on suggestions from the literature (e.g., Barsalou, 1999; Collins & Loftus, 1975; Mohr et al., 1996).

If, for whatever reason, coactivation accounts are *not* appropriate, a race model might explain any redundancy gains found in semantic categorisation. This would require that the categorisation process can be performed on multiple stimuli in parallel, as displayed in Figure 1.2a–c — where each model locates the finishing point of the race at a different level. As stated earlier, semantic categorisation seems an intuitively more complex task than many in which redundancy gains have been shown in the past, and there is no guarantee that dual categorisation processes can occur simultaneously. Though evidence from semantic priming studies and flanker tasks suggests that dual semantic processes can occur to some extent, these tasks generally rely on this processing being automatic. By contrast, using a redundant targets paradigm, both stimuli are task-relevant on redundant trials, and so processing is likely to be more effortful. Take Barsalou (1999)’s suggestion that categorisation occurs by way of attempts to simulate the entity whose category membership is in question. If there is only a single target category, a race model would require the one simulator to be able to simulate two items in parallel. In Collins and Loftus’ (1975) framework, it would require the assessment of two sets of “intersections” of spreading activation at once. Due to the likelihood of capacity limitations in categorisation, there is no certainty that either of these feats would be possible.

A third option could be that categorisation of redundant stimuli occurs in a serial manner. As discussed earlier, serial processing can also lead to redundancy gains, and even violations of the race model inequality, as outlined by Townsend and Nozawa (1997). For instance, if processing is exhaustive and non-target processing is slower than target processing, responses to redundant targets should be faster than responses to single targets. In addition, serial processing would seem to avoid any potential conflicts between processing mechanisms

being used to categorise redundant targets. For instance, serial processing would avoid any requirements for the parallel functioning of multiple simulators (Barsalou, 1999). Thus, it will be worthwhile to assess the validity of serial explanations for any redundancy gains found in the following experiments. This might be done by examining factors such as RT variance (which non-exhaustive models would predict to be smaller in trials with no targets), or differences between RTs for positive and negative responses (where exhaustive models such as that offered by Townsend & Nozawa, 1997 would predict universally faster positive than negative RTs). However, as results from Mohr and Pulvermüller's LDT experiments (e.g., Mohr et al., 1996) showed redundancy gains with varying display size, serial exhaustive models (which would predict no gain in such an instance) would seem intuitively unlikely; thus, I focus predominantly on the distinction between race and coactivation accounts.

Finally, given their adaptation from the experimental design used by Mohr and Pulvermüller (e.g., Mohr, Pulvermüller, Rayman, & Zaidel, 1994) in their lexical decision task studies, the semantic categorisation experiments described in this thesis bear many similarities to that task. This means that the results obtained in my experiments have the potential to shed light not just on semantic categorisation with redundant stimuli, but also on the lexical decision task findings themselves. As such, I will endeavour to use Mohr and Pulvermüller's work as a point of comparison in the following chapters.

In adapting Mohr and Pulvermüller's task there were a number of factors I had to consider to ensure that semantic categorisation was actually the main operation that participants performed. First, one of the difficulties in conducting redundant target experiments is ensuring that the right type of redundancy is being investigated. Take, for instance, those studies mentioned earlier which examined recognition of visually-presented emotional expressions (Tamietto et al., 2006, 2007). Tamietto et al. were interested in how their participants engaged in emotional processing, yet the stimuli they used were perceptually as well as emotionally redundant. Since there have been many findings of redundancy gain in simple perceptual tasks (e.g., Fischer & Miller, 2008; Schwarz & Ischebeck, 1994), this calls into question the extent to which gains resulted from redundancy in higher level processing.

If one wishes then — as I did — to ensure that any effects of redundancy are limited to processing of more complex features, one must endeavour to construct a task where other forms

of redundancy are limited, and ideally absent. In the realm of semantic categorisation, perceptual redundancy poses a particular problem in this regard, simply because items which are semantically similar also have a tendency to share physical characteristics: Trees often have the same general shape, mammals tend to have a recognisable configuration, and so on. This could allow such items to be categorised based almost entirely on these lower-level perceptual features.

Fortunately, a ready solution to this problem is available: abstract representations of concepts which bear minimal resemblance to the physical objects to which those concepts relate, which are familiar to vast swathes of the population, and the perceptual characteristics of which tend not to be strongly correlated with their semantic content. I am referring, of course, to words. The use of words in a semantic categorisation task with redundant targets allows a neat side-step of confounds caused by perceptual redundancy.

Using words in a semantic categorisation task could raise problems of its own, however. Mohr and Pulvermüller's work (e.g., Mohr, Pulvermüller, & Zaidel, 1994; Mohr, Pulvermüller, Rayman, & Zaidel, 1994; Mohr et al., 1996) showed that redundancy gains can occur in tasks involving lexical processing, meaning that this provides another potential source of redundancy needing to be discounted. Fortunately, unlike lexical decision, semantic categorisation tasks can involve non-target stimuli which are identical to targets in their "wordness", but differ in their category membership. In addition, both lexical decision and famous face judgement tasks (e.g., Mohr et al., 2002) using redundant stimuli have involved a confound between the familiarity of the stimulus and its status as a target (word/famous) and non-target (non-word/non-famous). By matching non-targets with targets on features such as word length and frequency, this confound can be obviated in a semantic categorisation task using word stimuli.

Another problem which has potentially affected previous attempts to investigate redundancy gain in tasks requiring higher level processing involves the highly repetitive pairing of stimuli with responses. As stated earlier, this could lead to a situation where automatic associations between the S-R pairs are formed, meaning that more complex processing is circumvented. Again, semantic categorisation allows a reasonably simple solution to this problem: provided the categories into which stimuli must be categorised are large enough, repetition of stimuli should be easy to minimise, meaning the formation of such associations

should be limited.

Finally, there is some evidence that the number of stimuli presented in redundant target experiments has an effect on responses (Mordkoff et al., 1996). That is to say, in some experiments there has been a confound between the number of targets, and the overall number of stimuli presented. This confound can be avoided in a semantic categorisation task by using distractor stimuli which are different from both targets (category members) and non-targets (category non-members). In a task using words, non-words would appear to be appropriate for this role, provided they are presented in such a way as to avoid biased contingencies of the type described by Mordkoff and Yantis (1991) — that is to say, provided they are uninformative about the stimuli with which they appear, and the responses which should be made.

1.5 Conclusion

Though redundancy gain is a well-established phenomenon, questions still exist regarding how and in what tasks it occurs. Evidence showing redundancy gain in higher order processing is somewhat limited by potential methodological issues with past experiments. Investigating redundancy gain in semantic categorisation thus has the potential to provide information about redundancy gain in higher order processing in general, and perhaps also about semantic processing and semantic memory. I aimed to use the experiments described next to shed light on both of these issues.

2. EXPERIMENTS 1–3

As outlined in the Introduction, redundant target experiments have predominantly used simple, abstract stimuli to shed light on perceptual, decisional, and response processing. However, there have also been some efforts to use such methods for investigating “higher level” processing. One example of this can be seen in the use of the redundant targets paradigm to investigate linguistic processing in a series of studies conducted by Mohr, Pulvermüller, and colleagues (e.g., Endrass, Mohr, & Pulvermüller, 2004; Mohr, Pulvermüller, & Zaidel, 1994; Mohr et al., 1996, 2007). Results of these studies showed significant redundancy gains for word but not non-word responses, which the authors attempted to explain by recourse to a Hebbian model of lexical representation (e.g., Pulvermüller, 1999).

Using Experiments 1–3 I aimed to determine whether similar findings would ensue in semantic categorisation. In Experiments 1 and 2, participants were asked to make speeded category membership decisions about briefly-presented word stimuli, displayed in a lateralised manner (that is, to the left and right of fixation). As the authors responsible for the LDT redundancy gain studies had suggested that hemispheric cooperation might be responsible for some of their findings (e.g., Mohr, Pulvermüller, & Zaidel, 1994), in Experiment 3, the same semantic categorisation task from Experiments 1 and 2 was used but stimulus presentation was non-lateralised (that is, stimuli were presented centrally).

2.1 Experiments 1 and 2

My first two experiments involved the brief presentation of lexical stimuli in the left and right visual fields. Participants were asked to make speeded responses indicating whether or not a word describing a member of a pre-specified semantic category was present (in either the left or right visual field, or both). In line with the cell assembly theory put forward by Pulvermüller

and colleagues (e.g., Pulvermüller, 1999), I hypothesised that a redundancy gain would be found, such that responses would be quicker and more accurate in trials where target words were presented redundantly (bilaterally) than when they were presented singly (unilaterally). Along with these variables I also measured response force. Force has been seen as indicative of coactivation in motor processing (e.g., Giray & Ulrich, 1993); if faster responses result from such coactivation, this should be evident in greater peak force in redundant trials.

The use of a semantic categorisation task allowed me to control for a number of potentially confounding non-semantic factors. First, any effects of visual redundancy could be controlled for by way of redundant trials where different words with the same category membership were presented to the left and right of fixation. Second, effects of stimulus familiarity could be controlled for by matching target and non-target stimuli on word frequency. Third, the use of a large pool of stimuli with limited repetition could ensure semantic processing. Finally, effects of stimulus number could be controlled for by presenting task-irrelevant non-word stimuli to the opposing visual field in both unilateral target and non-target trials.

Experiments 1 and 2 differed slightly in the semantic categories used, and in the instructions given to participants. In Experiment 1, words from one category — animals — were defined as targets; in Experiment 2, three categories — animals, parts of the human body, and parts of a building — were each used with separate groups of participants. This was aimed at ensuring that findings from Experiment 1 were generalisable across different categories (the experiments were planned and conducted sequentially, but are described together here because design and procedure was so similar). Also, in Experiment 1, the instructions encouraged participants to focus on the presence or absence of a member of the target category, and it was not explicitly stated that the presence of a single non-target word would be sufficient to conclude that the “target-absent” response was appropriate. In Experiment 2 the focus on the target category was removed, and more symmetrical instructions were given whereby participants were informed that the presence of a word from either within or without the target category was sufficient to conclude the “target-present” or “target-absent” response, respectively, was appropriate. Experiment 2 also contained one more block of trials than Experiment 1. In all other respects, these experiments were identical.

Method

Participants

Twenty-four undergraduate psychology students (19 female, five male; aged between 17 and 32 years; three left-handed, 21 right-handed) from the University of Otago completed Experiment 1, and another seventy-five (55 female, 20 male; aged between 18 and 31 years; four left-handed, 71 right-handed, and one who displayed no preference) completed Experiment 2, in return for partial course credit, as part of the Department of Psychology's experimental participation programme. Data from three participants in Experiment 2 (two female, one male) were removed from all analyses, due to one participant's failure to follow instructions, and the malfunctioning of force-sensitive response keys whilst the other two were completing the task.

Apparatus and stimuli

An IBM-PC Compatible 486 computer was used to run the experimental program, whilst instructions and stimuli were displayed on a CRT monitor, viewed at a distance of approximately 60cm. Behavioural responses made using the left and right index fingers were recorded by the use of two force-sensitive keys, located to either side of the computer's keyboard. These keys were linked to strain gauges, by way of which an analogue signal indicating response force was delivered to the experimental program. A response force of 100cN (approximately the force required to depress a key on a computer keyboard) was set as the threshold for assessing response times.

Three types of stimuli were used in these experiments. First, "target" stimuli were the names of animals, parts of the human body, and parts of a building. Animal names were taken from Battig and Montague's (1969) category norms. Body part and building words were taken from a norming study conducted with Otago psychology students (see Appendix A for a description of this study). This produced a pool of 82 names of animal species, ranging from three to ten letters in length, and 80 body part and building words, ranging from three to eight letters in length. Second, nouns that were neither animal names, body parts, nor building nouns were used as "non-target" stimuli. These were matched with the target words for length and frequency. For "animal" words, this was determined by way of word frequency lists

available online (Wiktionary, 2006); for “body part” and “building” words, the Wellington Written Corpus of New Zealand English (WWCNZE; Bauer, 1994) was used. Third, “distractor” stimuli took the form of pronounceable non-words generated via the ARC non-word database (Rastle, Harrington, & Coltheart, 2002). These were matched with the target and non-target words for length.

Stimuli were presented in lowercase, in serif font, with letter size of up to approximately 0.85° height and width, and with the centre of each word 3.7° to the left or right of fixation. Two stimuli were displayed on each trial, with non-words presented with equal frequency on target-present and target-absent trials, to control for any possible effect of inter-stimulus contingencies (Mordkoff & Yantis, 1991).

Procedure

Each participant completed an experimental session of between 45 and 55 minutes in length. This involved making responses to 216 target-present and 216-target-absent trials for Experiment 1, and 240 of each for Experiment 2, over the course of nine (1) or ten (2) blocks (48 trials per block). There were six trial types/conditions, tested equally often within each block: LVF target-present trials (with a target word presented to the left of fixation, and a non-word to the right), RVF target-present trials (with a target word to the right, and a non-word to the left), and BVF target-present trials (with different target words presented to both sides), as well as the target-absent counterparts to these (that is, identical to the above, but with non-target in the place of target words). Examples of stimulus displays are illustrated in Figure 2.1. This design controlled for interstimulus and non-target response contingencies (cf. Mordkoff & Yantis, 1991, Experiment 1).¹ Trial type presentation order was randomised within each block. A practice block of 24 trials preceded these, to allow participants to become accustomed to the task.

Each trial commenced with the presentation of a fixation cross for a period of 400 ms, following which stimuli were flashed to the left and right of fixation for 100 ms. Participants had 1500 ms from stimulus onset to make their response, before the trial was ended. The

¹ Though Mordkoff and Yantis (1991) described these contingencies in the context of a go/no-go task, in the context of a choice task the non-target response contingencies refer to non-target responses, rather than no-go “responses”.

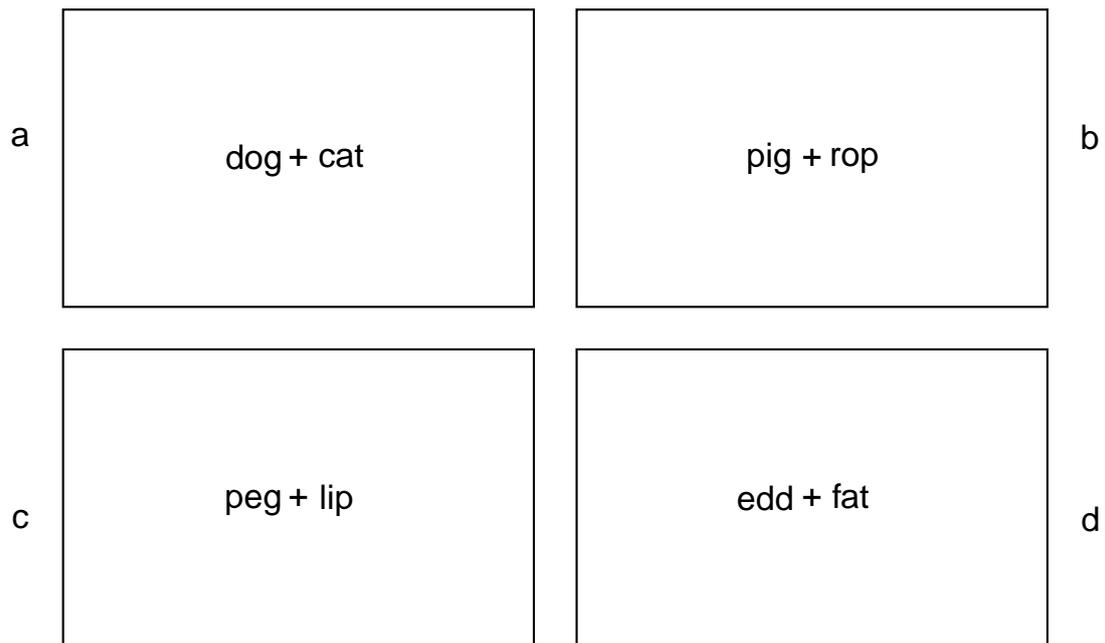


Fig. 2.1: Example stimulus displays from Experiment 1. Panel a shows a target-present trial with redundant stimuli (BVF-Present); panel b shows a target-present trial with a single target in the left visual field (LVF-Present); panel c shows a target-absent trial with redundant stimuli (BVF-Absent); and panel d shows a target-absent trial with a single non-target in the right visual field (RVF-Absent).

inter-trial interval was a constant (1000 ms) plus a randomly generated variable, uniformly distributed between 0 and 1000 ms; as such, mean ITI was 1500 ms. Participants received feedback on-screen when their responses were in any way erroneous (e.g., when they pressed the incorrect button, or when they failed to respond within the 1500 ms window). Left and right hand response assignment for target-present and target-absent responses was counterbalanced across participants (i.e., half of the participants made “target-present” responses with the left index finger and “target-absent” responses with the right, and this was reversed for the other half).

In Experiment 1, participants were instructed to press one key if a target (animal) word was present on either or both sides of the screen, and another if no such word appeared. In Experiment 2, they were instructed to press one key if a word from the target category (animals, parts of the human body, or parts of a building — varied on a between-subjects basis) was present on either or both sides of the screen, and the other key if a word from outside of the target category was present on either or both sides of the screen. In addition, in both experiments, participants were informed that non-words could co-occur with both target and non-target words, but that target and non-target words were never presented together. Finally, they were asked to respond “as quickly as you can without making too many mistakes”. Once any questions they had regarding the experiment had been answered, they commenced the practice block by pressing a foot-switch located underneath the computer desk. At the beginning of each block, written instructions appeared on-screen, and participants were required to press the foot-switch again to proceed to the following block of trials. Once all blocks had been completed, participants filled out a computerised version of the Edinburgh Handedness Inventory (Oldfield, 1971), were debriefed, and were thanked for their participation.

Data analysis

After data from practice blocks and response anticipations (RT <150 ms) were discarded, the remaining means of RT, peak force (PF), and accuracy (percent correct) for each participant were entered into separate 2×3 ANOVA with factors of Target (present versus absent) and Presentation condition (of target/non-target word stimuli: left visual field [LVF], right visual

field [RVF], or both visual fields [BVF or redundant/RED]) for each experiment.² For Experiment 2, a between-subjects Category factor was also used (with three levels: animal, body, and building). Tests of the race model inequality (Miller, 1982) were also undertaken for RT data using the RMITest program (Ulrich, Miller, & Schröter, 2007), which uses single-target RTs to determine whether redundant RTs are faster than would be predicted under a race model. Accuracy data were used to perform signal detection analyses, in which “target present” trials for each word presentation condition were considered to be signals (S), and “target absent” trials noise (N), and ANOVA were used to determine whether sensitivity differed across conditions (and categories, for Experiment 2). Criterion α -level for all statistical tests was set at .05, with Bonferroni corrections applied as noted. As a more stringent test of redundancy gain than the use of simple means, I followed Biederman and Checkosky’s (1970) process of taking each participant’s better (faster, more accurate) single-target condition and comparing this to performance with redundant targets. This method is designed to control for differences in individual locational preferences, which could otherwise present as a specious redundancy gain. Finally, a between-experiments ANOVA was conducted to compare performance in Experiment 1 with that for participants completing Experiment 2 with the target category “animals”.

Results

Analyses comparing male and female participants showed no consistent gender differences in these experiments, nor any of those reported in the following chapters. As such, I do not consider effects of gender henceforth.

Latency data

Experiment 1 Mean RTs for correct trials (approximately 75% of total trials) in each condition are displayed in Figure 2.2. Responses were significantly faster in target-present (mean RT = 624 ms) than target-absent trials (714 ms) [$F(1, 23) = 56.03$, $MSE = 5160.8$, $p < .05$], and responses in LVF trials (702 ms) were slower than those in RVF (659 ms) or BVF

² The ANOVA for accuracy were conducted on both raw and arcsin-transformed data. Except where noted, these analyses did not lead to qualitatively different results; thus the results of the raw data ANOVA are reported.

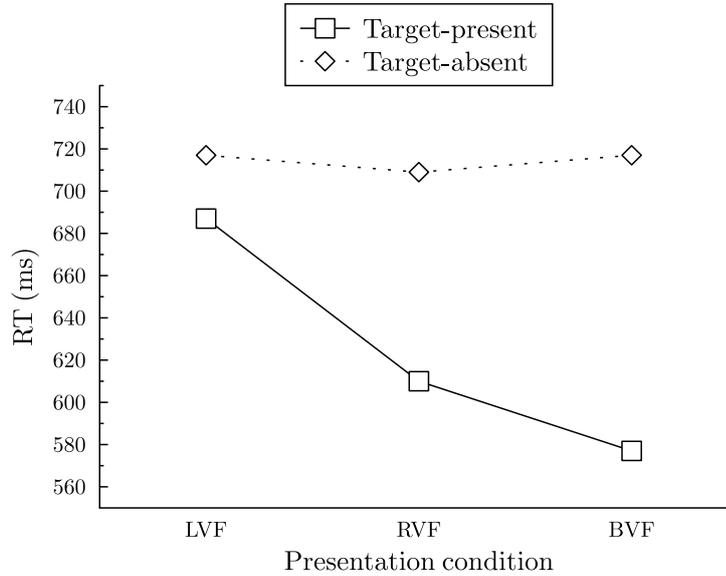


Fig. 2.2: Latency data for correct responses in Experiment 1.

(redundant; 647 ms) trials [$F(2, 46) = 12.26$, $MSE = 3257.8$, $p < .05$]. There was also a significant Target-by-Presentation interaction [$F(2, 46) = 15.64$, $MSE = 2371.8$, $p < .05$]. Separate one-way ANOVA showed a significant effect of Presentation for target-present trials [$F(2, 46) = 15.18$, $MSE = 5045.2$, $p < .05$], but not for target-absent trials [$F(2, 46) = 0.79$, $MSE = 584.4$, $p > .45$]. Two planned comparisons were also conducted on the target-present data. The first showed that responses to RVF trials were significantly faster than those to LVF trials [$t(23) = 2.85$, $p < .05$] (a right visual field advantage; RVFA). Following Biederman and Checkosky (1970), a second planned comparison showed that responses to redundant target-present trials (mean RT = 577 ms) were significantly faster than to the faster unilateral target-present RT for each participant (594 ms) [$t(23) = 2.67$, $p < .05$].

Given the high error rates, latency data for error trials were entered into a separate 2×3 ANOVA, after the removal of trials where no response was made (approximately 3.4% of error trials), and data from one participant with no erroneous responses for one condition. This showed no significant main effects, nor interactions (all $p > .27$).

Experiment 2 Mean latency data for correct responses (approximately 75% of total trials) for Experiment 2 are displayed in Figure 2.3, and it can be seen that the same patterns observed in Experiment 1 were present for all three target categories here. Responses were faster in target-present than in target-absent trials [659 vs. 728 ms, respectively; $F(1, 69) = 166.71$, $MSE = 3075.3$, $p < .05$]. A significant main effect of Presentation [$F(2, 138) = 31.47$, $MSE = 2924.6$, $p < .05$] was followed up with Tukey’s tests, which showed that responses were significantly faster in BVF trials (670 ms) than in RVF trials (689 ms), and significantly slower in LVF trials (720 ms) than either of the other two conditions. Similarly, a significant main effect of Category [$F(2, 69) = 5.64$, $MSE = 35904.4$, $p < .05$] was followed up with Tukey’s tests, which showed that responses were significantly faster when the target category was “parts of the human body” (650 ms) than when it was either “animals” (720 ms) or “parts of a building” (709 ms), which did not differ significantly from one another. The Category \times Presentation interaction was not significant [$F(4, 138) = 1.29$, $MSE = 2924.6$, $p > .25$, Greenhouse-Geisser corrected], but the Target \times Presentation interaction was [$F(2, 138) = 39.51$, $MSE = 1915.0$, $p < .05$]. This interaction is displayed in Figure 2.4. Separate one-way ANOVA for target-present and target-absent trials showed a significant effect in the former [$F(2, 142) = 38.92$, $MSE = 4295.1$, $p < .05$], but not the latter [$F(2, 142) = 0.75$, $MSE = 671.0$, $p > .45$, Greenhouse-Geisser corrected]. Tukey’s tests following the target-present ANOVA showed that all three conditions differed significantly from one another: BVF (616 ms) $<$ RVF (649 ms) $<$ LVF (711 ms). The Category \times Target interaction was also significant [$F(2, 69) = 7.24$, $MSE = 3075.3$, $p < .05$]. As shown in Figure 2.5, this appears to be due to the smaller present-absent difference when “parts of a building” was used as the target category (43 ms), compared to the other two categories (animals: 92 ms; body parts: 71 ms). Finally, the three-way interaction shown in Figure 2.3 was also significant [$F(4, 138) = 2.90$, $MSE = 1915.0$, $p < .05$, Greenhouse-Geisser corrected]. This appears to have been predominantly due to performance when the target category was “parts of a building”; unlike the other two target categories, where “target-present” responses were always faster than “target-absent” responses, for responses to Building trials LVF-present responses were slower than any of the absent responses.

As for Experiment 1, a comparison between the BVF target-present RT and the better

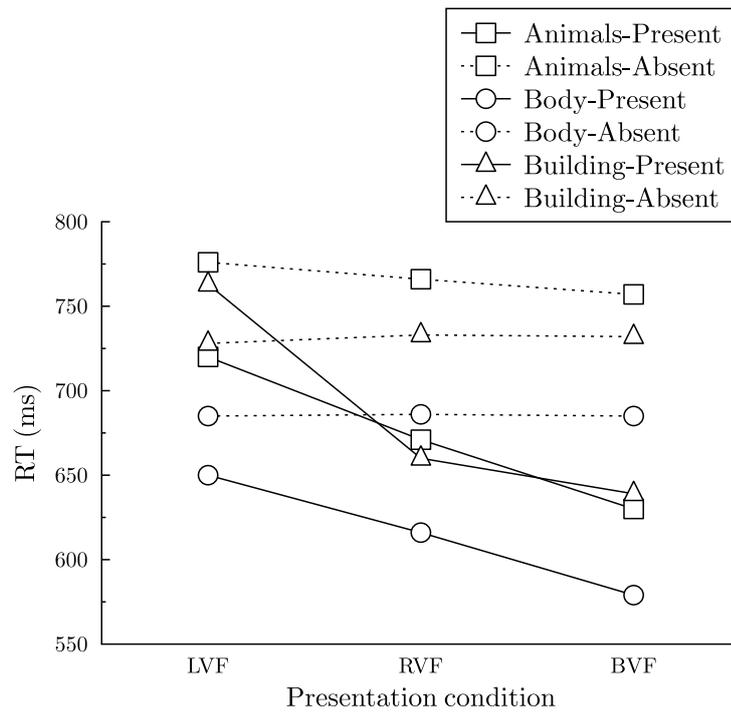


Fig. 2.3: Latency data for correct responses in Experiment 2.

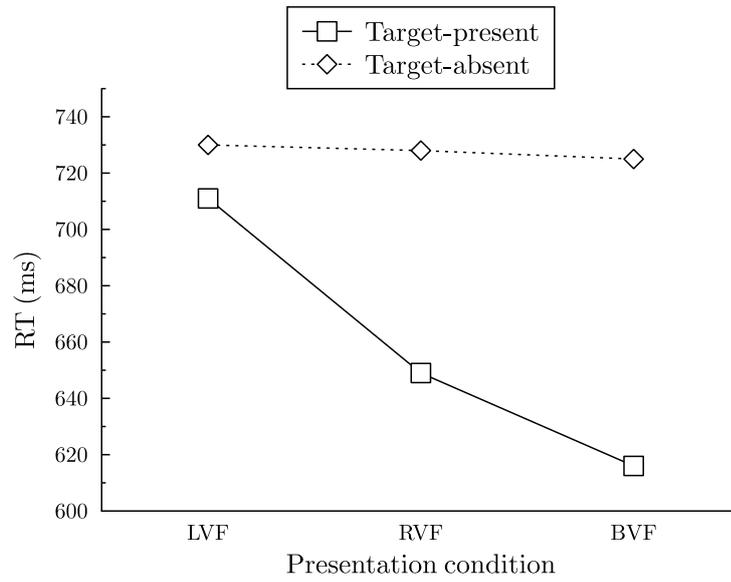


Fig. 2.4: Target \times Presentation interaction in RT data, Experiment 2.

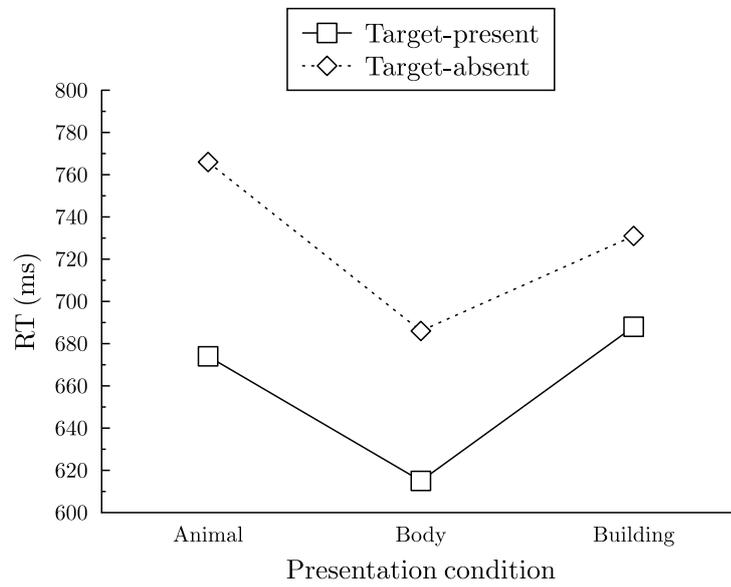


Fig. 2.5: Category \times Target interaction in RT data, Experiment 2.

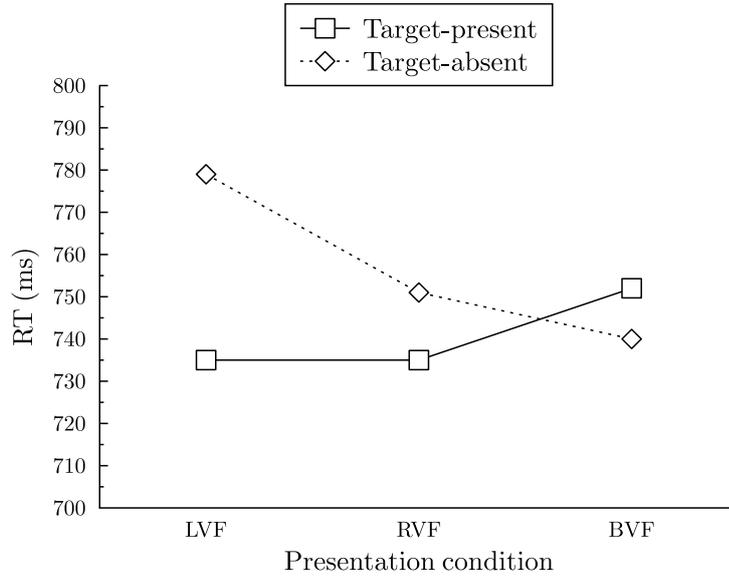


Fig. 2.6: Latency data for error responses in Experiment 2.

single-target RT for each participant was undertaken, using a 3 (Category) \times 2 (Presentation: BVF or single-target) ANOVA. This showed that the BVF RT was faster than the faster single-target RT (616 vs. 627 ms) [$F(1, 69) = 8.76$, $MSE = 548.88$, $p < .05$]. The Category \times Presentation interaction was not significant [$F(2, 69) = 0.30$, $MSE = 548.88$, $p > .7$], indicating that redundancy gain measured in this fashion was approximately equal for all three target categories.

Latency data for error trials were also entered into an ANOVA of the same design as that used for analysis of correct RTs, after the removal of data from a participant who made no errors in one condition. This showed a marginally significant main effect of Target [$F(1, 68) = 3.20$, $MSE = 8467.69$, $p = .078$], due to misses being faster (741 ms) than false alarms (757 ms). There was also a marginally significant Target \times Presentation interaction [$F(2, 136) = 3.08$, $MSE = 9356.35$, $p = .052$, Greenhouse-Geisser corrected]. As displayed in Figure 2.6, the present-absent difference was reversed for BVF trials (false alarms 12 ms quicker) compared to LVF (misses 43 ms quicker) and RVF (misses 16 ms quicker) trials. No other main effects or interactions approached significance (all $p > .1$).

Accuracy data

Experiment 1 Percentage correct for each trial type was calculated for each participant; mean values for individual conditions are displayed in Figure 2.7. Responses to target-absent trials (79.7% correct) were more frequently correct than those to target-present trials (70.3%) [$F(1, 23) = 16.87, MSE = 187.2, p < .05$], and responses to LVF trials (61.3% correct) were more frequently erroneous than those to RVF (79.8%) or BVF (83.9%) trials [$F(2, 46) = 43.33, MSE = 160.1, p < .05$]. There was also a significant interaction between these factors [$F(2, 46) = 42.27, MSE = 149.1, p < .05$]. Separate one-way ANOVA showed a significant effect of Presentation for target-present trials [$F(2, 46) = 48.24, MSE = 48.2, p < .05$], but not for target-absent trials [$F(2, 46) = 1.03, MSE = 35.5, p > .36$]. LVF-target-present trials led to correct responses on only 43.9% of occasions, a value not significantly different from chance (50%) [$t(23) = 1.29, p > .20$]. A planned comparison between RVF and LVF target-present conditions showed a significant RVF advantage [$t(23) = 5.42, p < .05$]. A second planned comparison between percentage correct for redundant-present trials (mean = 88.6%), and that for the more accurate single target-present trial type for each participant (mean = 81.0%), showed that responses to the redundant trials were significantly more accurate [$t(23) = 5.63, p < .05$].

The results of SDT analysis for Experiment 1 appear in Table 2.1. As can be seen, sensitivity was lowest in LVF trials, and highest in BVF trials, with RVF trials occupying an intermediate position [$F(2, 69) = 25.23, MSE = 0.62, p < .05$; LVF vs. RVF: $t(23) = 4.72, p < .016$; LVF vs. BVF: $t(23) = 7.58, p < .016$; RVF vs. BVF: $t(23) = 4.30, p < .016$; p-values Bonferroni corrected]. In addition, in LVF trials participants displayed a significant bias toward “target-absent” responses [$t(23) = 6.11, p < .05$], whilst the direction of this bias was reversed for BVF trials [$t(23) = 5.52, p < .05$]. Again, RVF trials occupied an intermediate position, and responses to these trials were not significantly biased [$t(23) = 1.09, p > .28$].

Experiment 2 Mean percentage correct data for Experiment 2 are displayed in Figure 2.8. Responses were more frequently correct in target-absent (78.9%) than target-present (70.2%) trials [$F(1, 69) = 46.15, MSE = 178.29, p < .05$]. There was also a significant main effect of

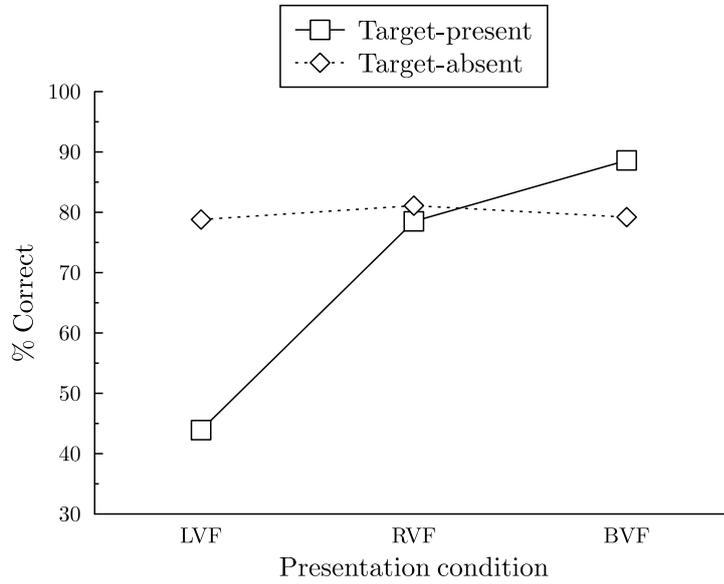


Fig. 2.7: Response accuracy data from Experiment 1.

Tab. 2.1: Sensitivity and bias for the three presentation conditions in Experiment 1.

	LVF	RVF	BVF
d'	0.67	1.85	2.21
c	0.52	0.04	-0.24

Presentation [$F(2, 138) = 91.81$, $MSE = 175.42$, $p < .05$]. Tukey's tests showed that all levels of the variable differed significantly from one another, with the most accurate responses in BVF trials (83.4%), followed by RVF trials (77.4%), and LVF trials (62.8%). There was no significant main effect of Category [$F(2, 69) = 2.19$, $MSE = 344.23$, $p > .1$], nor a significant Category \times Presentation interaction [$F(4, 138) = 1.06$, $MSE = 175.42$, $p > .35$, Greenhouse-Geisser corrected]. There was, however, a significant Target \times Presentation interaction [$F(2, 138) = 81.11$, $MSE = 159.46$, $p < .05$]. Separate one-way ANOVA were conducted for target-present and target-absent trials. In target-present trials, each of the conditions differed from both of the others, with highest accuracy in BVF trials (87.5%), followed by RVF trials (74.9%), and finally LVF trials (48.2%) [$F(2, 142) = 98.63$, $MSE = 293.16$, $p < .05$, followed by Tukey's tests]. In target-absent trials, there was only a marginally significant effect [$F(2, 142) = 3.00$, $MSE = 42.02$, $p = .056$, Greenhouse-Geisser corrected], which Tukey's tests showed was a result of lower accuracy in the LVF condition (77.5%) than in either the RVF (80.0%) or BVF (79.3%) conditions. The Category \times Target interaction was also significant [$F(2, 69) = 7.67$, $MSE = 178.29$, $p < .05$]. This interaction is displayed in Figure 2.9, and appears to be due to category-based differences between target-present and target-absent response accuracy, with a small discrepancy where animal words were targets, and a larger discrepancy where building words were targets. Finally, the three-way interaction was not significant [$F(4, 138) = 1.00$, $MSE = 159.46$, $p > .35$, Greenhouse-Geisser corrected].

A comparison between the percent correct value for the BVF target-present condition, and that for each participant's better single-target condition, was undertaken using a 3×2 ANOVA analogous to that used for the comparable comparison of latency data. This showed that responses to BVF trials were significantly more frequently correct (87.5% vs. 79.6%) [$F(1, 69) = 58.16$, $MSE = 38.63$, $p < .05$]. The Category \times Presentation interaction was not significant [$F(2, 69) = 0.74$, $MSE = 38.63$, $p > .45$].

