DOCTOR OF PHILOSOPHY

Semantic categorisation of actions and objects in monolingual and bilingual children

Tomos, Hedd

Award date:
2010

Awarding institution:
Bangor University

Semantic categorisation of actions and objects in monolingual and bilingual children

Hedd Tomos

submitted for the degree of Doctor of Philosophy

Bangor University

School of Psychology

September 2010

Acknowledgements

I would to take this opportunity to sincerely thank my supervisers, Prof. Virginia C.M. Gathercole and Dr Enlli Môn Thomas for all their help and support during this research study. I would also like to thank my colleagues in the School of Psychology Bangor University, particularly Hannah Morrish, for their friendship during my time at Bangor. I would also like to thank the chair of my supervisory committee, Dr Charles Leek, for his help and support in the completion of this study. I am grateful to the Mantais programme, especially Dr Ioan Matthews and Dr Dylan Phillips, for their assistance. Finally I would like to thank my wife Elizabeth for her support, patience and tolerance over the last four years – diolch o'r galon Lizzy.

Table of Contents

Appendices

Tables

List of Figures

Abstract

This study investigated categorisation processes by Welsh/English bilingual and monolingual English children. It focused specifically on the effects of complexity for group membership of a word group, and the amount of input received for both languages, influenced by the home language environment and developmental factors. The study adopted the constructivist model view that bilingual Welsh/ English children exhibit behaviour consistent with processing language using two separate sets of categorisation processes, a semantic system and a conceptual space in which the child develops cognitive abilities to make sense of the world. Children develop both languages in parallel to monolinguals but are delayed in their acquisition of lexical items due to reduced input.

The research consisted of four studies, with the main study involving two hundred and twenty eight monolingual and bilingual (Welsh/English) children aged 5-11 completing a picture naming task, either in Welsh or English, with referents presented in four word categories. The word categories were classical, homonyms, radial taxonomic and radial thematic with words always showing a 'wider' range in one language and 'narrower' in the other. The remaining three studies consisted of an analysis of vocabulary tests conducted using the BPVS and Prawf Geirfa Cymraeg; the use of a memory task to investigate the research question; and an analysis of errors in selecting related items in the naming task.

The lack of taxonomic bias across age groups and category groups was a notable finding as this supported the view that bilingual children rely more on cognitive skills to categorise rather than structured linguistic processes. In addition, there was strong evidence that children had greater difficulty with processing items from the classical word group. This

was seen in the poorer performance of bilingual children whose dominant language was English when tested in Welsh. A similar performance was found for the bilingual group whose dominant language was Welsh when tested in English across the three age groups. With children aged 5 to 7, 7 to 9 and 9 to 11 not performing as well as the other three home language groups on items from the classical words group.

The results of the study supports the constructivist theory model (Gathercole, 2007) and dual coding processing hypothesis (Paivio & Desrochers, 1980) that bilingual individuals use two separate language systems and a shared imagery system for categorising words. The results also concur with the assumption that input has an effect on improved acquisition and the better processing of complex language structures. This study marks an initial attempt to investigate semantic categorisation by bilingual Welsh/English children and would benefit from further research into the findings utilising supplementary testing methods.

Chapter 1. Introduction to the Study

Wales is unusual in the context of the United Kingdom as being a nation of three million people where around 20% of the indigenous population is naturally bilingual. The goals of this research study developed from the lack of a corpus of evidence from previous studies into the nature of word acquisition by Welsh bilingual children, researched through a study of semantic systems and associated word categorisation. The scientific study of bilingual language acquisition is a dynamic and rich research area with significant work being undertaken on a world-wide basis. However a detailed understanding of how bilingual Welsh/English speakers categorise words is still emerging (Gathercole ed, 2007). This study therefore aimed to add to the evidence base through a comparison of the experience of Welsh speaking children with their monolingual peers, and a more global assessment of the Welsh experience with studies carried out on other languages.

The interpretation of such semantic categorisation by monolinguals and bilinguals presents several generic challenges for researchers, not least in identifying key influences and interdependencies in an often complex developmental environment. For example, when the child receives input from parents speaking different languages or is exposed to a different language at school to the home language, what are the influential factors in ensuring that the individual child correctly categorises the differing semantic input and enhances his/her linguistic knowledge?

In investigating semantic categorisation as a feature of language acquisition by children, the study specifically focused on the acquisition of Welsh and English by primary school children. As part of the indigenous Celtic family of languages, Welsh is not spoken

uniformly across the country with higher percentages of first language Welsh speakers in north and west Wales. Children in these areas of Wales obtain an increasing exposure to Welsh through Welsh medium primary schools. Although there are interesting issues relating to influential social factors on the acquisition of Welsh by primary school children (Gathercole, ed. 2007), the design of this study chose to specifically investigate semantic categorisation represented by the acquisition of words relating to action and objects.

In attempting to better understand the cognitive processes involved, the study adopted a constructivist approach (Gathercole, 2007). This model assumes that in order to progress in language acquisition and the associated increased knowledge and understanding of the world, children rely on two separate features, a semantic system and associated cognitive knowledge which together enable children to successfully develop word categories.

As will be outlined later, the design included four tasks to assess whether input and language dominance have an effect on categorisation namely - a categorisation task, a taxonomic and thematic choices task, a memory task and a vocabulary task. The study design recognised the challenges of definitively asserting which factors ultimately influence the acquisition of novel concepts and words by children. It can be argued that similarities in categorisation across languages are accounted for by common developmental experiences where children encountered similar contexts. However more recent cross-linguistic and multi-lingual studies suggest that semantic categorisation by language learners of all ages requires complex processes to succeed (Bowerman & Choi, 2003).

1.1 The research goals

Based on an assessment of the results of previous research which are outlined in Chapters 2 to 5, and the expectations of the constructivist model, this study made the following predictions regarding the expected outcomes.

1.1.1 Word groups

It was expected that the performance of all of the bilinguals should be similar to each other and to the monolingual group with the homonym words (i.e. two words being identical in their phonic representation but having no relation in terms of meaning, e.g. lag which can mean to fall behind or covering a boiler). This is due to there being no conceptual link between the two words, only a homophonic relationship. Therefore there would be no reason to expect a carryover from one language to the other. In the English test condition, the bilinguals were expected to have the lowest performance in the classical word group, with the Only Welsh at Home (OWH) group having the greatest difficulty (therefore performing least like the English Monolingual Group). Conversely, the OWH group would perform best when the test condition was Welsh.

The performance of children on radial categories was expected to be somewhere between classical and homonym categories. This is due to the nature of radial categories as having members which are conceptually different but linked in the language concerned. With this in mind, performance would better in the radial taxonomic word groups compared to words in the radial thematic word group.

1.1.2 Participants

It was anticipated that monolinguals in general should perform better than the other bilingual groups due to the lack of interference from the other languages (in this case Welsh) in their decision making. Performance of bilinguals was expected to be better depending on their dominant language, for example, the Only English at Home (OEH) bilinguals would be expected to perform better than the other two bilingual groups. (i.e. closer to the performance of the monolingual group) when tested in English. It was expected that the Only Welsh at Home (OWH) groups should generally perform better than the Welsh and English at Home (WEH) bilinguals and the OEH groups when tested in Welsh. It is possible that bilinguals may overextend their choice in the narrower language as a result of the influence of the category in the other (wider) language being wider. The influence of the wider language on the narrower language may be more likely to be present in instances when it is the child's dominant language. For example, OEH group children tested in Welsh, are more likely to choose both targets when Welsh is narrower. From the other perspective, bilinguals may under extend their choice when the text item is wider in the language tested as a result of the influence of the category in the other (narrower) language being narrower. The performance of WEH group children could vary as a result of not having a dominant language.

1.1.3 Width

It could be argued that it would be easier to broaden a narrower category in going from a narrow category in the dominant language to a wider category in the non-dominant language, than to go from a broad category in the dominant language to a narrower category

in the non-dominant language. In going from a narrower category in the dominant language, the child will only need to receive exposure to the expanded use in the non-dominant language to understand the wider use of the referent in the non dominant language. Conversely when going from a broader category in the dominant language to a narrow category in the non-dominant language, the child would specifically need to be told or to notice that they were mis-applying the broader category in the dominant language when incorrectly using it in the non-dominant language. In addition, another possibility for preferences by children towards narrower categories may be as a result of the adoption of a cautious approach to word meaning. The child may prefer a narrow application of the use of the word until more input is received and assimilated indicating that broader applications is warranted.

Another possibility draws on the evidence from studies highlighting children's preferences for separately naming objects. Clark's (1993) principle of contrast proposed that children assume that no two words will possess the same meaning. It will be expected that it would be easier for children to understand narrower categories in the non-dominant language than it would be from going from a narrow category in the dominant language to a broad category in the non-dominant language. This prediction would also support Markman's (1991) theory of mutual exclusivity which states that children assume that objects can only have one name.

1.1.4 Age group

It was predicted that any effects of home language group would decrease with age with the degree of exposure from the two languages increasing (especially for the OEH group and the WEH group in the input of Welsh) over time. Therefore difference across language background groups may be neutralised by the time the children reach the oldest group.

1.2 Thesis structure

The thesis is structured as follows. Chapter 1 provides the rationale for the research study and outlines the research goals. Chapter 2-5 include a review of the literature relevant to the study. Chapter 2 examines the nature of word categories and the principles which may relate to the mapping of conceptual knowledge to words. Chapter 3 reports on factors relating to the acquisition of word meaning and examines the evidence from cross-linguistic studies. Chapter 4 investigates evidence to explain how bilinguals organise categories and the possible relationships between their two languages. Chapter 5 examines the acquisition of words in a bilingual context. Chapter 6 comments on the study goals, the lack of previous research relating to the bilingual Welsh context, and reports the findings of the language background, vocabulary tests, and the questionnaire data. Chapter 7 reports on the findings of the study's categorisation task. Chapter 8 reports the findings of the taxonomic and thematic choices task. Chapter 9 reports the findings of the memory task. Chapter 10 summarises and discusses the main findings of the study and its conclusions.

Chapter 2. The Nature of Categories

2.1 Introduction

This Chapter outlines how differing approaches to explain the process of semantic categorisation raise issues and further questions relating to how children map the real world onto their internal category systems. The chapter examines relevant research into children's categorisation behaviour in order to better understand the factors involved in how children form and acquire categories. Although the study adopts the view of the constructivist model, this examination provides an opportunity to review other viewpoints.

This section of the literature examines the nature of category membership for objects and the differing processes that are required for different types of categories. This is a complex set of processes which potentially require a child to look beyond simply identifying a universal set of visual features for an object. The current study identifies four different word category groups which have distinct requirements for membership. These groups differ in the emphasis on features for membership, with some groups requiring essential features, and others having more metaphoric or cultural references. Within this context, there is a need to understand the process used by children for the placing of novel objects into categories using pre-existing organisation of concepts in memory. The different category types and theories about group membership are discussed in the context of the requirements of the main tasks for the study. Specifically, the main interest relates to the nature of the formation of those categories which are more difficult for a child to incorporate an object into the correct group.

A category can be defined as a group of objects both in the natural world and as

artificial objects (Murphy, 2002). It is not difficult to grasp that without the cognitive process of categorisation it would be difficult for children to make sense of the world as they would constantly have to define new objects as they encounter them. Through the categorisation process, it may only be necessary to know the main properties of the category to infer whether the novel object is part of a category already identified by the child (Labov,1973). However this simplistic view raises immediate questions of how categories are structured by children to successfully interpret novel objects and attempt to add new words to their lexicon.

The historical view of the nature of word categories and how they are formed developed from Aristotelian theories of the organisation of knowledge. From this standpoint, individual categories are required to have defining features which are necessary, that is, the features must be present, and sufficient and no other elements are required to be present (Medin,1989). The model holds that there will be no specific feature that would be more important than any other as all would have to be present in the structure of the category (Murphy & Medin,1985). However this world view of categories has not been sufficient in itself to definitively explain the complex processes which enable children to categorise objects and acquire words. In addition to the historical view, other approaches will need to be reviewed for the purposes of the study.

The cognitive perspective (Piaget,1954) minimised the role of linguistic elements in the identification of categories. The idea of linguistic relativity (Whorf, 1956) proposed that language structure influenced a world view so that speakers of different languages had differing experiences. However Lakoff (1987) identified aspects of language such as metaphors which he argued demonstrated that speakers of different languages think in

different ways. How should linguistic and cognitive input therefore be viewed if we are attempting to explain children's categorisation?

An approach which attempted to reconcile the linguistic and the cognitive input required to support the early acquisition of words was developed by Gopnik and Meltzoff (1986). Their longitudinal observation of young children identified that an understanding of concepts and words developed together. For example, at the age where the child became aware of the permanent nature of objects, words such as *gone* were used. This suggests that young children encode concepts and acquire the relevant words at the same time. Linguistic cues such as naming can be seen to possibly assist children at a very young age in their categorising of non obvious object categories (Nazzi & Gopnik, 2001). Infants as young as 13 months appeared to categorise objects more successfully if they were named (Waxman & Markov, 1995) and already known to the child.

This perspective adopted by Gopnik and Meltzoff (1986) on linguistic and cognitive processes being linked in the acquisition of categories and word meaning is therefore useful for the requirements of the main study tasks. The next section will consider how evidence from other studies has attempted to explain the complex task for a child to acquire language and the potential relevance of the findings of these studies for the main tasks.

2.2 How monolingual children develop word categories

2.2.1 The classical view of categorisation

As outlined on p.20, the classical view of categorisation was developed from the Aristotelian view that in order for two objects to be classified under the same category they

must both possess similar characteristics which define them as being part of that particular group (Medin & Smith, 1981). There are significant features which can be identified to distinguish an object as belonging to a certain group. For example, taking the category *triangle*, all triangles have a geometric form consisting of three sides joined together forming interior angles to the sum of 180 degrees. Without all of these properties this form is not a triangle (Rips & Medin, 2005).

The point of the theory is that all aspects of a category have equal importance as all need to be present to determine the inclusion of an object in a given category. Although the individual aspects characterising the category need to be present they are insufficient on their own to be able to define that category. For example, the term *bachelor* needs to include the individual features *not married*, *is mal*e, and *is an adul*t. If they are not all present, for example if *not married* was not included, it would not be possible to determine whether a given adult male belonged to the bachelor category (i.e. he could equally be a husband or a bachelor if *not married* was not included as a feature) (Evans & Green 2006).

Typicality is also an important element as certain category members do appear to be more typical than others. For example, for the category *bird*, many people would perceive the *robin* or *canary* as being better exemplars than *ostrich* (Smith, Shoben & Rips 1974). Studies have shown that people can respond faster to good typical exemplars of categories over poor exemplars of members of a group (Rosch,1975). The issue of clarity of membership of certain objects within a group, potential disagreement between individuals as to whether an object is categorised in a particular group, and also inconsistencies from individuals categorising the same objects on different occasions emerge to challenge the classical view (Barsalou, 1989;

1996).

2.2.2 Semantic memory and categorisation

The theoretical descriptions of categories do not necessarily explain how children categorise novel objects or how categories support the acquisition of word meaning leading to the development of learning and social communication. In designing the main study task therefore other approaches need to be examined for their potential relevance to the study. Attempting to demonstrate what mechanisms could be available to the child, Collins and Quillian (1969) developed the hierarchical network model. This model proposes that concepts are stored in semantic memory in a hierarchical structure according to cognitive economy. The concept of cognitive economy assumes that the properties of categories are stored at the highest level with the concepts stored closest together having a stronger association than those stored furthest apart. The model was formed from the findings of response time tests to questions such as 'Is a canary a bird?' and 'Do canaries swim?' When compared, the second question took longer to validate than the first – the assumption being that concepts *bird* and *canary* were stored more closely together due to their association. Although this model provided a potential explanation for how concepts were stored in memory, it is still a theoretical construct which does not provide a full explanation for how the child identifies and retains clear definitions for novel categories.

Collins and Loftus (1975) suggested the spreading activation model which represented concepts and their properties as network nodes. This model moved away from the hierarchical model as each node is connected to every node although only one node is activated at any one

time. When activated, nodes connect with other nodes with the strength of association based on the distance between them. However this type of model does not help to explain how children make sense of sentences or phrases; as the classical view noted, categorisation is assumed to speed up the process not force children to have to learn every sentence as a separate concept.

Retaining the significance of semantic memory, Smith, Shoben and Rips (1974) proposed the semantic feature comparison model which outlined feature lists for different concepts which were held in memory. Using this model, the relationships between categories are not directly retrieved but instead the features are compared within the 'feature sets' and selected as appropriate. Each concept is assigned defining features which determine meaning and characteristic features which determine typicality. The hypothesis assumes that recognition is faster when an item is a typical member of a category as the comparison process is shorter (Rips,1975).

This approach however does not account for a more 'fuzzy' structure for categories. Hampton (1979) criticised the semantic feature comparison model as requiring both defining and characteristic features. He proposed that polymorphous concepts in semantic memory provided an alternative explanation which did not require this distinction. Polymorphous in this instance referring to a concept characterised as potentially one with differing forms. Hampton (1981) outlined polymorphous concepts as ones where category membership was associated with set features none of which need be held by all the category members. Taking the case of *sweetness* and *fruit* as a case in point. Many members of the category fruit would share the feature of sweetness but not lemons. Thus the category may have features which

however are not required for membership nor guarantee membership but nevertheless are significant in defining the category. McCloskey and Glucksberg (1979) attempted to address this problem via the feature comparison model which also removed the distinction between defining and characteristic distinctions. Sloman, Malt and Fridman (2001) found that feature sets correctly predicted the category names given by adults to 85–92% of objects in three large (60 objects apiece) sets of pictures representing objects. This suggests that features appear to capture a substantial part of the knowledge upon which adult naming choices are based whether explicitly or only implicitly represented in the mental lexicon. A comparative picture naming/drawing study of five year old children suggested that functional and physical properties were key aspects in object representation in the semantic lexicon (McGregor, Friedman, Reilly, & Newman, 2002). Where there are good exemplars which could represent prototypes of the concept, there is some evidence that properties such as centrality (core to membership category) and gradience (rank of importance, not clear cut membership) enable an individual to categorise a novel concept (Rosch,1965).