SDT analyses were also undertaken for Experiment 2, the results of which appear in Table 2.2. A 3×3 ANOVA on d' showed a significant main effect of Presentation [$F(2, 138) = 75.96$, $MSE = 0.48$, $p < .05$, Greenhouse-Geisser corrected], which Tukey's tests revealed was a result of all levels of that factor differing significantly from one another. There was also a significant main effect of Category [$F(2, 69) = 3.20$, $MSE = 0.87$, $p < .05$], though none of the

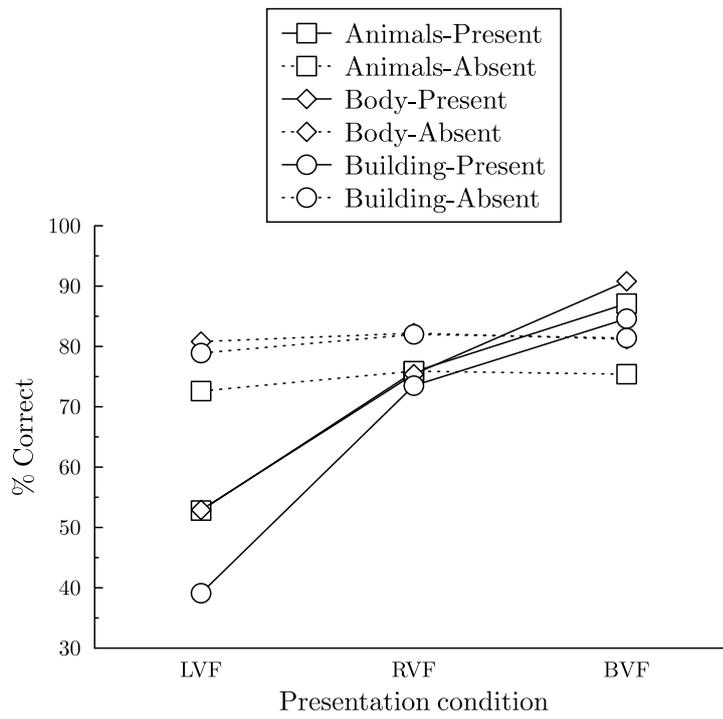


Fig. 2.8: Response accuracy data from Experiment 2.

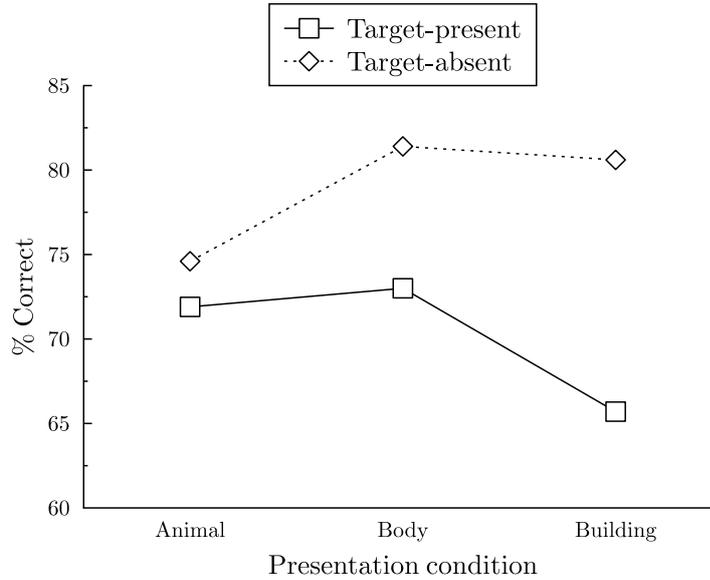


Fig. 2.9: Category \times Target interaction for % response accuracy data, Experiment 2.

pair-wise between-category differences reached the threshold for significance according to Tukey’s tests. The Category \times Presentation interaction was not significant [$F(4, 138) = 0.80$, $MSE = 0.48$, $p > .5$, Greenhouse-Geisser corrected]. A comparable analysis was carried out on c , which also showed significant main effects for Presentation [$F(2, 138) = 76.49$, $MSE = 0.09$, $p < .05$, Greenhouse-Geisser corrected] and Category [$F(2, 69) = 5.84$, $MSE = 0.11$, $p < .05$]. Tukey’s tests showed that all levels of Presentation differed significantly from one another (stronger bias toward “target-present” response in BVF than RVF and LVF; stronger bias toward “target-absent” response in LVF than RVF and BVF). The main effect of Category resulted from a significantly greater “target-absent” bias in Building than in Animal trials; none of the comparisons involving Body trials were significant.

Force data

Experiment 1 Analyses of peak force data showed no significant main effects or interactions for Experiment 1.

Tab. 2.2: Sensitivity and bias for the three presentation conditions and three categories in Experiment 2.

		LVF	RVF	BVF
Animal	d'	0.74	1.49	1.94
	c	0.28	0.02	-0.22
Body	d'	0.96	1.81	2.47
	c	0.45	0.10	-0.29
Building	d'	0.54	1.69	2.02
	c	0.58	0.13	-0.07

Experiment 2 For Experiment 2, there was a significant main effect of Presentation [$F(2, 138) = 6.32$, $MSE = 634.03$, $p < .05$, Greenhouse-Geisser corrected], which Tukey’s tests showed was a result of more forceful responses in BVF trials (494 cN) than in LVF (484 cN) or RVF (485 cN) trials. No other main effects or interactions were significant (all $p > .1$).

Tests of the race model inequality

Experiment 1 To test the race model inequality shown in Equation 1.1 (Miller, 1982), RT data for target-present trials were entered in to the RMITest application (after the removal of data from one participant for whom the accuracy rate in one condition was too low to provide sufficient correct trials for testing). T -tests showed no significant violations of the RMI occurred at any of the eight quantiles spread evenly across the CDF (0.0625, 0.1875, . . . , 0.9375), implying that a race-model explanation could not be ruled out as an account for the redundancy gain.

Experiment 2 Separate tests of the race model were performed on RTs from each category used in Experiment 2. There were sufficient correct trials to test Animal trials over 10 quantiles (0.05, 0.15, . . . , 0.95), Body trials over 9 (0.055, 0.165, . . . , 0.935), and Building trials over 13 (0.038, 0.115, . . . , 0.962); however, no significant violations of the RMI were found for any of the categories.

Between-experiments comparisons

To check for effects of the minor procedural changes between Experiments 1 and 2, I ran separate 2 (Experiment: 1 vs. 2) \times 2 (Target: present vs. absent) \times 3 (Presentation: LVF vs.

RVF vs. BVF) ANOVA for RT, percent correct, and force data, to determine whether the results of Experiment 1 and those of the participants in Experiment 2 who completed the task with “animals” as the target category differed in any meaningful way. None of the interaction terms involving the Experiment factor achieved statistical significance, though there was a marginally significant Target \times Experiment interaction for percent correct data [$F(1, 46) = 3.24$, $MSE = 244.0$, $p = .078$], due to a smaller percentage of errors in target-absent trials for Experiment 1 (79.7% correct) than Experiment 2 (74.6% correct). All other interaction terms had $p > .1$.

Discussion

As hypothesised, responses to redundant target-present trials were both quicker and more accurate (but not more forceful) than those to their single-target counterparts. By contrast, responses to redundant and single non-target trials (that is, target-absent trials with two or one non-animal word stimuli presented, respectively) showed similar values for both latency and accuracy. In some respects, the pattern of results found here is very similar to that found by previous researchers using the redundant targets paradigm with lexical decision tasks. For instance, accuracy data from Mohr et al.’s (1996) first experiment showed a redundancy gain for words but not for non-words, and an RVF vs LVF advantage for both words and non-words. In fact, as far as target-present responses are concerned (i.e., “word” responses for Mohr et al.; “animal” responses in my experiments), Mohr et al.’s results are almost identical to my own for trials with BVF and RVF presentation. However, a substantial discrepancy exists between LVF accuracy rates in their experiment, and those found here: approximately 64%, compared to 43.9% (Experiment 1) and 48.2% (Experiment 2), respectively.

This discrepancy could be informative in two respects. First, it seems to suggest that the left-hemisphere dominance evident in lexical decision tasks (as witnessed by the superior performance in RVF trials) not only persists in tasks involving semantic rather than lexical processing, but may potentially increase. This seems open to a reasonably simple explanation. Presumably determination of a word’s semantic identity first requires its lexical processing, so if lexical processing is degraded with LVF presentation, semantic processing will also suffer. When both tasks are time-sensitive (that is, participants have only a certain temporal window

in which to respond), any extra processing required for semantic categorisation beyond that for lexical decision would be likely to further degrade performance, and the poorer LVF accuracy found here could reflect that. Second, the fact that a redundancy gain was still present despite near-chance performance in LVF trials, suggests that semantic information in the LVF can still be used in the decision-making process even when it is not sufficient (or not sufficiently fast) to evoke a “target-present” response on its own. This conclusion is further supported by the signal detection analyses, where sensitivity was higher on BVF than RVF trials, implying a contribution of information presented to the LVF.

The SDT analysis also showed that response bias differed between the three conditions: there was a bias toward “target-absent” responses in LVF trials, and a bias toward “target-present” responses in BVF trials, with RVF trials somewhere in between. This might have been an artefact of different levels of sensitivity in the three conditions. If participants responded in a certain manner when not explicitly identifying a target — say, making a “target-absent” response 80% of the time, and a “target-present” response 20% of the time — then the facility with which targets are identified across different conditions will show up as a difference in response bias: those conditions where targets are easily identified will appear to have a “target-present” bias, and those where target identification is difficult will appear to have a “target-absent” bias. I return to the issue of bias in discussing later experiments.

Given the conservative assumptions of the race-model inequality and the paucity of data available for the LVF target-present condition, the fact that no significant RMI violations were found is not particularly surprising. Ideally, the RMI would be tested using a large number of correct responses for each participant in each condition, so that estimates of RT quantiles are more accurate.³ Nonetheless, the lack of RMI violations makes it impossible to rule out a statistical facilitation explanation for our results in this manner. In addition, I am unable to apply the logic used by Mohr and Pulvermüller (e.g., Mohr et al., 1996) to dismiss a race model explanation, as I have neither comparable data from split-brain patients nor any experimental condition in which multiple target stimuli were presented in a single visual field.

For the most part, the patterns of results for Experiments 1 and 2 were very similar.

³ Though note that previous research has found RMI violations with a similar number of trials. For example, Miller (1982) found RMI violations in his Experiment 1, which used only 20 trials per participant in each condition.

One of the goals in undertaking Experiment 2 was to ensure that the asymmetry between target-present and target-absent conditions in Experiment 1 (that is, VF effects and a redundancy gain for the former, but not the latter) was not simply due to the instructions in that experiment, where it was not made explicit to participants that a single non-target word was sufficient to indicate that a “target-absent” response was appropriate. Given that the same present/absent asymmetry occurred in Experiment 2 despite the change in instructional emphasis, it seems reasonable to conclude this was not the case. A second goal, to determine whether the results for the category “Animals” would also apply to other categories, also seems to have been answered effectively. Unlike the study of Smith, Shoben, and Rips (1974), where the category “Animals” appeared to be a special case leading to different results than when other categories were used, here the category defined as a target did not appear to have a large effect on the overall pattern of results.

One question remaining unanswered is the extent to which the finding of redundancy gain in a semantic categorisation task is reliant on the sort of hemispheric processing targeted by the presentation of stimuli to the left and right visual fields. Previous research has shown that the corpus callosum may be important for redundancy gain in simple detection tasks (e.g., Corballis, 1998), and there have also been efforts to determine the extent to which processing of lexical stimuli is influenced by hemispheric dominance (e.g., Hasbrooke & Chiarello, 1998). As such, evidence regarding whether the present semantic redundancy gain was a result of the presentation of stimuli to the left and right visual fields would provide further information about the hemispheric specificity of semantic processing. This question was addressed in the following experiment by presenting non-lateralised stimuli.

2.2 *Experiment 3*

My third experiment was identical to Experiment 1 except that stimuli were presented directly above and below fixation (HI and LO, respectively). If redundancy gain in semantic categorisation relies on the two hemispheres acting as separate entities, then this non-lateralised presentation should at least reduce the extent of such gains. On the other hand, if a redundancy gain persists at a similar magnitude with this presentation format, this would

imply that its origin is likely to be at least partially independent of the interaction between cerebral hemispheres.

The decision to present stimuli directly above and below fixation was made predominantly because it allowed each stimulus to be located as close to fixation as possible, thus avoiding presentation to separate hemispheres. This form of presentation represents a departure from previous redundant target experiments involving upper- and lower-visual field presentation of word stimuli, where stimuli have been presented off-centre (e.g., Mohr et al., 1996) to assess the effects of redundant presentation within a single visual field. Here, though, I was more interested in seeing whether a redundancy gain would obtain without the requirement that stimuli be presented separately to the left and right visual fields, than in assessing the contributions of the two separate hemispheres to task performance.

Method

Participants

Twenty-four undergraduate psychology students (16 female, eight male; aged between 18 and 48 years of age; two left-handed, 22 right-handed) from the University of Otago completed this experiment, in return for partial course credit, as part of the Department of Psychology's experimental participation programme. None had participated in Experiments 1 or 2.

Apparatus and stimuli, procedure, and data analysis

Apparatus and stimuli were the same as in Experiment 1, with the exception of stimulus location. In this experiment, stimuli were presented with their centre 14 pixels above or below fixation (for an example of a stimulus display, see Figure 2.10). Procedure, instructions, and data analysis were the same as in Experiment 1.

Results

Latency data

Mean correct RTs (approximately 87% of total trials) for each condition are displayed in Figure 2.11. Responses to target-present trials (mean RT = 641 ms) were faster than those to

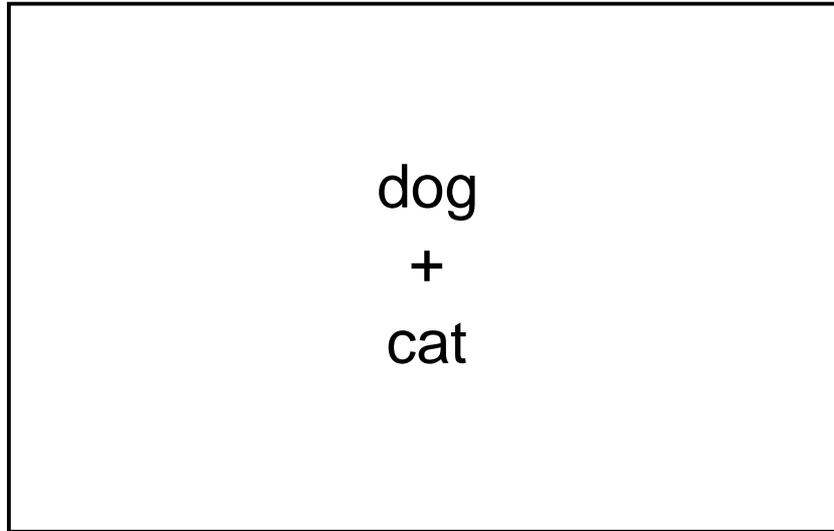


Fig. 2.10: Example of a stimulus display from Experiment 3, showing a redundant target-present trial.

target-absent trials (750 ms) [$F(1, 23) = 173.55$, $MSE = 2477.8$, $p < .05$], and responses to LO trials (727 ms) were slower than those to HI (685 ms) and redundant (RED; 675 ms) trials [$F(2, 46) = 48.21$, $MSE = 761.3$, $p < .05$]. There was also a significant Target-by-Presentation interaction [$F(2, 46) = 64.61$, $MSE = 584.5$, $p < .05$]. Separate one-way ANOVA showed a significant effect of Presentation for target-present trials [$F(2, 46) = 70.90$, $MSE = 1047.1$, $p < .05$], but not for target-absent trials [$F(2, 46) = 0.77$, $MSE = 298.7$, $p > .46$]. Again, two planned comparisons were undertaken for the target-present trials. The first showed that responses in HI trials were significantly faster than those in LO trials [$t(23) = 7.07$, $p < .05$]. The second, between the target-present redundant RT (mean = 596 ms) and the fastest single target-present RT for each participant (mean = 622 ms), showed that responses to redundant trials were significantly faster [$t(23) = 4.41$, $p < .05$]. Latency data for error trials were also analysed, after removal of data from trials where no response was made (approximately 6.4% of error trials), and data from nine participants who made no errors in at least one condition. Data from the remaining 15 participants showed that erroneous “target present” responses (mean RT = 676 ms) were significantly faster than erroneous “target-absent” responses

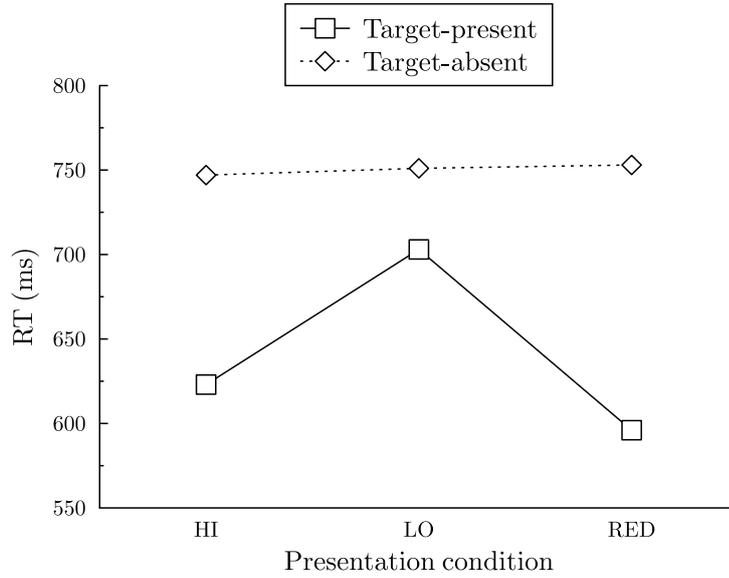


Fig. 2.11: Latency data for correct responses in Experiment 3.

(733 ms) [$F(1, 14) = 6.96$, $MSE = 10331.7$, $p < .05$]. There were no other significant main effects or interactions.

Accuracy data

Mean percent correct values for each condition are displayed in Figure 2.12. More accurate responses were made in target-absent (88.7% correct) than in target-present (86.0%) trials [$F(1, 23) = 5.51$, $MSE = 46.50$, $p < .05$], and responses to LO trials (80.7%) were less accurate than responses for HI (88.6%) and redundant (92.7%) trials [$F(2, 46) = 32.10$, $MSE = 55.41$, $p < .05$]. There was also a significant target-by-presentation interaction [$F(2, 46) = 35.21$, $MSE = 45.54$, $p < .05$]. Separate one-way ANOVA showed a significant effect of Presentation for target-present trials [$F(2, 46) = 37.98$, $MSE = 88.09$, $p < .05$], and a marginally significant effect for target-absent trials [$F(2, 46) = 2.94$, $MSE = 12.87$, $p < .1$]. Planned comparisons for target-present trials showed that responses were significantly more frequently correct in HI than LO trials [$t(23) = 4.04$, $p < .05$], and that redundant trials (97.3% correct) were significantly more frequently correct than the most accurate single target-present trial type for

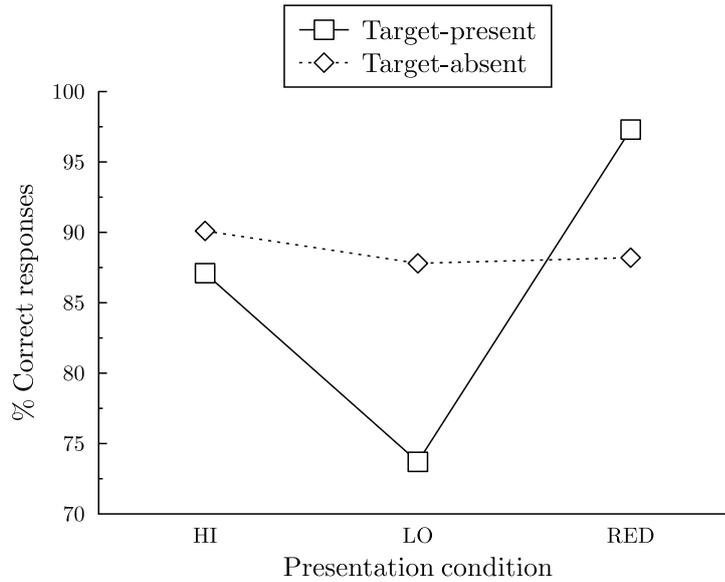


Fig. 2.12: Response accuracy data for Experiment 3.

Tab. 2.3: Sensitivity and bias for the three presentation conditions in Experiment 3.

	HI	LO	RED
d'	2.56	1.92	3.58
c	0.08	0.24	-0.57

each participant (mean accuracy = 88.5%) [$t(23) = 9.18, p < .05$].

Results of the SDT analysis appear in Table 2.3. Six participants had 100% accuracy rates in BVF-present trials; thus, I used the “adaptive estimator” method described by Hautus and Lee (2006) in calculating sensitivity and bias for that condition. As can be seen, sensitivity was lowest in LO trials, and highest in redundant trials, with HI trials occupying an intermediate position [$F(2, 69) = 29.07, MSE = 0.58, p < .05$; LO vs. HI: $t(23) = 4.49, p < .016$; LO vs. redundant: $t(23) = 7.93, p < .016$; HI vs. redundant: $t(23) = 4.87, p < .016$; p -values Bonferroni corrected]. In addition, in LO trials participants displayed a significant bias toward “target-absent” responses [$t(23) = 5.00, p < .05$], whilst this bias was reversed for redundant trials [$t(23) = 5.90, p < .05$]; again, HI trials occupied an intermediate position, though a significant bias was still evident [$t(23) = 2.11, p < .05$].

Force data

Analysis of peak force data showed no significant main effects or interactions.

Test of the race model inequality

RT data for target-present trials were entered in to the RMITest application. No significant violations of the RMI occurred at any of the twenty quantiles evenly spaced across the CDF (0.025, 0.075, . . . , 0.975), implying that a race-model explanation could not be ruled out in accounting for the redundancy gain.

Discussion

The redundancy gain found here provides evidence to suggest that the findings of Experiments 1 and 2 were not simply a result of the lateralised presentation used there, but instead an illustration of a general property of the processes involved in extracting semantic information from written words. In fact, the magnitude of the gain in Experiment 3 (26 ms faster and 8.8% more accurate responses on redundant trials than the average of the fastest and most accurate single-target condition for each participant) was numerically greater than the gains found in Experiments 1 and 2. Nonetheless, the lack of significant RMI violations in the RT data in this experiment means that a statistical facilitation explanation for these results cannot be ruled out on this basis.

The numerically higher overall accuracy rate in this experiment (approximately 87%) compared to Experiments 1 and 2 (approximately 75% each) suggests that participants found the task easier with stimuli presented above and below fixation, rather than to the left and right. This is perhaps not surprising; the offset from fixation to stimulus centre here was approximately 1° , compared to 3.7° of visual angle in Experiments 1 and 2, meaning that stimuli appeared less peripherally in Experiment 3.

2.3 General Discussion

The results of Experiments 1–3 indicate that a redundancy gain can occur in processing of word meanings when the redundancy in question is entirely semantic. Although the findings of

a number of previous studies have been consistent with the presence of a redundancy gain in higher-order processing (e.g., Baird & Burton, 2008; Laurienti et al., 2004; Mohr et al., 2002), to my knowledge none has provided unequivocal evidence in this regard, for reasons reviewed in the Introduction. As such, this study represents the most conclusive demonstration of a redundancy gain in higher-order processing to date, as well as the first demonstration of a redundancy gain in the processing of semantic information from word stimuli.

Possible explanatory models

The lack of any race model inequality violations in either experiment leaves open the possibility of a statistical facilitation-based (Raab, 1962) account of these results, and both race model- and coactivation-based accounts of data from Experiments 1–3 appear plausible. The primary pattern of results to be explained relates to the asymmetrical findings for target present and target absent trials; that is, a redundancy gain for the former, and no such gain for the latter. I outline possible explanations of this pattern within both race and coactivation frameworks next.

A race model explanation of the findings could invoke some form of “stopping rule” where both stimuli are processed fully only where the first to finish indicates a “target-absent” response. In other words, with a model of this sort, a “target” classification of either stimulus leads to an immediate “target-present” response; otherwise, processing continues until a target is identified, or until both stimuli are classified as non-targets. A similar model was mentioned in the Introduction, as an alternative explanation for Mohr and Pulvermüller’s (e.g., Mohr, Pulvermüller, & Zaidel, 1994; Mohr et al., 1996) lexical decision experiments. The source of the redundancy gain, according to such a model, is that on trials with redundant targets the two engage in a “race” to activate the appropriate response, whereas on single-target trials, as only one of the stimuli is a target, no such race can occur. This is a textbook example of statistical facilitation (e.g., Raab, 1962). Given that no targets are present on any of the target-absent trials, RTs are equal across the three target-absent conditions, as both stimuli are processed in their entirety on each trial (assuming that processing times for non-words and non-target words are similar). This aspect of the account is somewhat analogous to Ratcliff’s (1978) theory of memory retrieval, where negative responses involve a form of exhaustive processing.

Though this model requires parallel processing, it is worth noting that a similar model in

which stimuli are processed serially could also be forwarded to explain the redundancy gain found here, and the target-present/target-absent asymmetry. This would require processing to be self-terminating when a target is identified, and exhaustive otherwise. Redundancy gains in target-present trials would result from the increased probability — compared to single-target trials — that the first stimulus to be processed is a target (and thus leads to a response); and the present/absent asymmetry would result from processing always being exhaustive in target-absent trials, irrespective of what the stimuli are or where they are presented. This model would be more restricted; for instance, the extent of redundancy gains would necessarily equal the average difference in processing times between a distribution made up of single stimulus processing times (for redundant trials), and one combining instances where the first stimulus processed is a distractor and others where the first stimulus processed is a target (for single-target trials). In addition, if not artefactual, the slower LVF target-present RTs for “building” trials in Experiment 2, compared to target-absent RTs, would require that processing LVF target stimuli of that sort take far longer than processing non-targets in either visual field. As such, a serial explanation of these results may be less plausible, but pending further data it can not be ruled out entirely.

An explanation relying on some form of evidence summation — in other words, a coactivation model — would also appear to offer an account of most aspects of the results of these three experiments. Given the qualitative similarities (e.g., asymmetrical redundancy gain for targets and non-targets, RVF advantage) between my results and those of Mohr and Pulvermüller (e.g., Mohr et al., 1996), the particular explanation outlined here relies on Pulvermüller’s (1999) cell assembly theory; though other coactivation accounts (e.g., those mentioned in the Introduction) are also possible. Essentially, the proposed model — which corresponds to the coactivation version of Figure 1.2a — works as follows: target stimuli activate the semantic representation for the target category (in this case, that for “animals”, “parts of the human body”, or “parts of a building”), and then the participant uses the level of activation in this cell assembly as the guide for choosing what response to make. When this level of activation reaches a threshold, the participant concludes that a target was presented, and responds accordingly. After Mohr and Pulvermüller’s suggestions, a redundancy gain in target-present trials is explained by the more rapid ignition (that is, the more rapid onset of

self-sustaining activity in a network of neurons; Pulvermüller & Mohr, 1996) of the target cell assembly with redundant than with single target presentation. The lack of a symmetrical gain for target-absent trials may be due to the lack of a coherent cell assembly for the category “things that are not animals/body parts/building words”. This is analogous to Mohr and Pulvermüller’s (e.g., Mohr, Pulvermüller, & Zaidel, 1994) explanation for the lack of a redundancy gain for non-words in their LDT experiments. Another possibility is that participants respond only on the basis of activation in the target cell assembly, so that activity in other, non-target assemblies is essentially ignored. Either way, in this instance the decision process posited is thus: Over time, in the absence of activation in the target assembly, there is a gradual drift toward a target-absent response, which is arrested and reversed by the accumulation of evidence that a target is present (in the form of increased activation of that cell assembly). This explanation appears to account for the results of Experiments 1–3 quite well.

RVF/HI advantages

Regardless of whether redundancy gain in semantic categorisation is interpreted in terms of race or coactivation models, it is also necessary to consider potential explanations for the advantage of RVF (versus LVF) and HI (versus LO) conditions found in these experiments. First, in accounting for their finding of RVF advantage in lexical decision tasks, Mohr et al. (1996) posited a lateralisation of lexical cell assemblies, such that words are represented unequally across the cerebral hemispheres, and information presented to the right visual field is initially processed by the hemisphere possessing the majority of the assembly, thus leading to better responses in that condition. This suggestion could plausibly be extended to my experiments, and there have been suggestions that semantic representation also tends to be left-lateralised (Binder et al., 2009). However, whilst the possibility of faster lexical processing in the left hemisphere would appear to account for the RVF-advantage in Experiment 1, it is difficult to see why this would also lead to the HI-advantage found in Experiment 2.

An alternative to a hemisphere-specific account is an explanation that relies on preferential attention to one or another stimulus location, which would seem consistent with any of the models I have offered above. For instance, Mishkin and Forgays (1952) showed that words presented to the right of or below fixation were more easily recognised by those with

English as a first language than when words were presented to the left or above; this pattern was different for bilingual participants who also read Yiddish (which reads from right to left) when words in that language were presented. Heron (1957) then extended these results to show that this sort of preferential attention was quite variable, and could favour different visual locations depending on the manner in which stimuli were presented.

Conclusion

In sum, the experiments described here provide evidence for a redundancy gain in semantic categorisation of word stimuli; one which does not require the presentation of stimuli to the left and right visual fields, and thus one which appears not to rely on the two cerebral hemispheres processing separate stimuli. Whilst it is difficult to decide, on the basis of these results, between the race model and coactivation-based explanations offered above for these findings, some form of parallel processing of semantic stimuli seems to occur in this context. This issue will be investigated further in the coming chapters. Next, however, I present an experiment aimed at determining the source of the poor performance when stimuli were presented to the left visual field.

3. EXPERIMENT 4

A surprising statistic from Experiments 1 and 2 was the low level of accuracy in participants' responses to LVF target-present trials. One potential explanation for this is that the processing of stimuli presented in the left visual field was too slow given the demands of the task. This could be due to a left hemispheric dominance for language processing of the sort proposed by Mohr et al. (1996), to difficulties at an earlier (pre-linguistic) processing stage, or to some combination of the two. For instance, in a review of research conducted using lateralised presentation of linguistic stimuli, Chiarello (1988) suggested that the generally poorer performance for stimuli presented in the LVF than the RVF might be because encoding of visually presented words occurs serially on a letter-by-letter basis — and thus more slowly — when presented to the LVF (and thus processed in the right hemisphere). This would provide one potential mechanism behind a speed-of-processing related LVF inferiority for word stimuli, stemming from the absence of specialised word encoding mechanisms in the right hemisphere. In any case, whatever might be behind it slow processing of LVF stimuli could mean participants often reached the conclusion that no target was present and made the “target-absent” response before post-perceptual processing of the stimulus was complete. This account relies on the assumptions that participants (a) considered absence of target evidence as evidence of target absence, and (b) were aware of the pressure to make a response of some sort prior to the trial's end. Feedback spontaneously offered by participants from Experiment 1 suggested that, on a number of occasions, they perceived their response as erroneous almost as soon as they had made it (that is, prior to receiving the automated feedback to that effect). This could be conceived of as anecdotal evidence supporting this “slow processing” account. In addition, previous research suggests that, when participants are aware of time pressure — as they presumably would have been in Experiment 1 — they will sometimes respond prior to the

conclusion of stimulus evaluation (Kutas, McCarthy, & Donchin, 1977).

One potential method of determining the veracity of this account would be to reduce any effects of time pressure by allowing participants longer to respond. In the experiments described thus far, the response window (from stimulus onset to forced trial end) was 1500 ms, and each participant was instructed to respond “as quickly as you can without making too many mistakes”. In Experiment 4, this response window was doubled to 3000 ms, and response accuracy was emphasised in the instructions and in the rewards/penalties for correct/incorrect responses. If the “slow processing” account is valid, this should reduce the discrepancy in accuracy between the RVF and LVF conditions.

In addition, if increasing the response window provides more of an opportunity for LVF stimuli to be properly processed, this could inform questions regarding the source of redundancy gain in semantic categorisation. If information from both stimuli is used in responding to redundant trials, then improvements in LVF performance should also lead to increased redundancy gains — particularly in accuracy. Further, greater proportions of correct responses to LVF trials should allow for better tests of the race model inequality (Miller, 1982), which were limited in Experiments 1 and 2 due to a lack of data from LVF trials. Between a stronger effect and more precise RMI tests, if LVF performance improves with reduced time pressure, the chance of observing violations of the inequality should be increased.

Method

Twenty-four undergraduate psychology students (20 female, 4 male; 19 right-handed, 5 left-handed; aged 18-48 years) took part in this experiment. None had participated in any of the three previous experiments. Participation occurred through the Department of Psychology’s experimental participation programme, and participants received partial course credit for their time.

Apparatus and stimuli were the same as in Experiment 1. Procedure was altered in only a few respects. First, as mentioned, the response window was increased from 1500 to 3000 ms. Second, error messages appearing after incorrect responses were presented for an extended duration, slowing participants’ progress. These two factors were designed to give participants both the opportunity and the incentive to respond as accurately as possible. Participants were

also informed of the length of the response window and the time penalty for errors, and instructed to focus on the accuracy — rather than the speed — of their responses. Third, given the longer response windows, the number of blocks was reduced to 8 (plus a practice block to begin with), so that each session could still be completed in roughly 45 minutes. Data analysis was also similar, with the addition of comparisons between the results of Experiment 4, those of Experiment 1, and those of Experiment 2 where participants responded to the same target category (i.e., animals).

Results

Latency data

Participant mean RTs for correct trials (approximately 78% of total trials) in each condition are displayed in Figure 3.1. These individual participant means were entered into a 2 (Target: present or absent) \times 3 (Presentation: LVF, RVF, BVF) ANOVA. This showed a significant effect of Presentation condition [$F(2, 46) = 8.98, p < .05$]. Follow-up Tukey's tests showed that responses in BVF trials (mean RT = 839 ms) were non-significantly faster than those in RVF trials (860 ms), and that both BVF and RVF responses were significantly faster than those in LVF trials (894 ms). Responses in target-present trials (752 ms) were also faster than responses in target-absent trials (977 ms) [$F(1, 23) = 55.85, p < .05$]. A significant target-by-presentation interaction was also apparent [$F(2, 46) = 13.71, p < .05$]. Separate one-way ANOVA showed a significant effect of presentation condition for target-present trials [$F(2, 46) = 13.96, p < .05$], but not for target-absent trials [$F(2, 46) = 0.83, p > .05$]. A planned comparison between LVF and RVF target-present RTs showed a significant RVF advantage [$t(23) = 2.33, p < .05$]. A second planned comparison between the redundant target-present RT (mean = 693 ms), and the faster single target-present RT for each participant (mean = 707 ms), showed that responses to redundant targets were not significantly faster [$t(23) = 1.65, p = 0.11$].

Latency data for error trials were also analysed, after the removal of no-response trials (approximately 2.3% of error trials). False alarms in target-absent trials (mean = 944 ms) were faster than misses in target-present trials (1015 ms) [$F(1, 23) = 4.71, p < .05$]; there was no significant main effect of presentation, nor a significant interaction.

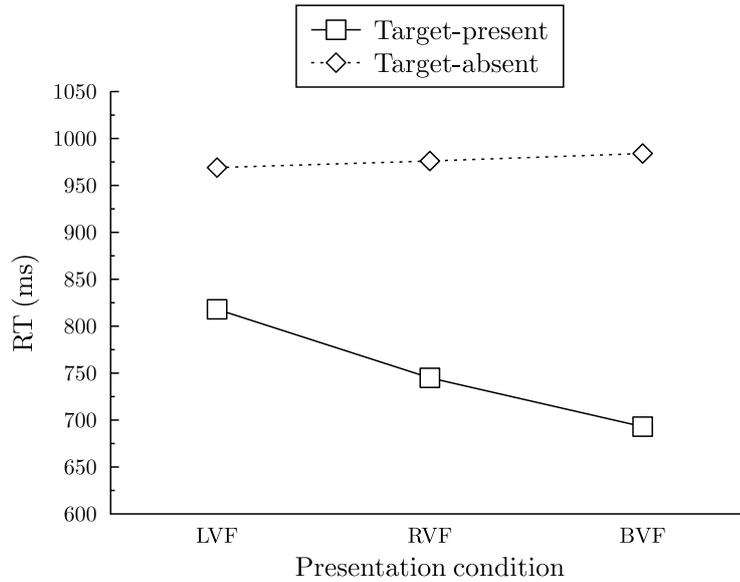


Fig. 3.1: Latency data for correct responses in Experiment 4.

Accuracy data

Of main interest in this experiment, percentage correct values for each trial type are displayed in Figure 3.2. Averages were calculated for each participant, and then entered into a 2×3 ANOVA. There was a significant main effect of Presentation condition [$F(2, 46) = 30.22, p < .05$]. Post-hoc Tukey's tests showed that responses in LVF trials were significantly less accurate (68.1% correct) than responses in RVF (81.2%) and BVF (84.6%) trials, for which percent correct values did not differ significantly. Responses to target-absent trials (80.5% correct) were marginally more accurate than responses to target-present trials (75.4%) [$F(1, 23) = 2.99, p = .097$]. There was also a significant target-by-presentation interaction [$F(2, 46) = 33.30, p < .05$].

Comparisons between the target-present and target-absent conditions for each visual field showed that LVF responses were more frequently correct in target-absent trials [79.7% vs. 56.6% for target-present trials; $t(23) = 4.21, p < .05$], BVF responses were more frequently correct in target-present trials [91.4% vs. 77.8% for target-absent trials; $t(23) = 6.26, p < .05$], and accuracy of responses in RVF trials did not differ significantly from target-present to

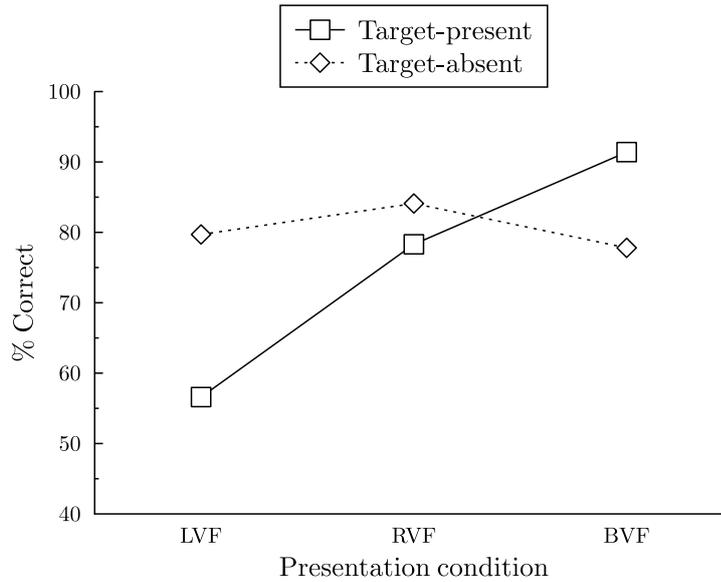


Fig. 3.2: Response accuracy data from Experiment 4.

target-absent trials [78.3% vs 84.1%, respectively; $t(23) = 1.71, p = .10$]. A planned comparison showed that responses to RVF target-present trials were significantly more frequently correct than those to LVF target-present trials [$t(23) = 3.81, p < .05$]. A second planned comparison between the accuracy rate for the BVF target-present condition, and each participant's better single target-present condition (mean = 82.3% correct) showed that responses in the BVF condition were more accurate [$t(23) = 8.07, p < .05$].

As per the previously described experiments, a signal detection analysis was also undertaken, the results of which appear in Table 3.1. As can be seen, sensitivity was the lowest in LVF trials and highest in BVF trials, with RVF trials occupying an intermediate position [$F(2, 69) = 23.94, p < .05$; LVF vs. BVF: $t(23) = 8.52, p < .016$; RVF vs. BVF: $t(23) = 2.28, p = .03$; LVF vs. RVF: $t(23) = 4.51, p < .016$; critical p -values Bonferroni-corrected].

In a similar pattern to Experiments 1 and 2, there was a significant bias toward “target-absent” responses in LVF trials [$t(23) = 4.61, p < .05$], and a significant bias toward “target-present” responses in BVF trials [$t(23) = 7.35, p < .05$]. There was also a smaller, but

Tab. 3.1: Sensitivity and bias for the three presentation conditions in Experiment 4.

	LVF	RVF	BVF
d'	1.00	1.95	2.26
c	0.42	0.14	-0.32

still significant, bias toward “target-absent” responses for RVF trials [$t(23) = 2.31, p < .05$].

Force data

Mean peak force (PF) values from correct responses for each participant in each trial type were entered into an analogous 2×3 ANOVA. There was a significant main effect of presentation [$F(2, 46) = 5.44, MSE = 411.21, p < .05$], which Tukey’s tests showed was due to LVF responses (mean peak force = 494 cN) being less forceful than BVF responses (508 cN); peak force for RVF responses (502 cN) did not differ significantly from that in either of the other two conditions. Responses were more forceful in target-present than target-absent trials [540 vs. 463 cN; $F(1, 23) = 4.26, MSE = 50632.0, p < .05$]. There was also a significant interaction [$F(2, 46) = 9.25, MSE = 254.42, p < .05$], as displayed in Figure 3.3. Separate one-way ANOVA for target-present and target-absent trials with Tukey’s tests showed that responses in LVF target-present trials (525 cN) were less forceful than those in RVF (545 cN) and BVF (550 cN) target-present trials (which did not differ significantly from one another) [$F(2, 46) = 10.59, MSE = 414.18, p < .05$], but a similar difference was not apparent in target-absent trials [$F(2, 46) = 0.81, MSE = 251.45, p > .05$; LVF, RVF, and BVF values of 464, 459, and 465 cN, respectively].

Test of the race model inequality

Latency data for target-present trials were entered into the RMITest application (Ulrich et al., 2007). This showed no significant RMI violations across any of the sixteen quantiles (.031, .094, .156, . . . , .969) tested.

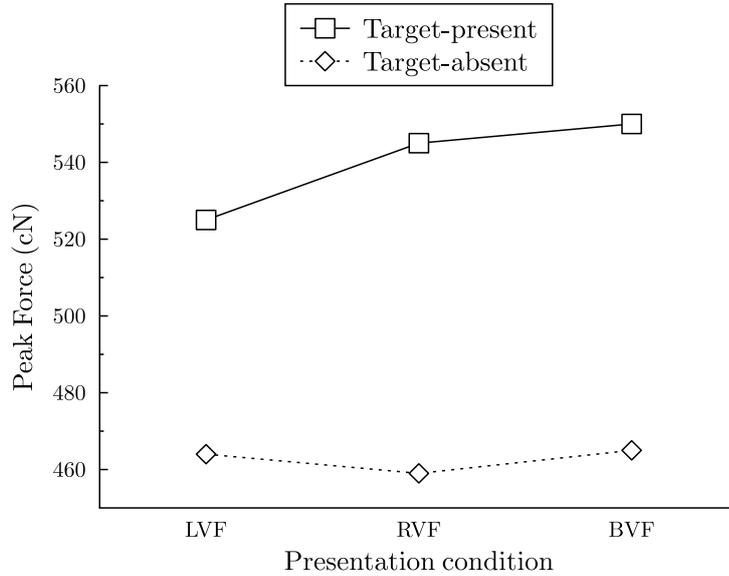


Fig. 3.3: Mean peak force from correct trials in Experiment 4.

Between-experiment comparisons

Figure 3.4 shows the mean RTs (panel “a”) and percent correct values (panel “b”) in Experiments 1, 2, and 4. Analyses of both these variables are detailed next.

Between-experiment RT differences were assessed by way of a 3 (Experiment: 1, 2, 4) \times 2 (Target: present, absent) \times 3 (Presentation: LVF, RVF, BVF) mixed ANOVA. Responses in Experiment 4 (mean RT = 864 ms) were slower than those in Experiments 1 (669 ms) and 2 (720 ms), leading to a significant main effect of experiment [$F(2, 69) = 21.51, p < .05$].

Bonferroni-corrected pairwise comparisons showed that the difference between Experiments 1 and 2 was not significant, but that responses in Experiment 4 were significantly slower than those in either of the other experiments. There was also a significant Target \times Experiment interaction [$F(2, 69) = 15.67, p < .05$], due to a substantially larger discrepancy between target-present and target-absent RTs for Experiment 4 (225 ms) than for Experiments 1 and 2 (90 and 93 ms, respectively). Presentation \times Experiment, and Target \times Presentation \times Experiment interactions were not significant.

An analogous analysis using percent correct data showed a marginally significant main

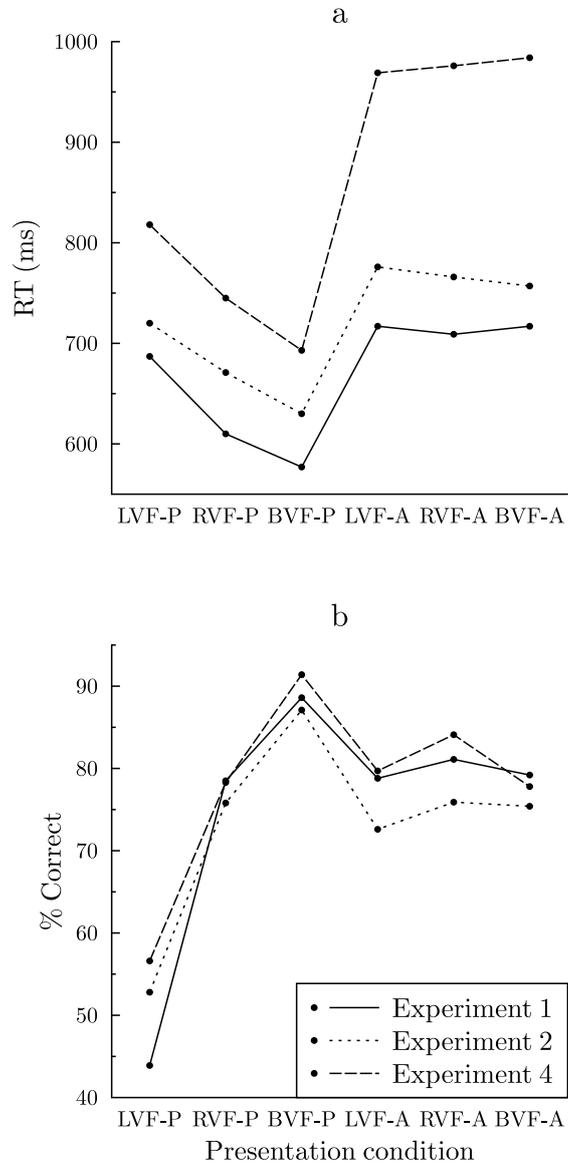


Fig. 3.4: Reaction time (a) and percent correct (b) for Experiments 1, 2, and 4, across all six conditions. Suffix “-P” refers to target-present trials; suffix “-A” to target-absent trials.

effect of Experiment [$F(2, 69) = 2.90, p = .062$]. Tukey's tests showed that the difference between percent correct values from Experiments 2 and 4 was responsible, with the test value (4.73) just shy of the critical value (4.76). None of the Experiment \times Presentation, Experiment \times Target, or — most importantly — the three-way interaction were significant (all $p > .1$).

A 3 (Experiment) \times 3 (Presentation) ANOVA conducted on sensitivity (d') also showed a marginally significant main effect of Experiment [$F(2, 69) = 3.09, p = .052$], and no significant interaction. Tukey's tests again showed that the difference between Experiments 2 and 4 was responsible for the marginally significant main effect, with d' higher in the latter (1.74) than the former (1.39).

Given the significant effects on PF detailed earlier, we also conducted a between-experiments analysis of force data, structured identically to the RT and percent correct analyses. This showed that PF was marginally higher in Experiment 1 compared to Experiment 2, with Experiment 4 falling between the two [$F(2.69) = 2.97, p = .058$]. PF was significantly higher in BVF trials than in either RVF or LVF trials, leading to a main effect of Presentation [$F(2.138) = 7.62, p < .05$; Tukey's tests for specific comparisons]. None of the interaction terms involving the Experiment factor approached significance. As displayed in Figure 3.5, the summed data also showed a significant Target \times Presentation interaction [$F(2, 138) = 5.27, p < .05$]. Separate Experiment \times Presentation ANOVA for target-present and target-absent trials followed by Tukey's tests were used to determine the source of the interaction. In target-present trials, PF was lower in the LVF condition than either the RVF or BVF conditions, which did not differ significantly from one another [$F(2, 138) = 7.10, p < .05$]. In target-absent trials, PF was higher in the BVF condition than the RVF condition, neither of which differed significantly from the LVF condition [$F(2, 138) = 4.56, p < .05$].

Discussion

The main aim of this experiment was to see what effect increasing the response window, and encouraging participants to respond accurately rather than quickly, would have on the low proportion of correct responses participants had made to LVF-present trials in earlier experiments. This issue is detailed below, followed by some discussion of the force effects also found.

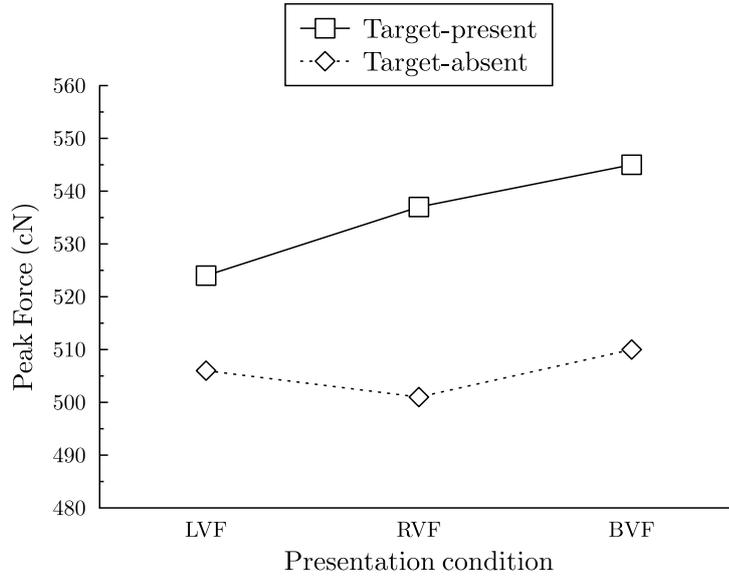


Fig. 3.5: Peak force data, summed across Experiments 1, 2, and 4.

LVF target processing

Results of the between-experiment comparisons showed that responses in this experiment were slower overall than those in either of Experiments 1 or 2, but not significantly more accurate (other than a marginal improvement in overall accuracy compared to Experiment 2).

Additionally, neither of the 3-way interaction terms (for latency or accuracy) were significant, implying that any differences that arose were not specific to the LVF-present condition. This can be taken as evidence against the “slow processing” account, according to which the scope for improvement with the decreased emphasis on fast responding should be greater in this condition. As such, the poor performance in the LVF-present condition still requires explanation. This issue will be further investigated in the experiments reported next.

Changes to the instructions, incentives, and response window did appear to alter overall performance. For instance, correct responses to target-present and target-absent trials were slower than responses in the previous three experiments. By contrast though, responses were only minimally more accurate than in either of the previous two experiments with lateralised presentation. One possible explanation for this is that participants in the earlier experiments

may have already been responding as accurately as they were able, such that the incentives and longer response window here only served to make them more cautious, but not greatly more accurate, in their responses. This would imply that processing of LVF stimuli was data-limited (Norman & Bobrow, 1975). As such, the small accuracy improvements and substantial response slowing in this experiment relative to its predecessors might simply be due to changes in the way participants responded (e.g., less fast guesses), rather than in the processing resources they were able to allocate to the task.

Of course, an unfortunate implication of the lack of improvement in LVF trial performance here is that more extensive tests of the race model inequality were not possible. In fact, the data obtained were sufficient to test the inequality over less quantiles than in Experiments 1 and 2. As such, results from Experiment 4 do not appear to offer any further information regarding the source of the redundancy gains in semantic categorisation.

Force effects

One distinct feature of the results from this experiment was the effect seen in peak force measurements, an effect which had been predominantly absent in the previous three experiments described. Here, main effects of both factors were at least marginally significant, and there was a significant interaction as well. This was reinforced by the results of the between-experiments comparison, where the summed data from Experiments 1, 2, and 4 produced a similar finding, indicating that it might simply have been low power in previous experiments which prevented any significant effects involving this dependent variable.

The summed data show that peak force was higher in the BVF condition than in either of the unilateral conditions. This implies that something about the presence of two words in a stimulus display led participants to make more forceful responses than where there was a single word accompanied by a non-word. Since both types of word stimuli in principle provided sufficient evidence for a participant to determine which response was appropriate, on its own this increased response force for redundant trials seems fitting: there are more stimuli evoking that response in redundant than in single-target trials. However, things are complicated by the presence of an interaction: BVF responses were more forceful than LVF (but not RVF) responses in target-present trials, and more forceful than RVF (but not LVF) responses in

target-absent trials. Further, data from each of the three experiments show that this pattern is consistent (if not statistically reliable in the experiments separately). Initially, there seems no clear explanation for this pattern on the basis of the number of response-informative stimuli presented in a trial. Chiarello (1988) suggested that, when word stimuli are presented laterally in lexical decision tasks, there is a bias to respond “non-word” in LVF trials, and a bias to respond “word” in RVF trials. It is possible that similar biases to respond “target-absent” and “target-present” occurred for LVF and RVF trials, respectively, in our experiments. This could be responsible for these varying differences: the “target-present” response bias with RVF presentation leads to stronger “target-present” responses, and weaker “target-absent” responses; thus, peak response force for RVF target-present trials is not significantly different from that in BVF trials, whereas the difference is sufficiently large in target-absent trials. The reverse would be the case for LVF presentation.

Conclusion

In sum, the results of this experiment provide evidence counter to the “slow processing” account of poor LVF performance. In addition, as well as VF effects in accuracy and RT, reasonably clear and commensurate effects on peak force were also present in this experiment, and in the summed data from Experiments 1, 2, and 4. Though the issue of force effects will receive more treatment in the General Discussion chapter, an explanation of the RVF advantage (or perhaps more accurately, the LVF disadvantage) is still pressing; thus, the following chapter describes two further experiments aimed at shedding light on this.

4. EXPERIMENTS 5a AND 5b

Because the “slow processing” account of poor LVF performance appears to receive no support from the results of Experiment 4, the origin of the RVF advantage remains unknown.

Explaining the RVF advantage for target-present trials is important, because this result represents a pronounced point of similarity between performance on the semantic categorisation task, and previously described performance on lexical decision tasks (e.g., Mohr et al., 1996). If the discrepancy between RVF and LVF performance here is not a result of some artefact of the experimental procedure or stimulus display, this would increase the confidence with which the theory and logic of Mohr and Pulvermüller’s cell assembly hypothesis might be applied to a task such as this, particularly given evidence for left-lateralised semantic representation (see e.g., Binder et al., 2009). On the other hand, should the RVF advantage in the semantic categorisation task prove to be an artefact of display geometry, this might contradict the idea that the cell assembly theory is appropriate for explaining performance in this task, or even call into question the RVF advantage as supporting evidence for this account of redundancy gain in lexical decision.

As mentioned in the discussion of Experiments 1–3, one possible explanation of the differing performance on RVF and LVF target-present trials was that the former involve the initial letters of a word being presented closest to fixation, whereas it is the final letters of the latter that appear in a comparable position. Mishkin and Forgays (1952) discounted a similar explanation of their finding of an RVF advantage in a word identification task, on the basis of results from an experiment where eight-letter words were presented with four letters to either side of central fixation. Participants in that experiment were no more or less successful in identifying the word when either the four letters to the left or the right were blurred, leading Mishkin and Forgays to conclude that both the beginning and ending of a word are of equal

value in its identification. However, this conclusion is contradicted somewhat by other research (e.g., Broerse & Zwaan, 1966) showing that words are better identified from their initial letters. Consequently, this factor cannot be ruled out as an explanation for the RVF advantage found in Experiments 1, 2, and 4.

In an attempt to shed light on whether this was the case, two further experiments were undertaken in which stimuli were presented in a lateralised manner, but with letters presented vertically, so that words read from top to bottom (rather than left to right). Given that this removes the difference between LVF and RVF stimuli in terms of the proximity of initial characters to fixation, if an RVF advantage were still present in the results from these two experiments it would indicate that another factor than proximity is responsible. For instance, Mohr et al.'s (1996) explanation for a similar RVF advantage in lexical decision tasks — hemispheric asymmetry of neural representations — could be applicable. On the other hand, if the RVF advantage were to disappear with stimuli presented in this manner, it might suggest that the differential placement of initial and later letters in the lateralised stimulus locations is to blame for its existence in the first place. This would imply either that the analogy between the LDT and semantic categorisation tasks can not be made easily, or that the RVF advantage does not provide the support for the cell assembly theory which Mohr et al. indicate.

Method

Participants

Twenty-four undergraduate psychology students participated in Experiment 5: 12 each in 5a and 5b. None of the participants had completed any of Experiments 1–4. Demographically, nine female and three male students (11 aged between 17 and 20 and one who failed to indicate age), 10 of whom were right-handed, participated in Experiment 5a, whilst twelve female students (11 aged between 18 and 20 and one of whom failed to indicate age), all of whom were right-handed, participated in Experiment 5b. All participants received partial course credit for taking part.

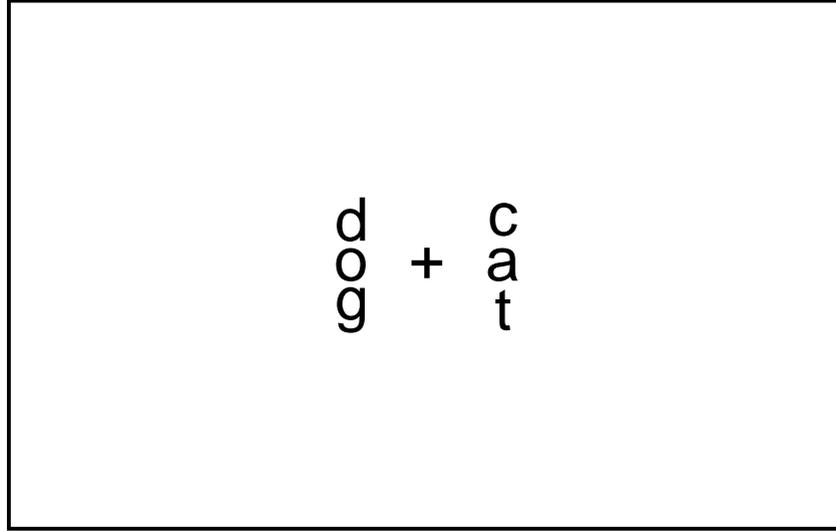


Fig. 4.1: Example of a stimulus display from Experiment 5b, showing a BVF-Present trial.

Apparatus and stimuli, procedure, and data analysis

Apparatus, procedure, and data analysis were predominantly as described in Experiment 1. Stimuli were also identical, except for the manner of presentation. In Experiment 5a, font size was halved, and both words and non-words presented to read vertically, from top to bottom, with all letters upright. The center of each word was horizontally offset from fixation by 20 pixels (approximately 1° of visual angle). Experiment 5b was identical to 5a, except that the original sized font was used, horizontal offset was 40 pixels (approximately 2° of visual angle), and stimulus duration was 150 ms, rather than 100 (for an example of a stimulus display, see Figure 4.1). This modification was a result of preliminary analyses for Experiment 5a revealing noticeably fewer correct responses than in preceding experiments; the increased font size and presentation duration were aimed at making the task easier. Consequently, analyses involved data for each DV from both experiments being first entered into ANOVA with Experiment as a factor, as well as Target and Presentation.

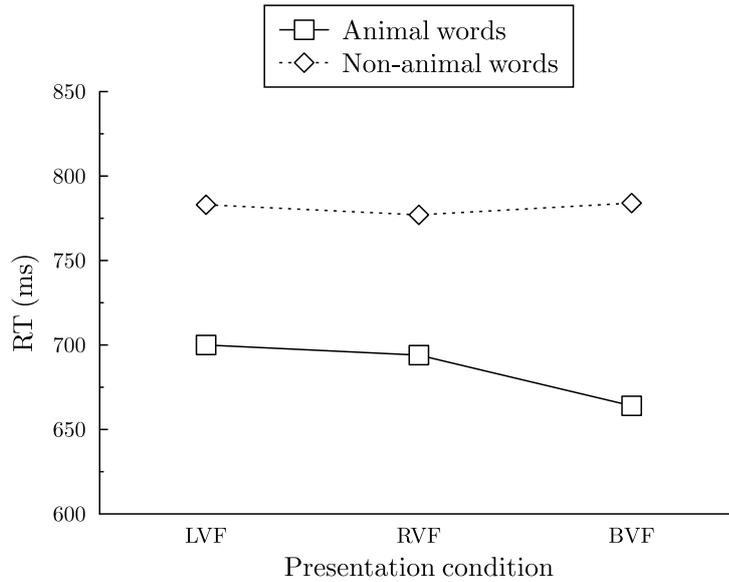


Fig. 4.2: Latency data for correct responses in Experiments 5a and 5b (combined).

Results

Latency data

Mean correct RTs for each participant in Experiments 5a (66.1% of total trials) and 5b (67.7% of total trials) were first entered into a 2 (Experiment: 5a vs. 5b) \times 2 (Target: Present vs. Absent) \times 3 (Presentation condition: LVF vs. RVF vs. BVF) ANOVA. This showed no significant main effect of Experiment [$F(1, 22) = .34, p = .57$], nor any significant interactions involving that term (all $p > .31$). Consequently, further analyses were conducted on the combined data from 5a and 5b. These combined data are displayed in Figure 4.2.

“Target-present” responses (mean RT = 686 ms) were significantly faster than “target-absent” responses (mean RT = 781 ms) [$F(1, 22) = 83.96, p < .05$]. There was also a significant main effect of presentation condition [$F(2, 44) = 4.45, p < .05$]. Bonferroni-corrected post-hoc comparisons showed that responses in the BVF condition (724 ms) were significantly faster than those in the LVF condition (742 ms; $p < .016$), and marginally faster than those in the RVF condition (736 ms; $p = .026$); the LVF and RVF conditions did not differ significantly.

A significant Target \times Presentation interaction was also found [$F(2, 21) = 7.04, p < .05$]. Separate one-way ANOVA for target-present and target-absent trials showed a significant VF effect for the former [$F(2, 22) = 10.70, p < .05$], but not the latter [$F(2, 22) = .48, p > .05$]. A planned comparison between LVF and RVF target-present RTs (means = 700 and 694 ms, respectively) showed no significant visual field advantage [$t(23) = 0.52, p > .05$]. A second planned comparison between the BVF target-present RT (mean = 664 ms), and the faster single target-present RT for each participant (mean = 671ms), showed no significant difference [$t(23) = 1.28, p > .05$].

Force data

Mean correct PF values for each participant were entered into a 2 (Experiment: 5a vs. 5b) \times 2 (Target: Present vs. Absent) \times 3 (Presentation condition: LVF vs. RVF vs. BVF) ANOVA. This showed no significant main effects or interactions (all $p > .09$).

Accuracy data

Percent correct values for each participant were entered into a 2 (Experiment) \times 2 (Target) \times 3 (Presentation) ANOVA comparable to those used for analyses of latency and force data. As there was no significant main effect of Experiment, nor any significant interactions involving that factor (all $p > .54$), further analyses were conducted using the combined data from Experiments 5a and 5b, displayed in Figure 4.3. These analyses showed no significant difference between the proportion of correct responses for target-present and -absent trials [$F(1, 22) = 2.45, p > .05$], but revealed a significant main effect of presentation condition [$F(2, 44) = 13.13, p < .05$], which Bonferroni-corrected post-hoc comparisons showed were due to more frequently correct responses in the BVF condition (mean = 72.8%) compared to the LVF and RVF conditions, which did not differ significantly from one another (means of 63.5% and 64.6% correct, respectively). The Target \times presentation interaction was also significant [$F(2, 44) = 20.12, p < .05$]. Separate one-way ANOVA showed significant VF effects for both target-present trials [$F(2, 46) = 19.60, p < .05$] — in which participants responded correctly more frequently in BVF trials (77.5% correct) than in LVF (56.7% correct) or RVF (56.4% correct) trials — and target-absent trials [$F(2, 46) = 4.99, p < .05$] — in which participants less

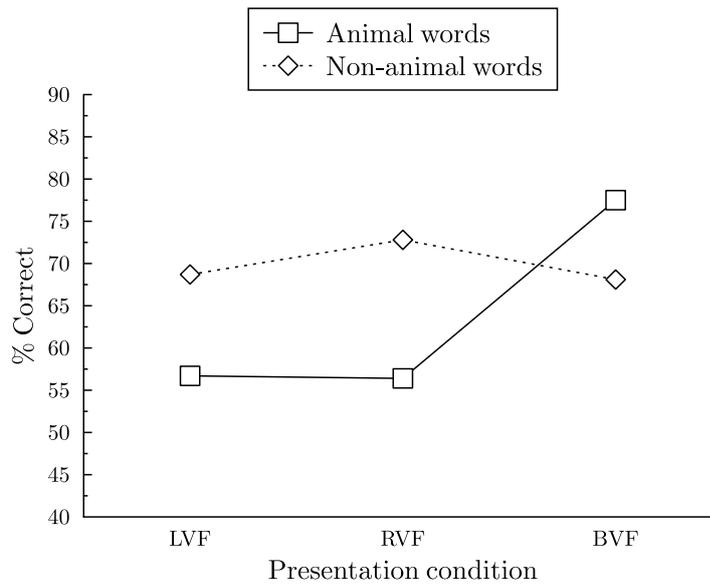


Fig. 4.3: Response accuracy data from Experiments 5a and 5b (combined).

frequently made correct responses to BVF trials (68.1% correct) than to LVF (70.3% correct) or RVF (72.8% correct) trials.

A planned comparison of percent correct values in LVF and RVF target-present trials showed no significant difference [$t(23) = .07, p > .05$]. A second comparison, between the percent correct value in the BVF target-present condition, and each participant's better unilateral target-present condition (mean = 65.4% correct) showed that responses were significantly more frequently correct in the BVF condition, $t(23) = 7.34, p < .05$.

As in Experiments 1–4, SDT analyses were also undertaken; results of these appear in Table 4.1. Participants' individual sensitivity values for the three conditions were entered into a one-way ANOVA, followed by post-hoc t -tests (Bonferroni corrected criterion $\alpha = .016$). A significant omnibus F -statistic [$F(2, 46) = 11.07, p < .05$] resulted from d' being higher for the BVF condition than for either of the unilateral conditions ($p < .016$), which did not differ from one another ($p > .3$). One-sample t -tests using c as the dependent variable showed a significant “target-absent” bias for both the LVF [$t(23) = 2.24, p < .05$] and RVF conditions [$t(23) = 3.50, p < .05$]; the BVF condition showed a significant “target-present” bias

Tab. 4.1: Sensitivity and bias for the three presentation conditions in Experiment 5.

	LVF	RVF	BVF
d'	0.81	0.85	1.30
c	0.21	0.26	-0.14

$[t(23) = 2.58, p < .05]$.

Test of the race model inequality

No race model violations were found for any of the ten quantiles (.05, .15, . . . , .95) tested.

Discussion

The goal of the two experiments reported here was to provide some indication of whether the RVF advantage over LVF performance found in Experiments 1, 2, and 4, might have been a result of the proximity of different parts of LVF and RVF word stimuli to fixation (and, thus, foveal vision). By presenting stimuli vertically, rather than horizontally, and thus ensuring that initial and final letters of both LVF and RVF stimuli were located equidistant from fixation, a similar advantage here would have provided evidence against an explanation of this sort.

As can be seen from both latency and accuracy data, and SDT analyses, no RVF advantage was apparent. Whilst RVF-present RTs were numerically faster than LVF-present RTs, this difference did not approach statistical significance; for accuracy, there was a non-significant numerical superiority for LVF- over RVF-present performance. At first glance, this would seem to unequivocally support the suggestion that RVF advantages found in Experiments 1, 2, and 4, were not a result of greater representation of word or semantic stimuli in the left hemisphere — as the extrapolation of Pulvermüller and Mohr’s cell assembly theory might suggest — but were instead dependent on the manner in which the stimuli were presented in those experiments.

However, two notable factors should be considered with reference to this conclusion. First, the accuracy values for both LVF and RVF target-present trials were extremely low, neither exceeding 57%. The apparently greater difficulty of the task (when compared to the earlier experiments), then, might have led participants to adopt different performance

strategies — for instance, deliberately attending (or even shifting their gaze) to stimuli in one or the other location, so as to maximise the possibility of correctly identifying at least one of the two stimuli appearing on each trial. The use of such a strategy would imply serial processing, which could plausibly be incorporated into the models previously outlined to account for performance in this task. On redundant target trials, attention to either stimulus would allow the participant to make the “target present” response, whereas this would only be the case on those single-target trials where attention happens to focus on that single target (rather than the non-word distractor). Thus, responses to redundant target trials would be expected to be more frequently correct than responses to single-target trials. On the other hand, as the models suggest participants do not draw any conclusions from the identification of a single non-target, a similar gain would not occur for redundant non-target trials (compared to single non-target trials). Even if this is accepted as an accurate representation of participants’ performance strategies in Experiments 5a and 5b, though, it still raises problems for the cell assembly explanation of RVF advantage: irrespective of whether stimuli were processed serially or in parallel, the supposed superior left hemisphere representation should be evident in faster and more accurate responses to RVF stimuli. This was not the case.

Second, there is some debate about the extent to which information presented in the central 2–3° of vision is projected bilaterally, or is split and projected contralaterally (e.g., Ellis & Brysbaert, 2010; Jordan & Paterson, 2009). As stimuli in this experiment were presented 1–2° from fixation, this placed them close to, or within the debated zone. Should they have not been projected contralaterally, this would undermine the ability to make any inferences about hemispheric processing from the results of Experiments 5a and 5b. However, this seems unlikely, given the similarity of the results from these two experiments. Though the stimulus display in Experiment 5a was potentially within foveal vision — with stimulus centres being approximately 1° from fixation — this is less plausible for Experiment 5b (in which stimulus centres were approximately 2° from fixation). As such, though it is worth noting the possibility that stimuli projected bilaterally, this seems unlikely.

If, then, it may be tentatively suggested that the RVF advantage present in earlier experiments and absent here is unlikely to have been a result of the greater proportion of cell assemblies associated with stimuli being located in the left cerebral hemisphere, there are two

immediately obvious factors which may have been responsible for such findings. First, as described above and in the introduction to this chapter, it may be that the difference in proximity of the informative initial letters of LVF and RVF stimuli to fixation made it easier for participants in the previous experiments to determine the identity of RVF stimuli. Second, the regular, horizontally-displayed text may have led participants' attention to focus to the right of fixation, as English-readers have been found to do (e.g., Heron, 1957).

The lack of an RVF advantage here presents several implications not only for any models of the semantic categorisation task, but further, for the lexical decision task used by Mohr and Pulvermüller (and perhaps also for other target-detection experiments with bilaterally displayed visual stimuli). First, it implies that, for these semantic categorisation tasks, the cell assembly explanation should not be preferred on the basis of differential results for stimuli appearing in different visual fields (since, had this been the reason for the RVF advantage in earlier experiments, one would have expected the advantage to persist irrespective of stimulus orientation).¹ Whilst this account still provides a seemingly plausible biological mechanism for redundancy gains in a task like this, given its flexibility and the lack of precise predictions such a model seems to make, it may even be considered less suitable than a parallel self-terminating race model (as described previously), which seems to clearly account for most aspects of the results found in all of Experiments 1–5.

Second, the lack of an RVF advantage prompts the question of whether some of the results — namely, the RVF advantage — from Mohr and Pulvermüller's experiments could also be explained as artefacts of stimulus presentation (e.g., the proximity of RVF word initial letters to fixation, or a tendency for participants to focus their attention to the right of fixation). If this were the case, this would represent a threat to their hypothesis, as it removes a reasonably major piece of the evidence (RVF advantage due to greater left hemisphere representation of stimuli) they used to support it. However, the differences between the categorisation task and the LDT with redundant stimuli mean it would not be prudent to confidently draw such a conclusion without a more direct experimental test. On the other hand, the results described here offer another possibility: Mohr and Pulvermüller's proposed lexical cell assemblies might

¹ Unless, of course, the cell assembly's activation is in some way dependent on a stimulus — or at least, a word — being presented at its "normal" orientation. This issue is discussed further in the following paragraph.

be restricted to the specific, horizontal configuration of letters making up a word (rather than reflecting all configurations designed to evoke the word's referent). This would seem to obviate the problems our findings raise for their hypothesis, though it narrows its scope substantially.

This possibility is anticipated in Chiarello's (1988) suggestion that the usual left hemisphere mechanisms for word recognition are unable to function when text is presented vertically, and that in such cases processing of words presented to the RVF occurs in the same manner theorised to occur for LVF stimuli — that is, serially, letter-by-letter. On a similar note, Bradshaw and Nettleton (1983) referred to the vertical presentation of text as “unnatural” (p. 88), and suggested that the unfamiliar form of words presented in that manner could result in different findings because hemispheric differences can also be found for the perceptual processing of unfamiliar stimuli — possibly due to the involvement of the right hemisphere in spatial transformation. This could lead to cases where a left hemisphere (RVF) linguistic advantage is counterbalanced by the fact that stimuli require right hemisphere processing, irrespective of where they are presented. Bradshaw and Nettleton's review and discussion of this literature dealt mainly with tasks involving single or multiple (sub-word) letters; nonetheless, they did mention evidence from a study where an RVF advantage for words in a lexical decision task appeared when stimuli were identifiable, but disappeared when presentation duration was shorter (Bradshaw & Nettleton, 1983). Should either of Chiarello or Bradshaw and Nettleton's explanations be valid, under the assumption that vertical presentation alters lexical processing rather than also disrupting the semantic processing to follow, it would suggest that the RVF advantages found in Experiments 1, 2, and 4 were simply a result of the use of horizontally-presented lexical stimuli rather than any asymmetry in the semantic representation of the entities to which the stimuli referred. This is because even if vertical word presentation does disrupt lexical processing, were the RVF advantages in those experiments due to an asymmetry in semantic representations the advantages should have persisted in Experiment 5 despite the interference to normal lexical processing mechanisms. That the advantages did not persist thus provides evidence against the idea that their presence in earlier experiments was a result of an asymmetry in semantic representation.

Conclusion

The two experiments described here lend support to the idea that the RVF advantages found in Experiments 1, 2, and 4, were a result of the horizontal presentation of lexical stimuli in those experiments, rather than being due to a predominantly left-hemispheric representation of the semantic entities to which those stimuli correspond. Whilst the findings of Experiments 5a and 5b are not decisive in this regard — due, for example, to the far greater difficulty participants apparently had in identifying stimuli presented vertically — they do raise questions for the cell assembly-based coactivation account of redundancy gain in semantic categorisation tasks. The cell assembly explanation of LDT results on which this account is based relies on — amongst other evidence — support from the RVF advantage Mohr and Pulvermüller (e.g., Mohr, Pulvermüller, & Zaidel, 1994) consistently found. If the inferences those authors drew from this evidence are problematic, this need not call their entire case into question. However, if combined with uncertainties in the other arguments used to prefer the cell assembly account over a race model explanation (as outlined in the Introduction to this thesis) an assessment of the account's overall validity might be required. Determining whether a similar cell assembly-based account is appropriate for the semantic categorisation task will require the manipulation of variables other than the purely perceptual; the following experiments were aimed at this end.

5. EXPERIMENT 6

In each of Experiments 1 through 5, a noticeable feature of the results was the asymmetry between the target-present and target-absent conditions. Target-present trials showed strong visual field effects and redundancy gains in both latency and accuracy, whereas target-absent trials had no such gains, and any visual field effects present were of a much smaller magnitude. Both the race model and coactivation accounts tentatively offered of these experiments suggest that asymmetrical redundancy gains are due to different precursors to “target-present” and “target-absent” responses: stimuli indicating that the latter response is appropriate (e.g., non-target words) do not seem to possess the same ability to trigger an immediate response as do target stimuli. According to the race model-based account, this is a result of the parallel, self-terminating nature of processing, which only ceases once a “target” has been identified, or once both stimuli have been processed. On the other hand, the coactivation account explains this by relying on the idea that hypothetical cell assemblies which allow a summation of activity for semantically related stimuli do not provide the same opportunities for “non-target” stimuli, which are dispersed across numerous semantic categories.