2.2.3 Prototypes and the family resemblance model

The idea of a prototype as the most central representative of a category forms the basis of the family resemblance view of categorisation (Osherson & Smith, 1981; Krascum & Andrews, 1998). The theory was described by Rosch (1973) who proposed the prototype view that each category represented the summary description of that category expressed as a best exemplar. Rosch studied the way the Dani Language of New Guinea apparently categorised focal colour terms (defined prototypically as the best examples chosen of colour categories)

structurally in the same manner as an English speaker, although it only uses two terms; *mili* (dark, black, blue, and green) and *mola* (light, white, yellow and red). The choice of best exemplars is a significant characteristic, for example, for the noun *pet*, dogs and cats are appropriate but tigers do not fit the profile of a best exemplar. A subsequent study investigated two hundred students' categorisation of exemplars for the category *furniture* (Rosch,1975). On a scale of 1 to 7 students ranked chair as the best exemplar with telephone as the weakest exemplar. The family resemblance model therefore helps to clarify thinking on the acquisition of both nouns and actions.

There is evidence that children select category membership on the basis of similarity rather than analysing features (Kemler-Nelson, 1984). These findings suggest that children use the family view for natural categories i.e. everyday categories at an early stage (5 years) whereas older children (10 years) and adults find classical categories easier to learn.

Markman (2003) proposed that there are three possible explanations as to why the family resemblance view provides a plausible explanation for the child's categorisation of novel objects. First, individuals may have to extend the simplicity of the classical view when they encounter exceptions; for example a *bird* is characterised by wings and uses flying for locomotion, therefore is a penguin a *bird*? It would be logical to categorise the penguin as a bird, however this would require an extension of the typical definition of a *bird* as an object with wings which also flies. The second explanation suggests that individuals are capable of defining their own categories (Murphy & Medin,1985) which they organise according to their own internal logic. Finally, Wattenmaker, Nagamura & Medin (1988) alternatively suggested that using a non-analytical approach for identifying category membership based on exemplars

is sensible in that individuals are not forced to reject information at an early stage therefore retaining information which might be helpful in future categorisation tasks.

Both the classical view and the family resemblance view assume that category terms have features in common which define their properties, and extensions which identify the objects associated with the term. The causal theory of reference (Kripke, 1971) suggests that previous explanations of categorisation were too specific. A better analogy would be the approach used for proper names where the defined term is useful for identification but is not totally dependent on common criteria.

Categorisation is a useful foundation for gaining a better understanding of the complex and flexible processes which occur in the creation of taxonomies with a distinction between natural categories (what an object is) and arbitrary categories (what an object is like). (Markman, 2003). Natural categories will develop richer taxonomies than arbitrary categories. Murphy (2002) further developed Rosch's work emphasising the organisation of the hierarchy. At the centre is the basic level. In the case of the superordinate category *furniture,* for example, *chair* is at the basic level The single relationship between the basic category and the superordinate category is expressed as "IS-A", with superordinates above and subordinates below that level. This relationship is asymmetric; for example, all cats are animals but not all animals are cats, and subordinate categories inherit properties of the superordinate, for example all cats have hearts therefore Siamese cats have hearts.

According to Lakoff (1987), the family resemblance model can be used as the basis for an extension of types of category that of radial categories. The relationship between these categories revolves around related meanings of words which bear family resemblances as

'meaning chains'. The objects in the meaning chain have a shared attribute which link them. A category can then be structured radially through a number of sub-categories which are extended in an ordered way. For example the word *window* can have three meanings: 'an opening in the wall', 'a frame fitting into the wall' and 'the glass filling the frame fitting into the wall' with the three representations linked through the chain of meanings.

Interestingly recent research has attempted to generalise the approach of the family resemblance model and another graded structure model, Hampton's polymorphous concepts (Hampton, 1979), to predict the degree of commonality across the feature sets of category members. Day & Storm (2009) found that both common and distinctive feature information were relevant in predicting typicality across semantic categories. In category judgments, common features were weighted more for within-category predictions and distinctive features weighted more heavily for contrast category predictions.

2.2.4 From category representation to category judgments

With the influence of models such as the family resemblance model and the emphasis on prototypes rather than rigid classification, cognitive theories placed an emphasis on examining semantic categorisation as a judgement process (Ramscar et al, 1997). The concept of similarity has already been demonstrated to underpin many theories of semantic categorisation. Gentner & Markman (1997) argued that both similarity and analogy can be used to map representations. Analogy is useful in mapping relations between seemingly different objects including acknowledging causal relationships which may not be immediately apparent. This approach proposed that three constraints define the relationship – that the

relationship should be structurally consistent so that alignment between representations is parallel and on a one to one basis that the representations have a common focus even if they do not share common object descriptions and that the representation must demonstrate systematicity or coherent connectedness.

Within these constraints both analogy and similarity can be viewed as contributing to the categorisation of new concepts – similarity also requires shared relationships (Gentner & Markman 1997). When the child compares new concepts which are relationally similar but without object similarities, analogy is more helpful. When the comparison involves concepts with object attributes, it is more helpful to call on similarity. Within this model, metaphors can be used to bridge between analogy and similarity.

The usefulness of this approach for semantic categorisation extends to a potential better alignment between similarity models of categorisation as illustrated by Rosch (1975) and theory-based models. Difficulties with representations e.g. although a bat flies like a bird it is classed as a mammal, can be resolved as a child logically moves from a features-based comparison of similarity to a deeper comparison based on encyclopaedic knowledge. In addition, the approach has been used to investigate decision-making (Lindemann & Markman, 1996) where studies demonstrate that alignable differences (where the identification of difference first of all requires a commonality between objects, e.g. a car and a motorbike both have wheels) appear to be given more weight in choice decisions than non-alignable differences (where characteristics of one object are not found in the other object, e.g. cars have seatbelts whereas motorbikes do not).

2.3 Chapter summary

The representation of word categories has moved from a historical view which emphasised the classical theoretical construct to a more complex cognitive approach which relies on making comparisons and choice decisions in both classifying and interpreting novel concepts. Boundaries for these decisions do not appear to be absolute; however models such as Gentner & Markman's (1997) structure mapping engine provide a basis for investigating further the relationship between semantic systems and cognitive organisation. This helps to address the question relating to the nature of the formation of specific categories which are more difficult for a child to incorporate an object into the correct group. The next chapter will move from the categorisation of concepts to the acquisition of words.

Chapter 3. The Acquisition of Word Meaning

3.1 Introduction

It is possible to view word learning as building on three components: the learner's existing knowledge, what data is required and available, and what inference process takes place (Xu & Tenebaum, 2000). This Chapter will build on the understanding of the cognitive organisation the child needs to identify and manage novel concepts discussed in Chapter 2 to investigate how he/she acquires an understanding of word meaning.

The literature focuses on the task that children complete in mapping meaning onto novel words and whether there is a differing level of difficulty depending on language structure. Understanding the process of how children obtain word meaning is crucial to understand how they also categorise words in the context of this study's tasks. For example, does a child see a novel word as a new group category or part of a group category? With regard to this question, the processes which children follow in grouping items together (for example taxonomically and thematically) has also relevance for the study. The selections of children beyond the target may illuminate the cognitive processes undertaken by children in the categorisation process. Children who are less confident on category boundaries may focus on the properties of an items or items which go with another item. Children may also have greater difficulty in understanding the meaning of a word when this meaning is less transparent (e.g. partial meanings). There may be an improvement in the ability of children to categorise words as they grow older as an exposure to and an understanding of word meaning increases. This is relevant for the current study design as it suggest that performance should improve in the older age groups on word groups which have less clear identifiers for

membership inclusion.

3.2 Clark's research on the move from categorisation to word acquisition

This section draws heavily on the work of Clark into early language acquisition. When a child is faced with a novel concept, there needs to be some guidelines which he/she can draw on in the process of mapping conceptual knowledge to words. A useful starting point are the fundamental ground rules which attempt to explain the interaction between semantic systems and cognitive organisation (Clark, 1987). The principle of contrast (Clark,1987) for example, proposes that on hearing a new term, a child assumes that the speaker means something different from what has already been represented. Following on from this, the principle of conventionality (Clark, 1983, 1993) proposes that children prefer to use conventional forms to express meanings as this reflects norms within their communication circle. Observations have shown that children's first fifty words and word combinations tended to be similar and consistent across children in relation to content. This consistency could be explained by a similar process of adoption of conceptual categories which is consistent across languages (Clark, 1971). Children acquire an understanding of how conceptual categories are appropriately used through cues received from their parents and carers. However, evidence for a universal conceptual representation is found in the way that children express these categories without necessarily receiving input from parents or other external sources. These categories can be viewed as 'emergent' as they are likely to appear first across languages and are refined in the child's second year (Clark, 1983). As the child receives language input, the categories become 'robust', represented by the use of nouns, verbs

and adjectives. The process of moving from emergent to robust categories includes experimentation by the child. These criteria could be used to identify the universal categories which underpin language development. The criteria operate through an analysis of what the child mapped or did not map to language (Clark, 1997).

The use of overextensions by the child in early language development can be mapped to categories which become robust categories in other languages, for example, numerical and shape classifiers found in some Asian languages (Clark, 1997). Although the overextensions may only be used for a short period of time, they are not random, for example, the classification of shapes used by children during early language development. As differing instances of the word are experienced, the child analyses the features found in each instance and gradually builds up a set of features which characterise the word leading to the more adult representation of the category represented by the individual word (Greenberg & Kuczaj,1982). This could lead to systematic errors until the child comprehends the appropriate use in her/his own language. Where a language consistently maps emergent categories, it is likely to lead to a more straightforward acquisition process, creating robust categories. The support from the input received by children is essential in this process of moving from emerging to robust categories. Idiosyncrasies can be seen in different languages which result in children of one language receiving support whereas children learning other languages may experience errors where the mapping of categories is not as seamless (Holowka, Brossea-Lapre & Petitto, 2002; Campbell, 2007).

The nature rather than the number of aspects associated with words is key to distinguishing child and adult lexical knowledge (Mervis, 1987). Young children often do not

share an adult's knowledge of culturally appropriate functions of objects and the correlated form attributes (i.e., what objects do and how they are used). This leads children to de-emphasise aspects of an object that are important from an adult perspective. At the same time, children emphasise aspects that are unimportant to adults. As a result, besides both over- and under-extensions, "child-basic" lexical categories may also partially overlap adults' basic level categories such that the child's category may include objects that are excluded from the adult category while excluding objects that are included for adults. For example, a child may use *ball* for round piggy banks, but not for (oval) footballs. Through growing experience with objects and their place within the culture, children learn to attach the same weights to the same aspects used by adults. For example, the most important aspect of a piggy bank is the slot on top, which determines its function. To attend to this aspect in choosing names for objects, the child needs to grasp what the slot is used for and attach this cultural knowledge to the appropriate lexical item.

The child's initial word meanings do not necessarily match those of an adult. (Clark,1979). A study of overextensions used by one year old children identified that only a third of the first seventy five words acquired were overextended with the majority of overextensions associated with a small number of high frequency words (Rescorla, 1980). Young children's over-extended categories are based primarily on perceptual or physical properties, such as shape, size and texture. Children do not just add perceptual information to the meaning of a word, but must also learn functional or cultural roles before they can learn the distinction between meanings of words, for example, *chair* and *throne*. Therefore over-extensions gradually disappear when children narrow down the initially very general

meanings of the overextended terms by adding more features to them (Clark,1979).

Young children acquire new terms in a domain through initially fixing the reference of a new term. This may be achieved through making some immediate inferences about an adult's intended use given the constrained context presented by the adult and child's focus of attention (Clark,1993). Clark argued that children learning the words of a first language need to isolate word forms, to identify potential meanings and to assign these meanings to the newly isolated words. The potential meanings may be based on conceptual categories already represented in memory (Clark, 1995, 2004) and it appears that children need only minimal exposure to a new form before assigning some meaning to it. As soon as a possible meaning is assigned, the word is ready for use or 'fast mapped'. Fast mapping allows children to add words to their vocabulary at a rapid rate during the first years of language learning. By age two, children are able to produce 50–500 or 600 words. In the period from age two to six, they are estimated to acquire around 14,000 words, at a rate of ten words a day (Carey, 1978; Carey & Bartlett, 1978). The typical age at which word groups are learnt appears to vary. For example, Clark (1979) observed that the meanings of orientational terms (*top, bottom, front* and *back*) are not fully understood until the age of 5. Similar findings were reported for deictic terms, such as *this, that, here* and *there* (Clark & Sengul, 1978). Verb meanings for even relatively frequent verbs, such as *pour* and *fill*, are not fully understood until eight or nine years of age (Gropen, Pinker, Hollander and Goldberg, 1991). Tomasello (1995) proposed that the process depended on joint attention. In addition, the child would need to infer that the object or event under scrutiny was the relevant item being referred to by the adult and that the child was able to use any additional available information about the referent

of the new term (Clark & Wong, 2002). However, the meaning assigned through fast mapping, often consists of only a fraction of the meaning adults attach to a word (Casenhiser & Goldberg, 2005). Learning the full conventional meaning of a word may take months or even years.

This narrowing-down process takes place in conjunction with the introduction of new words that take over sub-parts of an initially overextended semantic domain (Rescorla, 1980). Due to the addition of other features to word meanings and the acquisitions of new vocabulary, individual word meanings become more specific and semantic domains are considerably restructured. For instance, children might initially use the word *ball* for a variety of round objects including balloons, apples and a round lamp, since they might have characterised the word ball as meaning round. If they next acquire the word *apple,* they must add features to make this word distinct from *ball*. For instance, they may add the feature stalk. At the same time, they will probably add more features to the meaning of *ball*, such as bouncing, to contrast this meaning to the meaning of the newly learned word. Young children overextend the use of the word to other objects, while children who are a little older will concentrate on perceptual features resulting in the creation of many subcategories. The oldest children take information from not only the perceptual cues of objects but also what the functional features are in deciding the extension of a word (Andersen,1975).

Kuczaj (1982) queried whether the use of overextensions by young children in production is reflected in comprehension. When young children up to two and a half were given a series of choices during a comprehension task for an individual word, the choice made was usually an appropriate one given the target word. However the subsequent choices were

likely to include overextended objects. This could be argued is evidence for a prototype account for early object word meaning acquisition.

3.3 Difficulties with partial and multiple meanings in categorisation

An ability to understand partial meanings appears to be age related. Adults and older children are able to relate new nouns to known referents, for example 'An oak is a kind of tree' (Clark & Wong, 2002). An adult will also understand that the terms *beech, ash, oak, rowan* and *elm* refer to types of trees without necessarily knowing what each type of tree looks like or what size it grows to (Clark, 2006). Similarly, terms which describe colour such as *maroon, ecru and cyan* would be recognised as colour terms by adults without necessarily fixing their reference. Therefore children need to develop an appreciation of semantic groupings where similar terms are collected together with any relevant associated lexical expressions.

Young children find it difficult to understand that colour is a similar property of an object as is size and texture, and as a consequence, they find it difficult to fix the reference term (Schwartz, 1977; Kowalski and Zimiles, 2006). Colour terms represent a large range of features including brightness and shade. It is therefore harder for the young child to accurately describe colour terms. This suggests that children find it difficult to categorise objects correctly when there is nothing concrete (or when it is more opaque) to map meaning onto. Part of the acquisition process for colour terms is the ability to organise the relationships between new colour terms and any super-ordinates and subordinates and to learn how such relationships describe the function and properties usual for that type (O'Hanlon and

Roberson, 2006).

Polysemy and homonymy represent two additional difficulties for children to correctly understand category boundaries. Polysemy is the term used to describe a situation where a word has two or more related meanings. However a clear cut definition is not always possible when both physical senses and abstract meaning are involved in the analysis. As well as polysemy, the place of homonymy potentially complicates the organisation of semantic systems by the child. Homonymy is defined as two lexical items that happen to have the same phonological form. Lyons (1977) pointed out the difference between the polysemous use of *mold* with multiple unrelated meanings (vessel for liquids, fungus) and *line* with multiple related meanings (fishing line, line of people). For native English speakers the words *pupil* 'student' and *pupil* 'iris (of the eye)' are not usually considered related.

Polysemy introduces a range of representations which usually possess their own typicality. Fillmore & Atkins (2000) investigated the polysemous meanings of the verb *to crawl.* The central meaning relates to moving near the ground. However many of the peripheral meanings are related to higher level metaphors – *time crawling by* or *the table crawled with flies*. The various meaning appear to be processed as part of a chaining system, although Caramazza & Grober (1976) queried whether the different polysemous meanings are in fact stored separately. Research studies in this area highlight the importance of acknowledging that word meanings do not necessarily conform to standard patterns and that care needs to be taken in addressing polysemy and homonymy when investigating word acquisition organisational systems.

3.4 The influence of interactions with adults and their surroundings on word acquisition by children

Evidence shows that there is a direct correlation between the vocabulary of a child prior to attending school and the amount of interaction with an adult, and that this pattern is consistent across different languages (Hart and Risley, 1995). It therefore appears that children listen to adults which enables them to construct a language schema for objects and events. The schematic maps for each language differ in detail, for example, in the grammatical categories including gender, number and tense and in the number of basic terms used for colour and whether they are used as adjectives or as verbs. Where a language has several terms for various shades of blue, children concentrate on acquiring all the terms instead of just one single term, whereas children find it harder to acquire an additional term if they have no previous exposure (O'Hanlon and Roberson, 2006).

According to some researchers, the use of correction by adults when children make lexical, syntactic, phonological or morphological errors is also a feature of the acquisition process. In correcting the child, adults present another form which is the appropriate way for expressing the child's intended meaning (Chouinard & Clark, 2003). Children follow the lead of the adult and eventually adopt the adult form (Clark (1987). Similar effects have been found in the response of children up to six years of age (Saxton,1997). A comparison of the results of picture naming tasks of children up to six years and adults showed the influence of age differences in naming and category membership with the children relying more on shape to determine categories than adults (Bloomquist, 2007).

A concept could be formed through a child's interaction with his or her surroundings,

and his or her representation of the object is closely tied to this interaction (Nelson,1974). When children focus their attention on a new object, for example a new bike, this object maps onto a concept consisting of relations to self, to other possible actors (e.g., father), to locations (e.g., garden) and actions (e.g., ride, fall off) and the effects of actions over time. The bike does not exist for the child outside of one of these relations. When an instance of the concept bike comes to be named, it would be named *bike* only if these relations hold. Therefore, in contrast to Clark (1979), this view suggests that children initially attend to more features than adults do, rather than fewer when assigning names to objects. This explains why some of the child's initial lexical categories are narrower than the corresponding adult's categories. To name objects independently from their involvement in a set of relations, some relational features need to be identified as irrelevant to the concept's functional core, such as location of the activity. For example, the functional core of bike will end up containing the functional features - ride, fall off . In addition, as instances are added to the concept, a set of perceptual features are attached to the functional core, for example, *race* for *bike*. Eventually, the meaning of a word will consist of the specification of obligatory (core) and optional relations.

Nelson's claim that word use is tied to specific activities rather than being connected to objects independently was challenged by Huttonlocher and Smiley (1987) who found that even young children can learn names outside any interactive context. However, the possibility that children begin with more features than the adult set provides a useful alternative to Clark's view.

3.5 The naming of novel objects

The existence of several mechanisms may cause individual complexity in naming that can not be explained by the presence of particular features associated with a name (Malt et al, 1999). A particular name can be used for an object by linguistic convention rather than because of specific similarity relations to other objects associated with the category name. For example, the name can be introduced by a manufacturer through advertising. People may avoid calling an object by a particular category name because using that name would lead to ambiguity or confusion with another type of object that already receives the name, referred to by them as 'pre-emption'. This can be shown by the example of a soup *tureen* which may be called 'tureen' even though it shares many features with objects called *bowl* or *pot*. Calling it a *soup bowl* or *soup pot* would create referential confusion with vessels for eating or cooking soup. The use of *bowl* or *pot* for the serving container may therefore be "pre-empted" by the other uses of these names. They further argued that the idiosyncratic mappings can only be learnt through growing experience with the naming of individual objects.

Therefore the differences between child-based and adult-based categories may be due to differences in the features to which children attend when assigning names to objects. One possibility is that these differences are purely lexical—children understand the domain non-linguistically in much the same way was as adults but have not yet fully mastered the mapping of non-linguistic knowledge to lexical items (Halliday & Hasan, 1976). As a result, they use the wrong features or the wrong set of feature weights to direct their use of words. On the other hand, these differences may be due, in whole or part, to differences in the child's conceptualisation of the objects themselves (Morris & Hurst 1991).

Clark (1979), Nelson (1974) and Mervis (1987) all take conceptual development as an important part of early lexical development, suggesting that children acquire semantic knowledge through growing cultural and functional knowledge about objects and the world. Waxman and Braun (2005) and Booth et al (2005) demonstrated that both perceptual and conceptual information are acquired by young children in early word learning. On the one hand, children between 5 and 14 years old are still developing their understanding of many basic aspects of their world. Alternatively, Malt and Sloman (2003) found that second-language learners presumably had a full conceptual grasp of the objects showed substantial lexical learning in their use of English nouns for common containers and other house-wares over years of immersion in an English-language environment. This finding suggests that later semantic development could be a purely lexical process.

Gopnik and Sobel (2000) devised a novel method to investigate causal inference in children using what they called a blicket detector. Blocks were placed on the blicket detector machine which had the causal power to make the machine light up and play music when a blicket was placed upon it, some blocks were blickets while other were not. They did not have visual features that would indicate membership as a blicket. The researchers then altered patterns of evidence about the machine and the blocks looking at casual inferences made by the children. Results suggested that young children were able to categorise objects by understanding the casual relation between objects (Gopnik & Sobel, 2000).

Evidence suggests that children as young as two can use causal features of objects to help them with their categorisation (Gopnik, 2001). Booth (2008) supports this view highlighting that children were able to categorise objects due to their causal features even

when they did not have similar visual features. This is not to suggest that perceptual information does not play a part in categorisation by children. Although when there appears to be conflict between visual and causal relationship children tend to use causal features for deciding membership of that item within the group (Yoachim, Gopnik, Meltzoff, & Blumenthal, 2007). This finding links back to Gentor & Markman's (1997) work in Chapter 2 on how individuals manage 'deep' relationships between categories.

3.6 Early language acquisition and processing

The natural partition hypothesis and the associated relational relativity hypothesis attempt to explain the more rapid acquisition of nouns by a young child (Gentner,1982). There are only a small number of verbs which appear frequently, compared to a greater number of nouns which appear on a less frequent basis. It would be expected that verbs would dominate a child's early vocabulary because of the likely increased exposure to verbs. The natural partition hypothesis suggests a solution for this unexpected finding. It proposes that object nouns can be visually seen as concrete objects and therefore a young child will probably encounter these objects on a regular basis. Verbs, prepositions and adjectives on the other hand are more likely to include conceptually more complex abstract meanings (Gentner & Boroditsky, 2001). As a consequence there is a logical reason for the acquisition of common nouns as a faster process than verb acquisition because common nouns refer to entities easily segregated from their context, whilst verbs convey relationships among entities (Gentner, 2005). The relational relativity hypothesis suggests that there is a relational shift in cognitive development requiring the child to make the relational links between entities. This appears to

be a developmental process as older children appear to be able to make these links which are harder for younger children. As the meanings of verbs are harder to comprehend from the real world mapping of concrete entities, the relational links will be acquired later (Gentner, 2005).

Gentner's hypotheses imply that the meanings of common nouns will be relatively easy to learn. Her proposals are in line with other constraints already identified such as the whole object constraint (Markman, 1990) and the taxonomic constraint (Markman, 1990).The hypotheses were tested on word acquisition by children speaking Navajo to establish whether the findings could be replicated in a language other than English with consistent results (Gentner & Boroditsky, 2009). This approach is also compatible with Rosch's proposal that common nou ns such as those for objects such as *chair* and *bowl* capture sets of items sharing many inter-related properties (Rosch,1975). Hunn (1977) and Berlin (1978) suggested that some common nouns for plants and animals conformed to Rosch's definitions. If common nouns in general label groupings of objects made obvious to the individual by the distribution of properties across them, then naming patterns should be similar across languages, and they should be predicted by judgements of the similarity among the entities. De Groot (1993) and Kroll (1993) further proposed that the meanings of common nouns tend to be generally equivalent across languages. In contrast, Andersen (1975) asked English-speaking children aged 3 through 12 to name ordinary drinking vessels. She found that it was only at the age of twelve that children's use of terms such as *cup* and *glass* fully matched adult usage. This evidence suggests that the acquisition of common nouns may be more complex and implies that the knowledge required to use common nouns as adult speakers also takes a significant amount of time.

Evidence against this view has been presented in cross-linguistic studies where children did not demonstrate a noun bias in their early vocabulary (Gopnik & Choi 1995). An alternative view emphasises the child's ability to understand certain types of nouns, for example count nouns, as being part of the process of cognitive recognition of the relationship between words and meaning (Waxman, 1994). Children may be able to make sense of new language elements through assumptions which enable narrower predictions to be made in interpreting novel utterances.

Markman (1991) highlighted the apparent paradox for children faced with acquiring new words. On the one hand it would be natural for children to develop separate categories for each new object as it would be easier to remember categories with only one member i.e categories are mutually exclusive. On the other hand the grouping together of categories into classes is a better reflection of reality. Young children using mutual exclusivity would therefore be expected to find the organisation of categories in a hierarchy of classes difficult. Thus a child who had learned the word *doll* would struggle with the concept that a *doll* was also a *toy*. Markman suggests that the differences between the two concepts of classes and collections helps to address the child's problem with mutual exclusivity. Collections being the referents for collective nouns are organised as simple whole-part hierarchies, for example, a tree is part of a forest. This association with a 'whole' enables the child to be comfortable with collections as the mutual exclusivity rule is maintained, the tree is part of a forest not the same label as a forest. Testing children with novel terms in a made up language, the prediction was that children would attempt to make whole-part collections from class-inclusion hierarchies as this would be easier to understand. The results demonstrated that, when given minimal

information, even children as old as sixteen made errors in identifying class-inclusion hierarchies as collections. Further research (Markman, 1994) identified that the use of mass nouns may assist the young child to develop an understanding of class hierarchies through being a term having whole-part and class characteristics. This use of mass nouns as superordinate categories is not confined to English and Markman's research demonstrates that although young children have difficulty in learning hierarchical classes, using mass nouns is helpful in the learning process.

Davidson and Tell (2005) investigated monolingual and bilingual children's use of mutual exclusivity with the expectation that bilingual children would be less likely to retain mutual exclusivity on the basis that they would be familiar with two terms for an object. The results did not support this projection and they found that both monolingual and bilingual children retained mutual exclusivity for the naming of whole objects. The main difference was found between older children where bilingual children were less inclined to use mutual exclusivity.

Healey and Scarabella (2008) investigated the response of monolingual and bilingual children to two word learning strategies, the mutual exclusivity constraint and a socio-cognitive cue. Children completed a picture matching task which introduced three novel labels for familiar objects and a treasure hunt which provided a representational task to test their willingness to accept the objects as referents for the novel labels. Although both monolinguals and bilinguals used mutual exclusivity for the picture matching task, bilinguals were more willing to ignore the mutual exclusivity constraint for the second task where a socio-cognitive cue was provided. These results differed from Davidson and Tell (2005) in

that bilinguals performed differently from monolinguals when presented with socio-cognitive cues suggesting that the mutual exclusivity constraint is used differently by bilingual and monolinguals. Healy and Scarabella (2008) postulate that this difference may be due to the necessity of bilingual children dropping mutual exclusivity between languages making them more likely to drop the constraint within a language.

3.7 The shift in categorisation preference from thematic to taxonomic

This section considers the assumption that there is a shift in preference in categorisation from thematic to taxonomic from childhood to adulthood. Vygotsky (1962) proposed that children have a preference for thematic categorisation whereas older children and adults have a preference for taxonomic categorisation. Further research has identified a more complex situation where children do use limited taxonomic categorisation. Children may prefer to associate new words with objects that are taxonomically related rather than thematically related (e.g. *cat* + *cat food*) (Markman,1989). In addition, adults may demonstrate a preference for thematic categorisation (Lin & Murphy, 2001). In a series of ten experiments, college students grouped items thematically at the superordinate level when presented with a forced choice. For example, when asked to choose a related associate to group with *bee*, honey was often chosen in preference to flies. Overall, the study identified that 49% - 70% of responses provided thematic responses. Other studies have identified a stronger thematic response in less-educated individuals (Cole & Lave, 1979) and elderly participants (Smiley & Brown, 1979).

It can be argued that the use of thematic categorisation by adults may not have been

identified in earlier studies due to the type of stimuli used which did not present a strong thematic opportunity for grouping (Lin and Murphy, 2001). As a result the studies did not provide an equal opportunity for comparison. In addition there is some evidence (Denney, 1975) that instruction during experimental studies may have influenced the consequent grouping. Instructions to children to 'find things which go together' may suggest to the participant that they should group items thematically. The results suggested that a tendency for researchers to present sets of items taxonomically in past studies may partially explain why adults appear to have a preference for taxonomic over thematic groupings. When presented equally, adults expressed a preference for thematic groupings. As a consequence Lin and Murphy (2001) suggest that the shift from thematic to taxonomic groupings from young children to adulthood is not conclusively demonstrated by previous studies where the method adopted may have influenced the outcome. They accept that there is a difference in children's ability to group items taxonomically (Markman, 1989), however they suggest that adults are probably adept at categorising both thematically and taxonomically to suit the required action.

When learning novel words, children aged 2 and older reject taxonomic relationships in favour of thematic relationships (Markman, 2003). The mutually exclusivity assumption predicts that children assume that different words refer to different categories of objects (Markman et al, 2003) and helps to address the potential problem for children when faced with a novel word for a known object. According to the hypothesis, children appear to make sense of new language elements through assumptions which enable narrower predictions to be made in interpreting novel utterances. Where high levels of translation equivalents are found,

it could be argued that bilingual children have two distinct language systems as this suggests that mutual exclusivity is being ignored. The tendency to ignore mutual exclusivity may be acquired as a consequence of understanding cues about word meaning (Deuchar & Quay, 2000).

Young children appear to be able to be flexible in their categorisation (taxonomic/thematic) of the same object depending on the context. In the case of an *apple*, it may be a *fruit* (taxonomic), an *object in a lunchbox* (thematic) or *round* (perceptual) Studies by Blaye and Bonthoux (2001) challenged the view that young children progress from schematic to taxonomic (Nelson, 1985). However they concluded that although several types of relations may be observed in the categorisation of three year old children, their ability to change from one relation to another from available cues is limited. This conclusion challenges the perception that taxonomic representations are well developed in four year old children.

3.7.1 The Generative Lexicon

The 'Generative Lexicon' was proposed by Pustejovsky's (1995) to address the problem of lexical ambiguity, to the multiplicity of word meaning and to the question of people being able to give an infinite number of senses to words using finite means. The main argument is that a core set of word senses is used to generate a larger set of word senses when individual lexical items are combined with others in phrases and clauses. Pustejovsky argued that former approaches to natural language semantics had ignored either the problem of how words are used in novel contexts or the creation of such new senses. In language, words can have more than one meaning, but the means in which this extension of meaning is carried out

can vary.

3.8 Cross-linguistic studies

This section examines the usefulness of evidence from cross-linguistic studies of semantic categorisation in monolinguals. Associative priming effects can be seen in word identification tasks, with the target words being identified faster if they are preceded by a semantically related word; for example, the word *bread* is identified quicker when it is preceded by *butter* than when it is preceded by *chair* (Meyer and Schavaneveldt, 1971). Research has also shown that this finding also appears cross-linguistically with the level of facilitation between the word pairings similar in the same language pairs and the mixed language pairs (Schwanenflugel & Ray, 1986). When bilingual participants switch languages in a task requiring them to recognise words a clear cost is seen in their performance. The introduction of language specific orthographic cues is found to reduce this cost leading to the possible explanation that these switch costs arise from within the bilingual lexicon (Grainger and Beauvillain, 1987).

The cognitive processes which accompany semantic categorisation, from the internal recognition and encoding of the object as being related to other known entities to the articulation of a word or the external recognition of the word spoken by someone else, were highlighted by the work of Malt, Sloman, Gennari, Shi, and Wang (1999). The study which investigated the naming of sixty terms for containers across American English, Mandarin Chinese, and Argentinian Spanish showed a marked difference in the extensions used across the three languages. They proposed that the relationship between encoding and naming was

not necessarily straightforward across languages. Malt et al. (2003) conducted a detailed examination of the terms used for the sixty containers. The authors noted that some categories were similar across the three languages. However, some containers grouped in one language were seen as part of different categories in the other languages. Objects which were perceived to be similar across different languages were therefore not similarly grouped linguistically. They found substantial differences in the linguistic category extensions across the speakers of the three languages, while the perception of similarity across the languages was much the same. These results indicate that the linguistic categories of a language are not strictly formed around similarity-based clusters.

Malt, Slomen and Gennari (2003) examined in more detail the relationship among the linguistic categories of the three languages for the 60 containers and found a complex pattern. Some of the categories were grouped around similar prototypes across the three languages, whereas others were not. Some categories of one language were nested within those of another language and others were not. In some cases cross-cutting was found, where pairs of objects were put into a single category by one language but into different categories by another language. Taken together, these findings show that the groupings captured by common nouns do not consist simply of obvious groupings of objects having many interrelated properties in common. This conclusion was further supported by Ameel, Storms, Malt and Sloman (2005). Their study replicated Malt's study in 1999. The findings revealed a dissociation between naming and similarity using French- and Dutch-speaking monolinguals in Belgium. These groups are similar in culture and familiarity with particular types of objects. As expected their similarity judgements were highly correlated, yet they displayed

distinctly different naming patterns for two sets of household objects.

Direct evidence for the complex learning development of common nouns was provided by Malt and Sloman, (2003). They found that second-language learners retained discrepancies from native speakers in their use of English nouns for common containers and other household objects even after more than thirteen years of immersion in an English-language environment, and even after they acquired the appropriate prototypes. These discrepancies appear to reflect non-equivalences of meaning between languages and the resulting complexity of the learning task. The findings of these studies indicate that learning common nouns may present greater challenges to children than proposed by Gentner, and that the process learning may extend well past the early years of language acquisition. This observation supports Kronenfeld et al (1985) who noted this effect when studying English, Japanese and Hebrew speakers' categorisation of drinking containers. The study used eleven different drinking containers and compared the terms used. For example, American English speakers used the same term *cup* for a plastic container and a pottery tea cup whereas the Hebrew speaker did not.

A further perspective, relating to environmental influences on the development of lexicons was highlighted by a study of Russian and Greek-speaking children (Stephany & Voeikova, 2003). The study investigated the category of aspect in Greek and Russian as there are important differences in the way aspect is treated by the two languages. The research attempted to identify how far the development of aspect was influenced in the two languages. The authors concluded that children pick up language-specific cues from environmental influences in extending their lexicon rather than using innate lexical tools.

Spatial actions in particular appear to be coded according to the requirements of individual languages (Bowerman & Pederson, 1992). English uses the prepositions *in* and *on* as in *put in* and *put on* to express containment and support. In contrast, the Korean language does not have an equivalent preposition, with the important relationship being between 'loose' and 'tight' fit. Results from Bowerman & Choi's (2001) study of American English and Korean children aged 16 to 20 months showed differences in the children's responses to the same pictures expressed in two spatial representations of containment. These were represented by the English preposition *put in* and the Korean verbs *nehta* (loose fit) and *kkita* (tight fit). American English children preferred the language specific expression of containment whereas Korean children preferred the use of 'loose' and tight fit' verbs. The authors proposed that these results suggest that lexical interpretation is language specific by 18 to 23 months (Bowerman & Choi, 2001).

3.9 Categorisation of events

Cross-linguistic research has also provided evidence of a difference in the categorisation of events. Repetitive actions are experienced by most children regardless of language and therefore provide a mechanism to compare how individual languages express basic actions. It has been proposed that *cutting* and *breaking* verbs have semantic similarities across languages (Guerssel et al. 1985). However, other studies suggest that cultural differences lead to differences in usage depending on the action and the context, for example, English *break* can be used with a variety of object nouns. In contrast, the Maya K'iche language uses a range of verbs depending on the type of object (Pye, 1996).

Majid, Van Staden, Boster and Bowerman (2004) investigated two basic actions of *cutting* and *breaking* in a substantial study across twenty eight languages. They concluded that,

> "Speakers of a variety of typologically, genetically and a really diverse languages agree to a surprising extent in their linguistic categorization of events of material destruction of objects ('cutting and breaking' events)." (Majid, Van Staden, Boster & Bowerman, 2004: 889).

The results confirmed a degree of similarity between the representation of controlled actions of *cutting* and imprecise actions of *breaking*. However languages varied in the specific categorisation of sub-categories representing 'cutting and breaking' actions. Languages with similar roots such as English and Dutch varied, with Dutch not differentiating cutting actions to the same degree as English. 'Cutting' actions in Dutch are limited to the type of object used (knife, scissors). Despite the degrees of similarity therefore, the results also showed the range of words available to speakers of individual languages to describe event actions.

Associative learning appears to be a successful internal mechanism for individuals to expand their understanding of the meaning of novel words. Repeat encounters may enable the child to learn from the environment, building on previous exposure to an object or event. This leads to a gradual development in understanding dependent on the influence of previous encounters (Colunga & Smith,2005). When applied cross-linguistically, studies demonstrate both similarities and differences. English and Japanese children were observed to have similar responses to solids when completing novel noun tasks although English contains mass-count nouns whereas Japanese does not (Imai & Gertner,1997). Although another study found that Japanese children aged two and a half showed a bias towards non solid objects earlier than

English-speaking children (Samuelson & Smith, 2000). Colunga & Smith (2005) concluded,

> "that some universals might be the product of similar learning environments. In particular, Japanese and English speaking children's novel nouns generalizations may be similar because both are products of generalizations over the nouns children know and because the nouns children know in the two languages have similar category structures." (Colunga & Smith,2005: 27).