One way to test the general hypothesis that the asymmetry shown in Experiments 1–5 is a result of the asymmetry in the stimuli themselves (that is, having a clearly defined category as targets in opposition to a more disparate non-target definition) is to remove non-target trials from the picture by opposing one category to another, thus circumventing the target versus non-target dichotomy. This was the aim of Experiment 6. Participants completed a task similar to those used in Experiments 1–5, but which involved a decision about which of the categories “parts of the human body” and “parts of a building” was represented on each trial. If results were to show an asymmetry between the two responses similar to those found in earlier experiments, this would imply that such asymmetry is not a result of the broad and

ill-defined “things which are not targets” counterpoint to the target categories used. Instead, the asymmetry might reflect a general tendency for participants to focus on a single category, and define the task in terms of membership or non-membership thereof.

To that end, two methods were used in assessing the extent to which an asymmetry remained. First, the presence of redundancy gains and visual field effects in both conditions were sought. If these were absent from data pertaining to either of the two categories, this would be evidence against an explanation of asymmetrical responses based on the heterogeneity of stimuli associated with the “non-target” response in prior experiments. Of course, even without this heterogeneity, if the asymmetry had been a result of participants defining the task in terms of membership or non-membership of a particular category, it is possible that some participants in this task could thus focus on the Body category, and others the Building category. The result of this could be data showing a “compromise” between the patterns associated with target and non-target trials in the preceding experiments. Thus, to properly enable determination of whether this asymmetry changed or remained, the second method of assessing its extent involved results from those participants in Experiment 2 who had been presented with the same target categories being used as a “baseline” measure. Deviations from the patterns shown in “target-present” trials there were consequently used as an indication of whether or not the response asymmetry was altered in Experiment 6. Specifically, interactions implying smaller VF effects (that is, a more generally “flat” pattern of responding across the presentation conditions) in the present experiment than in Experiment 2 would suggest that some asymmetry had persisted here.

Method

Participants

25 first- and second-year psychology students (12 male; 2 left-handed) participated in Experiment 6. One participant failed to follow instructions, and his data were removed from all analyses. Participants were recruited through the psychology department’s participation programme, and received extra credit toward their psychology grades in return for their efforts. None of the participants had completed any of Experiments 1–5.

Apparatus and Stimuli

Word stimuli were the “parts of the human body” and “parts of a building” words as described in Experiment 2. Non-word stimuli were a selection of the non-word stimuli used in that experiment, which had been matched with the Body and Building words for length. No “non-target” words were presented in this experiment. Stimulus presentation and the apparatus used were identical to Experiment 2.

Procedure

Procedure was identical to Experiment 2, with the exception that the decision participants had to make was now between the two categories used, rather than between members and non-members of a single category. Response assignment was counterbalanced so that half of the participants made “Body” responses with their left hand, and “Building” responses with their right, and the remaining participants the reverse.

Data Analysis

Data analyses were similar to Experiment 1. The main difference was that the factor Target from previous analyses was here replaced by the factor Category. Further, to test changes in asymmetry, data from responses to each of the categories in this experiment were compared to data from both target and non-target trials from Experiment 2.

Results

Latency data

RTs from correct responses (approximately 79.2% of total responses) are displayed in Figure 5.1. Participant means were entered into a 2 (Category: Body or Building) \times 3 (Presentation: LVF, RVF, or BVF) ANOVA. Responses in Body trials (687 ms) were significantly faster than responses in Building trials (732 ms) [$F(1, 23) = 73.01$, $MSE = 1029.0$, $p < .05$]. Responses were significantly slower in LVF trials (781 ms) than in RVF (689 ms) or BVF (658 ms) trials [$F(2, 46) = 38.44$, $MSE = 5129.7$, $p < .05$, followed by Tukey’s tests]; RVF and BVF RTs did not differ significantly. The Category \times Presentation

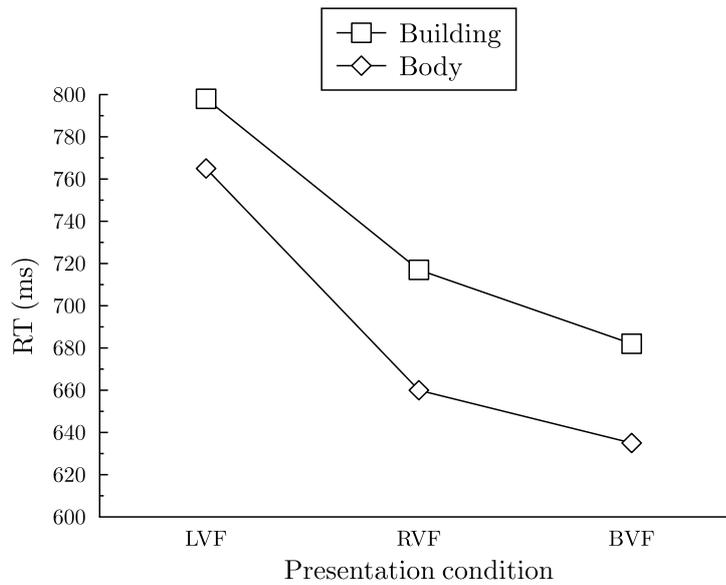


Fig. 5.1: Latency data for correct responses in Experiment 6.

interaction was not significant [$F(2, 46) = 1.21$, $MSE = 1281.5$, $p > .25$]. As a result of this, and because Tukey’s tests had already determined which of the three Presentation conditions differed from one another, planned comparisons between LVF and RVF RTs were not carried out for each category separately.

A comparable analysis was undertaken on trials with erroneous responses, after the removal of data from one participant who made no errors in one condition. The remaining data are displayed in Figure 5.2. Responses were faster in BVF trials (757 ms) than in LVF or RVF trials (both 834 ms) [$F(2, 44) = 7.20$, $MSE = 12440.7$, $p < .05$]. Error responses were not significantly quicker for one category than the other [Body: 804 ms; Build: 812 ms; $F(1, 22) = 0.39$, $MSE = 5707.0$, $p > .5$]. However, there was a significant Category \times Presentation interaction [$F(2, 44) = 3.45$, $MSE = 6722.5$, $p < .05$, Greenhouse-Geisser corrected]. Bonferroni-corrected post-hoc comparisons (revised $\alpha = .016$) showed that this was a result of erroneous responses to BVF Body trials (i.e., where a “Building” response was given) being marginally faster (728 ms) than those to BVF Building trials (i.e., where a “body” response was given; 787 ms) [$t(22) = 2.47$, $p = .022$], whereas the differences for LVF and RVF trials did

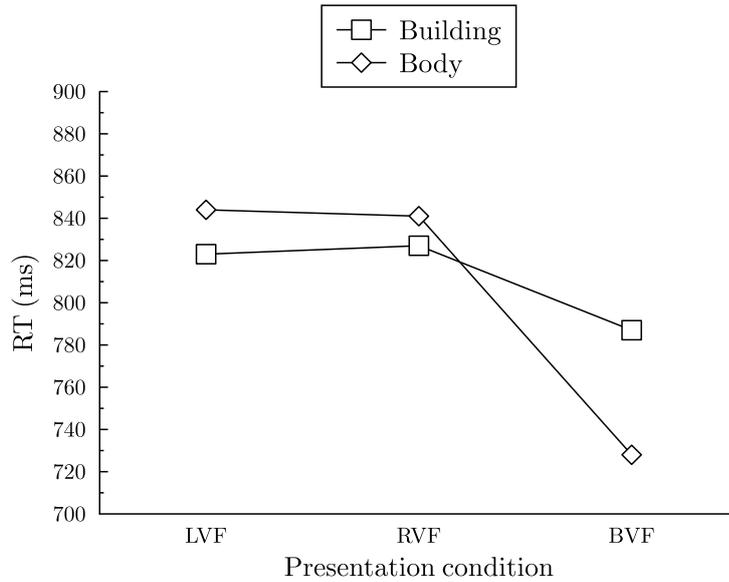


Fig. 5.2: Latency data for error responses in Experiment 6.

not approach significance (all $p > .3$).

Accuracy data

Percent correct values are displayed in Figure 5.3. Participant means were entered into a 2×3 ANOVA, which showed that responses to BVF trials were more frequently correct (90.4%) than those to RVF trials (84.8%), and that responses to LVF trials were less frequently correct (62.4%) than either of the others [$F(2, 46) = 80.11$, $MSE = 131.69$, $p < .05$, followed by Tukey's tests]. Percent correct values did not differ between Body (79.4%) and Building (79.0%) trials [$F(1, 23) = 0.09$, $MSE = 80.34$, $p > .75$]. A significant Category \times Presentation interaction [$F(2, 46) = 7.27$, $MSE = 70.91$, $p < .05$] was followed by Bonferroni-corrected post-hoc comparisons (revised $\alpha = .016$) between the Body and Building percent correct values. These showed that responses were significantly more frequently correct for Body (92.6%) than Building (88.3%) BVF trials [$t(23) = 4.61$, $p < .016$], non-significantly more frequently correct for Body (86.8%) than Building (82.7%) RVF trials [$t(23) = 2.06$, $p = .051$], and non-significantly more frequently correct for Building (66.0%) than Body (58.8%) LVF trials

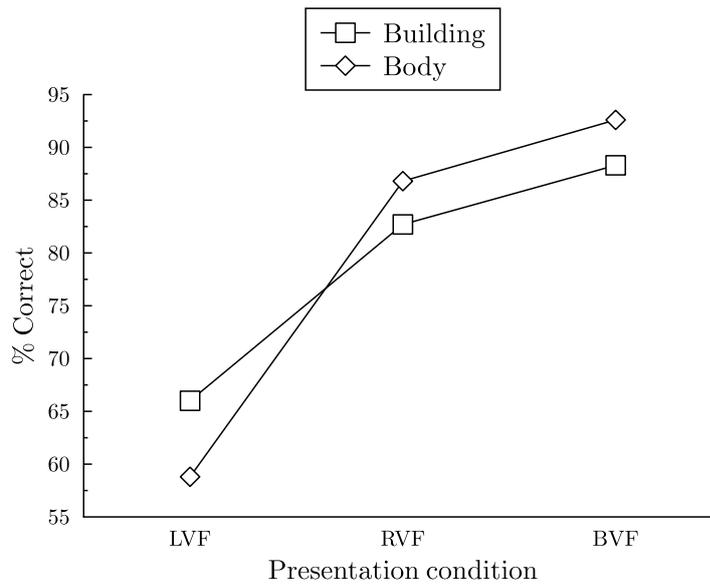


Fig. 5.3: Response accuracy data from Experiment 6.

[$t(23) = 1.93, p = .067$]. Finally, planned comparisons between BVF percent correct values and each participant's best single-target condition showed that responses were more frequently correct in the BVF condition for both Body and Building trials [Body: 92.6% vs. 88.3%, $t(23) = 3.11, p < .05$; Building: 88.3% vs. 83.9%, $t(23) = 3.22, p < .05$].

Results of SDT analysis are displayed in Table 5.1. Body trials were arbitrarily designated S, and Building trials N. As one participant made no errors in the Body BVF condition, the adaptive estimator method (Hautus & Lee, 2006) was used to determine d' and c for that condition. A one-way ANOVA, followed by Tukey's tests for specific comparisons, showed that sensitivity was significantly higher in the BVF condition than the RVF condition, and that it was significantly higher in both of these than in the LVF condition [$F(2, 46) = 75.42, MSE = 0.39, p < .05$]. Bonferroni-corrected one-sample t -tests showed that there was a non-significant bias toward the "Building" response in LVF trials [$t(23) = 1.87, p = .07$], and significant biases toward the "Body" response in RVF trials [$t(23) = 2.75, p < .016$] and BVF trials [$t(23) = 3.55, p < .016$].

Tab. 5.1: Sensitivity and bias for the three presentation conditions in Experiment 6.

	LVF	RVF	BVF
d'	0.67	2.23	2.81
c	0.09	-0.13	-0.17

Force data

ANOVA on each participant’s mean peak response force from correct trials showed a marginally significant main effect of Presentation [$F(2, 46) = 3.49$, $MSE = 1234.5$, $p = .07$, Greenhouse-Geisser corrected], with less forceful responses to LVF (450 cN) than RVF (463) or BVF (469) trials. There were no other significant main effects or interactions.

Tests of the race model inequality

RMI tests were conducted for both “Body” and “Building” trials. No significant violations were found across any of the 20 quantiles tested (.025, .075, . . . , .975) in either case.

Between-experiment comparisons

To determine whether asymmetrical responding had persisted despite the absence of heterogeneous “non-target” stimuli, latency, accuracy, and force data from Experiment 6 were compared with analogous data from Experiment 2. These analyses are described next.

Latency RT data from Body and Building trials were entered into 2 (Experiment) \times 3 (Presentation) ANOVA with data from trials involving the same category taken from Experiment 2. For Body trials — displayed in Figure 5.4 — responses were slower for Experiment 6 (687 ms) than 2 (615 ms) [$F(1, 46) = 10.12$, $MSE = 18410.9$, $p < .05$], and the interaction was also significant [$F(2, 92) = 4.47$, $MSE = 3888.3$, $p < .05$]. This appeared to be due to a greater discrepancy between the two experiments for LVF trials (115 ms) than in the other Presentation conditions (44 and 56 ms for RVF and BVF, respectively). For Building trials — displayed in Figure 5.5 — responses were again slower for Experiment 6 (732 ms) than for Experiment 2 (688 ms) [$F(1, 46) = 4.94$, $MSE = 14732.7$, $p < .05$]; however, the interaction term did not approach significance [$F(2, 92) = 0.42$, $MSE = 3566.4$, $p > .5$].

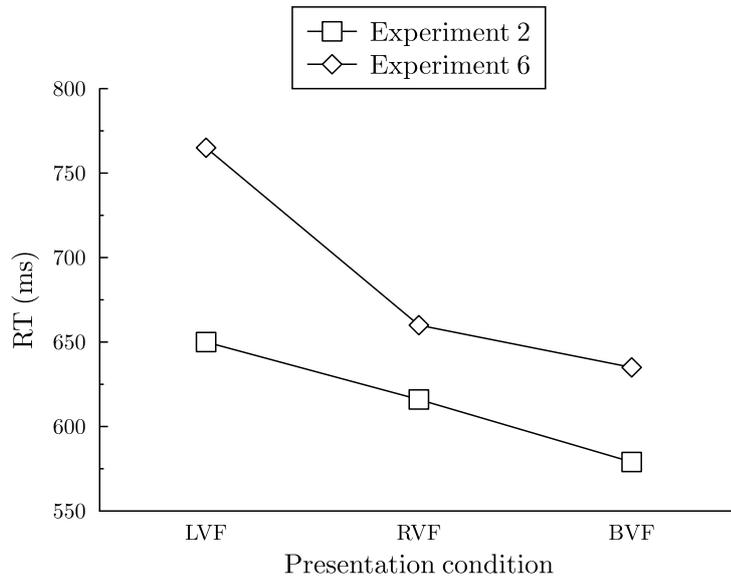


Fig. 5.4: Latency data for Body trials from Experiments 2 and 6.

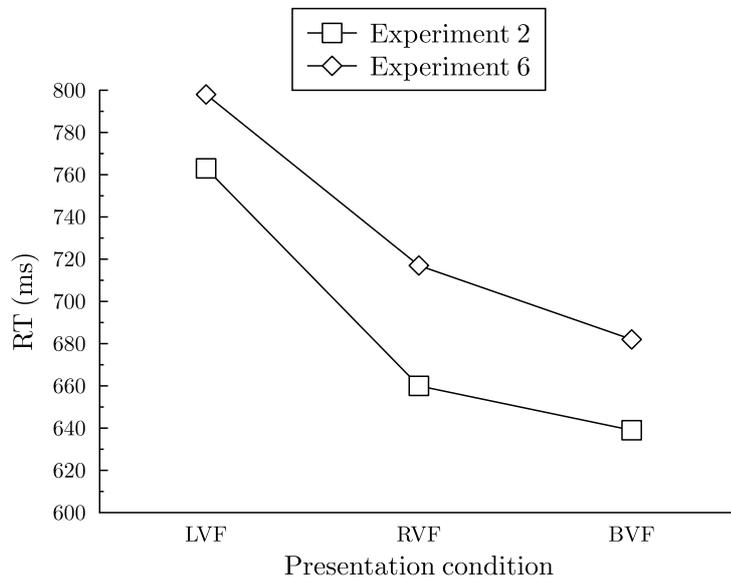


Fig. 5.5: Latency data for Building trials from Experiments 2 and 6.

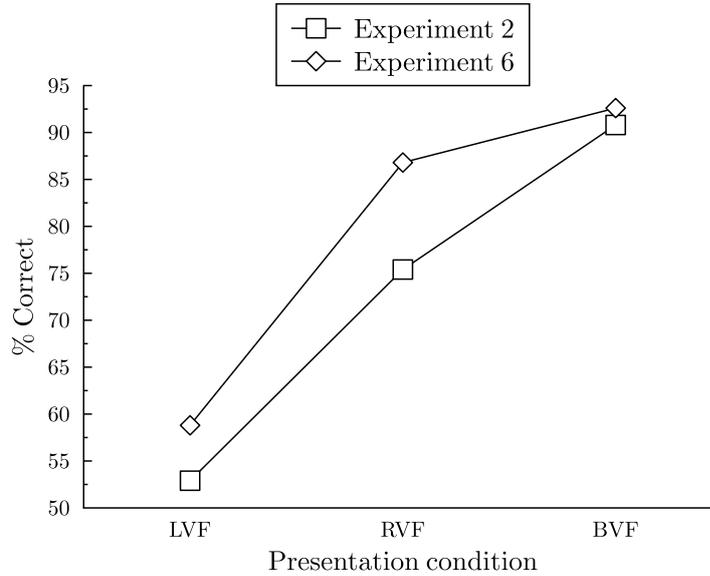


Fig. 5.6: Accuracy data for Body trials from Experiments 2 and 6.

Accuracy ANOVA analogous to those used for latency data were also carried out on percent correct values. For Body trials — displayed in Figure 5.6 — responses were more frequently correct in Experiment 6 (79.4%) than Experiment 2 (73.0%) [$F(1, 46) = 7.65$, $MSE = 193.2$, $p < .05$], but the Experiment \times Presentation interaction was not significant [$F(2, 92) = 1.13$, $MSE = 242.0$, $p > .3$]. For Building trials — displayed in Figure 5.7 — responses were more frequently correct in Experiment 6 (79.0%) than Experiment 2 (65.7%) [$F(1, 46) = 51.66$, $MSE = 122.7$, $p < .05$], and the interaction was also significant [$F(2, 92) = 9.97$, $MSE = 177.1$, $p < .05$]. This appears to have been a result of a larger between-experiments discrepancy in LVF trials (26.9%) than in RVF or BVF trials (9.2% and 3.7%, respectively).

Force Analogous ANOVA using peak force data showed no significant main effects or interactions for either Body or Building trials, though there was a marginally significant Experiment \times Presentation interaction for Building trials [$F(2, 92) = 2.85$, $MSE = 848.3$, $p = .083$, Greenhouse-Geisser corrected]. As displayed in Figure 5.8, this appears to have been a result of smaller between-experiment discrepancies for LVF trials (2 cN) than for RVF or BVF trials (28 and 25 cN, respectively).

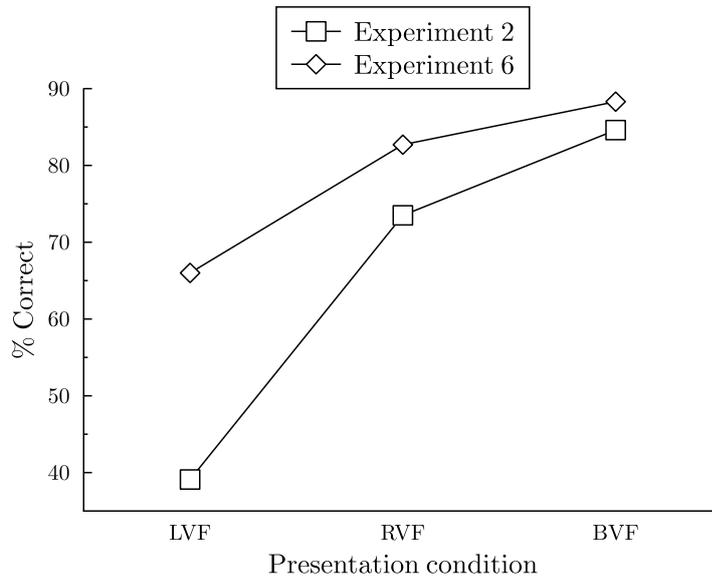


Fig. 5.7: Accuracy data for Building trials from Experiments 2 and 6.

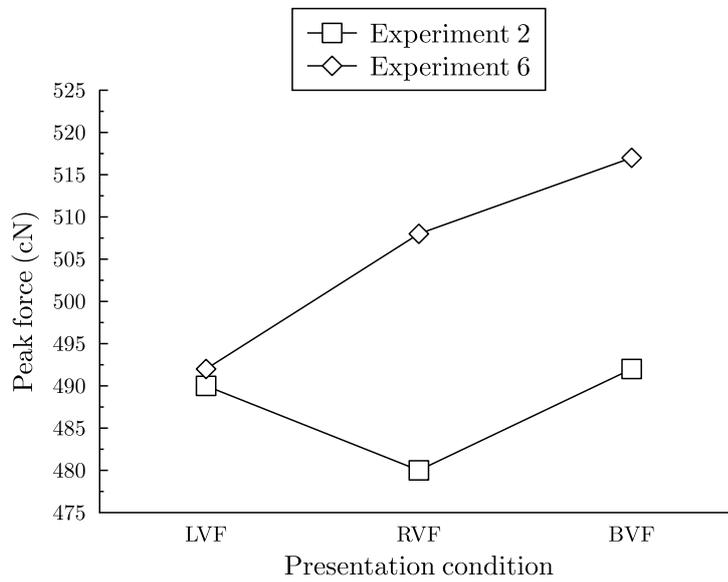


Fig. 5.8: Response force data for Building trials from Experiments 2 and 6.

Discussion

The aim of Experiment 6 was to determine whether the asymmetry found in Experiments 1–5 was a result of the non-target response category in those experiments being overly diffuse. This was accomplished by opposing two more narrowly defined categories to one another to see whether such an asymmetry would persist. The results show that it did not: VF effects were present for both “Body” and “Building” words, with responses to LVF trials being generally slower and less frequently correct than responses to RVF trials, which were in turn slower and less frequently correct than responses to redundant (BVF) trials. This seems to imply that such an asymmetry is specific to the “target” versus “non-target” opposition used in our previous experiments.

It is worth noting that two significant interactions — in RT data for Body words, and percent correct data for Building words — were found in the between-experiments comparisons. Both of these interactions seem to have been driven by performance in LVF trials differing less markedly from RVF and BVF trials in Experiment 6 when compared to Experiment 2. Though this could be considered as evidence for a slight reduction in VF differences, overall results of these analyses did not support the proposition of a “compromise” pattern of responding, as the magnitude of VF effects was generally maintained for both Body and Building responses. Given that results from Experiments 1, 2, and 4 suggested that information is less likely to be obtained from stimuli presented in the LVF compared to the RVF (or BVF) when presentation is bilateral and horizontal (as it was for this experiment), one explanation for the between-experiment interactions would be that participants in Experiment 2 had a tendency to respond “non-target” as a default when they did not perceive a target stimulus; as the response options were less asymmetrical in Experiment 6, no such “default” was available, which could have led participants to guess which was the correct response when unsure. In Experiment 2, responses were comparatively fast in LVF Body trials relative to LVF Building and Animal responses, and comparatively inaccurate in LVF Building trials relative to LVF Body and Animal trials in the same experiment. As such, guessing of this sort would be expected to lead to responses in Experiment 6 which were slower for Body trials (assuming the “guess” is only made once a participant is unsuccessful in identifying the appropriate response for that trial),

and more accurate for Building trials (since the percent correct value for Building trials in Experiment 2 was below 50%).

Also worth noting are the differences between patterns of responding to Body and Building trials in Experiment 6: “Body” responses were significantly faster than their “Building” counterparts, and accuracy data showed that “Body” responses were at least marginally more frequently correct for RVF and BVF trials, whereas “Building” responses were marginally more frequently correct for LVF trials. This point is reinforced by results of the SDT analysis, which are worth considering in light of the patterns seen in previous experiments. Where Experiments 1, 2, and 4 had shown lower sensitivity and “target-absent” biases for LVF trials, and higher sensitivity and “target-present” biases for BVF trials (with RVF trials occupying a middle ground), in Experiment 6 there was lower sensitivity and a bias toward “Building” responses in LVF trials, and higher sensitivity and a bias toward “Body” responses in BVF trials (with, again, RVF trials occupying a middle ground). In short, though the comparison is not perfect, to an extent it seems that “Building” responses here were being treated somewhat similarly to “target-absent” responses in the earlier experiments (with “Body” and “target-present” responses showing a similar parallel). One explanation for this could be that the “Building” response *was* perceived by participants as a “default” in this experiment — the response one makes if one is unsure which response is appropriate. It appears that the “Building” categorisation was more difficult than “Body” categorisation, both here and in Experiment 2. If this is the case, then it seems plausible that in cases of uncertainty and under time pressure, participants use their very lack of certainty as evidence for the appropriateness of the response associated with the more difficult categorisation. This might be because the more difficult categorisation tends to take longer, and thus participants more often find themselves in such a state of uncertainty when making it.

Such an argument could also provide an explanation for the “Body” bias in BVF trials — and by extension, the “target-present” bias found for BVF trials in earlier experiments. If the more difficult categorisations tend to take longer, then by definition easier categorisations are quicker. Using the same logic as above, there would be a bias toward the response associated with the “easier” categorisation when this categorisation is achieved more rapidly. Of course, the ease of the categorisation itself probably implies that the correct response is

made on most occasions, irrespective of this bias. Nonetheless, this line of reasoning provides one account of the differing biases from one Presentation condition to the next. In addition, Chiarello (1988) notes that in lexical decision tasks there tends to be a bias toward the “non-word” response when a stimulus is presented in the LVF, and a bias toward the “word” response when a stimulus is presented in the RVF. This finding could be explained in the same manner, if categorisation of a stimulus as a word is assumed to be easier than categorisation as a non-word.

One way of testing this idea would be to manipulate the difficulty of the categorisation associated with opposing responses. If the logic above is sound, there should be a stronger bias toward the more “difficult” response in LVF trials (where, for whatever reason, uncertainty seems to be greater), and a stronger bias toward the “easier” response in BVF trials. This issue might be worth addressing in future.

Conclusion

Experiment 6 provided evidence that the asymmetry between target-present and target-absent trials found in Experiments 1–5 was likely the result of the opposition of a specific category to a diffuse non-target space. When participants were asked to determine which of two more symmetrically defined categories was represented on each trial, the asymmetry disappeared. Further, this suggests that participants are not limited to defining the semantic categorisation task with reference to a single target category. This feature is relevant in light of the experiment which follows next, wherein two categories were defined as targets.

6. EXPERIMENT 7

The experiments described thus far have all provided evidence of redundancy gains in the semantic categorisation of lexical stimuli. The results of Experiment 3 suggest that such gains are unlikely to be wholly — if at all — reliant on interaction or cooperation between cerebral hemispheres because gains of a similar magnitude were found with non-lateralised stimulus presentation. The results of Experiment 5 suggest that Mohr and Pulvermüller's (e.g., Mohr et al., 1996) claims of hemispheric dominance in LDTs do not extend to semantic categorisation, as no visual field advantage was present when stimuli read vertically, rather than horizontally.

Nonetheless, a major unanswered question is how best to generally account for redundancy gain in the semantic categorisation task. Two plausible models, extending from those outlined in the introduction to this thesis, were offered in the discussion of Experiments 1–3: A race model, where stimuli are processed exhaustively unless a target is identified (as illustrated in the race version of Figure 1.2c); and a coactivation model based on Pulvermüller's cell assembly hypothesis (e.g. Pulvermüller, 1999), similar to that used by Mohr et al. (1996) to account for performance in lexical decision tasks with redundant stimuli (as illustrated in the coactivation version of Figure 1.2a). Experiment 7 was aimed at beginning to distinguish between these models, in terms of which best describes performance in the redundant semantic categorisation task.

One of the most common methods for determining which of race and coactivation models best accounts for the results of redundant targets experiments is by the use of Miller's (1982) race model inequality (RMI). According to race models, statistical facilitation can only enhance response times to a certain extent; the fastest responses to redundant trials should not be faster than the fastest responses to single-target trials. Instead, the mean redundancy gain is a result of these fast responses being more frequent, and slow responses less frequent. The race model

inequality uses response time data from both single target conditions to provide a boundary for the RT distribution that could be expected if statistical facilitation were operating at its potential peak; if redundant RTs are faster than that modelled distribution — something which tends to occur most frequently at the fast part of the distribution — the RMI is violated, and a race model can be ruled out.

In each of Experiments 1–6, the race model inequality was tested, and no significant violations found. In fact, only two experiments — Experiments 3 and 6 — showed even numeric (non-significant) violations. Whilst a race model requires the inequality to be satisfied, coactivation models are not invalidated by a lack of RMI violations. This means that the tests of the RMI in Experiments 1–6 provide no conclusive evidence for or against either of the models, and another method of distinguishing between the explanations is required.

Fortunately, the principles of cell assembly theory itself, as interpreted by Mohr et al. (1996), might offer such a method. Key to those authors' explanation of redundancy gains in lexical decision was the idea that presenting multiple copies of the same word resulted in stronger stimulation of the cell assembly by which that word was represented neurally. This was reliant on the idea that these assemblies form interconnected wholes, within which activation reverberates until they become “ignited”. Extrapolating these principles to semantic categories, rather than words, has played a substantial part in the cell assembly-based coactivation account of redundancy gain in semantic categorisation found in the experiments described thus far. One way to test whether this explanation holds, then, could be to compare redundancy gains in a task where redundant trials involve the presentation of two items from a single category, to those found when multiple categories are used as targets. The use of multiple target categories should not allow for the same level of neural summation as would a single target category, since there is unlikely to be a cell assembly for the union of separate categories (e.g., “tree” and “aeroplane”). As such, if redundancy gains in this task rely on summed activation within a cell assembly, they should decrease (or disappear entirely) when stimuli from differing categories — both defined as targets — are presented on redundant trials.

Experiment 7 was aimed at determining whether this was the case. Using categories which had previously been used in Experiments 2 and 6 — “parts of the human body” and “parts of a building” — here they were united and participants were instructed that both

categories should elicit the same “target” response (opposed to non-targets, which — as in all the previously described experiments bar Experiment 6 — were defined as all items from outside of those categories). To assess what effect the use of multiple target categories had on redundancy gain, results in this experiment — where redundant trials always involved one item from each of the two different categories — were compared to a “baseline” taken from results using same categories as sole-targets in Experiment 2. According to the cell assembly-based coactivation account, redundancy gains should be smaller in Experiment 7 than they were in Experiment 2, if not absent entirely.

One complicating factor is that the use of multiple target categories would seem likely to increase the difficulty of the task, which might encourage participants to use different strategies. For instance, as the definition of “target” is broader and more disparate, participants might consider target-present and target-absent trials in a more symmetrical manner, as occurred in Experiment 6. If this were the case, extrapolating the accounts for various aspects of performance on the simpler tasks used in Experiments 1– 5 to this task could be problematic. As such, also of interest were any broader changes in the patterns of results from Experiment 2 to this experiment.

Method

Participants

24 University of Otago first- and second-year psychology undergraduates (22 female; 20 right-handed) participated in Experiment 7 in return for partial course credit. All were recruited through the psychology department’s experimental participation programme. None had participated in any of Experiments 1– 6.

Apparatus and Stimuli

Stimuli were 40 words from the category “parts of the human body”, 40 words from the category “parts of a building”, 80 matched nouns from outside these categories, and 80 matched non-words. These were a random selection of the target stimuli used for these categories in Experiment 2, and their matched non-targets and non-words. The inclusion of only 40 words

from each of the target categories was designed so that the small amount of repetition occurring in previous experiments was not altered here. Redundant trials always included one word from each of the two target categories (or non-targets matched to words from each category, in target-absent trials). Other than these changes, all apparatus was as described in Experiment 1.

Procedure

Procedure was as essentially as described in Experiment 2, with the obvious exception that participants were informed that words from either of the categories “parts of the human body” and “parts of a building” should be met with the “target-present” response.

Data Analysis

Data analyses were somewhat different from those conducted in the previous experiments. First, in preliminary (within-subjects) analyses of the data from Experiment 7, an additional factor in the design, Category, was included so that single-target trials where the target stimuli were Body words could be distinguished from those where target stimuli were Building words. As a result of this, an ANOVA involving the factors Category, Target, and two levels of the factor Presentation (both those relevant to single-target trials: LVF and RVF) was included. This was accompanied by a separate ANOVA for redundant trials. This redundant-trial ANOVA was a 2 (Target: present or absent) \times 2 (LVF stimulus: Body or Building) design. Given the “flat” patterns of results for responses to target-absent trials in previous experiments, to simplify all analyses target-absent data were initially analysed separately, and if no effects were found further analyses were conducted on target-present data exclusively. Second, in the key (between-subjects) analyses data were summed across both levels of the Category factor and comparisons were made between the results of Experiment 7 and those of Experiment 2 which pertained to the categories included here. The latter were also summed across categories.

Results

The following within-experiment analyses of Experiment 7 were conducted as a preliminary step prior to the key, between-experiment comparisons to follow.

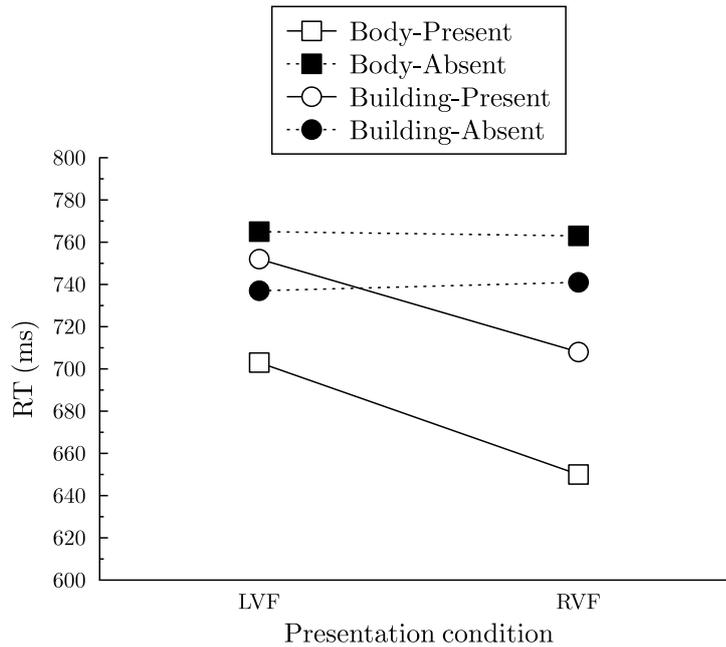


Fig. 6.1: Single-target trial latency data for correct responses in Experiment 7.

Latency data, Experiment 7

Single-target trials RTs from correct responses (approximately 66% of total responses) are displayed in Figure 6.1. First, a 2 (Category) \times 2 (Presentation) ANOVA on the data from target-absent trials showed that responses were faster where non-target stimuli matched to Building words (739 ms) than when they matched to Body words (764 ms) [$F(1, 23) = 24.63$, $MSE = 597.9$, $p < .05$]. Though there was no effect of Presentation, nor any significant interaction (both $p > .5$), the significant effect of Category meant that the different types of non-target words were not equivalent and so the distinction between the two types of non-target word categories was preserved by including data from target-absent trials in the remaining analyses.

The resulting 2 \times 2 \times 2 ANOVA showed that responses to RVF trials (715 ms) were faster than responses to LVF trials (739 ms) [$F(1, 23) = 6.74$, $MSE = 4050.1$, $p < .05$], that “target-present” responses (703 ms) were faster than “target-absent” responses (752 ms)

[$F(1, 23) = 13.23$, $MSE = 8470.4$, $p < .05$], and that responses to Body trials (720 ms) were faster than responses to Building trials (734 ms) [$F(1, 23) = 5.21$, $MSE = 1874.7$, $p < .05$]. The Category \times Presentation interaction was not significant [$F(1, 23) = 0.33$, $MSE = 2257.7$, $p > .55$], but a larger discrepancy between LVF and RVF RTs for target-present (LVF: 728 ms; RVF: 679 ms) than target-absent (LVF: 751 ms; RVF: 752 ms) responses led to a significant Target \times Presentation interaction [$F(1, 23) = 6.27$, $MSE = 4733.4$, $p < .05$]. A larger discrepancy between RTs for “present” and “absent” responses for Body (present: 677 ms; absent: 764 ms) than Building (present: 730 ms; absent: 739 ms) trials led to a significant Category \times Target interaction [$F(1, 23) = 28.72$, $MSE = 2546.4$, $p < .05$]. The 3-way interaction term was not significant [$F(1, 23) = 0.00$, $MSE = 2275.9$, $p > .95$].

Error RTs were assessed with an analogous ANOVA, after the removal of data from one participant who made no errors in one condition. This showed that error RTs for target-present trials (that is, misses; 762 ms) were significantly faster than those for target-absent trials (that is, false alarms; 819 ms) [$F(1, 22) = 8.63$, $MSE = 171118.2$, $p < .05$], and that error RTs were faster on Body (778 ms) than Building (804 ms) trials [$F(1, 22) = 6.09$, $MSE = 4950.5$, $p < .05$]. No other main effects or interactions were significant.

Redundant trials RTs from correct responses are displayed in Figure 6.2. A one-way ANOVA on target-absent data showed no significant difference between trials where non-target stimuli matched to Body or Building words were presented in the LVF [$F(1, 23) = 0.09$, $MSE = 879.9$, $p > .75$], so the remaining analysis was conducted on target-present data exclusively. This showed that responses were faster when a Building word was presented in the LVF (632 ms) than when a Body word was presented in the LVF (674 ms) [$F(1, 23) = 12.07$, $MSE = 1694.7$, $p < .05$].