3.10 Chapter summary

The cognitive organisation of word acquisition has attracted many studies which point to a series of tendencies when children acquire words. Children appear to assume that new concepts are different to previous ones (Clark, 1987) and are influenced both by adults and their surroundings in the way that they make sense of novel words. As might be expected, children initially consistently use overextensions in early language development, mapping them to known categories at an early stage. The ability of young children to make causal relationships when faced with new objects and grasp noun meanings is interesting and appears to link with assumptions that categorisation is present at an early developmental stage. This complexity is balanced with observed difficulties with the presence of polysemous and homonymous words where meaning is complicated. Cross linguistic studies provide an opportunity to examine communality between languages in the context of the requirements of the current study and leads to the main area of study in Chapter 4 where the complexity of the organisation of categories by bilinguals will be discussed.

Chapter 4. The Organisation of Categories by Bilinguals

4.1 Introduction

Children with access to two languages have to manage semantic systems to accommodate the features of both languages. This could be achieved by treating them as a single system or be managed as two separate systems (Paradis & Genesee,1996). If the two languages develop interdependently, what is the extent of the influence of one language on the other and how does this affect the rate of language acquisition in relation to the rate of acquisition by monolinguals? If the languages are acquired autonomously, what are the mechanisms which enable the bilingual child to manage the separate language structures of each language? The added challenges posed by the acquisition of two languages has been viewed by some as leading to a deficit in developmental terms when bilinguals are compared to monolinguals (Cook, 2002). This perspective has been challenged by the view that the linguistic development of bilinguals should be studied and assessed from a specific bilingual standpoint (Grosjean, 1997). The current study investigated the categorisation processes of bilingual children from differing levels of language dominance in both English and Welsh from the constructivist viewpoint. The review of the literature below was made in the context of justifying this approach for the study.

4.2 Bilingual linguistic development

This section reviews the evidence on semantic categorisation by bilinguals. Although bilingualism is a common occurrence from a world standpoint, early research into bilingual language acquisition did not necessarily have a positive view of the process, perceiving the

need to acquire a second language as potentially a handicap to the fluent acquisition of the dominant language. In the case of Welsh, a positive view of education through the medium of the Welsh language only appeared in the last fifty years. However there now appears to be greater interest in the successful acquisition of Welsh by young children. The results of a survey of 2,000 people with Welsh associations by Coupland et al (2006) identified language competence as a strong indicator of subjective affiliation to Wales.

Historically researchers have adopted differing language proficiency parameters in describing bilinguals. Bloomfield (1935) took the view that bilingual speakers must develop with similar levels of exposure of both languages, and be fluent and proficient to the ability of native speaker in both languages. Haugen (1956) proposed that speakers could be classified as bilingual if they could produce substantive utterances (no matter how small) in the less dominant language. Grosjean took a more moderate view, terming bilingualism as, "the regular use of two (or more) languages, and bilinguals are those people who need and use two (or more) languages in their everyday lives" (Grosjean 1992, p.51). According to Grosjean, the lexicon of a bilingual speaker would be more than just an amalgam of two monolingual lexicons. Simply having two different languages in the brain poses the question of how much carry over will there be from one language to the other.

Evidence has shown that the nature of different structures of a language influences the sequence in which those structures are learned (Choi 2000; Gathercole 2007). Complex structures are often acquired later than more transparent ones. For instance nouns are often learnt quicker than verbs (similar to the evidence produced by Clarke, 1987 for monolinguals). Similarly, as words are acquired within a context they will be acquired in a

piecemeal fashion.

The degree to which categorisation in one language is separate from the other and whether there are overlapping mechanisms poses a particularly interesting question (Gathercole & Moawd, 2010). In terms of dynamic interaction, there is also the issue of whether both semantic systems function in parallel or whether one shuts down to accommodate the other (Dijkstra & Van Hell, 2003). How far the system for one language can utilise the other requires an assessment of the interdependency of the cognitive mechanism used to categorise novel concepts in one language on the other language. It has already been established that semantic categorisation by monolingual children requires a complex cognitive effort. Does the existence of two semantic systems necessarily create an even more complex matrix for the child to navigate, or do the two systems support one another? Kroll & Stewart (1994) proposed that interaction between the dominant L1 language and L2 occurred at the conceptual level (although lexical interaction flowed from L2 to L1).

As individual languages map novel words onto categories according to language dependent requirements, it is not possible to exactly map between languages as a single process. Each language may have several words mapping onto a concept which are not comparable (Gathercole & Moawad, 2010). In relation to English and Spanish, Prior et al (2007) estimated that around 67% of words were potentially ambiguous in translation terms between languages. Where the concept only maps to one word in one language, the other language may produce two concepts with two accompanying words. For example, the Welsh word *dysgu* is represented by both *teach* and *learn* in English. Similarly, the English words *borrow* and *lend* correspond to one word in Welsh, *benthyg*.

4.3 The constructivist model

The constructivist model (Gathercole, 2007) assumes that the general patterns of the acquisition process in bilinguals for the two languages is a parallel one to that of monolinguals. For example, a bilingual Welsh/English individual will acquire the language specific structures for English in the same manner as an English monolingual. Although the timing of acquisition of these language specific structures will be dependent on the amount of exposure to input in that language. Evidence from Gathercole (2002) demonstrated that Spanish-English bilinguals acquired language specific structures such grammatical gender in Spanish, mass count structures in English, and that-trace in both languages, according to the amount of input received in that language. Bilinguals who received proportionally more Spanish input gained an earlier understanding of grammatical gender, whilst bilinguals who had received more English input had an earlier command of mass/count structures. It was also the case that bilinguals who had received a greater amount of Spanish or English still remained behind their monolingual counterparts. Similar findings were obtained for Welsh/English bilinguals, with language structures specific to Welsh e.g. grammatical gender acquired earlier when the amount of input in Welsh was greater (Gathercole & Thomas, 2005).

Research into the categorisation of collections of objects by bilingual Welsh/English children and monolingual English children demonstrated that language structure appears to influence the categorisation of novel words and that a greater understanding of language systems appears to improve with age. Eight-year-old Welsh speaking children took significantly longer to categorise than their English speaking peers. Eleven year old Welsh

speaking children categorised novel objects as collections more than English speaking children of the same age (Gathercole & Roberts, 2005). This finding did not occur with younger age groups. The finding indicates that there might be greater understanding of the individual language structure by children as a result of increased exposure (i.e with age).

The bilingual child appears to have the ability to switch off irrelevant information received for one language whilst accepting relevant input (Choi, 2006). The bilingual child appears to call on techniques such as overextensions to manage the complex cognitive process of acquiring two semantic systems. Their errors help them navigate their way as they develop their ability to make sense of new concepts (Choi, 2009). The constructivist model explains the visible success of bilingual children to eventually become proficient in more than one language to their ability to maintain two separate systems which are sensitive to the relevant input from language being accessed. Gathercole (2006) argued that a separation between semantic systems and cognitive knowledge enables the child to manage the complexity of language acquisition. Without this separation, the need to map two non-identical systems for the semantics of two languages with the associated social and cultural conventions would inevitably result in poor language acquisition with each language providing a distraction to the rapid understanding of novel concepts.

The constructivist model proposes that if any transfer is to occur between languages it is more likely to be present in balanced bilinguals who do not have a dominant language. This transfer would also occur at a later stage of development (Gathercole, 2007). The early bilingual may make necessary adjustments but will have the advantage of practising the cognitive function from an early stage. Whilst the late bilingual would benefit from a more

stable semantic and mapping system in the dominant language, this may then interfere with the required mapping of the L2 language. The constructivist model provides a perspective to further assess how monolingual and bilingual categorisation by children occurs as a separate but related process.

4.4 The advantages of two language systems for categorisation

An important consideration in relation to the question of how bilinguals are able to create an apparent monolingual language mode is whether there is cross language facilitation in all languages when an item is produced, or activation restricted to the specific target language. If the latter is correct, there would not need to be a control mechanism in bilingual speakers to inhibit the influences of other languages as they would have been activated. Language selectivity requires the speaker to repress the influence of another language to be able to respond correctly. One of the most notable abilities of bilingual speakers is that of being able to choose words from one language while preventing significant interference from the other language.

In this context, the fundamental question is whether the representations of the unused language affect the production process in the language in use. For example, if lexical activation is restricted to the representations belonging to the language in which the speaker chooses to use, there is no requirement to consider any control mechanism specific to bilingualism which operates over these lexical representations. There are many occasions in which bilinguals need to restrict access to the lexicon to only one language since the use of words from their other language may disrupt communication, given that the recipient may not

know the second language. In such circumstances, and given that the speaker is the one who decides in which language to carry out the communicative act, the semantic system may only activate representations of lexical items in the target language. In such a framework, the bilingual would be functionally equivalent to a monolingual and no control mechanism specific to cases of bilingualism that operates over lexical representations would be needed.

4.5 Evidence of language interference in bilingual categorisation

However despite the obvious benefits of restricting activation to one language, models of bilingual speech production have been presented which suggest that conceptual representations spread activation to the lexical representations of both languages of a bilingual (De Bot, 1992; Poulisse and Bongaerts, 1994; Costa, Miozzo & Caramazza, 1999; Costa, 2005). These studies have attempted to identify the type of controlling mechanisms which operate over lexical representations that guarantee that lexical selection is achieved in the intended language, while preventing increased interference from the non-response language. De Bot (1992) postulated an extended subsystem hypothesis where there is no systematic division between each language system leading to functional interference. The dominant language of bilinguals may influence the second language either functionally where items from the dominant language interfere with the use of the second language or representationally where the dominant language interferes at the point of acquisition of the second language (Paradis and Navarro, 2003).

Evidence in support of the co-activation of the two lexicons of a bilingual has been claimed from the presence of cross-language Stroop-like effects in bilingual contexts. These

were found in relation to the effects of distractor words on picture naming latencies. In these studies participants name a picture while ignoring a distractor word. Evidence showed that there appeared to be a semantic interference effect where picture naming latencies were higher when the picture (*dog*) appeared along with a semantically (categorically) related distractor (*cat*) than with an unrelated distractor (*cap*) (Rosinski, 1977; Lupker, 1979; Glaser & Düngelhoff, 1984). This effect has been interpreted as revealing the larger lexical competition created by a semantically related distractor word in comparison to an unrelated distractor word, and therefore has been used to explore issues related to lexical selection in speech production (Schriefers, Meyer & Levelt, 1990; Meyer, 1996; Caramazza & Costa,2000; Costa & Caramazza, 2002).

In a bilingual version of this procedure, participants named pictures in one of their languages while distractors were presented in the other language. For example, a Dutch–English bilingual was asked to name the pictures in Dutch (e.g. *hond* (*dog*)) while ignoring the presentation of distractor words in English (e.g. *rabbit* or *hammer*). The majority of these studies reported semantic interference (Miller & Kroll, 2002; Altarriba & Mathis, 1997). The explanation for the presence of such an effect across languages is based on the premise that the level of activation of the lexical node of the semantically related distractor word in the non-response language (*rabbit)* is larger than that of an unrelated distractor (*hammer*). This differential level of activation arises because the related lexical node (*rabbit*) receives activation from two sources (the presentation of the distractor word *rabbit*, and the activation sent by the semantic representation of the target word *hond* (*dog* in Dutch) while the unrelated lexical node (*hammer*) receives only activation from one source (the presentation of the

distractor word) (Costa et al, 1999). Therefore the presence of semantic interference across languages was proposed as demonstrating that the two lexicons of a bilingual are activated in parallel during the course of language production and that the lexical representations of the non-response language compete during the selection of the target representation in the response language.

4.6 Evidence for cross-language interference

However it is also necessary to identify whether there are alternative explanations for the presence of the interference effect. The presence of cross-language interference does not necessarily originate because of the competition produced by the lexical representations of the non-response word (*rabbit*) in the selection of the target word in the response language (*hond)*. Instead, it may be exposing competition within the response language rather than between languages (Costa et al., 1999). If the distractor word presented in the non-response language (English: *rabbit*) is automatically translated into the response language (Dutch: *konijn*) then selection of the target name in the response language (*hond)* might be hampered by the high level of activation of the lexical node of the distractor in the response language. It is therefore not the case that *rabbit* competes with the selection of *hond* but rather its translation in the response language, *konijn*, does. Using this explanation, semantic interference across languages does not require an assumption either that there is parallel activation of the two lexicons of a bilingual or that there is lexical competition between the two languages of a bilingual.

Several researchers have questioned the notion that the semantic interference effect

actually indexes competition between lexical representations (e.g. Rosinski, 1977; Costa, Alario & Caramazza, 2005). Instead, it can be argued that this effect may have its origin at the level at which semantic representations are selected for production. If this interpretation of the phenomenon is assumed to be correct, the results can not be confirmed as indexing lexical competition either within language or across languages. It is also possible that the experimental context could affect the representations of the non-response language which are activated in the course of producing the target language. The non-response language could already be activated from the word recognition system, and hence it is difficult to separate this source of activation from that coming from the semantic system (Grosjean, 1998).

Costa and Santesteban (2004) attempted to address the significance of variables for speech production. For example, variables such as the similarity of the two languages of a bilingual, the age (and manner) at which L2 has been acquired, the proficiency achieved in L2, the recency and frequency of use of the two languages, and the discourse topic may affect whether or not the two languages become activated in parallel even in monolingual contexts. They argued that the control mechanism which guarantees lexical selection in the target language is dependent on the L2 proficiency of the bilingual speakers. They further argue that there is a qualitative shift from a reliance on inhibitory control to a reliance on language-specific selection mechanisms that is intimately tied to an increase in the L2 proficiency.

The influence between L1 and L2 is not equally bi-directional, with cross-language influences being much stronger from L1 to L2 rather than L2 to L1 (Costa and Caramamazza, 1999). This finding is corroborated with evidence (Kroll, Michael,Tokowicz, and Dufour, 2002) which showed that generally participants responded quicker and with greater accuracy

when responding in tasks in L1 rather than L2. This research addressed the issue of whether the same meanings are accessed for translation equivalents in each language and also whether bilinguals and second language learners are capable of accessing meaning independently for L2 words without mediation through L1.

4.7 Relationships between the languages

The concept that words in languages relate to discrete lexical meanings may result in different meanings in individual languages with the subsequent effect on the relationship between those languages. However there is currently a lack of research evidence to indicate the level at which such differences in meaning are found. It is assumed that bilinguals use resources from both languages when undertaking complex high level cognitive processing. However, accessing common meaning does not imply that knowledge about which language was actually used to communicate that information is necessarily lost. Recent autobiographical memory studies suggest that language is a potentially powerful cue for memory retrieval (Schrauf & Rubin, 2003).

Cross language priming has been observed during early cognitive processes for semantic priming. L1 words being more likely to prime L2 words than the other way round (Keatley et al, 1994). This asymmetry in priming supports the projections of the hierarchical memory model where L1 words are more likely to stimulate their respective meanings than L2 words as more effective primes (Kroll & Stewart, 1994). In relation to this aspect however, competence in L1 and L2 appears to be an influence on semantic priming. The usual scenario is for the bilingual speaker to have less ability in L2 which in turn could influence the

interaction between languages. The evidence that semantic representation is held in common across the two languages has generally been based on picture representation tasks. This type of activity inevitably restricts the words used to those relating to objects. De Groot (1992) proposed that even if concrete objects shared the same conceptual meaning in both languages, it did not necessarily follow that more abstract concepts are shared. The evidence suggested that bilinguals performed consistently faster on word translation and word association tasks for more abstract concepts (De Groot, 1992).

As bilinguals receive input in two languages, any measurement of lexical development requires investigations across both languages. Single word vocabulary tests do not provide sufficient data to measure performance (Crystal,1997). A functional perspective would assume that semantic categorisation would vary between the two languages as learners receive a variable input from their environment. A child speaking mostly one language at home, yet another at school should demonstrate differing levels of performance in each language depending on their proficiency and exposure levels (Pena, Bedore & Zlatic-Guinta,2003; Gathercole, Thomas & Hughes,2008).

A study into the acquisition of English as L3 in schools in Barcelona where Catalan and Spanish were spoken investigated the interference on Catalan as L1 (Munoz et al, 2002). The results demonstrated that younger learners appeared to draw more heavily on L1 than older learners. Other studies in Spanish bilingual communities appear to show that younger learners use more borrowings than older learners (Celaya & Torras, 2001; Gost & Celaya, 2004). Cenoz (2003) suggested, in contrast, that the oldest learners appeared to transfer the most to L1. A study with teenage Dutch learners found that EFL learners transferred more

function than content words from their first language in an oral task. It concluded that low proficiency learners tended to transfer more elements from their first language than more advanced learners (Poulisse and Bongaerts, 1994). However, Sanz (2000) found the opposite result that the more proficient students transferred a greater number of borrowings.

Cenoz (2001) reported that the youngest and oldest children who were bilingual Basque/Spanish learners of English as an L3 transferred more content than function words but the differences were not so apparent in the middle age range of those tested, where similar numbers of content and function words were transferred. Some studies have suggested that monolingual subjects perform significantly better than bilinguals in specific tests, for example, the Boston Naming Test (Kohnert et al.,1998). Cenoz (2000) and Sanz (2000) concluded that the more balanced bilingual learners performed better in L3.

An investigation into Russian speakers who learned English concluded that the influence of their first language could be seen in oral responses in English and also that English structures began to influence the first language. The study acknowledged that the types of transfer are not uniform in the two languages (Pavlenko & Jarvis, 2002).

Two studies to investigate whether different languages map words onto referents in differing ways attempted to answer the question of whether bilinguals maintain separate mappings in their individual languages (Ameel et al, 2005). In an initial crosslinguistic study, monolingual Belgian French and Dutch speakers were studied using word to referent mappings for common words used in the home. The results were compared with a further study of how French/Dutch bilinguals mapped similar words. The results appeared to show that the bilingual speakers adopted a common naming pattern with the category boundaries in

the different languages moving closer together. This suggests that the bilingual speakers therefore do not keep the separate category boundaries distinct when processing the two languages.