Analysis of RT data from error responses showed no significant main effects or interactions (all $p > .4$).

Accuracy data, Experiment 7

Single-target trials Percent correct data are shown in Figure 6.3. Analysis of target-absent data showed that fewer correct responses were made in trials involving non-targets matched to

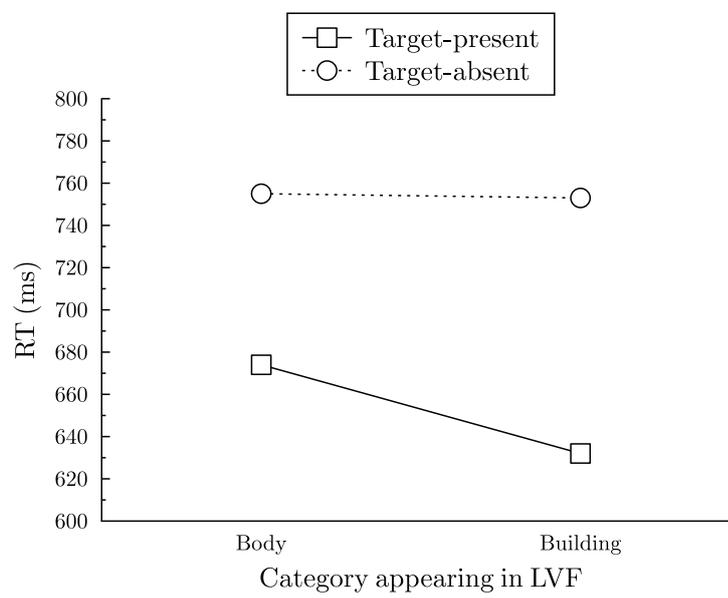


Fig. 6.2: Redundant trial latency data for correct responses in Experiment 7. Target-absent trials differ by category based on which set of targets they were matched with for word length and frequency.

Body words (73.4%) than those matched to Building words (79.2%) [$F(1, 23) = 29.90$, $MSE = 26.4$, $p < .05$]. As such, though none of the other effects were significant (both $p > .4$), data from target-absent trials were preserved in further analyses. The resultant $2 \times 2 \times 2$ ANOVA showed that responses were more frequently correct for RVF (71.3%) than LVF (60.8%) trials [$F(1, 23) = 9.31$, $MSE = 567.9$, $p < .05$], more frequently correct for target-absent (76.3%) than target-present (55.7%) trials [$F(1, 23) = 37.67$, $MSE = 540.7$, $p < .05$], and more frequently correct for Body (69.5%) than Building (62.5%) trials [$F(1, 23) = 23.15$, $MSE = 101.8$, $p < .05$]. A larger difference between LVF and RVF values for target-present (44.5% and 66.9% for LVF and RVF, respectively) than target-absent (77.0% and 75.6%) trials led to a significant Target \times Presentation interaction [$F(1, 23) = 7.50$, $MSE = 898.7$, $p < .05$]. A larger difference between target-present and target-absent values for Building (45.8% and 79.2%, respectively) than Body (65.6% and 73.4%, respectively) trials resulted in a significant Category \times Target interaction [$F(1, 23) = 145.81$, $MSE = 53.4$, $p < .05$]. The difference between LVF and RVF trials was larger for Body (62.4% and 76.6%, respectively) than Building (59.1% and 65.9%) trials, leading to a significant Category \times Presentation interaction [$F(1, 23) = 14.67$, $MSE = 42.9$, $p < .05$]. The 3-way interaction was also significant [$F(1, 23) = 7.78$, $MSE = 95.4$, $p < .05$], apparently because the difference between Body and Building percent correct values increased from the LVF to the RVF condition for target-present (Body 12.2% more accurate than Building trials for LVF, and 27.2% more accurate for RVF), but not target-absent trials (Body 5.6% less accurate for LVF, and 6.0% less accurate for RVF).

An SDT analysis was also conducted on the single-target trial data, which are displayed in Table 6.1.¹ 2 (Category) \times 2 (Presentation) ANOVA were carried out using d' and c . For d' , this showed that sensitivity was higher for RVF (1.30) than LVF (0.66) trials [$F(1, 23) = 10.88$, $MSE = 0.89$, $p < .05$], higher for Body (1.21) than Building (0.75) trials [$F(2, 23) = 23.22$, $MSE = 0.21$, $p < .05$], and that the discrepancy between LVF and RVF was greater for Body than for Building trials, leading to a significant interaction [$F(1, 23) = 17.44$, $MSE = 0.11$, $p < .05$]. For c , there was a stronger “target-absent” bias for LVF (0.48) than RVF (0.12) trials

¹ Due to one participant’s perfect accuracy in the Body-RVF trials, the adaptive estimator method (Hautus & Lee, 2006) was used to estimate sensitivity and bias in that condition.

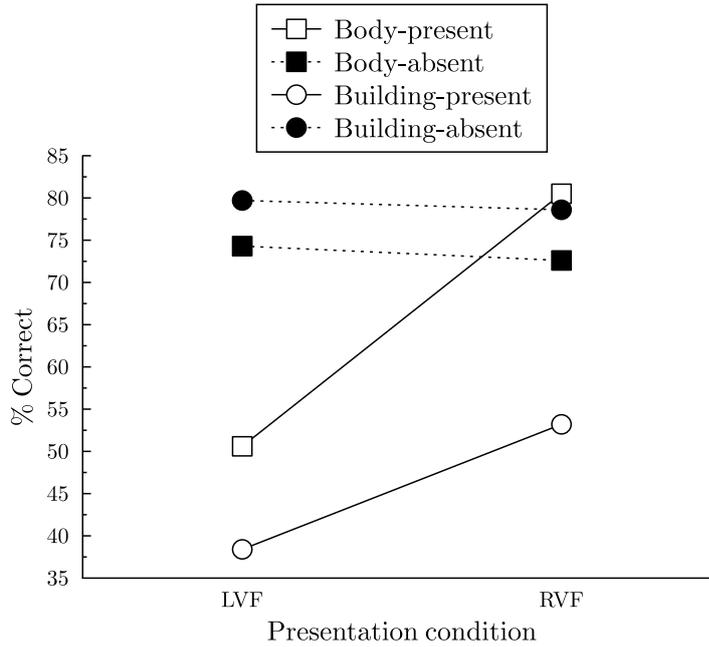


Fig. 6.3: Single-target trial response accuracy data from Experiment 7.

Tab. 6.1: Sensitivity and bias for single-target trials in Experiment 7.

		LVF	RVF
Body	d'	0.75	1.66
	c	0.35	-0.16
Building	d'	0.57	0.93
	c	0.62	0.40

$[F(1, 23) = 6.29, MSE = 0.51, p < .05]$, and a stronger “target-absent” bias for Building (0.51) than Body (0.09) trials $[F(1, 23) = 173.56, MSE = 0.02, p < .05]$. The difference in bias between LVF and RVF trials was greater for Body than Building trials, which led to a significant interaction $[F(1, 23) = 6.18, MSE = 0.08, p < .05]$.

Redundant trials Percent correct data are shown in Figure 6.4. Analysis of target-absent data showed no significant difference between trials with a Body-matched or Building-matched word presented to the LVF ($p > .2$), so the remaining analysis was conducted on target-present data only. This showed that responses were more frequently correct when a Building word was

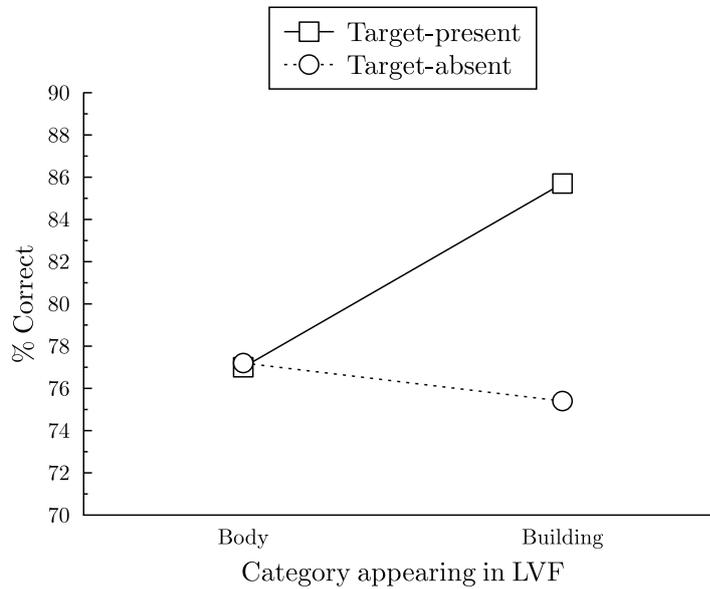


Fig. 6.4: Redundant trial response accuracy data from Experiment 7.

Tab. 6.2: Sensitivity and bias for redundant trials in Experiment 7.

	LVF-Body	LVF-Build
d'	1.55	1.81
c	-0.06	-0.25

presented in the LVF (85.7%) than when the LVF stimulus was a Body word (77.0%) [$F(1, 23) = 10.91, MSE = 84.2, p < .05$].

SDT values are displayed in Table 6.2. Comparisons between trials with Body words in the LVF and those with Building words in the LVF showed no significant difference in sensitivity [$t(23) = 1.76, p = .09$], but that there was a stronger “target-present” bias when Building words appeared in the LVF than when Body words did so [$t(23) = 2.64, p < .05$].

Force data, Experiment 7

Analyses of peak force data from single-target and redundant trials showed neither significant main effects nor interactions (all $p > .1$).

Tests of the race model inequality, Experiment 7

Tests of the race model inequality were carried out for both types of redundant trial (that is, Body-Building and Building-Body trials), after the removal of data from participants whose correct responses were too infrequent to allow the test to be conducted over a suitable number of quantiles (5 participants each for both types of trial). Neither showed any violations over the eight quantiles tested (0.0625, 0.1875, . . . , 0.9375).

Between-experiment comparisons

The key analyses, comparing results in the single-category and multiple-category experiments (Experiments 2 and 7, respectively), are reported below.

Latency RTs from Experiment 7 were collapsed across both categories to create a 2 (Target) \times 3 (Presentation) data set. This was then combined with Body and Building RT data from Experiment 2, also collapsed across categories, and a 2 (Experiment) \times 2 (Target) \times 3 (Presentation) ANOVA conducted. This showed that responses were marginally slower in Experiment 7 (717 ms) than in Experiment 2 (680 ms) [$F(1, 70) = 3.39$, $MSE = 39059.8$, $p = .07$], but tellingly none of the interaction terms involving the Experiment factor approached significance (all $p > .3$).

To assess changes in redundancy gain across experiments, difference scores from target-present trials (Redundant RT – best single-target RT) were calculated for each participant in each experiment, and entered into a t -test. Tellingly, this showed no significant differences between the two experiments [$t(70) = 0.52$, $p > .6$], implying that redundancy gain did not differ between Experiments 2 and 7.

Accuracy A comparable 2 \times 2 \times 3 ANOVA was performed on percent correct data. This showed that responses were more frequently correct in Experiment 2 (75.2%) than in Experiment 7 (70.3%) [$F(1, 70) = 6.95$, $MSE = 335.4$, $p < .05$], but again, none of the interactions involving the Experiment factor approached significance (all $p > .4$).

To assess changes in redundancy gain, difference scores were calculated for percent correct values as described for RT, and entered into a t -test. As with the latency analysis, this

showed no significant differences between the two experiments [$t(70) = 0.60, p > .5$].

2 (Experiment) \times 3 (Presentation) ANOVA were also performed using the SDT variables d' and c . The d' analysis showed that sensitivity was higher in Experiment 2 than in Experiment 7 [$F(1, 70) = 8.03, MSE = 0.9, p < .05$], and there was also a marginally significant interaction [$F(2, 106) = 3.17, MSE = 0.5, p = .06$, Greenhouse-Geisser corrected]. This was a result of the discrepancy between experiments being smaller in LVF trials (Experiment 2: 0.75; Experiment 7: 0.67) than in RVF (1.75 vs. 1.21) or BVF (2.24 vs. 1.71) trials. In the analogous c analysis neither the main effect of Experiment, nor the Experiment \times Presentation interaction approached significance (both $p > .5$).

Force An analysis analogous to those undertaken with RT and percent correct data was conducted using peak force values. This showed no significant main effects or interactions (all $p > .1$).

Discussion

The primary aim of Experiment 7 was to see whether the use of targets from different semantic categories in redundant trials would lead to a decrease in redundancy gain. If the cell assembly-based coactivation account of redundancy gain in semantic categorisation is correct, neural activity should be able to sum more effectively when targets come from the same category. Thus, coactivation (and consequently redundancy gain) should be decreased when redundant targets are drawn from disparate categories, compared to when they are from the same category.

Results from key analyses show that redundancy gains in Experiment 7 were not significantly different to those found using the same categories separately in Experiment 2. This was the case for both RT and percent correct data. The obvious conclusion is that the redundancy gains occurring in the semantic categorisation task are unlikely to be as a result of neural summation in semantic cell assemblies.

A secondary aim of Experiment 7 was to determine whether the use of multiple target categories modified participants' performance in the task in any substantive way. If this were the case, it might undermine any conclusions drawn from a comparison of this experiment and

those using single target categories. To that end, a few similarities and differences are worth noting.

First, one similarity is that the familiar qualitative patterns from earlier experiments, showing VF effects for target-present but not for target-absent trials, remained. Responses to LVF target-present trials tended to be slower and less accurate than responses to RVF target-present trials. In single-target trials, “target-absent” responses were generally slower and more accurate than “target-present” responses. These findings all have parallels in the results from the previous experiments which used single target categories. This is evidence against the idea that using multiple target categories causes fundamental differences to the semantic categorisation task.

Second, another similarity is that some of the differences between the Body and Building categories in Experiment 2 were retained in participants’ responses to the two types of words in Experiment 7. Most noticeably, “target-present” responses in LVF trials where Body words appeared were faster than the responses in the comparable target-absent trials, whereas the reverse was true for trials with Building words. There are two implications of this finding. First, it reduces the likelihood that the slow LVF target-present RTs for Building words in Experiment 2 were simply a Type I error. This is important as the occurrence of slower “target-present” than “target-absent” responses provides evidence against a serial model of task performance (as mentioned in the Discussion for Experiments 1–3). Second, it suggests that performance for trials involving each of the individual target categories maintains characteristics of performance using those same categories in tasks without multiple target categories. This is reinforced by the generally superior performance to Body than Building trials here, which parallels the superior performance for the same category in Experiment 2 (and also in the rather procedurally different Experiment 6, for that matter). This provides further evidence that the use of multiple target categories does not fundamentally change the semantic categorisation task.

A third point worth noting is that “target-present” responses to BVF trials were slower when the stimulus display involved a Body word in the LVF and a Building word in the RVF, than when they involved a Building word in the LVF and a Body word in the RVF. Given that responses to Body words were generally faster than responses to Building words, this suggests

that more information in BVF trials is obtained from the RVF stimulus than is obtained from the LVF stimulus. If true, this would align with the idea presented in relation to Experiment 5 that, with the horizontal presentation of word stimuli, participants obtain more information from the right of fixation (be it due to the proximity of a word's initial letters, or to a tendency to initially attend to the right).

Finally, one slight difference is that there is some evidence that having two target categories rather than one made the task more difficult, as was expected. Between-experiment comparisons for latency and accuracy showed that responses were slower and more errors made in Experiment 7 than for Body words in Experiment 2. The differences between Experiment 7 and Building words in Experiment 2 were not significant, but in each case performance in Experiment 7 was numerically worse (slower RTs, fewer correct responses). This implies that performance on the multiple category task is not simply an average of performance in separate tasks with the same categories as sole targets.

Before concluding, it might be informative to consider the implications of these results not just for the cell assembly model illustrated in the coactivation version of Figure 1.2a, but also for the other models shown in the same diagram. The finding that redundancy gain was not altered with a change from a single to multiple target categories also seems to be evidence against the coactivation model illustrated in 1.2b. This is because coactivation in categorisation itself would be expected to show effects of some sort when this categorisation process takes place with reference to two categories rather than one. By contrast, both race versions of each model, and the coactivation version of Figure 1.2c, seem compatible with the idea that redundancy gain was not decreased with the use of two target categories. For the coactivation model in 1.2c, it is only at the final, response-selection stage where redundancy is important (that is, where evidence is summed), and thus altering the number of target categories should not have an effect. For the race models too, there seems no reason to assume that changing from one to two target categories should alter the extent of redundancy gain, unless this somehow changes the “finishing line” of the race (e.g., by extending it from activation of semantic representation, to selection of a response).

In sum, the overall evidence suggests that performance in this experiment is comparable to performance in experiments using single target categories. As such, the lack of a significant

difference in redundancy gain between Experiments 7 and 2 remains problematic for the cell assembly-based coactivation account. This issue is investigated further next.

Conclusion

The results of Experiment 7 suggest that the redundancy gain found in the semantic categorisation task is not due to semantic coactivation (e.g., that illustrated in the coactivation version of Figure 1.2a). This is because the extent of redundancy gain did not appear to change from that found using the Body and Building categories in separate, single-category categorisation tasks in Experiment 2. This would also seem to be evidence against the coactivation model shown in Figure 1.2b, in which a change from one to two target categories might also be expected to have an effect of redundancy gains. Nonetheless, to ensure that this finding was not an artefact of some unknown task-based difference, the following experiment involved a comparison between single-category and multiple-category redundant trials being undertaken as part of a single task.

7. EXPERIMENT 8

By showing that redundancy gain was still of a similar magnitude when redundant targets were taken from different categories, the comparison between Experiments 2 and 7 provided evidence against the cell assembly-based account of these gains. However, as this involved comparing tasks using single target categories to one using dual target categories, it is possible this change in the requirements of the task was somehow responsible for this finding. Concordant evidence, obtained through another method, would allow for greater certainty regarding whether the cell assembly model can be ruled out as an account of redundancy gain in semantic categorisation.

The work of Miller (1991) may have provided just such a method, and one which is adaptable to tasks involving semantic processing. In a bimodal go/no-go task, Miller (1991) presented his participants with visual and auditory stimuli, which varied in presentation location and pitch, respectively. On some occasions, the two stimuli were congruent — for instance, when the auditory stimulus was high-pitched, and the visual stimulus was located at a high position on the screen. On other occasions, the relationship between the stimuli was incongruent (e.g., a high-pitched auditory stimulus with a “low” visual stimulus). Miller’s aim was to see whether the congruency of redundant targets affected response speed.¹

A similar logic might assist in determining which of the two models forwarded to explain redundancy gain in the semantic categorisation task best does so. By definition race models involve separate processing of redundant targets, and should thus be free from congruency effects.² In contrast, using the cell assembly-based coactivation model forwarded to explain the

¹ Miller’s interest was in distinguishing between what he called independent and interactive coactivation models. According to the former, whether or not targets are congruent should not affect redundancy gains, as the contribution of each stimulus to summed activation is independent of the relationship between them. By contrast, the latter type of model allows for congruency effects; and these effects were borne out in Miller’s results.

² An exception to this separate processing feature is found in Mordkoff & Yantis’ (1991) interactive race model, which allows contingencies that exist between redundant targets to affect processing speed. As it is unclear how this interactive race model could be distinguished from coactivation models in a situation where semantic relationships between stimuli are themselves a variable of interest, I here deal only with traditional race models.

results of the previously described experiments it could be suggested that the stronger the semantic relationship between stimuli, the greater the overlap between their respective neural representations, and the greater the likely coactivation.³ This implies that Miller’s (1991) manipulation of congruency might be applied to our purposes in distinguishing between race and coactivation models, by allowing a test of the predictions emanating from the latter.

By way of the same basic task that has been used for Experiments 1–7, there are two variables which could be manipulated within a task to investigate the effect of semantic relatedness on redundancy gains. (1) First, the relationship between the stimuli themselves in redundant trials could be altered. This could take two forms. In one incarnation (1a), it could involve seeing whether there is a difference between trials where the stimuli are closely semantically related — if the target category is “animals”, say, “dog” and “wolf” — and those where the relation is more distant — say, “dog” and “bird”. In the other incarnation (1b), targets could be drawn from more than one category and trials where both stimuli come from the same category (close semantic relationship) compared to those where they differ (more distant semantic relationship). This is analogous to how the differences between Experiments 7 and 2 were used to assess the cell assembly account. (2) Second, the relationship between categories could be manipulated in experiments involving more than one target category. For instance, the redundancy gain when targets are “animals” or “plants” could be compared to the gain when “animals” or “furniture” are targets.

Experiment 8 was designed to investigate the effects of semantic relatedness by using the first of these variables: the relationship between redundant stimuli. Given the difficulties inherent in manipulating semantic relatedness of items within a single target category (e.g., confounds involving frequency, imageability, etc.), a simple method of manipulating semantic relatedness of redundant targets is by using multiple target categories (that is, using method 1b). Consequently, Experiment 8 involved a direct test of the effect of semantic relatedness on redundancy gain, by the use of two disparate categories (“animals” and “musical instruments”) as targets. With a task mostly identical to that used in Experiments 1–7, on some redundant trials both targets were from the same category (“Redundant-Same” trials), whereas on other

³ Of course, this relies on a model of semantic representation in which shared semantic space between stimuli (e.g., shared features) implies stronger neural connectedness of their representations. Should this not be the case, the predictions of the model might be different.

trials one came from each (“Redundant-Different” trials). The effect this manipulation had on redundancy gain could then be used as evidence for or against the cell assembly-based coactivation account.

The predictions the cell-assembly based coactivation model makes in this situation seem reasonably straightforward. Again, under the assumptions that (a) semantically similar items share more overlapping neural space than semantically distant items (as is implied in a cell assembly-based implementation of Barsalou’s perceptual symbol systems theory; Barsalou, 1999), and (b) that cell assemblies exist for single categories, but not for arbitrary unions of disparate categories, it is likely that summation of activity will only occur, or will occur to a greater extent, when redundant targets from the same category are presented (compared to targets from different categories). This would be expected to lead to a more pronounced redundancy gain on Redundant-Same than on Redundant-Different trials.

Though this might seem to make the outcome of Experiment 8 relatively clear — whatever that outcome turns out to be — a complicating matter is exactly how best to measure the “strength” of any redundancy gains. There are three immediately apparent ways of doing this. First, the raw mean RTs for the two types of redundant target trials could be compared (e.g., Redundant-Same versus Redundant-Different). Second, using a method which owes its origins to the work of Biederman and Checkosky (1970), redundancy gains could be defined as the difference between redundant target RTs and the best single-target RTs for each participant, and these difference scores compared for Redundant-Same versus Redundant-Different trials. And third, divergences from the CDFs predicted by race models constructed from single-target trials could be compared for Redundant-Same and Redundant-Different trials. As each of these methods has its advantages and disadvantages (e.g., the potential for differences between trial types resulting from non-semantic factors, insufficient data for proper application), all three were used here in the hope that a clear picture would emerge.

Method

Participants

Twenty-four first- and second-year psychology students, recruited through the Psychology department's participation programme, took part in this experiment (seven male, one left-handed), in return for extra credit toward their psychology grades. None had participated in any of Experiments 1–7.

Apparatus and Stimuli

Target stimuli were 40 words from the category “a four-footed animal” (Animal), and 40 words from the category “a musical instrument” (Instrument). All words were between three and eight characters in length. Words were taken from the most frequent responses to each category from a norming study (described in Appendix A), and were matched between categories for length and frequency, as measured through the WCWNZE (Bauer, 1994). Non-targets were 80 nouns, matched with the targets for length and frequency. Distractors were 80 orthographically regular non-words, matched with target stimuli for length, taken from the ARC non-word database (Rastle et al., 2002). All other aspects of stimuli and apparatus were as described in Experiment 3, including the presentation of stimuli above (“HI”) and below (“LO”) fixation (rather than to the left and right visual fields). The decision to present stimuli in this manner was made in the hope that it would lead to a greater proportion of correct responses (as was evident in the higher accuracy levels for Experiment 3 compared to Experiment 2), and thus provide more RT data for analysis.

Procedure

Procedure was as described in Experiment 3, except that participants were informed that a “target” response was appropriate to any Animal or Instrument word, and a “non-target” response to any word from outside of these categories.

Data Analysis

As in Experiment 7, ANOVA for single-target and redundant target trials were conducted separately, as these conditions differed in their factorial structure. Most other aspects of the analyses were also similar to those of Experiment 7, including initial analyses being run on target-absent data separately. For redundant trials, these took the form of one-way ANOVA followed by post-hoc comparisons (Tukey's tests) where significant effects were apparent. An obvious departure from Experiment 7 was that comparisons with Experiment 2 were *not* included, and that the various tests of redundancy gain "strength" described in the introduction to this experiment *were*.

Results

Latency data

Single-target trials RTs for correct responses (approximately 80% of responses to single-target trials, after the removal of data from one participant who made no correct responses in one condition) are displayed in Figure 7.1. Initial analysis of data from target-absent trials showed no significant effects (all $p > .2$), so remaining analyses were conducted on data from target-present trials only. The resultant 2 (Category) \times 2 (Presentation) ANOVA showed that responses to HI trials (653 ms) were faster than those to LO trials (744 ms) [$F(1, 22) = 29.23$, $MSE = 6515.9$, $p < .05$], and responses to Animal trials (679 ms) were faster than those to Instrument trials (717 ms) [$F(1, 22) = 23.05$, $MSE = 1411.6$, $p < .05$]. The interaction between the two factors did not approach significance ($p > .5$).

A 2 (Category) \times 2 (Target) \times 2 (Presentation) analysis using RTs from error trials (after the removal of data from four participants who made no errors in at least one condition) showed that miss and false alarm RTs did not differ for Animal trials (786 and 789 ms, respectively), but they did for Instrument trials (756 and 815 ms, respectively), leading to a significant Category \times Target interaction [$F(1, 19) = 5.20$, $MSE = 6115.6$, $p < .05$]. No other main effects or interactions approached significance (all $p > .15$).

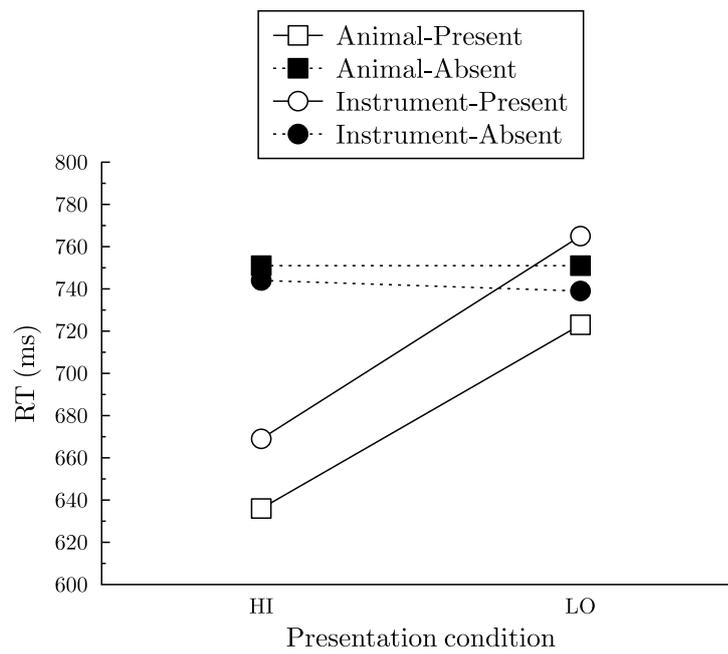


Fig. 7.1: Single-target trial latency data for correct responses in Experiment 8.

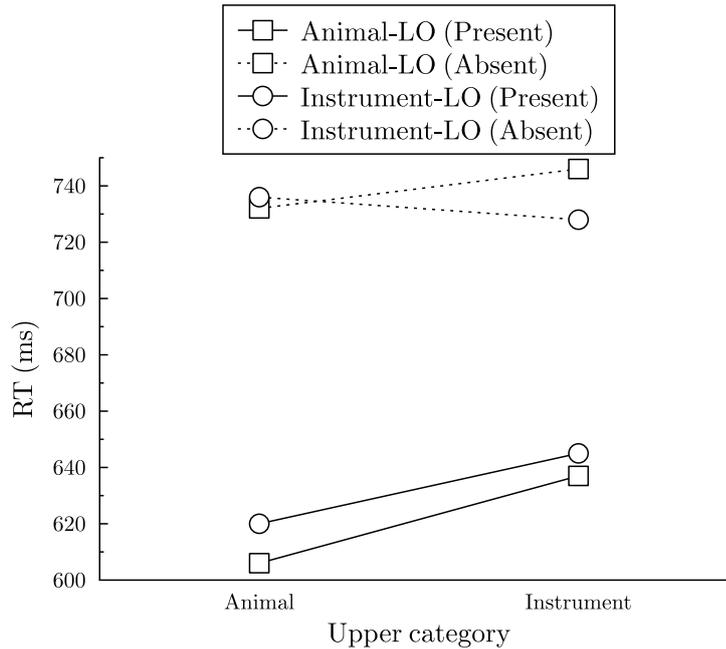


Fig. 7.2: Redundant trial latency data for correct responses in Experiment 8. Category (or matched non-target) appearing above fixation is on the x-axis; separate lines are used for the two categories of stimuli appearing below fixation. Target-absent trials differ by category based on which set of targets they were matched with for word length and frequency.

Redundant trials RTs for correct responses (approximately 90% of responses to redundant trials) are displayed in Figure 7.2. Initial analysis of target-absent trials was performed by way of a one-way ANOVA. This showed no significant differences between any of the four target-absent conditions [$F(3, 69) = 1.52$, $MSE = 995.1$, $p > .2$], so further analysis involved target-present data only. This took the form of a 2 (Upper category) \times 2 (Lower category) ANOVA, which showed that responses were significantly faster when an Animal word was presented above fixation (613 ms) than when an Instrument word was presented above fixation (641 ms) [$F(1, 23) = 14.48$, $MSE = 1313.4$, $p < .05$]. Importantly however, there was no significant interaction [$F(1, 23) = 0.15$, $MSE = 965.7$, $p = .7$], implying that RTs to Redundant-Same trials did not differ significantly from those to Redundant-Different trials.

As only four participants made at least one incorrect response in every condition, no analysis of error data from redundant trials was undertaken.

Accuracy data

Single-target trials Percent-correct values for single-target trials are displayed in Figure 7.3. Initial analysis of target-absent data showed that responses were significantly more frequently correct for Instrument (90.0%) than Animal (87.6%) trials [$F(1, 23) = 4.55$, $MSE = 30.29$, $p < .05$], so remaining analyses included both target-absent and target-present data. These showed that responses were more frequently correct in HI trials (82.4%) than LO trials (75.0%) [$F(1, 23) = 9.60$, $MSE = 275.6$, $p < .05$]; more frequently correct in target-absent (88.8%) than target-present (68.6%) trials [$F(1, 23) = 43.54$, $MSE = 449.0$, $p < .05$]; and more frequently correct in Animal (81.5%) than Instrument (75.9%) trials [$F(1, 23) = 35.96$, $MSE = 42.6$, $p < .05$]. The difference between HI and LO trials was more pronounced for target-present (76.9% and 60.4% for HI and LO, respectively) than target-absent trials (88.0% and 89.6%, respectively) [$F(1, 23) = 12.14$, $MSE = 326.7$, $p < .05$], and the difference between Animal and Instrument trials was more pronounced for target-present (75.5% and 61.8% for Animal and Instrument, respectively) than for target-absent trials (87.6% and 90.0%, respectively) [$F(1, 23) = 28.64$, $MSE = 108.5$, $p < .05$]. Neither the Category \times Presentation, nor the Category \times Target \times Presentation interaction approached significance (both $p > .4$).

SDT values for single-target trials appear in Table 7.1. As there was at least one participant in each condition who made no errors, Hautus and Lee's (2006) adaptive estimator method was used to calculate d' and c for all conditions. A 2 (Category) \times 2 (Presentation) ANOVA using d' showed that sensitivity was higher for HI (2.12) than LO (1.73) trials [$F(1, 23) = 5.18$, $MSE = 0.68$, $p < .05$], and higher for Animal (2.10) than Instrument (1.75) trials [$F(1, 23) = 16.79$, $MSE = 0.17$, $p < .05$]; however, there was no significant interaction [$F(1, 23) = 1.44$, $MSE = 0.19$, $p = .24$]. Similarly, an ANOVA using c showed that there was a stronger "target-absent" bias in LO (0.55) than HI (0.22) trials [$F(1, 23) = 12.15$, $MSE = 0.21$, $p < .05$], and a stronger "target-absent" bias in Instrument (0.54) than Animal (0.23) trials [$F(1, 23) = 22.8$, $MSE = 0.10$, $p < .05$], but no significant interaction [$F(1, 23) = 0.02$, $MSE = 0.06$, $p = .89$].

Redundant trials Percent-correct values for redundant trials are displayed in Figure 7.4. Analysis of target-absent data showed no significant difference between the four trial types

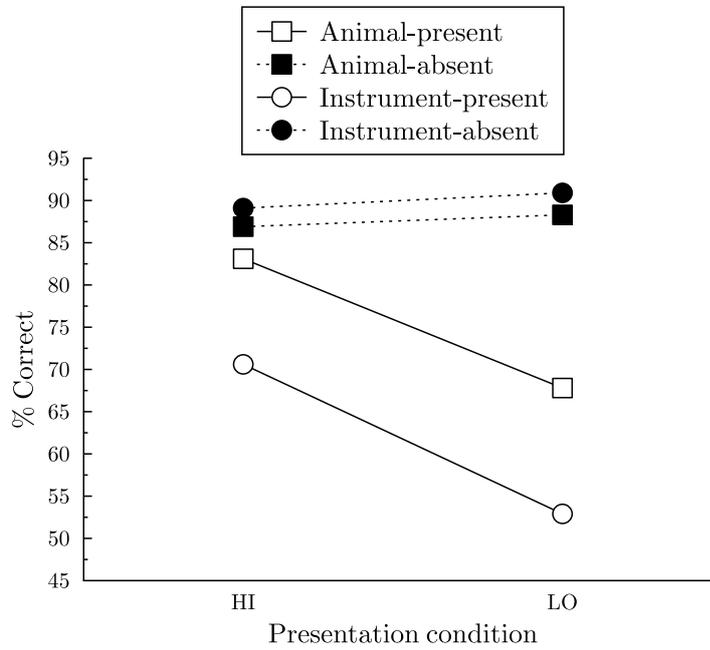


Fig. 7.3: Single-target trial response accuracy data from Experiment 8.

Tab. 7.1: Sensitivity and bias for single-target trials in Experiment 8.

		HI	LO
Animal	d'	2.34	1.85
	c	0.07	0.40
Instrument	d'	1.89	1.61
	c	0.38	0.70

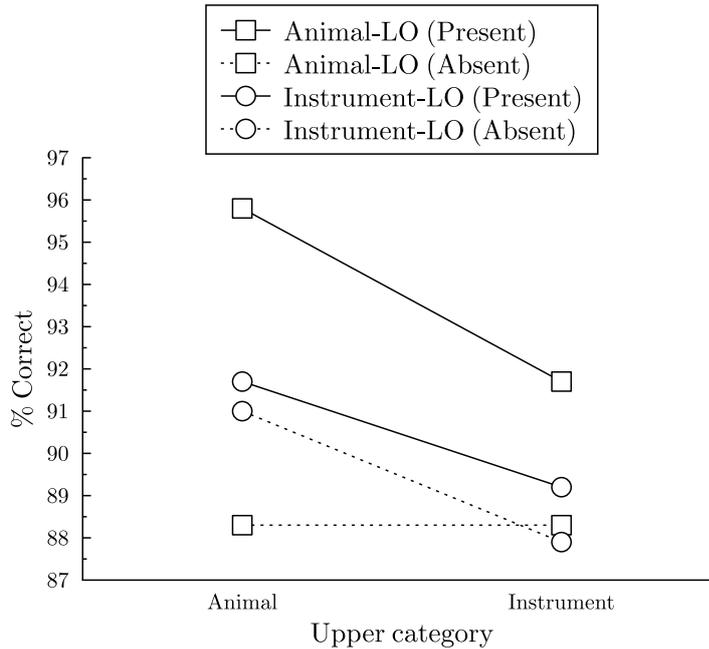


Fig. 7.4: Redundant trial response accuracy data from Experiment 8. Category (or matched non-target) appearing above fixation is on the x-axis; separate lines are used for the stimuli appearing below fixation.

[$F(3, 69) = 1.18$, $MSE = 42.14$, $p < .05$] so further analysis involved target-present data only. A 2 (Upper category) \times 2 (Lower category) ANOVA showed that responses were significantly more accurate when an Animal word was presented above fixation (93.8% correct) than when an Instrument word was presented above fixation (90.4% correct) [$F(1, 23) = 4.56$, $MSE = 58.51$, $p < .05$]. As with the latency analysis however, there was no significant interaction [$F(1, 23) = 0.91$, $MSE = 18.30$, $p = .35$], implying that redundant trials were not significantly more accurate than single-target trials.

SDT values for redundant trials appear in Table 7.2. Both d' and c were entered into separate 2 \times 2 ANOVA. The former showed that sensitivity was significantly higher when an Animal word was in the lower position (3.37) than when an instrument word was (3.07) [$F(1, 23) = 4.53$, $MSE = 0.49$, $p < .05$], but there was no significant main effect of Upper category, not a significant interaction (both $p > .2$). The latter showed that there was a significantly stronger “target-present” bias in Animal-LO trials (-0.28) than in Instrument-LO

Tab. 7.2: Sensitivity and bias for redundant trials in Experiment 8. A = Animal, I = Instrument.

	A (HI)/A (LO)	A/I	I/A	I/I
d'	3.49	3.23	3.24	2.90
c	-0.43	-0.06	-0.12	-0.04

trials (-0.05) [$F(1, 23) = 5.64$, $MSE = 0.21$, $p < .05$], that there was a marginally stronger “target-present” bias in Animal-HI (-0.25) than Instrument-HI (-0.08) trials [$F(1, 23) = 3.84$, $MSE = 0.17$, $p = .062$], and that there was a marginally significant interaction [$F(1, 23) = 3.07$, $MSE = 0.16$, $p = .093$]. This was a result of the difference between the two Animal-LO conditions being larger than that between the two Instrument-LO conditions.

Force data

Peak force data from correct responses were entered into separate $2 \times 2 \times 2$ ANOVA for single-target and redundant trials. Both analyses showed that “target-present” responses were marginally more forceful than “target-absent” responses [single-target: $F(1, 22) = 4.27$, $MSE = 132710.6$, $p = .051$; redundant: $F(1, 23) = 4.27$, $MSE = 131273.6$, $p = .050$]. There were no other significant main effects or interactions (all $p > .2$).

Tests of the race model inequality

RMI tests were conducted for each redundant target-present condition (Animal/Animal, Animal/Instrument, Instrument/Instrument, and Instrument/Animal). Initial tests were conducted after the removal of data from one participant, who made no correct responses in one condition. However, one further participant had too few correct responses in multiple conditions to perform tests across a reasonable number of quantiles (e.g., > 3). Consequently, data from this participant were removed also. Tests using data from the remaining 22 participants showed no significant violations for Animal/Animal (across 10 quantiles), Animal/Instrument (across seven quantiles), or Instrument/Instrument (across seven quantiles) conditions. The Instrument/Animal condition (tested across 10 quantiles) showed a significant violation at the fastest (.05) quantile, but not any of the others tested (see Figure 7.5).

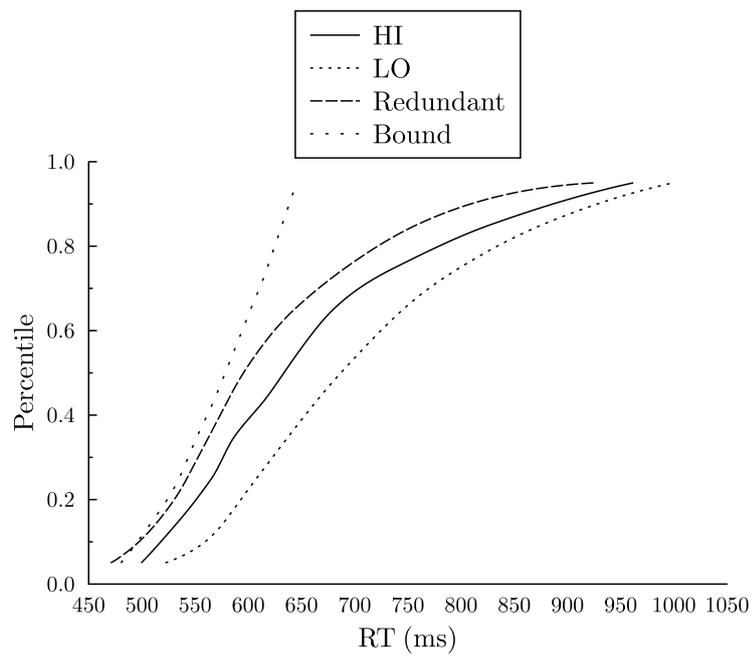


Fig. 7.5: CDFs for Instrument/Animal condition in Experiment 8. “Bound” refers to the boundary outlined by the Race Model Inequality, using the single-target data.

Redundancy gain comparisons

Comparisons between Redundant-Same and Redundant-Different redundant trials (target-present only) were conducted, using both latency and accuracy data for the raw mean and Biederman and Checkosky comparisons, and latency data only for CDF comparisons.

Mean comparisons Comparisons using the raw means for Redundant-Same and Redundant-Different trials involved the use of t -tests, with data from the two trial types averaged for each participant (e.g., the average of the redundant Animal and redundant Instrument means was taken to be their Redundant-Same RT). For the latency comparison, data from the participant who made no correct responses in one condition were removed. The resulting tests showed that responses to Redundant-Same trials (626 ms) were not significantly faster than those to Redundant-Different (630 ms) trials [$t(22) = 0.48, p = .64$], nor were they significantly more frequently correct [Redundant-Same: 92.5%, Redundant-Different: 91.7%; $t(23) = 0.95, p = .35$].

Biederman and Checkosky comparisons Differences were calculated between the redundant conditions and the best single-target conditions for each participant, for each of the four redundant conditions. Redundant-Same and Redundant-Different averages were then compared. Again, the removal of data from one participant who made no correct responses in one condition was necessary for the latency comparison. Neither accuracy nor latency comparisons showed significant differences in redundancy gains between the two redundant trial types [for latency, Redundant-Same redundancy gain = 19.8 ms, Redundant-Different = 18.6 ms, $t(22) = 0.19, p = .85$; for accuracy, Redundant-Same = 11.7%, Redundant-Different = 10.3%, $t(23) = 1.53, p = .14$].

CDF comparisons Data from the 22 participants used to test the RMI were placed into a modified version of the MATLAB script written by Ulrich et al. (2007). The modifications allowed a comparison of each participant's redundant CDFs with CDFs from an independent race model constructed from single-target data. Data from a further six participants were excluded from analysis, so that the comparison could be conducted across 10 quantiles using

data from the same participants in each redundant condition. These CDFs appear in Figure 7.6. For each quantile (.05, .15, . . . , .95), a difference score was calculated by subtracting the RT value from the independent race model CDF from the RT value from the relevant redundant condition. Difference scores for Animal/Animal and Instrument/Instrument CDFs were averaged to produce a Redundant-Same score for each participant. Difference scores for Animal/Instrument and Instrument/Animal CDFs were averaged to produce a Redundant-Different score for each participant. These scores were then entered into a 2 (Type) \times 10 (Quantile) ANOVA. This showed no significant main effect of Type [$F(1, 15) = 0.23$, $MSE = 3337.6$, $p = .64$], nor a significant Type \times Quantile interaction [$F(9, 135) = 1.01$, $MSE = 849.1$, $p = .38$, Greenhouse-Geisser corrected].

Discussion

The aim of Experiment 8 was to supplement the findings of Experiment 7, and determine which of the race model and cell assembly coactivation accounts of redundancy gain in the semantic categorisation task provided the best explanation for results. This was accomplished using the logic of Miller (1991), by seeing whether redundancy gain was more pronounced in trials where stimuli were from the same semantic category than in those where stimuli were from different categories. As outlined in the introduction to this experiment, greater redundancy gain in Redundant-Same trials would suggest a cell assembly-coactivation explanation, whereas no difference between Redundant-Same and Redundant-Different trials would be evidence against that account.

From the three types of comparisons conducted, it is clear that whether redundant trials were Redundant-Same or Redundant-Different in type had no observable effect on participants' responses. Neither raw means, nor differences from the best single-target condition, nor discrepancies from an independent race model CDF, showed significant differences between the two trial types. This is further evidence against the cell assembly account, and suggests that of the two models focused on in previous chapters, a race model explanation best accounts for the redundancy gains found in the semantic categorisation task.

However, in an unexpected twist, this experiment also showed the first significant violation of the RMI to be found in any of the experiments described to this point: In

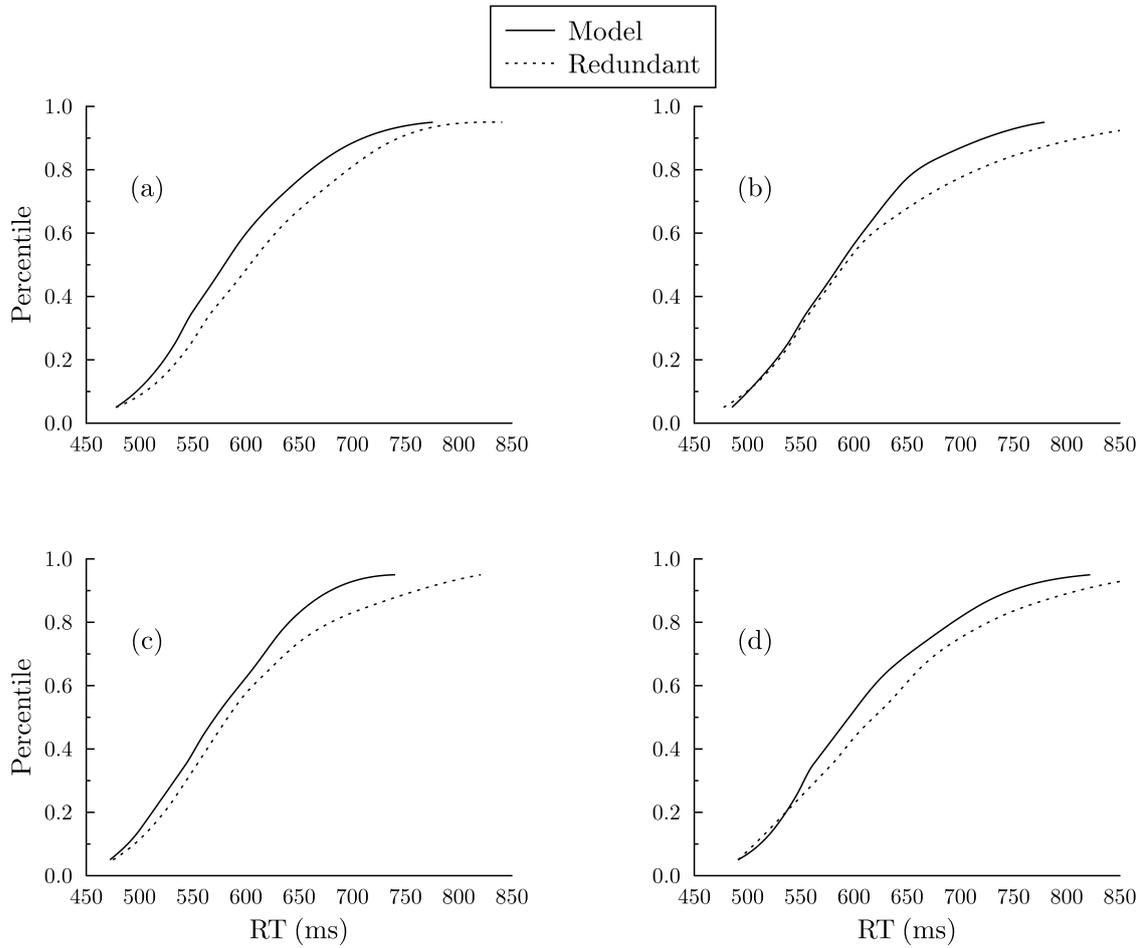


Fig. 7.6: CDFs for the four redundant trial types, along with independent race model CDFs derived from single-target trial data. Panel (a) shows Animal/Instrument trials, panel (b) Instrument/Animal trials, panel (c) Animal/Animal trials, and panel (d) Instrument/Instrument trials.

redundant trials where Instrument words were presented above fixation and Animal words below, responses were faster than could be accounted for by any race model. It is worth noting that this condition involved words from the slower of the two categories (Instruments) being presented in the faster position, and words from the faster category (Animals) being presented in the slower position. This means that the distribution of the individual stimulus processing times should have had the largest overlap, and thus the greatest scope for coactivation.

On the other hand, the redundancy gain predicted by a race model would also be greatest where the distributions have the largest overlap, meaning that the highest chance of an artefactual RMI violation would probably exist in this condition. RMI tests were conducted in all of Experiments 1–8, and across multiple quantiles, meaning that a significant violation would be expected by chance at some point, even if the bound were always satisfied with respect to the true underlying RT distributions. However, given that the RMI provides a conservative boundary, it could be imprudent to dismiss entirely even a single violation.

With the possibility that the RMI violation is merely a Type I error admitted, how else could the lack of a difference between Redundant-Same and Redundant-Different trials — suggested to support a race model explanation — be reconciled with the finding of a violation of the race model inequality, given no semantic coactivation appears to be present? One possibility is that coactivation occurred at a level other than the semantic (which was the focus of Experiment 8). In most tasks where a redundancy gain is found and coactivation assumed, there are a number of levels at which it is suggested this coactivation can operate. As shown in the coactivation versions of models illustrated in Figure 1.2, there are further possibilities for coactivation in the semantic categorisation task (i.e., coactivation of semantic representations, in categorisation itself, or in response selection). This task was deliberately constructed to preclude the possibility of coactivation at the perceptual level, by way of the use of non-repeated stimuli in redundant trials (i.e., two different words, rather than two copies of the same word). This leaves the categorisation and response levels shown in Figure 1.2, plus the motor level not shown there. The lack of significant effects on response force can be considered evidence against coactivation at the motor level (though it is not decisive in that regard). The lack of semantic effects would seem to argue against coactivation in categorisation (as shown in the coactivation version of Figure 1.2b) as much as it would coactivation of semantic

representations, given this form of coactivation would presumably be more likely to occur when categorisation of both stimuli relates to the same category than when it relates to different categories. This suggests response selection (as shown in the coactivation version of Figure 1.2c) as the most likely locus of any coactivation which might have occurred.

Conclusion

To summarise, the results of Experiment 8 provide strong evidence against the hypothesis that redundancy gain in the semantic categorisation task is a result of neural representations of a category being more rapidly activated by the presentation of multiple words from within that category. Further, the significant violation of the race model provides some evidence against a race model explanation of redundancy gain in the task, whilst the lack of any difference between Redundant-Same and Redundant-Different trials lends support to the race model and response-level coactivation explanations. These findings, along with those of Experiments 1–7, are considered in a broader theoretical context next.

8. GENERAL DISCUSSION

In carrying out the experiments described here, I aimed to shed light on a number of issues relating to redundancy gain in higher level processing in general, and to semantic processing and semantic memory. To that end, I was interested in seeking answers to a number of key questions. First, given the ambiguity of previous experiments where redundancy gains were found in tasks which may have required higher level processing, and given the differences between semantic categorisation and the processes involved in those other tasks, would a semantic categorisation task show any redundancy gain at all? Second, if a semantic categorisation task did show redundancy gain, how might this best be explained? Third, if a coactivation model can account for redundancy gain in semantic categorisation, at which level of processing does this coactivation occur? And fourth, what inferences about the cell assembly explanation for redundancy gain in lexical decision — which, as mentioned in the Introduction to this thesis, has also been invoked to explain redundancy gain in many other tasks involving complex stimuli — might be made on the basis of results from a semantic categorisation task?

I will now deal with each of these questions in light of the findings of Experiments 1–8, followed by discussion of some side issues which were raised over the course of these experiments.

8.1 Does a semantic categorisation task show any redundancy gain at all?

The answer to this question, as the results of Experiments 1–8 make clear, is a resounding “yes”. Redundancy gains were present in the results of each experiment, though their extents varied. By using the comparison of redundant means to each participant’s better single-target mean, gains varied across experiments from a low of 6 ms in Experiment 7 to a high of 26 ms in Experiment 3 (mean = 14 ms) for latency, and from a low of 4.4% in Experiment 6 to a high of

12.1% in Experiment 5 (mean = 8.6%) for accuracy. Experiment 4 also showed significant redundancy gains in the force of responses, but these were generally absent in other experiments.

In any case, this represents the first unequivocal demonstration of redundancy gain in a task requiring semantic processing, and perhaps the most unequivocal demonstration of redundancy gain in a task requiring higher level processing in general. As outlined in the Introduction, results from a number of previous studies had shown redundancy gains in tasks where such processing might have occurred, but few if any of these involved tasks where such processing could unequivocally have been said to be *required*. The semantic categorisation task used here required semantic processing, by definition.

These results allow the dismissal of those specific parallel models in which limited processing capacity is overwhelmed by the need to process two meaningful stimuli simultaneously, leading to interference effects. As mentioned in the Introduction to this thesis, such models would predict worse performance in redundant than single-target trials (i.e., a redundancy *loss*), which was not the case. As a result, I shall not discuss these models further.

It is worth noting that the results of Experiments 1–8 show a resemblance to results in a number of the tasks used by previous researchers. Most obvious, of course, is the similarity between results of the semantic categorisation task and those of Mohr and Pulvermüller’s lexical decision tasks (e.g., Mohr et al., 1996). Though the implications of my results for the cell assembly explanation of redundancy gains in lexical decision are addressed more thoroughly later in this chapter, I shall briefly summarise some of these task-related similarities now.

First, redundancy gains in the semantic categorisation task occurred only for affirmative (that is, “target-present”) responses. This was true across all experiments where alternative responses were associated with the presence of members and non-members of a target category (i.e., all experiments other than Experiment 6, in which two distinct categories were associated with opposing responses). This parallels results not only from lexical decision tasks with redundant stimuli, where gains appeared for “word” but not “non-word” responses (e.g., Mohr, Pulvermüller, & Zaidel, 1994; Mohr, Pulvermüller, Rayman, & Zaidel, 1994; Mohr et al., 1996), but also those from facial familiarity judgements (e.g., Mohr et al., 2002), where gains appeared for “familiar” but not “unfamiliar” responses.

Second, gains were not dependent on lateralised stimulus presentation. In fact, stronger gains were found in Experiment 3, where stimuli were presented above and below fixation, than in Experiment 1, where presentation was lateralised, though both used exactly the same procedure and stimuli (other than with regard to display, obviously). Whilst this might have something to do with a larger proportion of correct responses in Experiment 3, the maintenance of gains when stimuli are not lateralised mirrors similar findings of Mohr et al. (1996) in lexical decision. These results suggest that enhanced performance with redundant stimuli in both lexical decision and semantic categorisation is not reliant on a method of stimulus presentation designed to allow the cerebral hemispheres initial access to different stimuli.

Of course, though knowing that a task requiring semantic processing can show redundancy gains is theoretically useful, an understanding of how these gains come about is far more so. I discuss this issue next.

8.2 How can the redundancy gain in semantic categorisation best be explained?

In the brief overview of redundancy gain presented in the Introduction, I outlined the fact that two types of model are traditionally used to account for redundancy gain: race models, and coactivation models. Mohr and Pulvermüller's cell assembly account of gains in redundant lexical decision (e.g., Mohr, Pulvermüller, & Zaidel, 1994) — which was subsequently extended to explain gains in judgements of facial familiarity (Mohr et al., 2002) — can be considered a coactivation model, based on the fact that it posits a contribution from each redundant stimulus in eventual response activation. In this case, redundant targets provide dual stimulation to that target's neural representation, and a response is enacted more rapidly as a result. Though this explanation seems to provide an effective account of most aspects of the results from lexical decision tasks using redundant stimuli, my results suggest that at the very least it is not appropriate to explain similar findings in semantic categorisation. Rather, for reasons I shall now detail, a race model seems to provide a more suitable account. During this discussion, I deal predominantly with performance in redundant trials, but shall also incorporate other aspects of the findings from the semantic categorisation task which any model explaining redundancy gain would also need to account for. These findings are listed in

Tab. 8.1: Key findings from Experiments 1–8 and the ability of the cell assembly coactivation (CAC) and self-terminating race (STR) models to account for them.

Finding	CAC model	STR model
1. Redundancy gain in latency for “target-present” responses	Yes	Yes
2. Redundancy gain in accuracy for “target-present” responses	Yes	Yes
3. No gain for in latency for “target-absent” responses	Yes	Yes
4. No gain in accuracy for “target-absent” responses	Yes	Yes
5. Slower “target-absent” than “target-present” responses	No specific prediction	Yes
6. RVF advantage (vs LVF) in target-present trials	Yes	No specific prediction
7. Absence of semantic effects on redundancy gain	No	Yes

Table 8.1; I will refer to them by number (e.g., “Finding 3”) in the following sections.

Assessment of the cell assembly model

Table 8.1 lists the key findings from Experiments 1–8. As can be seen, the cell assembly-based coactivation account of redundancy gain in semantic categorisation can explain the majority of these. For instance, gains in positive (Findings 1 and 2) but not negative responses (Findings 3 and 4) for both latency and accuracy can result from only positive responses providing the opportunity for such coactivation; and the RVF advantage which occurred through most experiments (Finding 6) can be explained by lateralisation of either lexical or semantic representations.

However, a telling argument against the suitability of the cell assembly account comes from Finding 7 (i.e., the results of Experiments 7 and 8). According to the principles of cell assembly theory, and embodied theories of knowledge such as Barsalou’s (1999) perceptual symbol systems theory, neural representations of items from the same semantic category should not only have more overlapping components than those of items from different categories, but they should also be organised under a superordinate category-level cell assembly. Further,

Mohr and Pulvermüller (e.g., Mohr et al., 1996) suggested that the “strength” of cell assemblies (that is, the extent to which they form a defined and internally connected unit) should affect the magnitude of coactivation resulting from neural summation with redundant stimulation. By this logic, gains should have been larger when redundant trials involved two stimuli from the same category than when two stimuli from disparate target categories were presented. This is because the former allows the sort of coactivation illustrated in the coactivation version of Figure 1.2a, whereas the latter does not.

Results from Experiments 2, 7, and 8 showed that this was not the case. In Experiment 7, redundant trials involved one stimulus from each of the categories “parts of the human body” and “parts of a building”; however, redundancy gain in this experiment was no smaller than in Experiment 2, where both categories had been used as sole targets with separate groups of participants. This was reinforced by the results of Experiment 8, in which the categories “animals” and “musical instruments” were defined as targets, and redundant trials with two members of the same category compared to redundant trials with one member from each category. Again, no difference was found. Together, these findings cast substantial doubt on the suitability of a cell assembly-based coactivation explanation for redundancy gain in semantic categorisation.

It is worth noting that though such an account seems inappropriate in this context, this does not bring into question more general suggestions of cell assemblies as a basis for the neural representation of knowledge (e.g., Barsalou, 1999; Pulvermüller, 1999). Rather, the absence of any effect of semantic relatedness on redundancy gain simply implies that coactivation at the semantic level is not responsible for enhanced performance on redundant trials. Other accounts of these results might be consistent with cell assembly theory, even if it does not comprise an integral part of the explanations they offer.

In addition, the semantic relatedness used in my experiments pertained to the taxonomic categories from which target stimuli were taken. This is not the only form of semantic relatedness which exists, and might not even be the most promising form for any hypothetical neural coactivation. For instance, evidence from imaging studies (e.g., Sachs et al., 2008; Sass, Sachs, Krach, & Kircher, 2009) shows that neural priming effects are stronger for contextual than taxonomic relationships, implying that thematic similarity could be a better candidate for

revealing semantic effects in a categorisation task with redundant stimuli. However, the fact that semantic effects were not merely small in Experiment 8, but almost entirely absent even on a purely numerical level, suggests that this possibility is tenuous.

In any case, the apparent lack of semantic coactivation does not preclude coactivation at some other stage of processing, such as the response-level coactivation illustrated in the coactivation version of Figure 1.2c. In fact, results from Experiment 8 showed a violation of the race model inequality, which would ordinarily be used as evidence against a race model account, and in support of coactivation. However, as this represented the only RMI violation in any of Experiments 1–8, the extent to which such a violation is believable is questionable. I will return to this issue in the following section.

A tentative race model

A viable alternative to an account requiring coactivation may be a parallel race model of the form outlined in the discussion section of Experiments 1–3, and represented in the race version of the model illustrated in Figure 1.2c. According to the model I propose, a “target-present” response requires only that one stimulus is identified as a target; as soon as this occurs, the response is made, irrespective of the processing status of any other stimuli presented. This means that redundancy gain occurs due to statistical facilitation, as the stimuli race to activate the response. In contrast, a “target-absent” response is made only once the status of both stimuli as non-targets has been determined.

Together, the specifications of this race model seem sufficient to allow it to explain not only redundancy gain and its lack of variation from single-category to dual-category redundant trials, but most if not all of the other major findings of Experiments 1–8 (see Table 8.1). One important point such a model can account for is redundancy gains in accuracy as well as latency in target-present trials (i.e., Findings 1 and 2). This is because on each redundant trial, an erroneous “target-absent” response will only be made if *both* of the target stimuli are incorrectly identified. For target-absent trials, the story is different: irrespective of where stimuli are presented, or whether a trial involves single or redundant non-targets, both stimuli must be correctly identified as non-targets for the appropriate response to ensue, thus leading to no redundancy gain for latency or accuracy in target-absent trials (Findings 3 and 4).

Tab. 8.2: RT standard deviations for “target-present” and “target-absent” responses in Experiments 1– 5 (averaged across all participants). Column containing p -values refers to results of t -test comparisons, uncorrected.

Experiment	Present	Absent	p
1	136	139	.38
2	131	145	< .05
3	162	162	.89
4	229	287	< .05
5	155	153	.59

If this model is accurate, the lack of latency gains in target-absent trials (Finding 3) would seem to imply that both non-words and non-target words are processed as quickly as one another.¹ This would be strange however, both because the familiarity of non-target words would seem to make them intuitively more target-like than non-words, and because results of numerous lexical decision experiments show slower responses for non-words than for words. It is possible that this is an artefact of a temporal response threshold, whereby participants make the “target-absent” response by default if insufficient evidence for a “target-present” response has been obtained by some time t . Were this the case, RTs for “target-absent” responses should be less variable for each participant. However, as shown in Table 8.2, the opposite was the case: in both cases where standard deviations from “target-present” and “target-absent” response latencies differed, it was the latter which showed greater variance. Though not conclusive, this is evidence against the idea that participants were using a temporal response threshold for “target-absent” responses. If a temporal threshold is not used to determine when a “target-absent” response is made, and both non-target stimuli must instead be identified as such, this model would require that non-target stimuli are processed in a very similar fashion irrespective of whether they are words or non-words.

Another possible explanation of the lack of redundancy gains in latency for “target-absent” responses (Finding 3) is that over time, participants’ tendency to make this response gradually increases, and they enact the response once this tendency has reached a certain threshold. Evidence of target presence erroneously obtained from non-target stimuli

¹ And analogously, to explain the lack of any redundancy gains in accuracy of responses to target-absent trials (Finding 4), it must be comparably likely that participants will incorrectly identify non-words and non-target words.

(e.g., noise) could arrest or reverse this tendency, slowing “target-absent” responses accordingly. This would provide a further source of variability in target-absent RTs which is missing from the temporal threshold explanation, thus potentially explaining the similar variability in target-absent and target-present RTs: where this erroneous evidence of target presence is only weak, the drift toward the response threshold would proceed in a relatively linear fashion, whereas on occasion the evidence will be stronger and the threshold will take far longer to reach.

Another of the key findings from Experiments 1–8 that this tentative race model seems able to explain is that “target-absent” responses were generally slower than “target-present” responses (Finding 5). This would simply be due to the exhaustive processing in target-absent trials. The model would also suggest that visual field effects may occur in target-present trials (Finding 6) because a response is made as soon as a target is identified, meaning that if this occurs more quickly in one display location than another, latencies will tend to differ between the conditions. That said, this model provides no *a priori* reasons for assuming that target identification should occur more quickly in any display location than in any other.

Though a race model of this sort need not specify any stochastic dependence existing between processing on separate channels, it might be informative to compare the predictions of such a model with an assumption of channel independence to observed data from redundant trials (in a model analogous to that of, e.g., Mulligan & Shaw, 1980; Pirenne, 1943; Tulving, Mandler, & Baumal, 1964). Table 8.3 shows the predictions of accuracy results for each experiment, derived from just such a model, in comparison to the actual results of Experiments 1–8.² The model describes the probability of a target response, $\Pr(“t”)$, for each combination of the four possible stimuli — that is, a target in position one (T_1), a target in position two (T_2), a non-target in position one (N_1), or a non-target in position two (N_2). Each of these possible stimuli has a probability of being classified as a target, α ; for example, α_{T_1} is the probability that a target in position one is classified as a target. Critically, I assume that non-target stimuli are equally likely to be classified as targets, irrespective of what form they take or their display location. This assumption is based on the finding that accuracy seems

² With the exception of Experiment 6, as data from target-absent trials are required to test the model, and Experiment 6 had no such trials.

unaltered across target-absent trial types. Thus, it is assumed that $\alpha_{N_1} = \alpha_{N_2} \equiv \alpha_N$.

It is further assumed that the classification of the stimulus in one display location as a target or non-target does not affect the classification of the stimulus in the other display location (that is, this classification of the two stimuli is stochastically independent). According to this model, the observed probabilities of “target-present” responses in the four stimulus conditions are:

$$\Pr(\text{“}t\text{”}|N_1N_2) = \alpha_N + \alpha_N - \alpha_N \times \alpha_N \quad (8.1)$$

$$\Pr(\text{“}t\text{”}|T_1N_2) = \alpha_{T_1} + \alpha_N - \alpha_{T_1} \times \alpha_N \quad (8.2)$$

$$\Pr(\text{“}t\text{”}|N_1T_2) = \alpha_N + \alpha_{T_2} - \alpha_N \times \alpha_{T_2} \quad (8.3)$$

$$\Pr(\text{“}t\text{”}|T_1T_2) = \alpha_{T_1} + \alpha_{T_2} - \alpha_{T_1} \times \alpha_{T_2} \quad (8.4)$$

The model is testable, because it uses three free parameters (i.e., α_{T_1} , α_{T_2} , and α_N) to predict four observable probabilities [i.e., $\Pr(\text{“}t\text{”}|N_1N_2)$ et cetera]. In constructing Table 8.3, I first estimated α_N from the observed value of $\Pr(\text{“}t\text{”}|N_1N_2)$ using Equation 8.1, then estimated α_{T_1} and α_{T_2} from $\Pr(\text{“}t\text{”}|T_1N_2)$ and $\Pr(\text{“}t\text{”}|N_1T_2)$ using Equations 8.2 and 8.3 together with the previously-estimated value of α_N , and finally computed the estimated value of $\Pr(\text{“}t\text{”}|T_1T_2)$ from the two previously estimated values of α_{T_1} and α_{T_2} . As can be seen in the table, these predicted values are close to the observed values in every experiment, with the largest discrepancy being a 5.5% under-estimate of correct responses in Experiment 7. In short, even this uncomplicated version of a race model seems to provide a surprisingly accurate estimate of actual performance, albeit with consistent [and statistically significant: $t(6) = 5.06, p < .05$] slight underprediction of performance on redundant trials.

Can a serial model account for redundancy gain in semantic categorisation?

Both race and coactivation models imply that stimuli are processed in parallel. Could an account where processing is serial also provide an adequate explanation for redundancy gain in the semantic categorisation task? As mentioned in the Introduction to this thesis, under some circumstances such models are able to be dismissed in redundant targets experiments simply

Tab. 8.3: Observed redundant target-present percent-correct values from each experiment (bar Experiment 6, for which there were no “target-absent” responses to construct a prediction), compared to those derived from a race model with channel independence assumed.

Experiment	Predicted	Observed
1	84.9	88.6
2	83.5	87.5
3	96.2	97.3
4	88.3	91.4
5	73.2	77.5
7	75.9	81.4
8	89.7	92.1

because of the improved performance with two stimuli versus one: in serial models, the mere fact that two targets are present should not improve performance in comparison to single-target trials, because responses to both will occur after processing of the initially attended stimulus is complete (e.g., Townsend, 1990). However, this is only the case where no distractor or non-target stimuli are present on single-target trials; in my experiments two stimuli were presented even on single-target trials (the second stimulus being a non-word, of course). This means that redundancy gains cannot be used as evidence against serial processing, simply because the first stimulus to be processed will always be a target on redundant trials, but will only be a target some of the time on single-target trials. Some other method is necessary to rule out serial processing, then.

Initially, the asymmetrical results for “target-present” and “target-absent” responses (Findings 1–4) might be presumed to assist in this regard. In all of Experiments 2–8, participants were directly informed that a single non-target stimulus was sufficient to conclude that the “target-absent” response was correct. What could explain the lack of a redundancy gain in non-target trials then, if redundant non-target trials also involved a higher probability than their single stimulus non-target counterparts that the first stimulus to be processed would be informative? A serial model analogous to the race model offered earlier should lead to similar target/non-target asymmetries. To wit, in this model, processing only ends when a stimulus is identified as a target, or the processing of all stimuli is complete. Thus, in redundant target trials a response occurs once the first stimulus has been processed; in single

target trials a response occurs after processing of the first stimulus where it is a target, and after processing of the second otherwise; and in all non-target trials, a response is only made once processing of both stimuli is concluded.

However, such a model makes a prediction that seems to have been contradicted by the data from Experiments 2 and 7. In both of these experiments, responses were slower in target-present trials where Building words were presented to the LVF than in target-absent trials. According to the serial model outlined above, “target-present” responses should generally be faster than “target-absent” responses, as the latter are only ever made once processing of both stimuli is complete, whereas the former are made as soon as processing of the target is complete. Formally, RTs for the LVF-Building and target-absent trials can be broken down as follows:

$$\text{LVF-Building}_{RT} = a(T_L) + (1 - a)(T_L + NT_R) + x \quad (8.5)$$

$$\text{Absent}_{RT} = NT_L + NT_R + x \quad (8.6)$$

Here, a is the proportion of trials on which the LVF stimulus is processed first, $(1 - a)$ is the proportion of trials on which the RVF stimulus is processed first, T_L is the processing latency for the LVF target stimulus, NT_L and NT_R are the processing latencies for the LVF and RVF non-target stimuli, and x is the RT component unrelated to stimulus processing. Using these equations, it can be seen that for responses to LVF-Building trials to be slower than “target-absent” responses would require processing of LVF-Building stimuli to be substantially slower than processing of non-words and non-targets. This can be shown by a demonstration of what criteria LVF-Building stimulus processing would need to fit for RTs in such trials to be

even equal to RTs in “target-absent” trials:

$$\begin{aligned}
 \text{LVF-Building}_{RT} &= \text{Absent}_{RT} \\
 a(T_L) + (1 - a)(T_L + NT_R) + x &= NT_L + NT_R + x \\
 a(T_L) + T_L + NT_R - a(T_L) - a(NT_R) &= NT_L + NT_R \\
 T_L + NT_R - a(NT_R) &= NT_L + NT_R \\
 T_L &= NT_L + a(NT_R)
 \end{aligned}$$

In sum, obtaining equal RTs in LVF-Building trials and target-absent trials would require the processing of Building words in the LVF to be slower than the processing of LVF non-targets by a value equivalent to the product of RVF non-target processing and the proportion of trials where the LVF stimulus is processed first. Obtaining *slower* RTs for LVF-Building than target-absent trials would require the difference to be even more pronounced.³ This is far from impossible, but does seem counter-intuitive. Thus, the fact that “present” responses for LVF-Building trials were slower than “absent” responses is evidence against a serial account for performance in the semantic categorisation task.

A second argument against a serial explanation comes from the fact that, as described earlier, RT variance in target-absent trials across Experiments 1–5 was either no different to or significantly greater than variance in target-present RTs (see Table 8.2). In fact, this was the case both for redundant and single-target “present” responses, across all experiments (single-target trials: present SD = 163, absent SD = 165; redundant trials: present SD = 140, absent SD = 164; excluding Experiment 6, where there were no target-absent trials). In the serial model proposed, target-absent RTs should be less variable, as all responses result from exhaustive processing; by contrast, some target-present responses on single-target trials will occur after exhaustive processing, and others will occur after processing of only one stimulus is complete. Thus, larger or equal RT SDs for target-absent trials argue against a serial explanation.

³ Of course, this would be less problematic were the stimulus in the right visual field *always* processed first: in this case the LVF-Building stimulus would only need to be processed slower, on average, than the LVF stimulus on target-absent trials. This seems unlikely given that if RVF stimuli are always processed first, there should be no difference between RVF and BVF target-present RTs — which is clearly not the case.

In combination, these two pieces of evidence suggest that serial processing in the semantic categorisation task with redundant targets is unlikely to have occurred. This implies that some form of parallel processing was involved in performing the task — likely either by way of a summation of activity at a non-semantic level of processing, or through simultaneous category membership judgements being made about both stimuli — suggesting that the use of this relatively complex form of processing task-related information is not limited in its application to a single entity at a time. Either way, a serial model does not provide a convincing alternative to the tentative race model proposed earlier.

8.3 *Locus of coactivation*

Though I have stated above that a race model appears to offer the best explanation for the redundancy gains found in Experiments 1–8, I also mentioned that coactivation at a non-semantic level of processing might have been possible. This suggestion obtains support from the RMI violation found in Experiment 8, albeit this may simply have been a Type I error. If not the result of error though, at what level might any coactivation have occurred?

The use of word stimuli was designed to preclude potential coactivation at a perceptual level which — due to the physical similarities often shared by items from the same semantic category — might have occurred had picture stimuli been used. In addition, redundant trials involved two physically different words (unlike the lexical decision tasks used by Mohr and Pulvermuller, for instance; e.g., Mohr et al., 1996). When combined with the fact that all trials — redundant and single-target — involved the presentation of the same number of stimuli, these two facts make it seem unlikely that any perceptual coactivation could have occurred.

Coactivation at the level of lexical processing also seems implausible. There are two reasons for this. First, both target-present and target-absent redundant trials involved the presentation of two words. However, only target-present trials showed a redundancy gain. Second, as mentioned above redundant trials involved the presentation of two different words rather than two copies of a single word. As such, any coactivation of hypothetical lexical cell assemblies of the sort postulated by Pulvermüller (1999) seems unlikely.

One caveat to this is that participants did seem to use the presentation of two words on

a trial as evidence for a “target-present” response, as witnessed by the consistent bias toward that response in redundant trials. It is unclear why this should be the case, but one possibility is that the extra lexicality was misidentified as “targetness” (“misidentified” because the presentation of two words was equally as likely to occur in target-absent as in target-present trials). If this were true, lexical coactivation could somehow have led to enhanced “target-present” responses, albeit due to task-irrelevant factors.

As stated in the preceding section, the lack of any effects of semantic similarity on redundancy gains seems to be evidence against semantic coactivation. Of course, it is always possible that there is sufficient flexibility in semantic processing to allow items from distinct categories to provide the same level of coactivation as items from a single category. However, were this the case, it would be questionable as to whether any resultant coactivation could really be termed semantic. That is, if semantic variables have no effect on its magnitude, describing the effect as one which affects semantic processing seems counterintuitive.

Motor coactivation provides another possible source of enhanced responding in the semantic categorisation task. One piece of evidence suggests that this might have occurred: peak response force was consistently higher in redundant trials than in single-target trials. Though the difference was rarely significant within experiments, when redundant peak force values were compared to the highest single-target peak force value from each experiment, redundant responses were significantly more forceful [$t(7) = 3.19, p < .05$], by an average of 4.4 cN. Complicating matters, however, is that this was true for both “target-present” and “target-absent” responses, whereas only the former showed redundancy gains for latency and accuracy. The apparent implication is that though trials with dual word stimuli evoked more forceful responses, this seems to have been irrelevant to other aspects of task performance. As such, it is difficult to see how motor coactivation could have been responsible for, say, the RMI violation found for Instrument/Animal trials in Experiment 8.