4.8 Nesting tendencies

An interesting area of study has been progressed over recent years led by Pavlenko who outlined the concept of nesting to address how bilinguals and monolinguals potentially differed in managing their semantic systems. The existance of partial (non) equivalence relationships, of which nesting is one, provides explanations for the overlap between categories (Pavlenko, 2009). Nesting is expressed as how two or more categories in one language are placed into one category in the other language. Nesting can be expressed as a smaller category in one language being subsumed in a larger category in the second language. For example, the English category *jar* is seen to be nested within the category *frasco* in Spanish. This also includes items which would have been categorised as being *bottles* and *containers* in English (Malt et al., 2003).

Interesting results have also been reported for terms related to emotion. A study of naming tasks by monolingual Russian, monolingual English and bilingual Russian-English speakers for *revnost* (jealousy) and *zavist* (envy) found that monolingual Russians differentiated between the two terms whilst monolingual English speakers accepted both descriptions as appropriate for the term *envy*. Thus Russian speakers are clear about the boundaries between the two concepts whereas English speakers accept less distinct boundaries. Interestingly, bilingual Russian-English speakers followed the convention of

Russian speakers for the terms in Russian (that is they differentiated between *revnost* and *zavist*), but for English they chose only *envy* as a term to categorise both scripts (Stepanova and Coley, 2002).

The concept of nesting therefore suggests that bilingual speakers need to be able to make "fine grained distinctions" about linguistic markers which may be required in one language but not the other (Pavlenko, 2009). These distinctions may be to follow the conventions established by monolinguals in each language or to relate to the larger category in the dominant language as in the case of English-Russian bilinguals in the U.S.A. who were observed to be losing the distinction between *sinij* and *goluboj* to denote the colour *blue* in Russian due to the dominant influence of the use of the single term in English (Andrews, 1994).

4.9 Chapter summary

The challenges posed for research into bilingual categorisation revolve around the perceived complexity of semantic systems and cognitive function in more than one language. The constructivist model suggests that this complexity is managed through a separation of semantics and cognitive knowledge for each language. The two languages maintain separate spaces which are able to focus on the categorisation requirements of individual languages although bilinguals appear to use techniques identified as being used by monolinguals to rectify errors and categorise related concepts. Interesting evidence has emerged from previous studies of interference by one language when categorising in the other and that differing types of categories may influence this interference. Classical categories where members share essential features are more likely to be closely located whereas radial categories by their

nature, in not possessing all the essential features for membership may be more scattered. The consequences for potential cross-linguistic influences on a bilingual's semantic system may be projected to be that for the categorisation of classical categories which are conceptually close, the language may highlight the similarities between members and play down the differences (Gathercole & Moawad, 2010). Whereas for radial categories which are projected to be linked in multiple 'chains' (Lakoff, 1987), the link will be restricted to particular aspects which the related members share. A counter view was proposed by Elston-Guttler & Williams (2008) which argued that as young children who are bilingual are able to access the conceptual space at a very early age, constructing categories in a shared way across languages when needed could provide synergy (assuming that such a shared space exists). In such circumstances the shared process is more likely to occur when constructing classical categories where members are closer than when constructing radial categories (Gathercole & Moawad, 2010).

Chapter 5. The Acquisition of Word Meaning in a Bilingual Context

5.1 Introduction

Building on the evidence identified in Chapter 4, this chapter investigates further the management of the two semantic systems. For example, how do the lexicons for each language develop and how does this binary structure influence the bilingual child's perception of the world. The construction of a detailed mental lexicon is a key developmental milestone for language learners of all ages in their comprehension of the meaning of new words (Aitchison, 2003), with the categorisation process being a core aspect of the development of word acquisition (Elman, 2004).

Cross-linguistic and multi-lingual studies allow for a better understanding of lexical acquisition by children through contrasting and comparing the experiences of monolinguals and bilinguals. Adopting the constructivist model approach, bilinguals need to learn how to map a specific concept in each language onto labels for each individual language. As was observed in Chapter 4, these labels sometimes overlap in meaning, although not on all occasions. The cognitive knowledge required to manage this process will need to take account of the relationship between individual languages, starting with the issue of which language is dominant and how does the acquisition of the primary language influence the acquisition of the second language. The investigation into the degree to which the processing of words is done as an independent process or part of an integrated system involving both languages, is an essential component in achieving the main goals of the study. Also relevant to the current study is the issue whether there is any effect of language dominance on transfer (i.e. the effect of the non-test language on the processing of an item in the test language of the study), and

developmentally, whether an increase of exposure in T2 affects this transfer.

5.2 Language dominance

Language dominance has been examined by means of tests which measure the individual's performance in each of the two languages (Abrams, 2000), or by self-reported ratings in which the subjects assess their own abilities in both languages (Flege et al., 2002) or by language background scales, which extract bio-data of the subject's language use (Baker & Jones, 1998). Three factors - frequency, ambiguity and language dominance were analysed to explain transfer as reflected in the form of compound reversals by bilingual French/English children (Nicoladis, 2002). Results showed that language dominance did not affect the number of reversed compounds used in the non-dominant language, however, the balanced group reversed more French compounds than English compounds.

Other studies have identified specific examples of where the dominant language influenced word acquisition by the bilingual child. Evidence showed an effect of language dominance on French-dominant children's prosodic structure in English words (Paradis, 2001). Costa and Santesteban (2004) argued that differences may exist in the type of mechanisms involved in lexical access for low and high proficient bilinguals.

De Houwer (1990) concluded that bilingual children are able to differentiate between the two languages both at the syntactic and pragmatic level from an early age. However, Muller (1998) suggested that the separate development of bilingual children's grammar does not preclude the possibility that the two languages are in contact and therefore have an influence on each other. It is therefore implied that the two linguistic systems are separate but

non-autonomous. Similarly, the identification of the factors which determine cross-linguistic influence as well as its locus and direction are major issues in research on bilingual children, as studies need to demonstrate that cross-linguistic influence is a controlled and systematic phenomenon which does not occur at random. On the one hand, Paradis and Genesee (1996) claimed that dominance, an external factor, is the crucial factor that determines cross-linguistic influence and thus, the dominant language influences the weaker language. This view was supported by two other studies relating to an Italian-English bilingual child (Serratrice & Sorace, 2002), and other relating to a Spanish-English bilingual child (Paradis & Navarro, 2003). Both studies explored the distribution of overt and null subjects in Italian and Spanish, which is also a grammatical domain that belongs to the syntax-pragmatics interface, and it was found to be vulnerable to cross-linguistic influence from English on Italian and Spanish respectively.

According to Hulk and Muller (2000), cross-linguistic influence may be due to internal sources where two conditions occur. Firstly that the grammatical structure involved should belong to the syntax/pragmatics interface, and secondly, a certain syntactic construction in language A allows for more than one syntactic analysis from the perspective of child grammar, and language B provides evidence for one of the two analyses.

The phenomenon of object drop, which belongs to the syntax-pragmatics interface, is susceptible to cross-linguistic influence in the French and Italian grammar of three bilingual children simultaneously acquiring pairs of Romance and Germanic languages (German-French, Dutch-French and German Italian). In those cases language dominance could not explain the linguistic data since the target of cross-linguistic influence was the dominant

language (Romance languages) and not the weaker one (Germanic languages) (Müller& Hulk, 2001). However, the vulnerability of the syntax/pragmatics interface to cross-linguistic influence in bilingual language acquisition was not clearly shown, as only the frequency of object omissions was counted and there was no investigation of the pragmatic contexts in which those omissions took place. In this way, the pragmatics end of the interface was not tested.

Bates and McWhinney (1982) proposed the competition model for language acquisition which provided assumptions for bilinguals. They argued that bilinguals would interpret input using cues which were strengthened if they were similar in both languages and were weakened if both languages differed. Four strategies were identified - differentiation where different strategies were used for both languages, strategies where those for L1 are used for L2, strategies where those for L2 are used for L1 and amalgamation where one set is used for both languages.

Other models, for example, Long's interaction hypothesis (Long, 1983) highlighted the role of interaction in second language acquisition. Swain proposed the output hypothesis to emphasise a perceived increase in performance in bilingual language acquisition when it is accompanied by an increase in language use (Swain, 1995). The minimal trees approach suggests that initially only lexical categories are transferred from L1 with functional categories developed later in L2 (Vainikka & Young-Scholten, 1996). A refinement of this approach, the modulated structure building hypothesis suggests that functional projections from L1 transfer to L2 but only when syntactic structures are available (Hawkins, 2001). More detailed studies suggest that children quickly recognise differing requirements for each

language. Babies of 4 months with access to both languages were able to differentiate between Spanish and Catalan although both languages are prosodically similar (Bosch & Galles, 2001). When young bilingual children initially produce two word and multi-word utterances there is a difference in word order usage (Meisel (1997) in relation to French-German, Deuchar, (1992) in relation to Spanish-English). Although these studies introduce interesting theories in relation to language dominance, they do not fundamentally challenge the constructivist view that language is learned is a similar manner to that of a monolingual and is dependent on input for acquisition, rather than being boosted by any influences by the dominant language.

5.3 Transfer between languages

Language dominance theories offer an explanation for some instances of cross-linguistic transfer in non-overlapping morphosyntactic structures (Nicoladis, 2002). This transfer is proposed to occur when the language structure in the dominant language is more transparent (and is easier to learn), and is therefore learned at an earlier stage than the corresponding structure in the second language. Some evidence has been found suggesting that transfer is more likely to be seen when bilingual children use structures from their dominant language in their weaker language. Cantonese/English bilingual children who were dominant in Cantonese were found to use Cantonese structures in their use of English (the weaker language) (Yip & Matthews, 2000). The asynchronous development of the two languages may also be a factor with structures being acquired at different stages in the individual languages. Yip & Matthews (2003) noted that the acquisition of structures may

take place earlier in the weaker language than would be the case for the normal developmental period. This is due to the transfer from the dominant language structure which normally acquires those corresponding structures earlier.

However as identified in Chapter 4, Gathercole (2007) proposed that this was more likely to occur in balanced bilinguals at a later stage of development. Van Hell and Dijkstra (2002) found that the influence of L2 on L1 increased as proficiency in L2 increased. Their experiment using written stimuli tested trilinguals with Dutch as L1, English as L2 and French as L3. The differing proficiencies enabled assessments to be made where the L1 target words were cognates with the English and French translation terms or with non-cognates. The findings demonstrated that trilinguals who were very proficient in L2 and less proficient in L3 showed shorter word association and lexical decision times for the L1 cognate terms with L2 than for non-cognate terms. The response times for L1 terms which were cognate with L3 were no different than the response times for non-cognate terms. For those trilinguals who were equally proficient in L2 and L3, the response times were faster for L3 terms which were cognate with L1 than for non-cognate terms. They concluded that "the language processing system of multilinguals is profoundly non selective with respect to language" (Van Hell & Dijkstra,2002:786). The stimulation of the L1 terms activates words from the non native language as a parallel process. Therefore even when the experiments only focused on the native language, the results demonstrated that the participant's ability in another language was a factor. The study's findings therefore challenge Grosjean's (1997) proposal of language selectivity in the language mode, as control mechanisms believed to suppress the weaker language to create a monolingual language mode were being over-ridden, i.e. the weaker

language was activated.

Many proposals of bilingual representation have been put forward to attempt to address the intricate relationship between words, events, images and concepts (Haritos & Nelson 2001). Paivio's (1986) dual coding theory assumed that words are not merely encoded as verbal representations but as two subsystems, one responsible for the verbal information and the other in control of processing non-verbal information (imagery). The words relating to specific terms are accessed in one system by both languages although the two words for the specific concept are not always linked to a common referent. However abstract terms do not share a common referent and as a consequence are kept separate in the processing system. This is further discussed in the section 5.4 below. Paivio's work is particularly relevant for this current study in offering a plausible explanation for maintaining separate systems for each language.

5.4 Bilingual memory

Research results demonstrate that bilinguals use a complex set of relationships to achieve an efficient fluency of language use and understanding in both languages, including the effective management of memory processes. Grosjean (2001, p.3) defined language mode as ".. the state of activation of the bilingual's languages and language processing mechanisms at a given point in time". Since a bilingual will have access to two lexicons, one topic which has been of particular interest is whether bilingual speakers process their two languages separately (independent) or together (interdependent). A related issue is whether different languages are stored in separate or the same storage system. If languages are kept in the same

system this implies that information on events and objects is stored and processed as being general concepts rather than being language specific entities.

The common store hypothesis was proposed by Lopez and Young (1974) as result of studies conducted into transfer effects for translation equivalents for Spanish-English bilinguals. Individuals were shown a list of either Spanish or English terms. Three cohorts were created. One group learned the list presented; the second was given the translated terms of the list and the third group a new list of terms. When asked to provide a recall to the words presented aurally, the results demonstrated that there was a positive transfer effect for both English and Spanish which was uniform for both languages. Corroborative findings were found by Kirstner (1984) who identified that when bilinguals were presented with a stimulus in one language, this influenced their response to a corresponding stimulus in the other language. Similarly Green (1986) identified that bilinguals activated terms in both languages with the language not in use being suppressed.

In contrast, the independence hypothesis proposes that there are discrete memory stores for each language (Kolers,1963). As a consequence, any processing of one language does not necessarily affect the processing of the other language. Several different experimental paradigms have been proposed by researchers investigating the complexity in the workings of bilingual memory. The research paradigm adopted for this study concentrates on recognition and recall, word naming and association and language transfer and inference. Studies by Kolers (1963) led to two differing theories which can be described as the interdependence hypothesis where a common storage model is envisaged and the independent hypothesis where a bilingual uses a separate storage model.

The independent hypothesis suggests that individuals described as balanced bilinguals should report as monolinguals for each language as the two languages are stored separately. In addition, there should be no evidence for language transfer between languages. In contrast, the interdependence hypothesis should report findings that bilinguals demonstrate interlingual behaviour.

The dual coding theory proposed by Paivio (1971) is interesting in this context. According to this hypothesis a persons memory contains two systems, one to process verbal information and the other to process imagery or non-verbal information. As both are required to process knowledge there would need to be interconnection between the two systems even though they are perceived to be separate. To connect between the verbal and non-verbal systems, one system would need to be able to activate the other. Therefore words should generate images, and images should be capable of being named when recall is required. Paivio's theory suggests more than one level of processing. Initially a stimulus, which could either be a word or an image, stimulates a verbal representation or an imagery representation. This is followed by a referential process where there could be an interconnection between word and image to name the object. Finally an associative process examines connections between stored verbal representations and stored imagery in each storage system.

A further extension to this hypothesis was developed by Paivio and Desrochers (1980) to explain processing by bilinguals (Bilingual Dual Coding Theory). This approach postulates three representational systems, two verbal systems for each language and a third imagery system. The theory suggests that, as with the dual coding hypothesis for monolinguals, the three systems are separate and only interconnect at the referential level. Thus bilinguals can

remember non-verbal events and objects with interference from either verbal language systems. This bridging hypothesis appears to reconcile the opposing arguments for the shared and separate storage models.

Paivio and Lambert (1981) tested the hypothesis on a cohort of balanced English-French bilinguals. The findings supported the theory of independent verbal systems in that bilingual coding resulted in an effect on recall which was additive. Sung and Padilla (1996) further developed the hypothesis with a study of unbalanced bilinguals which included English-Korean bilinguals and English-Spanish bilinguals. The findings again supported Paivio's original hypothesis. These studies suggest that linguistic differences and similarities do not affect the basic premise of separate representational systems in bilingual memory.

In terms of morphosyntax, bilingual children appear to learn languages structures specific to their individual languages at the same time as their monolingual peers (Deucher & Quay, 2000). Paradis and Genesee (1996) noted that three year old bilingual English/French children acquired finite verb forms earlier in French than in English. This pattern of acquisition mirrored what would be expected of monolingual French children. This finding is line with the expectation of the constructivist view.

Repetition priming studies have investigated whether the processing of a target word facilitates the processing of the translation of the word. The assumption here is that if repetition priming is found across the languages then some underlying representation is shared. De Grot and Nas (1991) reported across language facilitation of lexical decision in some cases for Dutch-English bilinguals. Korean-English adult bilinguals demonstrated some priming effects for object terms but not abstract terms suggested a shared storage

representation for abstract terms.

The hierarchical hypothesis proposed that words are stored in separate languages at the lexical level although sharing a common representation at the conceptual level (Potter, So, Von Eckhardt & Feldman, 1984). The hypothesis draws on the inconclusive results for a shared store and a separate store model. The findings from experimental studies suggest that different processes may have been tested by individual studies (Gerard & Scarborough, 1989). The separate store hypothesis tends to be supported by tasks which highlight surface attributes whereas the common store hypothesis tends to be supported by tasks with semantic attributes.

The types of words used in bilingual memory tasks may influence the subsequent findings (Taylor and Taylor, 1990). Studies have a tendency to use frequent terms for common objects which have identifiable equivalents in both languages. In experimental conditions, findings showed that distinct objects and culturally similar terms appear to be stored in a common store with culturally distant words represented in a separate store.

There has been an established effect of interference in monolinguals when taking part in tests of the picture naming paradigm. Monolinguals are slower in naming a picture if it is paired with a distractor word which is associated categorically to the picture compared to a distractor which is unrelated (Glaser & Glaser,1989). This finding potentially demonstrates that the selection process incorporating semantic interference is a competitive one (which was previously discussed in section 4.5). Equivalent translation words (e.g. *bwrdd* and *table*) are associated even closer than related words within a language (e.g. *table* and *chair*). Costa and Caramazza (1999) in studying Spanish-Catalan bilinguals identified opposite findings. The authors found that "the interference effect reverses completely and translation equivalent

distractors lead to faster picture naming times relative to unrelated distractors" (Costa & Caramazza,1999:234). They proposed that this outcome was as a result of the equivalent distractor working as a prime across languages, and only lexical representations of the same (target) language were considered by the lexical selection mechanism. The results showed that the semantically-related distractors from both languages (e.g. for naming a table in Catalan; *Silla* -Spanish for chair, and *Cadira* - Catalan for chair) created an equal amount of interference (Costa & Caramazza 1999). The results of the study indicated that bilingual lexical involves the activation of both languages when a child speaks, but there is no competition between languages during lexical selection in speech production. This however does not explain why there are interference effects by categorical related distractors in monolinguals.

The issue of whether language proficiency and the age of acquisition influences parallel language activation was investigated by Silverberg and Samuel (2004) using an L1 lexical decision task. Spanish/English bilinguals were initially assessed to identify ability levels for speech writing and reading for both languages. The findings indicated that bilinguals with a high proficiency in L2 demonstrated L1-L2 connectivity at the form level although semantic representations did not appear to be shared. However, bilinguals with low L2 proficiency did not demonstrate shared semantic representations or have L1-L2 connectivity at the form level. Conceptual representations have also been studied on the premise that if there is shared representation at the conceptual level then semantic priming should be seen across L1 and L2 (Altarribba, 1992; Schwanenflugel & Ray, 1986). For example, a bilingual with L1 English and L2 Spanish, the prime L2 *cuerpo* (L1 *body*)

stimulated the response *arm*. Such results suggest that there is some shared representation at the conceptual level. The findings of a study of German/English bilinguals by Blumenfeld and Marian (2005) using eye tracking also supported the view that the parallel activation of the two languages is dependent on the relevant proficiency levels in L1 and L2.

The influence of a translation strategy (L2 prime translated to the L1 equivalent leading to priming from the activation of the L1 word) as a potential explanation for shared representation at the conceptual level was addressed by Keatley, Spinks and de Gelder(1994). Taking into account the influence of translation strategies, De Groot and Nas (1991) proposed that semantic priming for L1- L2 depended on cognate status.

5.5 Chapter summary

Further investigation into how bilinguals manage their semantic systems identified language dominance as a specific factor which appears to influence the management of the two systems. Studies have demonstrated however that although the dominant language can impact on the acquisition of words in the second language, the bilingual child also appears perfectly capable of understanding that categorisation mapped onto words is different for each of the languages. This is in line with the constructivist view used for this study.

The second factor which emerged from studies highlights the role of memory in the management of the semantic systems. Although this study has identified research which propose shared stores, they do not provide compelling evidence which challenges the constructivist model. Observing the rate of retrieval of individual words provides an opportunity to assess whether the retrieval of related categories is more efficient than the retrieval of unrelated categories (as proposed in Chapter 4). Having assessed the evidence

from previous research, the next Chapter will develop the goals of the current study, focusing on the particular circumstance of Welsh/ English bilingual children.

Chapter 6 The Goals of the Study

6.1 Introduction

The specific circumstances of Welsh/ English language acquisition by children have not been the subject of many previous studies. An overview of research at doctoral level into second language acquisition completed in Welsh universities between 2003 and 2008 identified a lack of studies on Welsh language teaching (Fitzpatrick, 2010). This situation represents a deficit in an important research area which the goals of this study begin to address. The study is able to build on the significant advances which have been made by previous studies to better understand semantic systems and related cognitive knowledge, as outlined in earlier chapters. One aim of the critical review of existing literature was to identify a relevant and suitable model which could assist in developing the study design. I considered that the constructivist model (Gathercole, 2007) provided a robust approach which particularly addressed the issues associated with semantic organisation by early bilinguals. Therefore as the study proposed to investigate the performance of this target group, the constructivist model was adopted for the study design.

The goals of the study were to investigate the semantic categorisation of actions and objects by Welsh/ English bilingual children aged between 5 and 11 in order to understand the role of semantic structures and cognitive knowledge in the categorisation of word categories. The study provided an opportunity to test the existence of two semantic structures as proposed by the constructivist model with the general patterns of the acquisition process in bilinguals for the two languages as a parallel process to that of monolinguals. For example, according to the model, a bilingual Welsh/English individual should acquire the language specific

structures for English in the same manner as an English monolingual. As the model proposes that the timing of the acquisition of these language specific structures will be dependent on the amount of exposure to input in that language, the study allocated participants to groups of differing home language backgrounds.

6.2 The Welsh bilingual context

According to the 2001 census, 20.8% of people living in Wales are bilingual Welsh/English. However this is not a uniform distribution across Wales. The percentage of Welsh speakers varies markedly from 69.0% in Gwynedd in north west Wales to 9.3% in Monmouthshire in south east Wales. 37.0% of children between the ages of 3-15 living in Wales are Welsh speakers. Children have the opportunity to acquire the Welsh language within four main home contexts: firstly homes where only Welsh is spoken, secondly homes where both Welsh and English are spoken, thirdly where only English is spoken and fourthly where another language is spoken. Research has shown that when both parents had themselves been brought up in a mainly Welsh home environment, 96.3% used only Welsh in speaking with their children (Gathercole & Thomas 2009).

Research activities into the acquisition of the two languages by Welsh/English bilinguals which can assist with identifying the scope for more detailed study are mostly confined to the experience of the last twenty years. The results of a study in early word recognition by bilingual Welsh/ English infants at 11 months suggested that the word acquisition process was different depending on the language or languages spoken to the child (Vihman et al, 2007). For the dominant language, the time taken for acquisition appears to be

similar to the monolingual context. However for monolinguals hearing only the minority language, there was a delay. This suggests a complexity in the acquisition process which may be related to the particular form of the Welsh language (Vihman, 2007).

6.3 The social context

Recent policy initiatives such as the Government of Wales Act 2006, which includes a requirement by the Welsh Assembly Government to protect and promote the use of Welsh, build on earlier initiatives, for example, the establishment of Bwrdd yr Iaith Gymraeg (Ihe Welsh Language Board) in 1988. The Board is the statutory body to promote the use of the Welsh language. Following a high profile political campaign, a Welsh language television channel, S4C, was established in 1982 to provide a mass medium for the language. However the pressures on the Welsh language in a world where the English language is a dominant global language continue. The Internet provides new opportunities for minority languages, however it also represents the latest mass medium which is dominated by the English language. As a result, children learning Welsh are increasingly exposed to the English language in Wales which potentially may have implications for the findings of the study.

Historically the Welsh language was not taught in schools until the middle of the twentieth century. In 1990 Welsh became a compulsory subject for all pupils in Wales at Key Stages 1,2, and 3 (up to the age of fourteen). In 1999, this was extended to Key Stage 4. As a consequence all pupils in Welsh state schools learn Welsh either as a first or second language from five years to sixteen years of age. In 2002/2003 over fifty one thousand children were taught in classes where Welsh is the main medium of instruction. The majority of these

children were taught in 448 primary schools. In parts of Wales, 95% of pupils in Welsh medium schools come from homes where only English is spoken (Welsh Language Board, 2007). Increasingly Welsh medium schools have adopted Welsh immersion teaching techniques to support children from English only backgrounds. As a result older primary aged children from English backgrounds attending Welsh medium primary schools will have received a greater exposure to Welsh than their older peers. Research has shown that the frequency of exposure to a language in the home and school context is a significant factor until a critical mass of understanding is achieved by the child (Gathercole & Thomas 2005).

6.4 Language transmission

The evidence to demonstrate how Welsh is transmitted in the home is mainly contained in studies funded by the Welsh Language Board on behalf of the Welsh Assembly Government (Jones & Morris, 2005; Gathercole ed. 2007). An increasing number of households are represented by only one parent speaking Welsh. Jones & Morris (2005) found that although the Welsh-speaking parents usually spoke Welsh to the child, both parents when together would speak in English in front of the child. The likelihood of the child speaking Welsh to siblings and parents appears to depend on the language of the dominant parent (Jones & Morris, 2005). However even where there is an English speaking parent, attitudes to the children learning and speaking Welsh are generally positive (Gathercole, ed, 2007).

Evidence for a link between socio-economic status and the educational background in Welsh-speaking and bilingual households is limited, although a study identified that parents' education and profession were not influential factors in determining where they spoke Welsh

to their children (Gathercole, ed, 2007).

6.5 Home language background vocabulary tests and questionnaire data

6.5.1 Introduction

As outlined in 6.2, the use of Welsh and English in areas of Wales where both languages are present is complicated by several factors which include social attitudes to the languages as well as linguistic competence. The wide variance in the percentage of bilingual speakers, even in areas traditionally perceived to be Welsh-speaking such as the north and west Wales results in communities and households with a differing linguistic mix. Increased social mobility and improved communication routes such as the A55 expressway in north Wales have resulted in major changes to the linguistic pattern of long-established communities as monolingual English families settling in areas which previously had a high percentage of Welsh speakers (Jones, 2003).

The use of a questionnaire included in the study, containing questions relating to the language context in which children acquire language, was used to facilitate placing children into the specific language background groups, and to further investigate any socio economic and/or educational link with language uptake (Gathercole, ed. 2007). Although the questionnaire and vocabulary tests were intended to be compared to identify whether there was a link between home language environment and vocabulary ability, these were not directly utilised to investigate categorisation behaviour in children.

Lyon conducted a survey using two questionnaires to investigate parental influence on language use in families in Wales where a second language was usually spoken (Lyon,1991).

The survey specifically examined individual differences between maternal and paternal influence on the language used by the child over a period of time. In cross language marriages the study found that there was a difference in the language spoken to the child depending on which parent was the Welsh speaker (first language Welsh speaker). Where the mother was the Welsh speaker, Welsh represented 72% of the language spoken to the child by its mother. In comparison, in couples where the father was the Welsh speaker, Welsh represented only 46% of the language spoken to the child by its father.

The results of this survey suggested that the mother's language was more influential than the father's in determining which language the child acquired. However, later research established that in cross language couples both parents' language background was equally influential (Gathercole, Thomas, Williams & Deuchar, 2006). The only exception was the youngest children of couples with a Welsh mother and English father where the mother's language possibly had more influence as a direct result of greater exposure when the mother spent more time with the child. In addition, no difference was found between mothers and fathers whose home language was Welsh in the amount of Welsh spoken to their children. The only group where the partner's language could influence the choice of language spoken to the child was of parents who themselves had come from homes in which both Welsh and English were spoken.

In that same study, influential factors in the transmission of Welsh to children, were found to include the parent's own actual or perceived level of ability in Welsh. However parental profession and education were not found to be influential in this process. This appears to be a surprising result considering that higher levels of education are often

perceived to be associated with bilinguals in Wales (Jones, 2003).

The current study aimed to test whether parental education or profession or vocab ability was a significant factor in the performance of categorisation tasks across the four language groups.

6.5.2 Bilingual questionnaire

A bilingual linguistic questionnaire was given to the parents of each child participant to assess in which language background group each child should be placed. The questionnaire was also designed with the intention of obtaining a detailed picture of the language transmission practices between parents and children. Specific questions noted the amount of exposure children received in both Welsh and English. Initial questions captured early language interaction between the child and adults. The starting age when a child began to speak English, Welsh or any other additional languages along with other personal information regarding date of birth, gender, place of birth, and the child's school year was noted. The percentage which corresponded to the amount of time a child spoke each language along with the percentage of how much time each language was spoken in the home was recorded. Developmental changes of individual language use in the home between the child and adult or siblings were captured for each two year intervals from birth to age 10. Participation in external activities involving both social organisations which are purely Welsh medium such as the Urdd and Eisteddfodau and potentially bilingual clubs and societies (sports clubs, music, scouts, girl guides) was included as a specific question.

Specific data relating to parental educational attainment levels and profession for

participating children was also collected, as well as the languages which parents spoke. This included details of the period of language acquisition, divided into, as long as they could remember, before starting school, during primary school, not until secondary school and whether they learnt the specific language as an adult.

6.5.3 Monolingual questionnaire

A questionnaire was distributed to the parents of participating monolingual children who attended schools in England to collect data on parental educational attainment and profession and to record any other linguistic input from parents in languages other than English.

As noted previously, the initial importance of the questionnaire was to identify the language background of children for the main categorisation task of the study. Although the questionnaire also enabled an investigation to be made into the findings of previous studies (Lyon,1991; Gathercole, ed. 2007)) relating to factors affecting a child's linguistic exposure and use. The analysis looked at the language spoken by mothers and fathers to their children depending on their language dominance. It was also investigated whether the predominant language spoken in the home had a bearing on the language spoken to the child regardless of language dominance of the parent. The language spoken in the home was also compared with parents' perception of the linguistic abilities of their children. The data obtained from the questionnaire was also compared with child's performance on the British Picture Vocabulary Scale (BPVS) and Prawf Geirfa.

6.5.4 Evidence from the home language questionnaires

The home language questionnaires provided an additional view of the socio-economic context of the children who participated in the study. Correlations between the responses of the adults/parents of children's to questions posed in the home language questionnaire along with the raw scores of both vocabulary tests. The profession of the mother was correlated significantly with the profession of the father (r =.405;p.001) and with the education level attained by the mother (r =.231;p.002). At the higher level, there was a contrast between the education level for example, mothers of children from monolingual English homes 23.3% had post graduate qualifications compared to 13% of mothers from only Welsh homes. Conversely 40% of mothers from Welsh only homes were educated only to GCSE level compared to 13% of mothers from monolingual English who were educated to GCSE level. The education attainment of the father was also highly correlated with educational level of a child (r =.573;p.001). A comparison between fathers of children from monolingual homes and those of fathers of Welsh only homes produced similar results to that of the mothers i.e. 43% of fathers from monolingual English homes were educated to graduate level compared to 22% of fathers from Welsh only at homes.

6.5.5 British Picture Vocabulary Scale (BPVS)

The British Picture Vocabulary Scale II test (Dunn, Whetton & Burley, 1997) was administered to determine the proficiency of the participants in each group. The test involved 14 sets of 12 items with each item consisting of 4 pictures with only one correct target word and three distractor pictures. These distractor pictures were intended to be plausible but also

would not logically be chosen as the correct word. As the participant worked through the test the items became progressively more difficult during each set, including less commonly used words. The participant's starting set was directly linked to his/her age, for example set 3 would be the starting point for participants aged 6-7 whilst set 6 would be the starting set for participants aged 11. If participants made any mistakes during the first (base) set, they then reverted to the previous set until they provided correct responses for all items in a set. The testing continued until the individual's ceiling level was established which was reached when 8 or more responses were wrong in a set of 12 items. The BPVS is a measure of receptive vocabulary for English and as such was a useful research tool to assess whether children may have been placed in a particular language background group as a result of information obtained from the language background questionnaire whereas in practice, they were significantly more or less proficient than would be expected.

6.5.6 BPVS results

The BPVS scores provided were standardised according to the age of the child when tested. The initial Prawf Geirfa data showed the scores of children which had been standardised by age and then standardised by home language x age x school year. These initial analyses showed the scores of the children with reference to their expected performance. The second set of analyses used the raw scores of children in both the BPVS and the Prawf Geirfa as the proficiency of the individual child.

6.5.6.1 Analysis of BPVS results by age group and by home language

Results showed main effects for age group F (2,198) = 5.02, p<.007; home language F (3,198)= 18.11, p<.001; and a two way interaction: Age x Home Language F (6,198) = 2.59, p<.020. A pairwise comparison was carried out on home language with significant differences found between the Monolingual English group and the OEH group, (*MD* 13.37), with the OWH group, (*MD* 12.98), and with the WEH group, *MD* 14.85 ps<.001. Pairwise comparisons of the age groups found significant differences between children aged 9 to 11 and those aged 5 to 7,: (*MD* 3.49), and between children aged 9 to 11 and those aged 7 to 9,: *MD* 5.30, *ps*<.005.

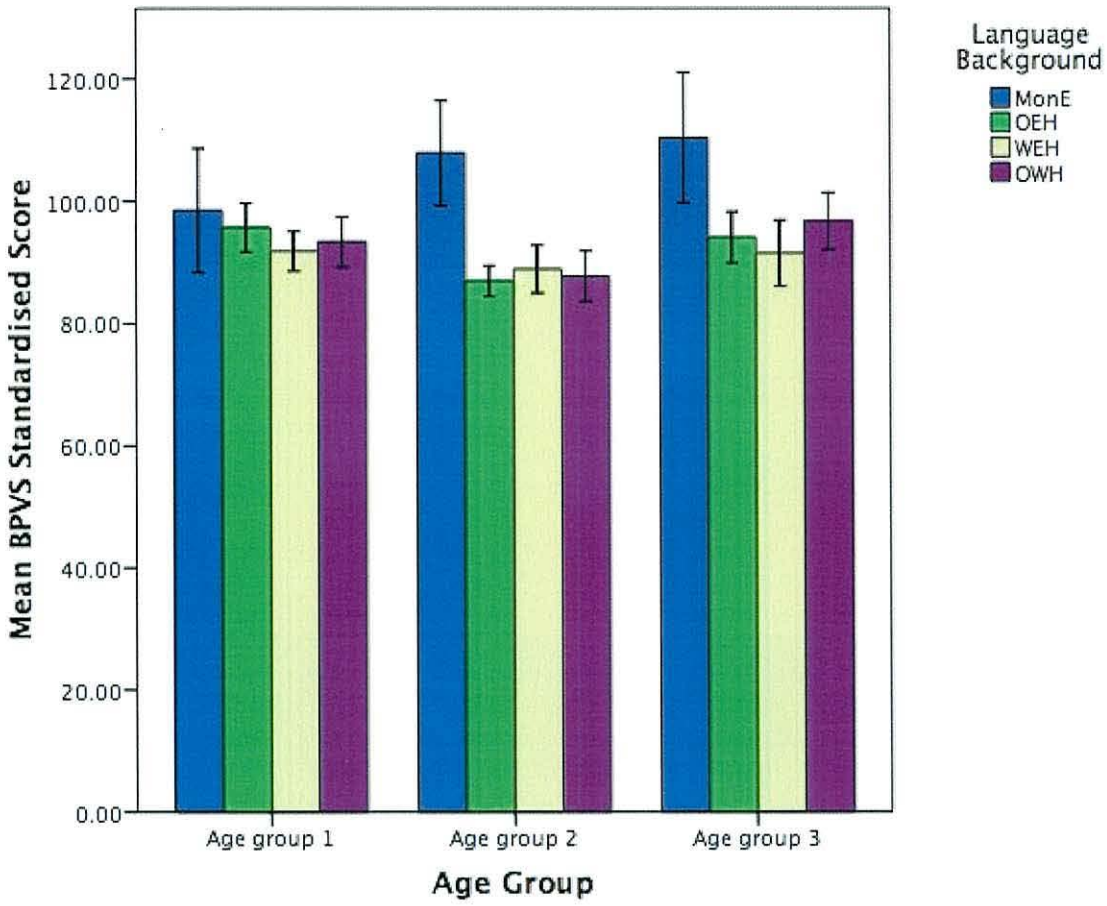


Figure 1: Standardised scores on the BPVS vocabulary task x Age x Home Language

6.5.7 Prawf Geirfa Cymraeg

The Prawf Geirfa Cymraeg ages 7-11 picture vocabulary scale was also administered to assess the level of fluency of Welsh in bilingual child participants. This test consists of 111 items which increase in difficulty as the test progresses. For the youngest age group (aged 5 to 7) in this study a different version of Prawf Geirfa suitable for ages 2-6 was presented using a computer to display the picture. Participants responded either verbally or by pointing at their chosen picture. Unlike the Prawf Geirfa 7-11, the test ends if a child participant answers incorrectly on 10 out of 12 items and the last item will be seen as the ceiling item. Incorrect

answers were then deducted from the ceiling item. A decision was made not to include their results in the analyses for comparison with children aged 7 to 9 and 9 to 11 as the range of the scales for the two versions of the Prawf Geirfa test were substantively different for comparison purposes.