By a process of elimination, this leaves the categorisation and response-selection levels as the remaining possible loci of any coactivation which might have occurred. Given the lack of semantic effects on redundancy gain, coactivation in categorisation appears unlikely. As such, coactivation in response selection — wherein separate category membership judgements are made for each stimulus and the overall evidence for a “target-present” response accumulates

centrally (as illustrated in Figure 1.2c) — appears the most plausible form any coactivation could have taken. This is analogous to what Miller (1991) referred to as an “independent” coactivation model, as opposed to an “interactive” model where processing of one target is influenced by processing of the other.

As to how seriously the RMI violation in Experiment 8 should be taken as evidence of coactivation, given the lack of similar violations in Experiments 1–7, this is uncertain. On the one hand, the RMI is a conservative tool by which to dismiss race models as it provides a boundary which should not be violated by any separate activation model, irrespective of the extent and direction of stochastic dependence that exists between single stimuli. Many researchers have chosen instead to test their redundant trial data against a specific race model which assumes stochastic independence, sometimes supported by an argument that this is the most appropriate model to test (e.g., Veldhuizen et al., 2010), and sometimes without such support (e.g., Laurienti et al., 2004; Molholm et al., 2004). On the other hand, here there is no clear reason to assume stochastic independence, so the conservative test seems satisfactory. Consequently, though this is certainly an issue which could be further illuminated with more research, on the basis of the present data it seems reasonable to conclude that coactivation is unnecessary to explain the results found across my experiments.

8.4 Implications for cell assembly accounts of redundancy gain in lexical decision

The semantic categorisation tasks I used were heavily based on the tasks used in lexical decision experiments run by Mohr and Pulvermüller (e.g., Mohr, Pulvermüller, & Zaidel, 1994; Mohr, Pulvermüller, Rayman, & Zaidel, 1994; Mohr et al., 1996). As such, perhaps it is not surprising that the pattern of results obtained here should also resemble theirs in a number of ways. What then — if anything — might be inferred from my results about the cell assembly-based coactivation account they offered to explain redundancy gain in lexical decision?

First and most obviously, it is worth noting that extrapolating the cell assembly model to explain redundancy gains in the semantic categorisation task does not appear possible. This is primarily due to the lack of semantic effects, as detailed earlier. Given that Mohr et al.

(1996) suggested that the strength of connections within an assembly should affect the ease with which neural summation occurs, and that neuropsychological evidence points to conceptual knowledge taking the form of feature-based, distributed cell assemblies (Kiefer & Pulvermüller, 2012), the prediction of such an account would be that coactivation occurs more strongly for semantically related than unrelated redundant targets.

In and of itself, of course, the unsuitability of the cell assembly-based coactivation account in explaining results of the semantic categorisation task does not imply that it is also a flawed explanation for findings in lexical decision tasks. There are enough differences between the two experiments to make such a direct inference problematic. For instance, in Mohr and Pulvermüller's experiments (e.g., Mohr et al., 1996) redundant trials involved simultaneous presentation of two copies of the same stimulus — be it word, or non-word. By contrast, redundant trials in my experiments always involved the presentation of two distinct members (target-present) or non-members (target-absent) of a target category or categories. It may be the case that, to evince coactivation, redundant stimuli must represent precisely the same concept, rather than simply being similar or components of the same superordinate category assembly; or it may be that any coactivation in the LDT was entirely perceptual, due to redundant stimuli being physically identical (which was obviously not the case in the semantic categorisation task).

In addition, there are a number of potential differences between the tasks of determining “wordness”, as in lexical decision, and category membership, as in semantic categorisation. As stated earlier, though there is unequivocal evidence that a variety of information can be used in lexical decision (for review, see e.g., Balota, Cortese, Sergent-Marshall, Spieler, & Yap, 2004), at its most basic all that lexical decision requires is a decision about whether or not the stimulus presented is familiar. By contrast, models of semantic categorisation tend to describe it as a matching process (e.g., Barsalou, 1999; Collins & Loftus, 1975). It seems plausible that the enhanced neural activation which is purported to result from the presentation of redundant stimuli can aid judgements about familiarity, but not category membership. For instance, one of the factors contributing to a familiarity judgement in a lexical decision task could be the overall activity level in perisylvian and fusiform regions proposed by Pulvermüller et al. (2009) as the neural site of written word form representations. Such information would not have been

useful in the categorisation task, where both target-present and target-absent trials involved the presentation of words.

With these between-task differences noted, some of the similarities in performance between the two tasks are striking. First, in those of my experiments which involved lateralised presentation, performance in RVF trials was far superior to that in LVF trials. This reflects similar findings in Mohr and Pulvermüller's experiments (e.g., Mohr, Pulvermüller, & Zaidel, 1994; Mohr, Pulvermüller, Rayman, & Zaidel, 1994; Mohr et al., 1996). Second, as mentioned previously, both the lexical decision experiments and my semantic categorisation experiments showed redundancy gains only for "positive" responses — that is, "target-present" or "word" responses. Third, gains were present both when stimuli were presented in a lateralised (that is, to the left and right visual fields) and a non-lateralised manner (above and below fixation in my experiments; two copies presented to a single visual field in those of Mohr et al., 1996). And fourth, calculations of redundant trial accuracy levels using single-target data in a stochastically independent race model framework (see Equation 8.4) provided close estimates of observed accuracy in the semantic categorisation task, and in lexical decision (Mohr et al., 1996).⁴

Given these similarities, it would be ideal if a single model could explain the patterns of results across both tasks. Mohr and Pulvermüller (e.g., Mohr et al., 1996) explained each of these findings as follows. For RVF superiority, they suggested that the contralateral processing of laterally presented stimuli, combined with a left hemisphere superiority for language processing, was responsible. For word-specific gains, Mohr et al. (1996) suggested this was a result of only familiar stimuli (words) having existing neural representations which would allow a summation of activation. For redundancy gains with non-lateralised presentation, they suggested such gains should occur with dual stimulation irrespective of where redundant stimuli were presented. Finally, Mohr et al. (1996) suggested that though close predictions of redundant trial accuracy could be obtained through a stochastically independent race model, accuracy was sufficiently high that ceiling effects would likely reduce the usefulness of any such comparison.

However, despite the ability of the cell assembly model to account for LDT findings, as

⁴ Though, as they did not have any distractor stimuli in single-target trials, the equation Mohr et al. (1996) used to make this estimate was simplified.

described earlier it seems unable to provide a sufficient explanation of results from the semantic categorisation task. Could the race model outlined earlier adequately then account for redundancy gain in lexical decision? It is possible, but for it to do so would require at least one modification. This is because of the difference between the two tasks in the format of single-target trials: in the semantic categorisation task I used distractors (non-words) to keep stimulus number constant between redundant and single-target trials, whereas in Mohr and Pulvermüller's lexical decision tasks single-target trials involved the presentation of only one stimulus.⁵ To account for the lack of visual field effects in target-absent trials in the semantic categorisation task, the race model outlined earlier relies on the presence of two stimuli on each trial, irrespective of whether or not they are targets. This is because if stimulus number changes from one target-absent trial to another the mean RTs would also be expected to vary, since it would presumably take longer to process two stimuli exhaustively than to process one. In short, if single-target trials lacked task-irrelevant distractors to keep stimulus number constant, the model would predict a redundancy *loss* for negative (target-absent/non-word) trials. No such loss was apparent in the lexical decision tasks with redundant targets (e.g., Mohr, Pulvermüller, & Zaidel, 1994; Mohr et al., 1996), thus implying that the model would require modification to account for this.

One possible modification could involve the application of a temporal response threshold for negative responses. Earlier I suggested that comparable or increased RT variability for “target-absent” responses in the semantic categorisation task was evidence against such a feature, and this point stands. However, one other piece of evidence lends support to the temporal threshold idea. In Experiment 4, when the length of time participants had to respond was increased from 1500 to 3000 ms, this slowed RTs for “target-absent” responses more extensively than it slowed “target-present” RTs. If a temporal threshold were being used, this would be expected due to the reduced time pressure on responses allowing the threshold to be more conservative (that is, participants feel “safer” waiting longer before concluding that no target was presented); by contrast, evidence for target-presence could still lead to an immediate response.

⁵ It is worth noting that this factor precludes a serial model explanation for gains in the two tasks, as such a model would predict no redundancy gain in the lexical decision task (cf. Townsend, 1990).

To summarise, though there might be more suitable variants of the race model to achieve the specific goal of explaining target-absent RTs from the semantic categorisation task, a model where — in the absence of evidence for a positive response — “target-absent” or “non-word” responses are made at a temporal threshold is better able to account for results from both semantic categorisation and lexical decision tasks with redundant targets. As such, this model can be considered preferable to the cell assembly coactivation account, which fails to adequately explain the lack of semantic effects on redundancy gain in semantic categorisation. Further research could investigate whether redundant stimuli in lexical decision must be identical words for a gain to result, or whether performance is still improved when different words are used. If the former, this would support either a cell assembly account of redundancy gain in the LDT or a model involving perceptual coactivation;⁶ if the latter, it would be evidence against both.

8.5 *Other issues*

Over the course of Experiments 1–8, a number of findings arose which did not directly bear on the key issues I aimed to investigate in this thesis. These were: the lack of VF advantages when words were presented lateralised and vertically; the poor performance when stimuli were presented to the left visual field; and the differences in response bias across different presentation conditions. I shall briefly touch on each of these issues in turn now.

First, I turn to the issue of the lack of any RVF advantage when words were presented vertically, reading from top-to-bottom, as in Experiment 5. Initially, this would appear to be evidence against the suggestion that RVF advantages are due to a left hemisphere superiority for lexical processing (which should presumably persist irrespective of the manner in which a word is displayed), and in support of an attentional or perceptual explanation. For instance, it could be that more information about stimulus identity is obtained from a word’s initial letters (e.g., Broerse & Zwaan, 1966), which are presented closer to fixation in the RVF when stimuli are horizontal and lateralised. However, as stimuli were presented more foveally (1–2° to either side of fixation), any inferences about visual field advantages might be inappropriate. This is

⁶ These two models could themselves be dissociated by an experiment comparing redundant trials where two copies of the same word are presented in the same, or varying fonts; perceptual coactivation would predict decreased gain with varying fonts, whereas cell assembly coactivation would predict no change.

because there is considerable debate regarding whether visual information presented foveally (within the central 2-3° of vision) projects bilaterally, or is split and projected contralaterally (e.g., Ellis & Brysbaert, 2010; Jordan & Paterson, 2009). As such, drawing firm conclusions about hemispheric differences from the findings of Experiment 5 is probably unwise.

Second, poor LVF performance was itself the focus of Experiment 4, in which the window participants were allowed for a response was increased to see whether this enhanced response accuracy. As it did not, this implies that the frequent errors in LVF trials are not a result of processing occurring too slowly, but rather may be due to stimuli often not being perceived to the extent that their identity can be determined. Whether this is a result of poor right hemisphere language processing (e.g., Mohr et al., 1996), poor right hemisphere semantic processing (e.g., Binder et al., 2009), or other factors, cannot be concluded from my results. However, an explanation of the findings based on the location of words' initial letters might be tested via an experiment where words are presented bilaterally and participants are asked to report those words, or as many letters of each as they can. If participants are able to report the foveal and parafoveal letters presented in the LVF (that is, the final letters of the words) but not the words themselves, this would imply that perceptual rather than linguistic factors are to blame for poor LVF performance; if not, this would be evidence against such an explanation.

Third, to the varying biases across presentation conditions. As mentioned in the discussion of Experiment 6, when presentation of stimuli is lateralised a consistency between lexical decision tasks and the semantic categorisation task is that the bias toward positive (target/word) and negative (non-target/non-word) responses seems to change based on whether relevant stimuli are presented in the left or right visual fields. Specifically, in the LVF there is a bias toward negative responding, which is reversed or lessened with RVF presentation. In the semantic categorisation task, of course, the presence of a word should not provide any evidence regarding a trial's "target-ness", meaning that despite the use of word stimuli the biases found in LDTs should not *a priori* have been expected to persist. What this might indicate is that there is a greater proportion of LVF trials — in both semantic categorisation and LDT — where participants are inclined to guess, and that these guesses tend to be weighted toward the negative response (that is, if a participant is unsure whether or not they perceived a target, they tend to assume they probably did not). Earlier I speculated that this could be a result of

participants finding it harder to conclude target absence (or “non-wordness”) with any certainty. This was because Experiment 6 showed a bias toward “Building” responses in LVF trials, and Experiment 2 had shown that the Building categorisation was more difficult (slower RTs, more errors) than Animal or Body categorisation. This implies that under conditions of uncertainty participants tend to make the response associated with the more “difficult” decision, perhaps simply because there are likely to be more instances of that decision which lead to uncertainty. This suggestion could be tested in future research by manipulating the difficulty of decisions associated with competing response options, and seeing how this affects response bias.

8.6 Conclusion

Experiments 1–8 have revealed that redundancy gain can occur in a semantic categorisation task, and that the most plausible explanation for this gain appears to be a parallel self-terminating race model. This implies that multiple categorisation decisions can be made in parallel. The lack of semantic effects on redundancy gain casts doubt on a cell assembly-based coactivation model of redundancy gain in semantic categorisation, and perhaps also on its suitability to explain previous results in lexical decision tasks (e.g., Mohr, Pulvermüller, & Zaidel, 1994; Mohr et al., 1996) and fame judgements for faces (e.g., Mohr et al., 2002; Schweinberger et al., 2003). However, if coactivation was responsible for redundancy gains in semantic categorisation, it appears most likely to have occurred at the level pertaining to response selection, rather than the lexical, semantic, categorisation, motor, or perceptual levels. Productive future research might involve further tests of the cell assembly explanation for LDT redundancy gain, or use of the semantic categorisation task with redundant stimuli to investigate factors which affect the ability of stimuli to be categorised in parallel (e.g., determining how many items can be processed simultaneously).

References

- Baird, L. M., & Burton, A. M. (2008). The bilateral advantage for famous faces: interhemispheric communication or competition? *Neuropsychologia*, *46*(5), 1581–7. doi: 10.1016/j.neuropsychologia.2008.01.001
- Balota, D. a., Cortese, M. J., Sergent-Marshall, S. D., Spieler, D. H., & Yap, M. (2004). Visual word recognition of single-syllable words. *Journal of Experimental Psychology: General*, *133*(2), 283–316. doi: 10.1037/0096-3445.133.2.283
- Banich, M., & Karol, D. (1992). The sum of the parts does not equal the whole: Evidence from bihemispheric processing. *Journal of Experimental Psychology: Human Perception and Performance*, *18*(3), 763–784.
- Barsalou, L. (1999). Perceptual symbol systems. *Behavioral and brain sciences*, *22*(4), 577–660.
- Barutchu, A., Crewther, D. P., & Crewther, S. G. (2009). The race that precedes coactivation: development of multisensory facilitation in children. *Developmental science*, *12*(3), 464–73. doi: 10.1111/j.1467-7687.2008.00782.x
- Barutchu, A., Crewther, S. G., Fifer, J., Shivdasani, M. N., Innes-Brown, H., Toohey, S., ... Paolini, A. G. (2010). The relationship between multisensory integration and IQ in children. *Developmental psychology*. doi: 10.1037/a0021903
- Battig, W., & Montague, W. (1969). Category norms for verbal items in 56 categories. *Journal of Experimental Psychology*, *80*(3), 1–46.
- Bauer, L. (1994). Introducing the Wellington Corpus of Written New Zealand English. *Te Reo*, *37*, 21–28.
- Biederman, I., & Checkosky, S. F. (1970). Processing redundant information. *Journal of Experimental Psychology*, *83*(3, Pt.1), 486–490. doi: 10.1037/h0028841
- 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), 2767–96. doi: 10.1093/cercor/bhp055
- Borowsky, R., & Besner, D. (1993). Visual word recognition: a multistage activation model. *Journal of experimental psychology. Learning, memory, and cognition*, *19*(4), 813–40.

- Bradley, D., & Garrett, M. (1983). Hemisphere differences in the recognition of closed and open class words. *Neuropsychologia*, *21*(2), 155–159.
- Bradshaw, J. L., & Nettleton, N. C. (1983). Evidence from normal subjects. In *Human cerebral asymmetry* (pp. 82–117). Englewood Cliffs, N.J.: Prentice-Hall.
- Broerse, A. C., & Zwaan, E. J. (1966). The information value of initial letters in the identification of words. *Journal of Verbal Learning and Verbal Behavior*, *5*(5), 441–446. doi: 10.1016/S0022-5371(66)80058-0
- Caramazza, A. (1997). How Many Levels of Processing Are There in Lexical Access? *Cognitive Neuropsychology*, *14*(1), 177–208. doi: 10.1080/026432997381664
- Casey, P. J., & Heath, R. A. (1988). Category norms for Australians. *Australian Journal of Psychology*, *40*(3), 323–339.
- Cavina-Pratesi, C., Bricolo, E., Prior, M., & Marzi, C. (2001). Redundancy gain in the stop-signal paradigm: Implications for the locus of coactivation in simple reaction time. *Journal of Experimental Psychology: Human Perception and Performance*, *27*(4), 932–941.
- Chen, Y.-C., & Spence, C. (2010). When hearing the bark helps to identify the dog: semantically-congruent sounds modulate the identification of masked pictures. *Cognition*, *114*(3), 389–404. doi: 10.1016/j.cognition.2009.10.012
- Chiarello, C. (1988). Lateralization of lexical processes in the normal brain: A review of visual half-field research. In H. A. Whitaker (Ed.), *Contemporary reviews in neuropsychology* (pp. 36–76). New York: Springer-Verlag.
- Chiarello, C., & Nuding, S. (1987). Visual field effects for processing content and function words. *Neuropsychologia*, *25*(3), 539–48.
- Collignon, O., Girard, S., Gosselin, F., Roy, S., Saint-amour, D., Lassonde, M., & Lepore, F. (2008). Audio-visual integration of emotion expression. *Brain Research*, *1242*, 126–135.
- Collignon, O., Girard, S., Gosselin, F., Saint-Amour, D., Lepore, F., & Lassonde, M. (2010). Women process multisensory emotion expressions more efficiently than men. *Neuropsychologia*, *48*(1), 220–5. doi: 10.1016/j.neuropsychologia.2009.09.007
- Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, *82*(6), 407–428. doi: 10.1037/0033-295X.82.6.407

- Collins, A. M., & Quillian, M. (1969). Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior*, 8(2), 240–247. doi: 10.1016/S0022-5371(69)80069-1
- Corballis, M. C. (1998). Interhemispheric neural summation in the absence of the corpus callosum. *Brain*, 121, 1795–807.
- Dalton, P., Doolittle, N., Nagata, H., & Breslin, P. A. S. (2000). The merging of the senses: integration of subthreshold taste and smell. *Nature Neuroscience*, 3(5), 431–2. doi: 10.1038/74797
- Dijkstra, T., Frauenfelder, U., & Schreuder, R. (1993). Bidirectional grapheme-phoneme activation in a bimodal detection task. *Journal of Experimental Psychology: Human Perception and Performance*, 19(5), 931–950.
- Dunlap, K., & Wells, G. (1910). Some experiments with reactions to visual and auditory stimuli. *Psychological Review*, 17(5), 319–335. doi: 10.1037/h0075154
- Ellis, A. W., & Brysbaert, M. (2010). Split fovea theory and the role of the two cerebral hemispheres in reading: a review of the evidence. *Neuropsychologia*, 48(2), 353–65. doi: 10.1016/j.neuropsychologia.2009.08.021
- Endrass, T., Mohr, B., & Pulvermüller, F. (2004). Enhanced mismatch negativity brain response after binaural word presentation. *The European Journal of Neuroscience*, 19(6), 1653–60. doi: 10.1111/j.1460-9568.2004.03247.x
- Endress, A. D., & Potter, M. C. (2012). Early Conceptual and Linguistic Processes Operate in Independent Channels. *Psychological Science*, 23(3), 235–245. doi: 10.1177/0956797611421485
- Fischer, R., & Miller, J. (2008). Differential redundancy gain in onset detection versus offset detection. *Perception & Psychophysics*, 70(3), 431. doi: 10.3758/PP
- Funnell, E. (1983). Phonological processes in reading: New evidence from acquired dyslexia. *British Journal of Psychology*, 74, 159–180.
- Giray, M., & Ulrich, R. (1993). Motor coactivation revealed by response force in divided and focused attention. *Journal of Experimental Psychology: Human Perception and Performance*, 19(6), 1278–1291.
- Gondan, M., Götze, C., & Greenlee, M. (2010). Redundancy gains in simple responses and go/no-go tasks. *Attention, Perception, & Psychophysics*, 72(6), 1692–1709. doi:

10.3758/APP

- Grice, G. R., Canham, L., & Boroughs, J. (1984). Combination rule for redundant information in reaction time tasks with divided attention. *Perception & Psychophysics*, *35*(1), 451–463.
- Grice, G. R., & Reed, J. M. (1992). What makes targets redundant? *Perception & Psychophysics*, *51*(5), 437–42.
- Hasbrooke, R. E., & Chiarello, C. (1998). Bihemispheric processing of redundant bilateral lexical information. *Neuropsychology*, *12*(1), 78–94.
- Hautus, M. J., & Lee, A. (2006). Estimating sensitivity and bias in a yes/no task. *The British Journal of Mathematical and Statistical Psychology*, *59*(2), 257–73. doi: 10.1348/000711005X65753
- Hebb, D. O. (1949). *The organization of behavior*. New York: John Wiley & Sons.
- Heron, W. (1957). Perception as a function of retinal locus and attention. *The American Journal of Psychology*, *70*(1), 38–48.
- Hershenson, M. (1962). Reaction time as a measure of intersensory facilitation. *Journal of Experimental Psychology*, *63*(3), 289–93.
- Jamet, E., & Lebohec, O. (2007). The effect of redundant text in multimedia instruction. *Contemporary Educational Psychology*, *32*(4), 588–598. doi: 10.1016/j.cedpsych.2006.07.001
- Jordan, T. R., & Paterson, K. B. (2009). Re-evaluating split-fovea processing in word recognition: a critical assessment of recent research. *Neuropsychologia*, *47*(12), 2341–53. doi: 10.1016/j.neuropsychologia.2008.07.020
- Juola, J., & Atkinson, R. (1971). Memory scanning for words versus categories. *Journal of Verbal Learning and Verbal Behavior*, *10*, 522–527.
- Kiefer, M., & Pulvermüller, F. (2012). Conceptual representations in mind and brain: Theoretical developments, current evidence and future directions. *Cortex*, *48*(7), 805–25. doi: 10.1016/j.cortex.2011.04.006
- Koppen, C., Alsius, A., & Spence, C. (2008). Semantic congruency and the Colavita visual dominance effect. *Experimental Brain Research*, *184*(4), 533–46. doi: 10.1007/s00221-007-1120-z

- Krummenacher, J., Muller, H., & Heller, D. (2002). Visual search for dimensionally redundant pop-out targets: Redundancy gains in compound tasks. *Visual Cognition*, *9*(7), 801–837. doi: 10.1080/13506280143000269
- Kuipers, J., & La Heij, W. (2008). Semantic facilitation in category and action naming: Testing the message-congruency account. *Journal of Memory and Language*, *58*(1), 123–139. doi: 10.1016/j.jml.2007.05.005
- Kutas, M., McCarthy, G., & Donchin, E. (1977). Augmenting mental chronometry: The P300 as a measure of stimulus evaluation time. *Science*, *197*(4305), 792–795.
- Laurienti, P. J., Kraft, R. a., Maldjian, J. a., Burdette, J. H., & Wallace, M. T. (2004). Semantic congruence is a critical factor in multisensory behavioral performance. *Experimental Brain Research*, *158*(4), 405–14. doi: 10.1007/s00221-004-1913-2
- Liederman, J., Merola, J., & Martinez, S. (1985). Interhemispheric collaboration in response to simultaneous bilateral input. *Neuropsychologia*, *23*(5), 673–683.
- Loftus, E., & Cole, W. (1974). Retrieving attribute and name information from semantic memory. *Journal of Experimental Psychology*, *102*(6), 1116–1122.
- Marshall, C. E., & Parr, W. V. (1996). New Zealand norms for a subset of Battig and Montague's (1969) Categories. *New Zealand Journal of Psychology*, *25*(1), 24–29.
- Martin, a., & Chao, L. L. (2001). Semantic memory and the brain: structure and processes. *Current Opinion in Neurobiology*, *11*(2), 194–201.
- McFarland, C., Kellas, G., & Klueger, K. (1974). Category similarity, instance dominance, and categorization time. *Journal of Verbal Learning and Verbal Behavior*, *708*, 698–708.
- Miller, J. (1982). Divided attention: evidence for coactivation with redundant signals. *Cognitive Psychology*, *14*(2), 247–79.
- Miller, J. (1986). Timecourse of coactivation in bimodal divided attention. *Perception & Psychophysics*, *40*(5), 331–43.
- Miller, J. (1991). Channel interaction and the redundant-targets effect in bimodal divided attention. *Journal of Experimental Psychology: Human Perception and Performance*, *17*(1), 160–9.
- Miller, J. (2004). Exaggerated redundancy gain in the split brain: a hemispheric coactivation account. *Cognitive Psychology*, *49*(2), 118–54. doi: 10.1016/j.cogpsych.2003.12.003

- Miniussi, C., Girelli, M., & Marzi, C. A. (1998). Neural site of the redundant target effect: Electrophysiological evidence. *Journal of Cognitive Neuroscience*, *10*(2), 216–30.
- Mishkin, M., & Forgays, D. (1952). Word recognition as a function of retinal locus. *Journal of Experimental Psychology*, *43*(1), 43–48.
- Mohr, B., Endrass, T., Hauk, O., & Pulvermüller, F. (2007). ERP correlates of the bilateral redundancy gain for words. *Neuropsychologia*, *45*(9), 2114–24. doi: 10.1016/j.neuropsychologia.2007.01.015
- Mohr, B., Landgrebe, A., & Schweinberger, S. R. (2002). Interhemispheric cooperation for familiar but not unfamiliar face processing. *Neuropsychologia*, *40*(11), 1841–8.
- Mohr, B., & Pulvermüller, F. (2002). Redundancy gains and costs in cognitive processing: effects of short stimulus onset asynchronies. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*(6), 1200–1223. doi: 10.1037//0278-7393.28.6.1200
- Mohr, B., Pulvermüller, F., Cohen, R., & Rockstroh, B. (2000). Interhemispheric cooperation during word processing: evidence for callosal transfer dysfunction in schizophrenic patients. *Schizophrenia Research*, *46*(2-3), 231–9.
- Mohr, B., Pulvermüller, F., Mittelstädt, K., & Rayman, J. (1996). Multiple simultaneous stimulus presentation facilitates lexical processing. *Neuropsychologia*, *34*(10), 1003–13.
- Mohr, B., Pulvermüller, F., Rayman, J., & Zaidel, E. (1994). Interhemispheric cooperation during lexical processing is mediated by the corpus callosum: evidence from the split-brain. *Neuroscience Letters*, *181*(1-2), 17–21.
- Mohr, B., Pulvermüller, F., & Zaidel, E. (1994). Lexical decision after left, right and bilateral presentation of function words, content words and non-words: evidence for interhemispheric interaction. *Neuropsychologia*, *32*(1), 105–24.
- Molholm, S., Ritter, W., Javitt, D. C., & Foxe, J. J. (2004). Multisensory visual-auditory object recognition in humans: a high-density electrical mapping study. *Cerebral Cortex*, *14*(4), 452–465. doi: 10.1093/cercor/bhh007
- Mordkoff, J. T., Miller, J., & Roch, a. C. (1996). Absence of coactivation in the motor component: evidence from psychophysiological measures of target detection. *Journal of Experimental Psychology: Human Perception and Performance*, *22*(1), 25–41.
- Mordkoff, J. T., & Yantis, S. (1991). An interactive race model of divided attention. *Journal of*

- Experimental Psychology: Human Perception and Performance*, 17(2), 520–538. doi: 10.1037/0096-1523.17.2.520
- Mordkoff, J. T., & Yantis, S. (1993). Dividing attention between color and shape: evidence of coactivation. *Perception & Psychophysics*, 53(4), 357–66.
- Mulligan, R., & Shaw, M. (1980). Multimodal signal detection: Independent decisions vs. integration. *Attention, Perception, & Psychophysics*, 28(5), 471–478.
- Murray, M. M., Foxe, J. J., Higgins, B. a., Javitt, D. C., & Schroeder, C. E. (2001). Visuo-spatial neural response interactions in early cortical processing during a simple reaction time task: a high-density electrical mapping study. *Neuropsychologia*, 39(8), 828–44.
- Norman, D. A., & Bobrow, D. G. (1975). On data-limited and resource-limited processes. *Cognitive Psychology*, 7, 44–64.
- Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia*, 9(1), 97–113.
- Pirenne, M. (1943). Binocular and unocular threshold of vision. *Nature*, 152, 698–9.
- Plato. (1892). *Dialogues, vol. 4 - Parmenides, Theaetetus, Sophist, Statesman, Philebus* (3rd ed., Vol. 4; B. Jowett, Ed.). London: Oxford University Press.
- Pulvermüller, F. (1999). Words in the brain's language. *Behavioral and Brain Sciences*, 22(2), 253–279. doi: 10.1017/S0140525X9900182X
- Pulvermüller, F. (2001). Brain reflections of words and their meaning. *Trends in Cognitive Sciences*, 5(12), 517–524.
- Pulvermüller, F., & Mohr, B. (1996). The concept of transcortical cell assemblies: a key to the understanding of cortical lateralization and interhemispheric interaction. *Neuroscience and Biobehavioral Reviews*, 20(4), 557–66.
- Pulvermüller, F., Shtyrov, Y., & Hauk, O. (2009). Understanding in an instant: neurophysiological evidence for mechanistic language circuits in the brain. *Brain and Language*, 110(2), 81–94. doi: 10.1016/j.bandl.2008.12.001
- Quillian, M. R. (1962). A revised design for an understanding machine. *Mechanical Translation*, 7(1), 17–29.
- Quillian, M. R. (1967). Word concepts: A theory and simulation of some basic semantic

- capabilities. *Behavioral Science*, *12*(5), 410–430.
- Quillian, M. R. (1969). The teachable language comprehender: a simulation program and theory of language. *Communications of the ACM*, *12*(8), 459–476. doi: 10.1145/363196.363214
- Raab, D. (1962). Statistical facilitation of simple reaction times. *Transactions of the New York Academy of Sciences*, *24*, 574–90.
- Rastle, K., Harrington, J., & Coltheart, M. (2002). 358,534 nonwords: The ARC Nonword Database. *The Quarterly Journal of Experimental Psychology: Section A*, *55*(4), 1339–1362. doi: 10.1080/02724980244000099
- Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, *85*(2), 59–108.
- Rips, L., Shoben, E., & Smith, E. (1973, February). Semantic distance and the verification of semantic relations. *Journal of Verbal Learning and Verbal Behavior*, *12*(1), 1–20. doi: 10.1016/S0022-5371(73)80056-8
- Sachs, O., Weis, S., Zellagui, N., Huber, W., Zvyagintsev, M., Mathiak, K., & Kircher, T. (2008). Automatic processing of semantic relations in fMRI: neural activation during semantic priming of taxonomic and thematic categories. *Brain Research*, *1218*(2), 194–205. doi: 10.1016/j.brainres.2008.03.045
- Sass, K., Sachs, O., Krach, S., & Kircher, T. (2009). Taxonomic and thematic categories: Neural correlates of categorization in an auditory-to-visual priming task using fMRI. *Brain Research*, *1270*, 78–87. doi: 10.1016/j.brainres.2009.03.013
- Savazzi, S., & Marzi, C. a. (2008). Does the redundant signal effect occur at an early visual stage? *Experimental Brain Research*, *184*(2), 275–81. doi: 10.1007/s00221-007-1182-y
- Schwarz, W., & Ischebeck, A. (1994). Coactivation and statistical facilitation in the detection of lines. *Perception*, *23*, 157–168.
- Schweinberger, S. R., Baird, L. M., Blümmler, M., Kaufmann, J. M., & Mohr, B. (2003). Interhemispheric cooperation for face recognition but not for affective facial expressions. *Neuropsychologia*, *41*(4), 407–14.
- Shallice, T., Warrington, E. K., & McCarthy, R. (1983). Reading without semantics. *The Quarterly Journal of Experimental Psychology*, *35A*, 111–138.
- Smith, E. E., Shoben, E. J., & Rips, L. J. (1974). Structure and process in semantic memory:

- A featural model for semantic decisions. *Psychological Review*, *81*(3), 214–241.
- Stolz, J., & Besner, D. (1996). Role of set in visual word recognition: Activation and activation blocking as nonautomatic processes. *Journal of Experimental Psychology: Human Perception and Performance*, *22*(5), 1166–1177.
- Suied, C., Bonneel, N., & Viaud-Delmon, I. (2009). Integration of auditory and visual information in the recognition of realistic objects. *Experimental Brain Research*, *194*(1), 91–102. doi: 10.1007/s00221-008-1672-6
- Tamietto, M., Adenzato, M., Geminiani, G., & de Gelder, B. (2007). Fast recognition of social emotions takes the whole brain: interhemispheric cooperation in the absence of cerebral asymmetry. *Neuropsychologia*, *45*(4), 836–43. doi: 10.1016/j.neuropsychologia.2006.08.012
- Tamietto, M., Latini Corazzini, L., de Gelder, B., & Geminiani, G. (2006). Functional asymmetry and interhemispheric cooperation in the perception of emotions from facial expressions. *Experimental Brain Research*, *171*(3), 389–404. doi: 10.1007/s00221-005-0279-4
- Todd, J. W. (1912). Reaction to multiple stimuli. *Archives of Psychology*, *21*(8).
- Townsend, J. (1990). Serial vs. parallel processing: Sometimes they look like Tweedledum and Tweedledee but they can (and should) be distinguished. *Psychological Science*, *1*(1), 46–54.
- Townsend, J., & Nozawa, G. (1997). Serial exhaustive models can violate the race model inequality: Implications for architecture and capacity. *Psychological Review*, *104*(3), 595–602.
- Tulving, E., Mandler, G., & Baumal, R. (1964). Interaction of Two Sources of Information in Tachistoscopic Word Recognition. *Canadian Journal of Psychology*, *18*(1), 62–71.
- Ulrich, R., Miller, J., & Schröter, H. (2007). Testing the race model inequality: an algorithm and computer programs. *Behavior Research Methods*, *39*(2), 291–302.
- Van Overschelde, J. (2004). Category norms: An updated and expanded version of the norms. *Journal of Memory and Language*, *50*(3), 289–335. doi: 10.1016/j.jml.2003.10.003
- Veldhuizen, M. G., Shepard, T. G., Wang, M.-F., & Marks, L. E. (2010). Coactivation of gustatory and olfactory signals in flavor perception. *Chemical Senses*, *35*(2), 121–33. doi:

10.1093/chemse/bjp089

Warrington, E. K., & Shallice, T. (1980). Word-form dyslexia. *Brain*, *103*(1), 99–112.

Wiktionary. (2006). *Wiktionary: Frequency lists*. Retrieved May, 2010, from

http://en.wiktionary.org/wiki/Wiktionary:Frequency_lists

Yonelinas, A. P. (2002). The Nature of Recollection and Familiarity: A Review of 30 Years of Research. *Journal of Memory and Language*, *46*(3), 441–517. doi:

10.1006/jmla.2002.2864

Yoshizaki, K. (2001). Effects of visual familiarity for words on interhemispheric cooperation for lexical processing. *Cognitive Brain Research*, *12*(3), 409–14.

Young, A. W., & Ellis, A. W. (1985). Different methods of lexical access for words presented in the left and right visual hemifields. *Brain and Language*, *24*(2), 326–58.

APPENDIX

A. CATEGORY NORMS

In the initial experiments described in this thesis, I used Battig and Montague’s (1969) category norms to create target stimuli. However, using these norms with participants from a present-day New Zealand student population is problematic, given the temporal and geographical differences between Battig and Montague’s sample and my own. Consequently, I set out to collect norm data from the very population I was testing, so as to ensure its appropriateness for use in the semantic categorisation task.

Over the course of three semesters I collected category norms for those categories originally used by Battig and Montague (1969), and with the addition of extra categories for which norms were collected by Van Overschelde (2004) and Casey and Heath (1988). I also added some further *ad hoc* categories to the list, with the idea of making some comparison between natural and *ad hoc* categories in future semantic categorisation experiments. This resulted in an overall list of 74 different categories (though early participants provided norms for an abridged set, as described below), responses to which not only provided useful data for stimulus construction, but also provided an interesting insight into University of Otago psychology students’ representations of the categories used.

When using these data to create stimuli for the categorisation experiments, responses in any of the categories consisting of more than one word, and those where category membership was potentially problematic, were removed from the stimulus set. Additionally, after pre-testing with a separate participant group, those stimuli classified correctly as members of one of those categories used in Experiment 1–8 at a below-chance level were also removed.

Method

Participants 231 first- and second-year psychology students from the University of Otago (179 female, 51 male, one who did not specify sex; aged between 17 and 51 years, mean = 20.4, SD = 4.6, four who did not specify age) participated in return for extra credit toward their psychology grades. 36 completed the task using an abridged set of categories, whilst the remaining 195 were presented with the full list of 74.

Apparatus and Stimuli Categories for which norms were collected consisted of the 56 categories used by Battig and Montague (1969), additional categories used by Van Overschelde (2004) and Casey and Heath (1988), and some further *ad hoc* categories not appearing in any of those studies. Of the Battig and Montague categories, a small number of modifications were made to several: “A member of the clergy” was changed to “A religious official”, “A type of footgear” was changed to “A type of footwear”, and “A girl’s first name” was changed to “A female’s first name” (so as to mirror the category “A male’s first name”). “A type of ship” was also altered to “A type of watercraft”, after some participants in the initial testing sessions misinterpreted “ship” as “sheep”, leading to a number of inappropriate responses.

Responses were collected on Excel spreadsheets with the column width adjusted so that only one category appeared on the screen at a time. Columns were labelled with category numbers (e.g., “Category 29”) but not names. A stopwatch was used to time response periods for each category.

The first 36 participants tested were presented with an abbreviated set of 45 categories, which consisted of the modified Battig and Montague (1969) categories, minus those categories for which NZ-specific norms had been collected by Marshall and Parr (1996). The full list of categories appears in Table A.1; items appearing on the abridged list are marked with an asterisk.