6.5.7.1 Analysis of Prawf Geirfa Cymraeg results

This section outlines the performance of participants from the different age and language background groups on the Prawf Geirfa vocabulary test. The first set of results shows the analysis for performance of children from the different home language groups and age groups with the scores standardised by age only. Only children in age groups 2 and 3 were included in the analysis as the number of items in the Prawf Geirfa 2-6 differs to the Prawf Geirfa 7-11 version which would not make it appropriate for comparison between age groups. Results did not show a main effects for age group but there was a significant effect for home language F (2,120)= 59.16, p<.001. The two way interaction: Age x Home Language was not significant. A pairwise comparison was carried out on language background with significant differences found between the performance of children from the OEH group and the performance of children from the WEH group, (*MD* -4.24) and the OWH group, *MD* -19.03 *ps*<.005. There was also a significant difference between the performance of children from the WEH group and those from the OWH group, *MD* -14.83 p<.001.

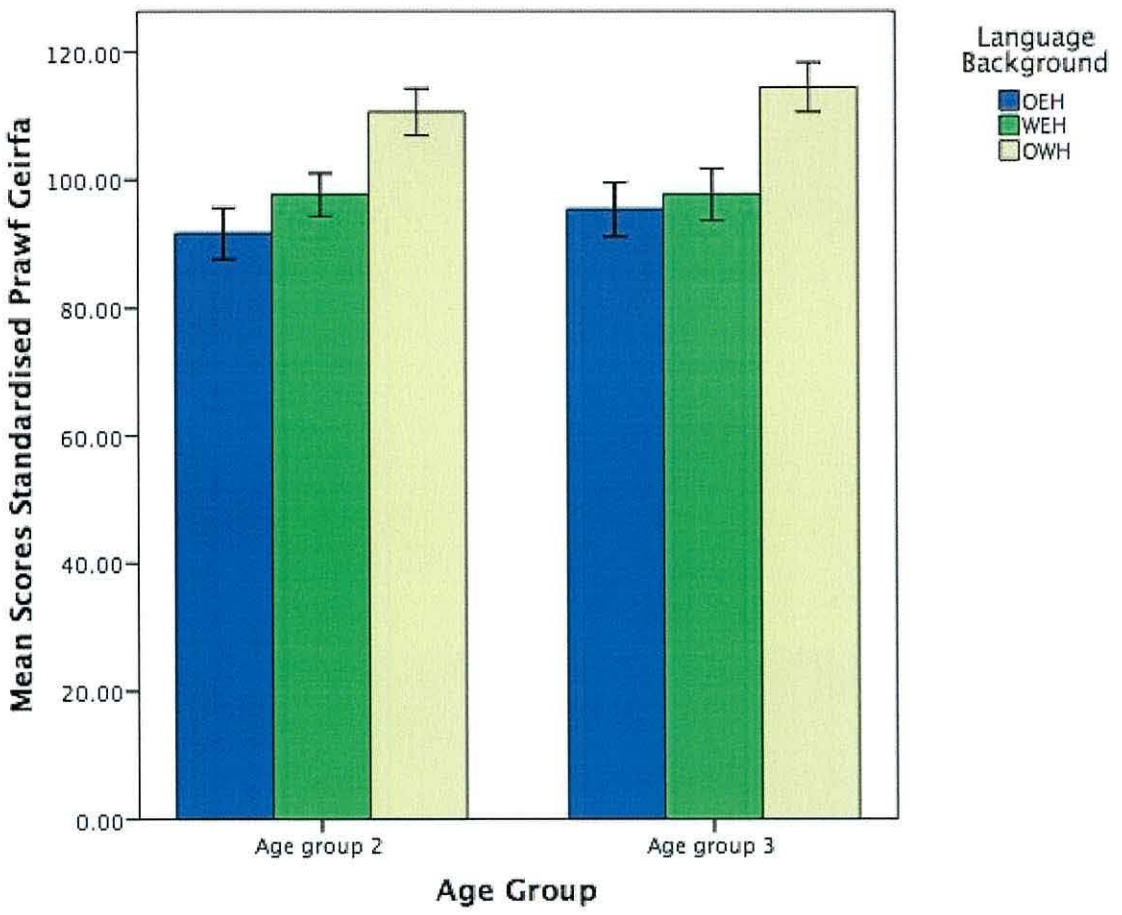


Figure 2: Participants' scores on the Prawf Geirfa using the standarised scale vocabulary task x Age x Home Language

Analysis was also performed on scores standardised by age, home language and school year group by the different age and home language groups on the Prawf Geirfa vocabulary test. Results showed a main effects for age group F (1,120) = 4.80, p<.030 with children aged 9 to 11 (M= 104.38) performing better than those aged 7 to 9 (M=100.93) ; but there was no main effect for home language. The two way interaction: Age x Home Language was also not significant. As there was no effect of home language, children were performing according to the standardised score for their home language background.

6.5.8 BPVS raw scores

Children from all language and background groups were tested on the British Picture Vocabulary Scale in order to investigate any differences in their vocabulary knowledge of English. The additional use of raw scores was essential for comparisons to be made between participants' scores in both BPVS and Prawf Geirfa. The use of age standardised scores would have potentially resulted in misleading conclusions to be made.

Results using BPVS raw scores showed a significant effect for age group F (2,198) = 217.17, p<.007; and also between the performance of children from different home language groups F (3,198)= 15.14, p<.001; there was also a two way interaction: Age x Home Language F (6,198) = 2.59, p<.020. As was predicted, children from the Monolingual English group performed better than the OEH group (*MD* 14.23), the OWH group (*MD* 12.97), and with the WEH group, *MD* 14.00 ps<.001. There were no significant differences observed between the performances of children from the other home language groups. Differences between the performance of children from different ages was seen between children aged 5 to 7 and 7 to 9, 2 (*MD* -18.75), and between children aged 5 to 7 and those aged 9 to 11, (*MD* -38.29), pairwise comparison all, *ps*<.005. Children aged 7 to 9 also performed worse than the older children in the study, *MD* -38.29, *ps*<.005. Simple effects analysis each age group showed that children from the monolingual English background were constantly more proficient in English than the other three groups in all three age group *ps*<.005. The other three groups performed similarly on the BPVS task with raw scores obtained by all three groups being almost identical.

Following on from the BPVS task, the raw scores were used to correlate with the

performance obtained from the categorisation task. Performance scores across the four word groups were collapsed to create one score for each width and then these were correlated with the BPVS score. This was done to obtain a comparable score for general performance in both tasks. The scores of children tested in English, tested by age group, were correlated with the performance of children in the categorisation task using stimuli which adopted the condition, English wider than Welsh (r =.335;p.035). Although not investigated in this study, it would be worth investigating further by possibly collapsing the age groups for all home language groups.

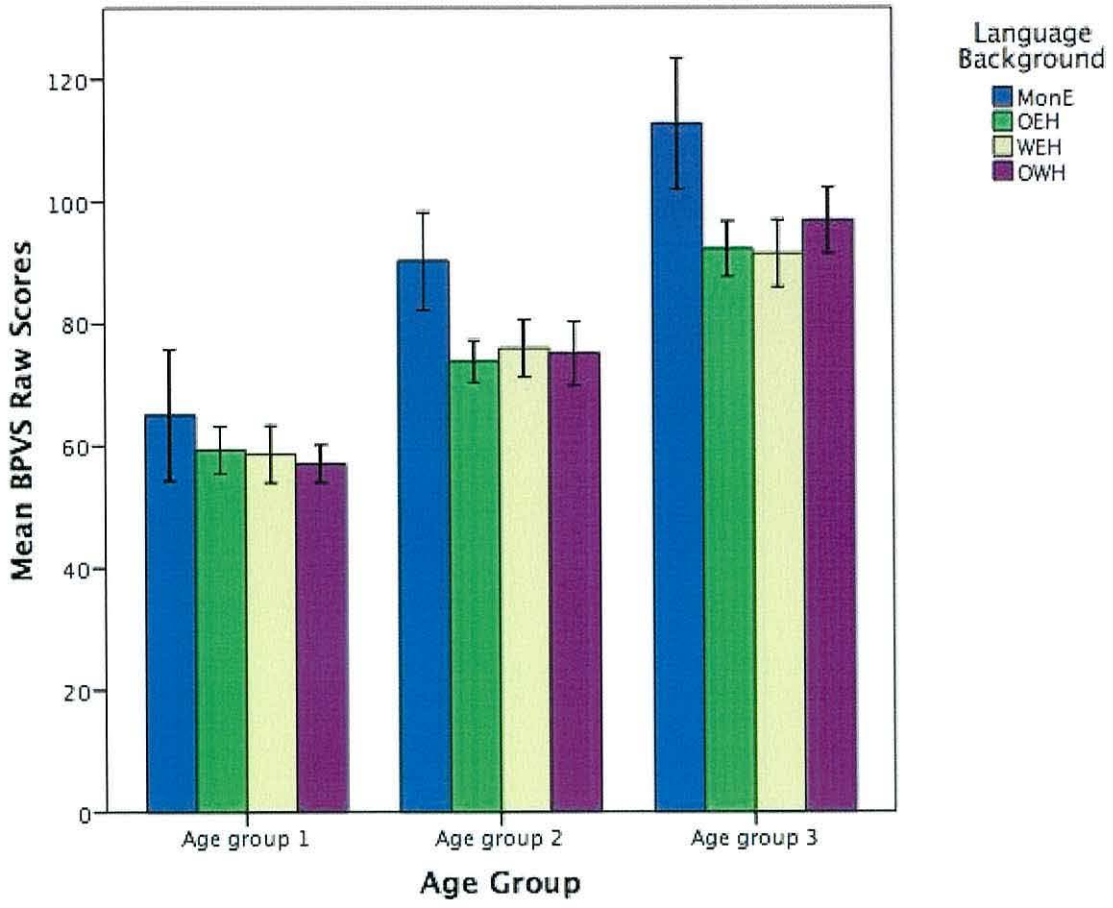


Figure 3: Raw scores on the BPVS vocabulary task x Age Group x Home Language

6.5.9 Prawf Geirfa raw scores

Differences between children from Age groups F (1,114)= 51.33, p<.001 and Home Language F (2,120)= 69.70, p<.001 were both found to be significant. Although the two way interaction Age x Home Language was not significant. As was expected, pairwise comparisons on the performance of children from different home language groups found the OWH performed better than both WEH, (*MD* 14.68), and the OEH groups, (*MD* 19.13 *ps*<.001. The WEH group also performed better than the OEH group on the Prawf Geirfa task. Children aged 9 to 11, also performed significantly better than children aged 7 to 9, p<.001.

Similar to follow up analysis made for the BPVS task, the raw scores were used to correlate with the performance of the Prawf Geirfa with categorisation task. The scores of performance across the four word groups were once again collapsed to create one score for each width and these were then correlated with the Prawf Geirfa score. The results did not provide any significant correlations, however the scores of children aged 9 to 11 tested in English on stimuli which English was wider than Welsh were nearing significance (*r* =.319;p.068) .

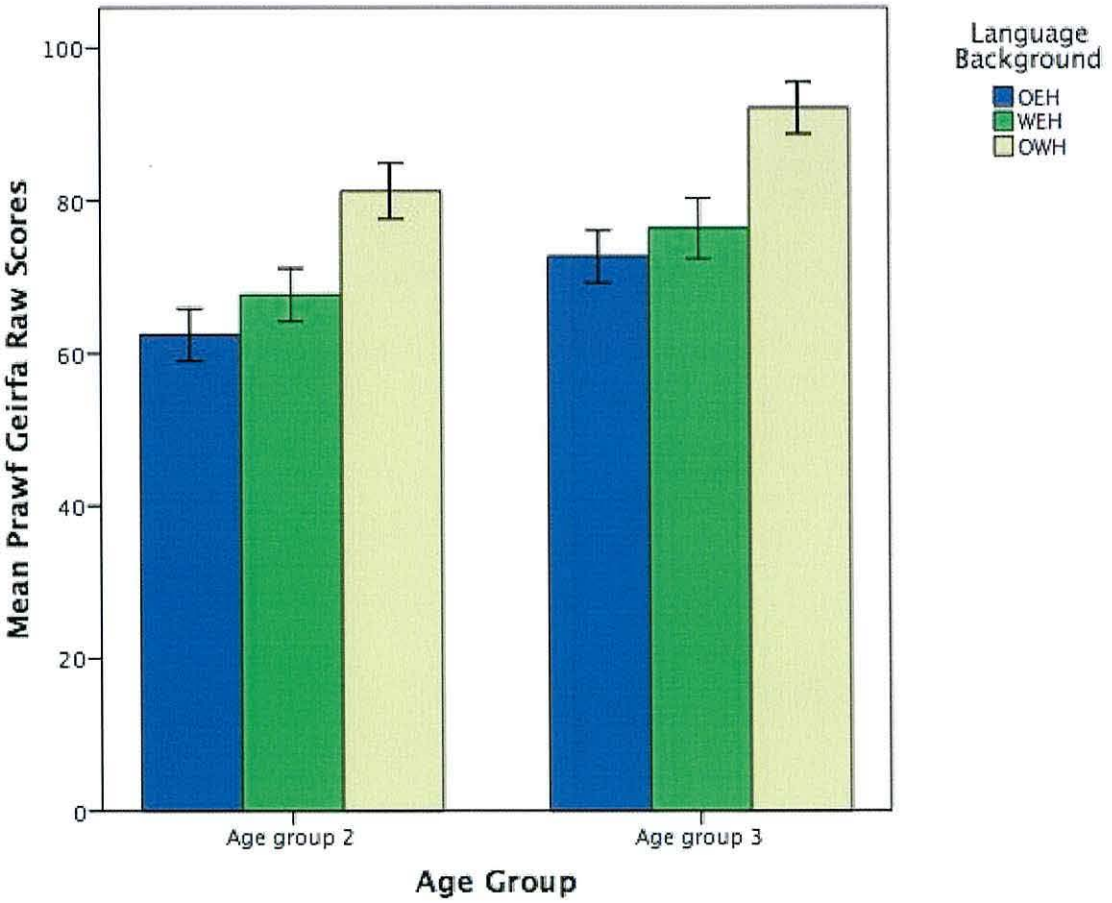


Figure 4: Raw scores on Prawf Geirfa test x Age Group x Home Language for age groups 2 and 3 (age group 1 was not included in the analysis)

Chapter 7. The Categorisation Task

7.1 Introduction

This chapter reports the categorisation task performed by monolingual and bilingual children from selected schools in locations across North and West Wales, Cheshire and Berkshire. The study was designed to identify influential factors involved in the categorisation of specific word categories (classical, homonym, radial taxonomic, and radial thematic) and to compare the performance of children from different language backgrounds. The study used a test which was performed by individual children in either English or Welsh.

7.2 Participants

Two hundred and thirty eight children (having no reported language disorders identified from a language background questionnaire) took part in this study. The children attended 14 primary schools in North and West Wales, one school in Cheshire and one school in Berkshire. Participants in the study were placed into one of four different groups depending on their home language. The selection process to place participants into these groups was made following an evaluation of the home language questionnaires which were completed by parents or guardians prior to the testing process. The four groups consisted of monolingual English speakers (MonE), only English at home (OEH), both Welsh and English at home (WEH), and only Welsh at home (OWH). Out of the two hundred and thirty eight participants tested in English, 37 were from MonE group, 34 from the OEH group, 32 from the WEH group, and 30 in the OWH group. From the participants tested in Welsh; 39 were from the OEH group, 32 from the WEH group, and 34 from the OWH group. Where participants were

subdivided by further age groups, there were still at least ten participants in each subgroup. Monolingual English speaking participants were defined as those whose linguistic background consisted of speaking only English at home and also receiving their education only through the medium of English. OEH participants were defined as those whose linguistic background consisted of both parents speaking only or mostly English with the child, with the child receiving his/her education mainly through the medium of Welsh. OWH participants were defined as those whose linguistic background consisted of both parents speaking only or mostly Welsh to the child, and the child receiving his/her education mainly through the medium of Welsh. Welsh and English participants were defined as those whose linguistic background consisted of either or both parents speaking both English and Welsh to the child, or had one parent speaking only Welsh and the other speaking only English, and the child receiving his/her education mainly through the medium of Welsh.

Child participants from school year classes 1 to 6 were placed into three different age groups, with age group 1 consisting of children ranging between 5 – 7 years (mean 6 years 2 months) from school year 1 and school year 2. Age group 2 consisted of children ranging between 7 – 9 years (mean 8 years 4 months) from school year 3 and school year 4. Age group 3 consisted of children ranging between 9 – 11 years (mean 10 years 6 months) from school year 5 and school year 6.

7.3 Stimuli

7.3.1 Linguistic Stimuli

The study was designed to include two language conditions, with the Welsh condition involving only Welsh instructions and the English condition only involving English instructions. The words identified for the categorisation task were wider in one language than the other (W>E for Welsh wider, E>W for English wider) i.e. the two targets for a test item could be named by a single word in the wider language, however would be named by two words in the narrower language. For example in the English condition, the word *nail* represents nail as a support to hang up a picture or in the anatomical sense. Therefore for the two target stimuli for the *nail* test item, the word nail could be used to select (or name) both target pictures. Whereas in the Welsh condition (narrower language) the word *hoelen* should only be used to name the first target (T1), as a support to hang up a picture, but it would not be appropriate to name the second target (T2), *nail* in the anatomical sense. The task was designed so that there were four different word category types with two nouns and one verb for each type and width in each language. The words used for the test stimuli can be seen in Table 1 below.