Procedure Data collection was undertaken in a computer lab, with up to 10 participants per session. The session commenced with the reading of instructions, which were a modification of those used by Battig and Montague (1969), taking account of the fact that responses were entered via computer rather than in a notebook. The first 30-second response period was then

initiated with the reading of the first category name. This procedure was repeated for all categories used. Order of category presentation was randomised so that each group of roughly 20 participants received them in a different order.

Data Analysis Responses were collated on a category-by-category basis. Where multiple spelling variations (or mis-spellings) of a word appeared, these were standardised to the most common variation (e.g., “aluminum” changed to “aluminium”). For the purposes of stimulus creation, frequency of occurrence was the only variable of interest. As such, other variables (e.g., within-participant rank of responses) were ignored. Responses which obviously pertained to a category other than that for which they had been given (e.g., when they described a member of the preceding or following category) were discounted.

Results

Results reported here are for a subset of the total 74 categories. This subset is composed of the categories used in Experiments 1–8, and those used by McFarland, Kellas, and Klueger (1974). Full collation of data for the remaining categories has not yet been completed.

Tables A.2–A.17 display the response frequency data for the following categories: “A precious stone”; “A metal”; “A military title”; “A four-footed animal”; “A kitchen utensil”; “A part of the human body”; “An elected office”; “A religious official”; “A part of a building”; “A musical instrument”; “A flower”; “A tree”; “A fish”; “A weapon”; “A bird”; and “A vegetable”.

Discussion

In this section, I will briefly note some of the differences between the data collected here, and those from the norms of Battig and Montague (1969), focusing on the categories used in my experiments (animals, parts of the human body, parts of a building, and musical instruments).

First, with reference to the category “A four-footed animal”, the majority of the examples offered most frequently by Battig and Montague’s (1969) participants also occurred frequently amongst Otago students’ examples. One notable exception is “squirrel”, which was the 17th most frequently cited animal in Battig and Montague’s study, but equal 64th here, with only two mentions. Other animals native to North America, such as skunks, raccoons, and

beavers, appeared in Battig and Montague's top 40, but were not mentioned by Otago students at all.

With regard to "parts of the human body", responses followed quite a similar pattern to those recorded by Battig and Montague (1969). In fact, the top ten responses were exactly the same between my study and theirs, albeit appearing in a slightly different order. Perhaps reflective of a decreased taboo, genitalia were more frequently mentioned by Otago students than Battig and Montague's sample: despite having a sample less than half the size, "penis" was offered as an example on more than twice as many occasions here (20 versus 9), and "vagina" seven times more (14 versus 2). Cross-referencing with the Van Overschelde (2004) norms, this seems likely to be more an effect of time rather than geography — though "vagina" did not make their response cutoff of being included by at least 0.05% of all respondents, "penis" received exactly the same rank as it did here (34th).

"A part of a building" also showed distinct similarities to earlier norms. For instance, the top six-ranked items were not only the same, but in the same order, as in the study of Battig and Montague (1969). The only obvious difference in the top-ranked items was the presence of "lift" in the Otago norms (rank 18); this was entirely absent from the Battig and Montague norms, indicating its Commonwealth-specific usage.

Finally, to responses for "A musical instrument". As opposed to being ranked 7th by Battig and Montague's (1969) participants, "guitar" was the most-frequently offered example from this category amongst Otago students, perhaps reflecting its ubiquity in popular music over the years since 1969. "Fiddle", ranked highly (27th) in Battig and Montague's study, was only mentioned by one of my respondents; and "bass fiddle" (29th in the Battig and Montague study) was not mentioned at all. These aside, responses tended to be broadly similar.

Conclusion

In sum, though there were some noticeable differences in response patterns from Battig and Montague's (1969) study to mine, overall responses tended to be quite similar. This suggests that the need to collect category norms from the particular population with which my experiments were conducted was not pressing. However, it is worth noting that the categories mentioned here are those which should display limited geographical effects — parts of the

body, for instance, are the same all around the world. Other categories not yet analysed, such as cities and countries, might show very different response patterns.

Tab. A.1: Categories used.

1. A precious stone*	38. A flower*
2. A unit of time*	39. A disease*
3. A relative*	40. A tree*
4. A unit of distance*	41. A type of watercraft*
5. A metal*	42. A fish*
6. A type of reading material*	43. A snake*
7. A military title*	44. A city*
8. A four-footed animal*	45. A college or university*
9. A kind of cloth*	46. An article of furniture
10. A colour*	47. A fruit
11. A kitchen utensil*	48. A weapon
12. A building for religious services*	49. A carpenter's tool
13. A part of speech*	50. A sport
14. A part of the human body*	51. An article of clothing
15. An elected office*	52. A bird
16. A type of human dwelling*	53. A type of vehicle
17. An alcoholic beverage*	54. A toy
18. A country*	55. A vegetable
19. A crime*	56. A state or province
20. A religious official*	57. A drug
21. A substance for flavouring food*	58. A type of car
22. A type of fuel*	59. A liquid
23. An occupation or profession*	60. A thing people wear
24. A natural earth formation*	61. A thing that flies
25. A weather phenomenon*	62. A thing that makes noise
26. A part of a building*	63. A thing that is green
27. A chemical element*	64. A thing made of wood
28. A musical instrument*	65. A thing taken from a burning home
29. A kind of money*	66. A gardener's tool
30. A type of music*	67. A herb
31. A non-alcoholic beverage*	68. A thing found at the beach
32. A science*	69. A thing found in a garage
33. A type of dance*	70. A thing found in a house
34. A type of footwear*	71. A language
35. An insect*	72. A derogatory term
36. A female's first name*	73. An emotion
37. A male's first name*	74. An exclamation

Tab. A.2: Responses for “A precious stone”.

Rank	Response	Frequency
1.	diamond	200
2.	ruby	170
3.	sapphire	138
4.	emerald	118
5.	amethyst	55
6.	gem	49
7.	topaz	48
8.	opal	46
9.	crystal	40
10.	gold	34
11.	greenstone	29
	jade	29
13.	pearl	28
14.	garnet	27
15.	quartz	25
16.	silver	21
17.	onyx	13
18.	amber	9
	jewel	9
	turquoise	9
21.	peridot	8
22.	gemstone	6
	lapis lazuli	6
	moonstone	6

25. marble, pounamu (5)
 27. hematite, obsidian (4)
 29. aquamarine, coal, granite, platinum, rock (3)
 34. black diamond, cubic zirconia, expensive, graphite, green, jasper, pink diamond, rare, ring, tiger eye, white gold (2)
 45. africa, amazonite, aventurine, birth stone, blue lace agate, blue stone, bronze, citrite, copper, extravagant, glass, gold nugget, gold stone, iron, limestone, malachite, mineral, money, oamaru stone, paua, pet rock, red diamond, red stone, rose quartz, royal, slate, sodalite, special, sunstone, tanzanite, tiger stone, white stone, wishing stone, yazmine, yellow diamond (1)

Tab. A.3: Responses for “A metal”.

Rank	Response	Frequency
1.	silver	146
2.	gold	141
3.	iron	139
4.	copper	137
5.	steel	131
6.	aluminium	123
7.	tin	53
8.	magnesium	50
9.	bronze	40
10.	zinc	38
11.	titanium	33
12.	platinum	32
13.	lead	30
14.	brass	24
15.	nickel	19
16.	stainless steel	16
17.	mercury	15
18.	lithium	8
	sodium	8
20.	cobalt	7
	white gold	7
22.	alloy	6
	calcium	6

- 24. potassium, tungsten (5)
- 26. chrome (4)
- 27. boron, carbon (3)
- 29. adamantium, fluorine, hard, manganese, selenium, solid, sterling silver, uranium (2)
- 37. alloyed steel, argon, armour, army, barium, battle, berillium, cast iron, coins, concrete, conductor, coverings, dagger, electricity, expensive, fibreglass, good conductor, grey, helium, hydrogen, iron ore, knife, medal, metallic, mt, neodinium, neon, nitrogen, osmium, periodic table, pewter, pot, protection, quartz, rare, roofing, rose gold, science, shiny, silicon, strong, sulphate, sword, timber, tinfoil, useful, valuable, viking, wire (1)

Tab. A.4: Responses for “A military title”.

Rank	Response	Frequency
1.	sergeant	138
2.	lieutenant	123
3.	captain	78
4.	general	74
5.	commander	69
6.	officer	68
7.	colonel	58
	private	58
9.	corporal	55
10.	soldier	52
11.	major	50
12.	sir	30
13.	chief	20
14.	cadet	16
15.	admiral	10

16. army, brigadier, comrade (5)

19. brigadier general, commanding officer, senior (4)

22. commander in chief, marine, navy, recruit, staff sergeant, superintendant (3)

28. air force, boss, constable, corps, doctor, governor, grunt, lance corporal, lieutenant colonel, major general, marshall, medic, midshipman, navy seal, petty officer, pilot, private first class, rank, seaman, sergeant major, war, warrant officer, warrior (2)

51. 2 I.C., admiral of the marine, admiral one star, admiral two star, attorney general, battle, boatswain, brave, brigade, camp, captain america, cardinal, chief petty officer, colonel-major, commandant, commando, commodore, dangerous, deputy, deputy captain, engineer, executive, first officer, first officer in command, first-lieutenant, footsoldier, guard, head, inspector, intendant, junior, leader, lieutenant commander, lord, loud, mam, marine unit, master sergeant, militant, military personnel, miss, missionaries, muscles, navy diver, navy officer, nco, officer in command, pathfinder, pilot officer, rating, runt, sailor, SAS, scout, second class captain, second in command, sheriff, shout, skipper, squad leader, sub lieutenant, superior, supervisor, tactician, territorials, trainee, warden, weapons, wing commander (1)

Tab. A.5: Responses for “A four-footed animal”.

Rank	Response	Frequency
1.	dog	217
2.	cat	204
3.	lion	143
4.	horse	128
5.	cow	122
6.	tiger	116
7.	sheep	112
8.	elephant	95
9.	giraffe	90
10.	mouse	60
11.	pig	55
12.	rat	51
13.	zebra	50
14.	bear	48
15.	goat	47
16.	cheetah	39
17.	deer	36
	donkey	36
19.	hippo	31
20.	rhino	30
21.	rabbit	26
22.	leopard	25
23.	panther	18
	wolf	18
25.	lizard	14
26.	hyena	12
	llama	12
28.	buffalo	11
	bull	11
	camel	11
	hamster	11
	monkey	11
33.	guinea pig	10
	turtle	10
35.	alpaca	9

	antelope	9
	jaguar	9
38.	fox	8
	gazelle	8
40.	possum	7
41.	dingo	6
	lamb	6
	moose	6
	polar bear	6

45. crocodile, ferret, frog, lemur, meerkat, stoat (5)
51. badger, elk, hare (4)
54. alligator, beaver, chinchilla, gerbil, mule, otter, panda, sloth, toad, weasel (3)
64. anteater, chimp, gecko, gorilla, hedgehog, jackal, liger, lynx, ocelot, ox, platypus, pony, puma, shrew, springbok, squirrel, stag, wildebeest, yak (2)
83. aleph, animals with feet??, armadillo, ass, baby horse, baby rhino, bandicoot, bird, boar, bobcat, calf, coyote, dolphin, dwarven donkey, farm, great dane, grizzly bear, husky, impala, kudu, marmoset, miniature pony, moa, mongoose, naked mole rat, quadrupeds, reindeer, shark, skink, slow loris, snake, snow leopard, sow, tortoise, ungulates, wapiti, wombat (1)

Tab. A.6: Responses for “A kitchen utensil”.

Rank	Response	Frequency
1.	knife	211
2.	spoon	194
3.	fork	193
4.	spatula	123
5.	pot	60
6.	wooden spoon	57
7.	whisk	55
8.	pan	50
9.	tongs	44
10.	plate	40
11.	bowl	36
12.	ladle	35
13.	grater	31
14.	can opener	29
	sieve	29
16.	chopping board	27
	fish slice	27
18.	peeler	23
	teaspoon	23
20.	cup	22
21.	frying pan	20
22.	measuring cup	17
23.	cutting board	15
24.	beater	13
25.	egg beater	12
26.	cheese grater	11
	masher	11
	mixer	11
	rolling pin	11
30.	blender	10
	chopsticks	10
32.	measuring spoon	9
	oven	9
	soup spoon	9
35.	scissors	8

	serving spoon	8
37.	jug	7
	stove	7
	wok	7
40.	butter knife	6
	potato masher	6
	strainer	6

43. bread knife, garlic crusher, kettle, microwave, saucepan, sink, tablespoon (5)
50. carving knife, fridge, glass, mixing bowl, potato peeler, soup ladle, spork, steak knife (4)
58. brush, butcher's knife, colander, dessert spoon, dish, flipper, frypan, juicer, mug, pizza cutter, saucer, sharp knife, slicer, tin opener (3) 72. cheese cutter, dishwasher, egg slicer, electric mixer, garlic press, grinder, measuring jug, oven mitt, oven tray, platter, prong, rice cooker, rubber spatula, slotted spoon, stirrer, toaster, tray (2)
89. baking, baking tin, baking tray, beaker, bench, bench cleaner, big knife, board, bottle opener, butter, cake fork, cake knife, cake slicer, cake tin, chair, cheese slicer, chopping block, cleaver, cloth, container, cooking pan, corer, corkscrew, crock pot, cup measure, cutlery, dicer, dishwashing brush, egg flip, egg slide, electric beater, fish flipper, fondue fork, food, food processor, freezer, frying pot, funnel, george foreman, griddle, groom, half cup measure, herb chopper, island, kitchen knife, knife sharpener, large spoon, lid, magic bullet, mandolin, measuring cylinder, meat knife, meat smasher thing, mincer, muffin tin, muffin tray, orange juicer, oven dish, pallet knife, paring knife, pasta scoop, pastry brush, pie dish, plastic bowls, pottle dish, pressure cooker, roller, salad server, salad spinner, sandwich maker, sandwich press, scale, scraper, skew, skewer, skillet, slow cooker, smasher, spatula, sponge, steamer, steelo, stir fry pan, sushi mat, table, tea towel, teapot, towel, vegetable peeler, whizz, zester (1)

Tab. A.7: Responses for “A part of the human body”.

Rank	Response	Frequency
1.	leg	188
2.	arm	181
3.	finger	152
4.	head	148
5.	eye	144
6.	toe	141
7.	foot	137
8.	nose	122
9.	ear	110
10.	hand	107
11.	mouth	96
12.	heart	89
13.	brain	83
	stomach	83
15.	hair	73
16.	knee	56
17.	lung	55
18.	liver	51
19.	shoulder	44
20.	tongue	40
21.	neck	36
	teeth	36
	torso	36
24.	thigh	32
25.	nail	29
26.	bones	28
27.	face	27
28.	chest	26
	kidney	26
30.	ankle	25
	elbow	25
32.	intestines	23
	lips	23
34.	penis	20
35.	muscle	17

36.	fingernail	16
	hip	16
	wrist	16
39.	spleen	15
40.	back	14
	vagina	14
42.	skin	13
43.	blood	12
44.	breast	11
	pancreas	11
	throat	11
47.	genitals	10
48.	bum	9
	skull	9
	veins	9
51.	calf	8
	forearm	8
	shin	8
54.	tummy	7
55.	butt	6
	chin	6
	spine	6
	thumb	6
	toenails	6

60. cheek, ligament, nipple (5)

63. anus, belly button, bladder, eyebrow, limb, patella, ribs (4)

70. abdomen, appendix, arteries, bicep, boobs, bottom, buttocks, eyelashes, femur, gallbladder, heel, humerus, knuckle, nerves, oesophagus, organs, radius, small intestine, tendons, tricep, ulna, waist (3)

92. abdominal, armpit, blood vessels, body, capillaries, cartilage, cells, colon, jaw, metatarsal, phalanges, spinal cord, sternum, tissue, trunk (2)

107. abs, amygdala, belly, cerebellum, choroid plexus, clavicle, clitoris, collarbone, core, deltoid, dick, digestive system, digestive tract, earlobes, eyeball, forehead, gi tract, harn, joints, jugular vein, kneecap, large intestines, larynx, left arm, left leg, mind, nasal cavity, nephron, nostril, pancreatic islets, pelvis, pore, prostate, rectum, right arm, right leg, saliva, scapula, scrotum, shoulderblades, skeleton, tarsal, tibula, tonsils, trachea, truck, ureter, urethra, voice box (1)

Tab. A.8: Responses for “An elected office”.

Rank	Response	Frequency
1.	president	118
2.	prime minister	100
3.	mayor	47
4.	parliament	45
5.	labour	27
6.	green	25
	national	25
8.	government	22
9.	MP	19
10.	vice president	16
11.	maori	13
12.	act	11
	cabinet	11
	councillor	11
15.	council	10
	governor	10
	member of parliament	10
	minister	10
19.	CEO	8
	senator	8
21.	OUSA	7
	secretary	7
	treasurer	7
24.	chairman	6
	chairperson	6

26. congressman, white house (5)

28. board of trustees, boss, city council, committee, executive, head girl, manager, presidential (4)

36. beehive, chancellor, class rep, elections, governor general, head boy, house of representatives, judge, major, minister of foreign affairs, PM, principal, representative, senate, student president (3)

51. board, cabinet minister, conservative, corporal, council member, D.A., dean, head of department, health minister, john key, management, minister of education, minister of finance, nz 1st, party, politics, pope, school board, secretary of state, sheriff, student rep, vice chancellor, vote, voting (2)

75. 11 downing street, alliance, ambassador, army, army official, assistant, bank director's office, barisan nasional, board member, board of members, books, boss' office, business, canberra, captain military, CEO of company, CEO of freezing works, chief executive, chief's office, choice, chris tremain, christian coalition, church committee, city councillor, committee leader, communist party, community centre, computer, dalai lama, democracy, democratic, democrats, department of labour, deputy, deputy prime minister, desk, diplomat, director, director on board, education minister, election booth, electorate, environment, executive officer, executive position, files, finance minister, general, general office, government official, gp, head master, head of a party, head of board of trustees, head of company, head of school, head of the board, head office, head surgeon, head teacher, health board, helen clark, house of cabinet, house of commons, in charge, international minister, jury, kim jong-il, leader of a party, legalise marijuana, life member, local body, local electorate, local government, local government building, manukau, military

position, minister of agriculture, minister of external affairs, minister of health, minister of inland, minister of maori party, minister of parliament, minister of transport, money, mt eden, navy official, news room, north dunedin, north shore, nzi, official, ohariu-belmont, opposition leader, OUSA president, parliament representative, parliamentary, party leader, pastor, phone, political, politician, power, president of the united states, presidentship, priest, principal's office, promotion, pta, queen, RBA board members, region, regional council, regional office, religious, republican, responsibility, right winged, room, royal house, ruling party, runanga, school council, school president, secretary general, secretary general of the UN, sergeant, smart, south dunedin, speaker, specific membership clubs, sports captain, state, student committees, student union representative, students club, supreme court justice, teacher union, teacher's office, the state government, town hall, union leader, union representative, united nations, various ministers, vice chairman, vice secretary, vice treasurer, wellington cbd (1)

Tab. A.9: Responses for “A religious official”.

Rank	Response	Frequency
1.	priest	182
2.	pope	166
3.	minister	76
4.	nun	63
5.	pastor	54
6.	bishop	52
7.	monk	43
8.	rabbi	40
9.	reverend	35
10.	father	33
11.	cardinal	26
12.	god	26
13.	chaplain	24
14.	jesus	17
15.	archbishop	14
16.	buddha	14
17.	vicar	13
18.	deacon	10
19.	brother	9
20.	clergyman	8
21.	elder	7
22.	preacher	7
23.	buddhist	6
24.	christian	6
25.	dalai lama	6
26.	prophet	6
27.	saint	6
28.	sister	6

- 29. imam, shaman (5)
- 31. muslim (4)
- 32. allah, baptist, catholic, clergy, cleric, disciple, guru (3)
- 39. buddhism, celebrant, church, dean, hindu, hinduism, islam, missionary, mullah, sheikh (2)
- 49. abbess, abbot, altar boy, altar server, alter boy or girl, arch deacon, B . . . , camerlengo, catholicism, chamberlain, chancellor, chinese, christianity, church goers, church leader, church minister, church pastor, communion to sick, cross, cult leader, cupid, daoist, deaconess, evangelist, exec pastor, exorcist, formal, friar, gandhi, godly, head, high priest, human, humble, in charge, indian, inquisitor, john, judaism, king of england, krishna, lay leader, leader, leader of the church, levite, liman, listen, lord, madman, magistrate, malay, matthew, matua, messiah, mohammed, morman, mosque, mother, mother abbess, mother theresa, ordained, padre, parent, parish council, people, pew boy, police, power, pray, priestess, pundit, sacreston, sacristan, sermon, sheik, shepherd, spinster, spiritual leader, teacher, vishnu, witch doctor, youth pastor, zeus (1)

Tab. A.10: Responses for “A part of a building”.

Rank	Response	Frequency
1.	window	171
2.	door	161
3.	roof	150
	wall	150
5.	floor	116
6.	ceiling	83
7.	stairs	59
8.	room	49
9.	elevator	40
10.	foundation	36
11.	bathroom	30
12.	basement	26
13.	office	25
14.	kitchen	23
15.	hallway	19
16.	bricks	18
	carpet	18
18.	lift	17
19.	attic	16
	bedroom	16
21.	chimney	15
22.	wood	14
23.	insulation	12
24.	balcony	10
	concrete	10
	frame	10
	ground floor	10
	lights	10
	lounge	10
	reception	10
	staircase	10
32.	entrance	9
	levels	9
34.	corridor	7
	lobby	7

	steps	7
37.	curtains	6
	plumbing	6
	skirting	6
	storey	6
	support beams	6
	tiles	6

43. cement, corner, entrance way, foyer, garage, glass, interior, paint, pillar (5)

52. beam, deck, door frame, electrical wiring, exit, fire escape, gutter, heating, stairwell, toilet, verandah, window sill (4)

64. carpark, dining room, door knob, eaves, electricity, exterior, fire exit, first floor, floors, framework, front door, gib, nails, pipes, rooftop (3)

79. annexe, atrium, awning, bottom floor, cellar, closet, door bell, door handle, downstairs, en suite, façade, flooring, furniture, garden, ground, hall, hinges, house, joist, laundry, ledge, lintel, living room, loft, metal, mezzanine, parlor, piles, porch, railing, roofing, scaffolding, screw, signs, storeroom, structure, struts, studs, third floor, top, top floor, wallpaper, weatherboards, whiteboard, window frame, wing, wires (2)

126. air conditioning, apex, auditorium, back door, base, batt, bearer, bracing, car, carport, chairs, cladding, coffee room, columns, concrete base, conservatory, contents, different floors, different levels, doorstep, doorways, drains, drapes, drawing room, dungeon, east wall, electrical, emergency exit, entry and exits, every other room, fifth floor, fixtures, flag pole, floor pad, floor plan, fourth floor, framing, front entrance, gable, gargoyles, ground structure, groundings, heater room, high tower building, holden, infrastructure, innerwalls, innings, jib, kith, knob, lamp, land, leaks, light fittings, lino, load bearing wall, locks, main entrance, masonry, meeting room, metal structure, mirror, north wing, outer walls, overhang, plans, plaster, plugs, pools, post, power point, property, reinforcing, restroom, revolving doors, rock cote, rook, scouture, second floor, shingle, shower, shutters, side, sink, skirting board, skylight, smoke alarm, socket, south wall, spire, spouting, stage room, steel, steeple, stone, storage, structural beams, structural support, supports, table, tall, terrace, theatre, threshold, toilet seat, tower, tv room, upstairs, vent, ventilation, walk in wardrobe, water pipes, west wing, window handle, wiring, workers (1)

Tab. A.11: Responses for “A musical instrument”.

Rank	Response	Frequency
1.	guitar	205
2.	piano	192
3.	drums	163
4.	violin	160
5.	flute	151
6.	trumpet	102
7.	cello	84
8.	saxophone	77
9.	bass	70
10.	triangle	67
11.	clarinet	64
12.	keyboard	62
13.	viola	55
14.	recorder	52
	trombone	52
16.	harp	42
17.	oboe	40
18.	double bass	30
19.	ukelele	25
20.	harmonica	23
21.	bassoon	19
	voice	19
23.	tambourine	18
24.	xylophone	16
25.	electric guitar	15
26.	bass guitar	14
	cymbals	14
28.	piccolo	12
29.	tuba	11
30.	organ	10
31.	accordion	9
32.	bagpipes	8
	banjo	8
	french horn	8
35.	cornet	6

harpsichord	6
horn	6
mandolin	6
synth	6
timpani	6

- 41. maracas, shakers, sitar (4)
- 44. acoustic guitar, bongos, pipe (3)
- 47. cow bell, keytar, lute, microphone, piano accordion, pipe organ, strings, tin whistle, whistle (2)
- 56. african bon bons, african drums, bane, bass clarinet, bells, box bass, calm, chimes, chords, computer synth, computer with music program, cor anglais, didgeridoo, dj set up, dj-deck, electric bass, electric keyboard, feet, fiddle, gamelan, hands, hearing, koto, loud, metalophone, mizmah, mouth organ, music, nice, notes, orchestra, panflute, panpipes, percussion, pots and pans, shamisen, sing, snare, sound, tabla, tam tam, tenor saxophone, upright bass, viol, voice box, wind instruments, wine glasses, wood, woodpipe (1)

Tab. A.12: Responses for “A flower”.

Rank	Response	Frequency
1.	rose	204
2.	daisy	137
3.	lily	113
4.	sunflower	103
5.	daffodil	100
6.	tulip	92
7.	pansy	44
8.	poppy	38
9.	dandelion	25
	lavender	25
11.	orchid	23
12.	gerbera	20
	hydrangea	20
14.	petunia	17
	rhododendron	17
16.	violet	16
17.	carnation	15
18.	iris	14
19.	blossom	13
	buttercup	13
21.	hibiscus	12
22.	jasmine	11
23.	frangipani	10
	marigold	10
25.	bluebell	9
	chrysanthemum	9
27.	camelia	8
	cherry blossom	8
	forget-me-not	8
	peony	8
31.	lotus	7
	magnolia	7
33.	agapantha	6
	geranium	6
	petal	6

36. freesia, honeysuckle, pohutukawa (5)

39. bird of paradise, pretty, water lily (4)

42. colourful, lilac, red, stem, tigerlily (3)

47. baby’s breath, bees, clover, dahlia, foxglove, garden, hyacinth, rata, snapdragon, summer, sweet pea, viola (2)

59. alstroemeria, amber, arum lily, azalea, ball, beautiful, chinese orchid, christmas lily, clematis, corsage, cosmos, crocus, daphne, dogwood, edelweiss, english rose, expensive, fern, flame, fuchsia, gardenia, gift, girls, girly, gladioli, gorse, grass, growing, ixora, kowhai, leaf, leaves, living, lupin, morning glory, nasturtium, nice smell, olive flower, orange blossom, peony rose, perfumed, perrigold, phlox, pink, plumeria, pollen, pomegranate, pring flower, protea, purple, queen of the night, ranuncula, romantic, roots, rosemary, scent, self raising flour, shamrock, smells good, snowbell, snowdrop, snowflake, spec, sun, tiger flower, tree, valerian, venus fly trap,

water, waterleaf, wedding, yellow, yoket (1)

Tab. A.13: Responses for “A tree”.

Rank	Response	Frequency
1.	pine	121
2.	oak	115
3.	pohutakawa	58
4.	willow	58
5.	rimu	57
6.	apple	47
7.	kauri	46
8.	kowhai	43
9.	palm	39
10.	beech	35
11.	cabbage	31
12.	maple	24
13.	fern	20
	rata	20
15.	totara	19
16.	evergreen	18
17.	redwood	17
18.	christmas	15
	gum	15
	orange	15
	pear	15
22.	coconut	14
23.	cherry	13
	leaves	13
25.	birch	12
	fir	12
	macrocarpa	12
	poplar	12
	walnut	12
30.	elm	11
	tall	11
32.	lemon	10
33.	acorn	9
	eucalyptus	9
	plum	9

	silver birch	9
37.	banana	8
	branch	8
	deciduous	8
	manuka	8
41.	blossom	7
42.	bark	6
	cherry blossom	6
	chestnut	6
	green	6
	weeping willow	6

47. cedar, fruit (5)

49. big, conifer, mahogany, mango tree, paper, peach, sycamore, yew (4)

57. apricot, ash, bonsai, magnolia, matai, plant, short, shrub, small, wood (3)

67. aspen, bamboo, baobab, bird, bush, camelia, douglas fir, elder, feijoa, fig, hazelnut, kanuka, lancewood, lime, linden, old, olive, ponga, puriri, rhododendron, roots, rowan, shade, spruce, teak, trunk, umbrella tree (2)

94. acacia, albizia, algae, autumn, avocado, bald, banyan, barley, berries, black pine, brown, bushy, carree, climbing, coconut palm, coniferous, conker, cut down, diagram, elderwood, everywhere, fallen, family, fern tree, flowers, forest, garden, ginkgo, gorse, grape, grass, hibiscus, holly, intelligent, jacaranda, jungle, juniper, kahikatea, kamahi, kiwifruit, konini, korfi, lacebark, large, lavender tree, lemonwood, living, mandarin, mangrove, mill, miro, monkey, mountain, mowgli, nectarine, neem, nest, nettle, nikau, norfolk pine, outside, oxygen, photosynthesis, pineapple, pittosporum, ponderosa pine, pungu, radiata, rakau, rewa, ribbonwood, sap, sapling, saw, scrub, seed, sequoia, shelter, silk, soil, stick, stump, summer, tarzan, teatree, tree, tree house, veins, wax, whomping willow (1)

Tab. A.14: Responses for “A fish”.

Rank	Response	Frequency
1.	salmon	116
2.	shark	90
3.	tuna	87
4.	trout	83
5.	goldfish	77
6.	snapper	76
7.	blue cod	63
8.	cod	45
9.	hoki	41
10.	clown	40
11.	flounder	38
12.	herring	35
13.	red cod	30
14.	catfish	25
15.	puffer fish	21
	swordfish	21
	tarakihi	21
18.	gurnard	19
19.	eel	18
	whitebait	18
21.	rainbow trout	16
22.	sole	14
23.	kingfish	13
	mackarel	13
	whale	13
26.	angel	12
	grouper	12
	john dory	12
	nemo	12
	sardine	12
31.	bass	11
32.	barracuda	10
	marlin	10
34.	dogfish	9
35.	rainbow	8

	sting ray	8
37.	brown trout	7
	flatfish	7
	kahawai	7
	starfish	7
41.	orange roughy	6
	sea	6
	spotty	6
	trevally	6

45. guppy, lemon, lobster, pike, piranha (5)
50. blow fish, carp, dolphin, elephant, flying, jellyfish , minnow, ocean, seafood, squid, water (4)
61. angler fish, butterfish, dory, fishing, haddock, hapuka, lionfish, moki, octopus, tropical, yellow fin, yellow tail (3)
73. battered, black moor, blue fin tuna, brill, fighting, gill, great white shark, kingklip, koi, lake, mud fish, mullet, parrotfish, paua, prawn, rock cod, scales, seahorse, shrimp, sprat, stonefish, sucker, swimming, swims, tigerfish, warehou, zebrafish (2)
100. ahi, ale, anchovies, anonymon fish, aqua, bacalao, banded wrass, blue, bream, breathing, brim, brown tailed, caper, celestial, chillad, chimaera, clam, clopper, cold, colour, colourful, comet, conger eel, crayfish, cuttle, deep sea, dorade, dover, eat, fins, fish tank, flake, flathead, food, gillet, giwan ruwa, globe eye, halibut, hammer head, hook, jaws, katon kai kifi, kipper, krill, kyle, leatherrish, lionhead, mahi mahi, mexican walking fish, monk, mussel, neon tetra, one sided, opa, orange, oyster, perch, pet, pigfish, plankton, poisonous, red herring, red snapper, river, sailing, salty, scarpy, schools, sea animal, sea urchin, silver, slime fish, slimy, smelt, snook, sun, tasty, tetra, tigershark, vampire, wahoo, whaleshark, white, white fin, yellow cod, yellow jacket, yuppie (1)

Tab. A.15: Responses for “A weapon”.

Rank	Response	Frequency
1.	gun	173
2.	knife	170
3.	sword	86
4.	bomb	48
5.	machete	31
	rifle	31
7.	spear	25
8.	bat	24
9.	dagger	23
	pistol	23
11.	bow and arrow	22
12.	hammer	21
13.	baseball bat	19
14.	fist	18
	shotgun	18
16.	axe	17
17.	stick	16
18.	nunchucks	15
	rope	15
20.	club	14
	grenade	14
22.	baton	13
23.	machine gun	12
	slingshot	12
25.	missile	11
26.	bazooka	10
27.	cannon	9
	grenade	9
	taser	9
30.	arrow	8
	crowbar	8
	revolver	8
	tank	8
34.	bayonet	7
	blade	7

	sniper	7
37.	chainsaw	6
	crossbow	6
	pepper spray	6
	rock	6

41. mace, pocket knife, scissors, spanner, stone (5)
46. bullet, car, fork, glass, hand gun, nuclear bomb, rocket launcher, spade, taiaha (4)
55. atomic bomb, bow, brass knuckles, candlestick, fire, flame thrower, golf club, hand, kitchen knife, nuke, tool (3)
66. 2 by 4, air rifle, ak47, blood, boomerang, broken bottle, catapult, chain, cricket bat, feet, garrote, gas bombs, glass bottle, katana, knuckle dusters, lance, lead pipe, legs, metal pole, nail, ninja stars, pen, person, plane, poison, razor, rpg, samurai sword, shield, short sword, shuriken, sledge hammer, whip, words (2)
100. 12 gauge, 20 gauge, A-bomb, air gun, AK, any object, anything metal and sharp, apache, arch, army, automatic rifle, ballista, banana, bar, bare hands, battering ram, BB Gun, belt, biological weapon, blackmail, blowdarts, body, book, brick, broad sword, bullying, cane, charm, chemical weapon, clash, copper wire, corkscrew, cutlass, dart, dirke, drugs, dynamite, electric shock, emotions, f12, fight, firearm, flick knife, food, gatling gun, gladius, glock, halberd, harpoon, heavy objects, heavy rock, helicopter, hunting knife, keys, killing, knitting needle, language, life, lightsaber, manipulation, martial arts, match, metal rod, missile launcher, mouth, MP5, musket, my body, my mind, nail clippers, needles, nuclear missile, nuclear weapon, p90, patu, penknife, piano wire, pike, pipe, plank of wood, plate, pole, potato gun, power, rail gun, razorblade, rocket, saber, safety pin, SAM, saw, screw driver, scythe, sharp detrimental object, sharp objects, shiv, shoe, shovel, shrapnel, slug gun, smoke grenade, speech, staff, stiletto, stun gun, suicide bomber, switchblade, teeth, tomahawk, torch, torpedo, torture, uzi, voice, war, warhead, wooden bat, writing (1)

Tab. A.16: Responses for “A bird”.

Rank	Response	Frequency
1.	sparrow	108
2.	seagull	93
3.	kiwi	79
4.	pigeon	74
5.	eagle	70
6.	tui	68
7.	parrot	57
8.	fantail	56
9.	duck	47
	hawk	47
11.	blackbird	44
12.	albatross	41
13.	kea	31
14.	penguin	27
15.	chicken	26
16.	budgie	24
	moa	24
	owl	24
19.	dove	22
	swan	22
21.	peacock	21
22.	kakapo	20
23.	ostrich	19
	pukeko	19
25.	finch	18
	magpie	18
	robin	18
28.	emu	17
	goose	17
30.	bellbird	16
31.	canary	15
32.	crow	14
33.	bluebird	12
	flamingo	12
35.	heron	11

	kaka	11
	pelican	11
	woodpecker	11
39.	cockatoo	10
	falcon	10
	hummingbird	10
	parakeet	10
	starling	10
	swallow	10
45.	weka	9
46.	hen	8
	kingfisher	8
	thrush	8
49.	cockatiel	7
	wood pigeon	7
51.	lovebird	6
	morpork	6
	takahe	6
	waxeye	6

55. raven, rooster, stork (5)
58. gull, kereru, saddleback, shag, toucan (4)
63. flight, oystercatcher, pheasant, quail, rifleman (3)
68. beak, blue tit, crane, fly, gannet, godwit, ibis, kakariki, kokako, kookaburra, macaw, mocking bird, mynah, piwakawaka, plover, rainbow lorikeet, tern, turkey, wings, wren (2)
88. animal, avian, bald eagle, bird of paradise, black back woodpecker, black robin, blue jay, chaffinch, cormorant, dodo, feathers, female, freedom, girl, goshawk, grey warbler, guinea fowl, haast eagle, hammock bird, harrier, huia, jay, kahu, kite, lark, lewis woodpecker, lorikeet, lute bird, lyrebird, mallard duck, mollymawk, nest, new zealand hawk, noise, oriole, osprey, paraki, phoenix, red headed sapsucker, redbird woodpecker, redevye, rook, saw, sea bird, seal, silvereve, sing, singer, small, sooty shearwater, starlet, swift, trees, tuatara, vulture, white headed woodpecker (1)

Tab. A.17: Responses for “A vegetable”.

Rank	Response	Frequency
1.	carrot	171
2.	potato	143
3.	broccoli	123
4.	lettuce	94
5.	peas	93
6.	kumara	71
7.	cabbage	67
8.	beans	64
9.	pumpkin	62
10.	cauliflower	59
11.	onion	54
12.	tomato	52
13.	cucumber	47
14.	corn	44
	spinach	44
16.	leek	37
17.	parsnip	30
18.	capsicum	27
19.	celery	22
20.	beetroot	21
	brussels sprout	21
22.	mushroom	20
23.	yam	19
24.	pepper	16
	silverbeet	16
	zucchini	16
27.	courgette	14
28.	eggplant	13
	turnip	13
30.	swede	12
31.	avocado	11
32.	asparagus	9
33.	sprout	7

- 34. radish (5)
- 35. aubergine, sprouts (4)
- 37. garlic, marrow, squash (3)
- 40. baby marrow, bak choy, beet, bok choy, broad beans, chickpeas, chilli, coriander, red onion, spring onion (2)
- 50. baobab leaves, basil, bean sprouts, butternut, causet, coalso, edamame, gherkin, green beans, green pepper, olives, pak choi, pickle, red cabbage, red pepper, rhubarb, rocket, sweet potato, watercress, wheat (1)