Items	Classical	Homonym	Radial Taxonomic	Radial Thematic
E>W				
	STICK(N)/FFON	**NAIL(N)/HOELEN**	**KEY(N)/GORIAD**	**DISH(N)/DYSGL**
Target 1	Walking stick	Nail	Door Key	Plate
Target 2	Stick	Fingernail	Computer	Food
Taxonomic 1	Walking frame	Screw	Lock	Cup
Taxonomic 2	Leaf	Lips	Disk	Wine
Thematic 3	Umbrella	Hammer	Door	Washer
Thematic 4	Tree	Nail Polish	Mouse	Microwave
	WALL(N)/MUR	**CHEST(N)/BREST**	**HAND(N)/LLAW**	**EARTH(N)/DAEAR**
Target 1	Castle wall	Anatomy	Hand on body	The Earth (planet)
Target 2	Interior wall	Chest box	Hand on clock	Earth/ground
Taxonomic 1	Outside gate	Arm	Foot	The Moon
Taxonomic 2	Window	Trolley for lifting	Digital clock	Gravel
Thematic 3	Knight	Vest	Mittens	Sky
Thematic 4	Paint	Gold coins	Sand timer	Spade
	FOLDING(V)/PLYGU	**BOXING(V)/PAFFIO**	**GRAZING(V)/PORI**	**SMOKING(V)/ YSMYGU**
Target 1	Origami	Boxing	Cow grazing	
Target 2	Arms	To put in a box	A grazed knee	Smoking a cigarette
Taxonomic 1	Tying shoe lace	Karate	Pig trough	Smoking (fire)
Taxonomic 2	Waving	Putting box on shelf	Some injury	Drinking
Thematic 3	Piece of paper	Boxing bag	Picture of a field	Burning fire
Thematic 4	Arm	Ball of string	Plaster	Lighter Fire Wood
W>E				
	CYSGOD(N)/SHADE	**YSGOL(N)/SCHOOL**	**PEN(N)/HEAD**	**UCHEL(A)/HIGH**
Target 1	Shadow	School	Head	High
Target 2	Shade	Ladder	Top of Mountain	Loud
Taxonomic 1	Lying in a pool	Library	Hand	Low table
Taxonomic 2	Sitting in the sun	Stairs	Lake bottom mountain	Whispering
Thematic 3	The sun	Teacher	Hat	Climbing gear
Thematic 4	Suntan Lotion	Paint	Flag	Ear protectors
	BAWD(N)/THUMB	**HWYL(N)/FUN**	**PWRS(N)/PURSE**	**FFORDD(N)/ROAD**
Target 1	Thumb	Fun	Purse	Road
Target 2	Big toe	Sail	Udder	Way
Taxonomic 1	Elbow	Getting told off	Handbag	River
Taxonomic 2	Knee	Powerboat	Cow's head	Childproof stairs
Thematic 3	Ring	Someone crying	Pound coin	Car
Thematic 4	Shoe	The sea	Bottle of milk	Road map
	CHWYRNU(V)/ SNORING	**BWRW(V)/RAINING**	**CODI(V)/LIFTING**	**COLLI(V)/LOSING**
Target 1	Snoring	Raining	Lifting a box	Loosing a race
Target 2	Growling Dog	Hitting	Getting out of bed	Spilling
Taxonomic 1	Someone sneezing	Sunny	Jumping high jump	Trophy
Taxonomic 2	Happy dog	Kissing	Washing face	Drinking
Thematic 3	Bed	Umbrella	Forklift truck	Chequered flag
Thematic 4	Dog food	Boxing ring	Alarm clock	Kitchen roll

Table 1: Matrix of word groups and test items with taxonomical and thematic associates

7.3.2 Word categories

For the purposes of this study, four word category types were used: classical, homonym, radial taxonomic, and radial thematic. Words in each group had multiple meanings (sometimes depending on which was the test language), although the extent of relationships between words varied depending on word type.

The classical category type included words that can be defined by critical features essential for membership in that particular class. For example, the word *folding* (represented as 'folding' a piece of paper or *folding* arms) in English, whilst the corresponding test item in the Welsh condition was *plygu* (represented only as folding a piece of paper). An example for the Welsh condition; the word *bawd* (represented as the *big toe* or a *thumb)* in Welsh, with the test item for participants tested in English test condition being the word *thumb* (represented only by the thumb). If a feature is missing, the concept is unlikely to be suitable for membership within the category group, e.g. if an object was constructed of rubber and consisted of tread but was square in shape, it would not be perceived as a wheel.

The homonym word type contains words that were phonetically the same in the wider language but their meanings are unrelated. For example, the word *chest* (represented as part of the body or a container made of wood) in English, and the word *hwyl* (*sail* or *fun*) in Welsh.

The radial taxonomic type included words that had radial structure with links based on similarity of shape or function or on metaphorical extension. The category has a central use for the word but the use of the word is associated with 'motivated' extensions (Gathercole & Moawad, 2010). For example, the word *hand* (for the anatomical part of the body and the hands of a clock) in English, and *pen* (represented as a body part 'head' or top as a spatial

orientation) in Welsh.

The radial thematic type included words that had radial structure with links based on thematic association. For example, *dish* (for a plate or for food) in English, and *colli* (for spilling or for losing) in Welsh.

Words were not specifically selected from any specific list in terms of frequency or word length, although numerous sessions were held with various peers in order to assess suitability of words for inclusion in the test and within an appropriate word group.

7.3.3 Non-linguistic stimuli

The picture stimuli consisted of 6 pictures for each test item. The pictures for each test item were chosen according to a clearly defined criterion. This consisted of two of the six pictures representing two target referents, two taxonomically linked (to the targets) distractors, and two thematically linked (to the targets) distractors.

The 'taxonomic links' associated with the 2 targets by belonging to the same superordinate as the targets. For example, for the word *nai*l (representing the body part) the taxonomic link was *lips* (another body part). For *nail* representing a support to hang up a picture, the taxonomic link was a *screw*. The thematic links were associated with the 2 targets by thematic association, for example, for the body part *nail,* the thematic link was *nail polish*; n*ail* as a support for hanging up a picture, the thematic link was a *hammer*. The participants were shown a PowerPoint presentation in which each slide represented one target word. The slide for each word contained the six pictures representing each category. Slides were shown to each participant with either the English or Welsh as the narrower language depending on

the particular test item.

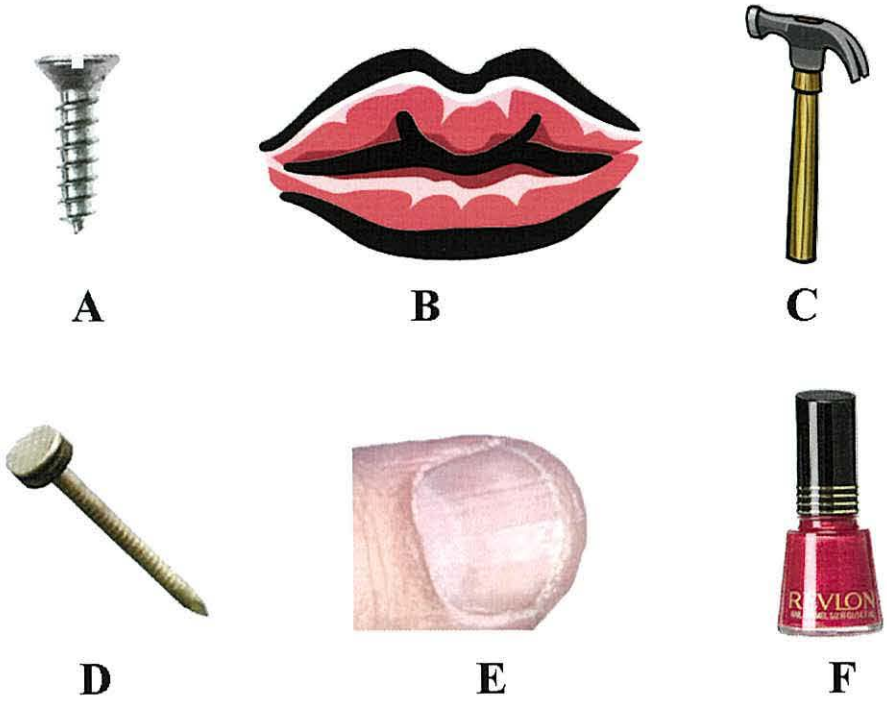


Figure 5: Example of a slide used in the categorisation task (Target: Nail)

7.4 Procedure

The selection of words was made during a series of consultations with various peers to assess item suitability for the test. This process was designed to ensure that words were placed into the appropriate word categories and that the items fit the criteria of the study. This process led to a number of items being rejected due to potential problems as a result of degree of difficulty and pictorial representation. Although words were not specifically taken from any database, an effort was made to ensure that the use of equivalents in both language were comparable. For example, no words were selected if the Welsh equivalent was not as well

known, and therefore potentially, individuals were likely to code switch and use the English equivalent when using the term. The Cronfa Electroneg (Ellis et al, 2001) does provide classifications by frequency of certain Welsh words but it was decided not to use the database as a source of items as it does not specifically compare the frequency of use in both Welsh and English.

All envisaged ethical considerations for the study were identified and submitted for approval by the Bangor University School of Psychology Ethics Committee. It was essential that ethical approval was obtained before any contact was made with schools to invite them to take part in the study. The use of child participants required careful ethical considerations. Children were always in the plain sight of more than one adult during the testing procedure and they only took part if the signed parental consent slip had been returned to the school. Children were constantly asked at appropriate times if they were comfortable to ensure that the task did not inadvertently cause distress and to also ensure that their performance did not significantly drop due to fatigue factors. It is important to note that the presentation of test items was randomised in both the order of presentation and in their positioning on the computer screen.

Protecting the identity of children was also a priority for the study to ensure the anonymity of participants. Every participant was given an individual ID number which was used when their data was inputted for purpose of analysis. Data protection requirements meant the list containing participant names and IDs was not shown to anyone else. This procedure was also useful as it allowed for the removal of participants from the study without causing confusion or leading to inadvertent errors. Each school was also allocated an unique

Id to further safeguard the identity the of the participants.

Children from reception/nursery school classes (3-4½ years) were initially included as being a possible age group but were not finally selected due to the length of the test. Their potential inability to concentrate was considered to be a risk factor for the reliability of the study. No formal pilot study was conducted. However, it was decided that after 3 participants from each group had been tested, there would be an evaluation of the design to ascertain if there were any general problems with the study design.

An initial contact was made with school head teachers across north and west Wales, Cheshire and Berkshire inviting participation by primary schools. Schools which responded and agreed to take part in the study received further information about the study. This information included linguistic background questionnaires, an explanatory letter and parental consent slips for distribution to parents and guardians. The linguistic background questionnaire for each child participating in the study included questions relating to details of linguistic background found in the home, at school and with friends. (see Appendix 1). Questionnaire responses were used to determine into which group the child was placed for the purposes of analysis.

Children were tested on a one to one basis, with a PowerPoint presentation being shown on a touch screen monitor. The software used for the touch screen monitor automatically recorded the time taken from when the child pressed the onscreen button to view an item, to when the button was pressed to move on to the next item. Despite the randomised positioning of the pictures for each item, the software recorded the unique identifier for each picture, not its position on the screen.

The responses of the child participants were recorded with the touch screen programme noting the pictures selected and the reaction times. The position of each of the six pictures for a test item was randomised by the computer program for presentation purposes. The order of presentation was also randomised. A blank screen was shown between each test slide. The participants were then told that the researcher would say a word out loud that would correspond to each slide they were about to see. The researcher would then ask the participants "Which picture or pictures, could be labelled as....". Participants were then instructed to choose their answers by pressing the relevant picture/s on the touch screen.

To ensure that the participants fully understood the nature of the task, the first six slides were included as a practice for the test procedures. The practice slides included items which did not meet the criteria of the word groups of the test (e.g. practice item blue which included pictures of six different colours). Although it was important to show the children examples of both practice items which had either 2 or 1 possible correct answers. This was included to ensure that participants understood that they were expected to select more than one picture on certain items without specifically cuing them to do so.

After they had completed their selection of pictures for a given slide, they were asked to press a small arrow on the right hand side of the screen to go to the next slide. The children were told that there may be more than one answer and that they could select as many as they thought were appropriate. After the completion of the categorisation task, children also undertook a memory task which will be discussed in Chapter 9.

7.5 Results

For the choice data task, initial analyses focused on the choice of target items. When tested in the wider language, participants were required to select both targets (T1 and T2) to obtain a correct score on an item. For example, when tested in English for the item *nail* it was necessary for participants to select the picture of a *nail* (fastener) and the second picture of *nail* (anatomical). Participants selecting only 1 of the target items received a score of zero for that item. When a participant was tested in the narrower language, participants were required to select only target 1 to obtain a correct score. (e.g. in Welsh participants were required to select only T1 when they heard *hoelen*). If a participant selected both targets in the narrower language they received a score of zero for that item (e.g. when tested in Welsh selecting both pictures of T1 and T2 for the item *hoelen*). Choices of distractor items were ignored for this first set of analyses. Thematic and taxonomic preferences in distractor items will be discussed in Chapter 8. Results were analysed separately for the Welsh and English data. The first section 7.5.1 reports the results for children tested in Welsh while the second 7.5.2 reports the results for children tested in English.

The method of analysis selected for the study was a complex factorial ANOVA which was selected as being the most appropriate way of achieving the study's goals. Using this method can lead to a more difficult analysis and interpretation but has the advantage of addressing all the research questions for the study. The study investigated the effects of frequency/exposure (age), language dominance (home language), and word group (membership criteria) on categorisation behaviour by the participants. The relationships between these factors were assumed to be interconnected, therefore investigating them

independently would not provide adequate answers for the research questions. The analysis of the data initially examined the main effects of the variables for the study, but also looked for any interactions between variables. This a particular strength of a factorial design as independent variables rarely occur in isolation. For example, when investigating the performance of participants on different word groups, developmental factors (i.e. an increase in exposure to Welsh or English with age), and the language background environment may all affect results independently, or be dependent on a combination between variables. This could not be achieved if the variables were analysed independently. Another benefit of using a repeated measures factorial Anova was the ability to address the goals of the study with a smaller number of participants. If there was a requirement for different participants to be recruited for each of the different independent variables (IV), significantly more participants would be needed. As the study required individual schools to agree to take part voluntarily (around 30% of the schools contacted agreed to participate) and for a parental consent form to be returned for each participant (around 15% of forms were returned), the logistics to recruit a sufficient number of participants would have potentially resulted in requiring more than three times the number of actual participants to achieve the goals of the study. This method of statistical analysis was also used for subsequent analyses in the other studies of this current research.

7.5.1 Welsh

Analyses of variance were performed on the data in which width (Welsh wider than English, and English wider than Welsh) and word category group (classical, homonym, radial taxonomic and radial thematic) were within subject variables, and age group (children aged 5 to 7, children aged 7 to 9, children aged 9 to 11) and home language group (OEH, WEH, and OWH) were treated as between group variables. The results of main effects analyses showed that showed significant effects for word group, F (3,288)= 13.68, p <.001 and Width, F (1,96) =163.74, p <. 001. Results of between-subject factors showed main effects for age group F (2,96)= 10.52, p <.001, and home language F (2,96)= 3.46, p <.036.

On the whole, participants from the OWH group performed significantly better than those from the OEH group with pairwise comparisons showing *MD* -.228, p <.032. Children generally performed worse on test items in the classical word group (word group 1) compared to all three of the other word groups with pairwise comparisons showing the homonym word group (word group 2) *MD* -.354, p<.001; with the radial taxonomic word group (word group 3) *MD* -.400, p<.001; and the radial thematic group (word group 4) *MD* -.223, p<.003. Children also performed better on the radial taxonomic stimuli when comparing with the radial thematic test stimuli *MD* .177, p<.038. Performance was typically worse on items when Welsh was wider than English *MD* .928, p<.001. The oldest age group (9 to 11) performed better than both the other two age groups and the youngest group of children (5 to 7) did not perform as well as the other two groups. Pairwise comparisons of the age groups show: children aged 5 to 7 vs children aged 7 to 9: *MD* -.225, p<.015; Children aged 5 to 7 vs children aged 9 to 11: *MD* -.412, p<.001, and children aged 7 to 9 vs children aged 9 to 11:

MD -.187, $p<.035$.

These main effects were modified however by two and three way interactions for Width x Age group, F (2, 96)= 4.65, $p<.012$; Word group x Width, F (3, 96)= 15.08, $p<.001$; Word group x Width x Age group F (6, 288) = 3.11, $p<.006$, and Word group x Width X Home Language F (6, 288) = 5.12, $p<.001$. and Age x Home Language F (4,96) = 2.66, p <.037.

Follow-up analyses indicate that on classical categories, width was significant, F (1,96) =11.44, $p<.001$, as was the Width x Home Language interaction , F (2,96) =3.15, $p<.047$. Age was also found to be significant, F (2,96) =6.18, $p<.003$. Children aged 9 to 11 (1.29) performed significantly better than children aged 5 to 7 (0.83) $p<.002$. When Welsh was wider than English there were significant differences between Age groups and between Home Language groups, with children aged 5 to 7 (0.48) once again performing worse than children aged 9 to 11 (1.24), $p<.001$. Children in the OWH group (1.18) performed best and were significantly better than the OEH group (0.63) $p<.014$. When English was wider than Welsh there was no significant difference between the different age groups or the different home languages.

Performance on the homonym categories indicated that width was significant, F (1,96) =290.88, $p<.001$. There was also two way interactions between Width x Home Language interaction, F (2,96) =9.75, $p<.001$, and Width x Age group F (2,96) =3.66, $p<.029$. Age group was also significant, F (2,96) =3.18, $p<.046$. When English was wider than Welsh there were significant differences between age groups and home language groups. Children differed in their responses F (2,96) $p<$ 5.20, with children aged 5 to 7 (0.55) performing

significantly worse than children aged 7 to 9 (0.90), $p<.001$, and children aged 9 to 11 (1.68) $p<.001$. The children in the WEH language group performed significantly worse than the OWH $p<.010$. When Welsh was wider than English there was no significant difference between the different age groups or the different home language groups.

For the radial taxonomic words width was significant, F (1,96) =165.02, $p<.001$, as was Age group F (2,96) =6.02, $p<.003$. Home Language was also significant, F (2,96) =6.28, $p<.003$. Width x Age group F (2,96) =3.29, $p<.042$ was significant. When English was wider than Welsh there were significant differences between age groups, F (2,96) =4.09, $p<.020$, and between home language groups, F (2,96) =8.67, $p<.001$. Performance of children aged 5 to 7 (1.89) was significantly worse than those children aged 7 to 9 (2.37), $p<.014$. Those in the OWH group (2.52) performed significantly better than both OEH (1.93) group $p<.002$, and WEH group $p<.001$ (1.88). There was also a significant difference on items where Welsh was wider than English between the age groups, F (2,96) =5.05, $p<.008$. Children aged 9 to 11 (1.09) were again the best performers, do significantly better than children aged 5 to 7 (0.56) $p<.006$. There were no differences between home language groups for radial taxonomic words when Welsh was wider.

For the radial thematic words width was significant, F (1,96) =16.81, $p<.001$, as was age group F (2,96) =5.29, $p<.007$. Children aged 9 to 11 (1.53) performed significantly better than children in aged 7 to 9 (1.19) $p<.047$, and children aged 5 to 7 (1.10) $p<.009$. Home language was not significant. Width x Age group F (2,96) =5.63, $p<.005$. was significant. When English was wider than Welsh there was no significant difference between the different age groups or the different home language groups. Although significant differences were

found between the age groups when Welsh was wider. Children aged 9 to 11 (1.53) once again performed significantly better than those aged 5 to 7 (0.57) $p<.001$, and also aged 7 to 9 (0.92) $p<.019$. There were no significant differences between home language groups.

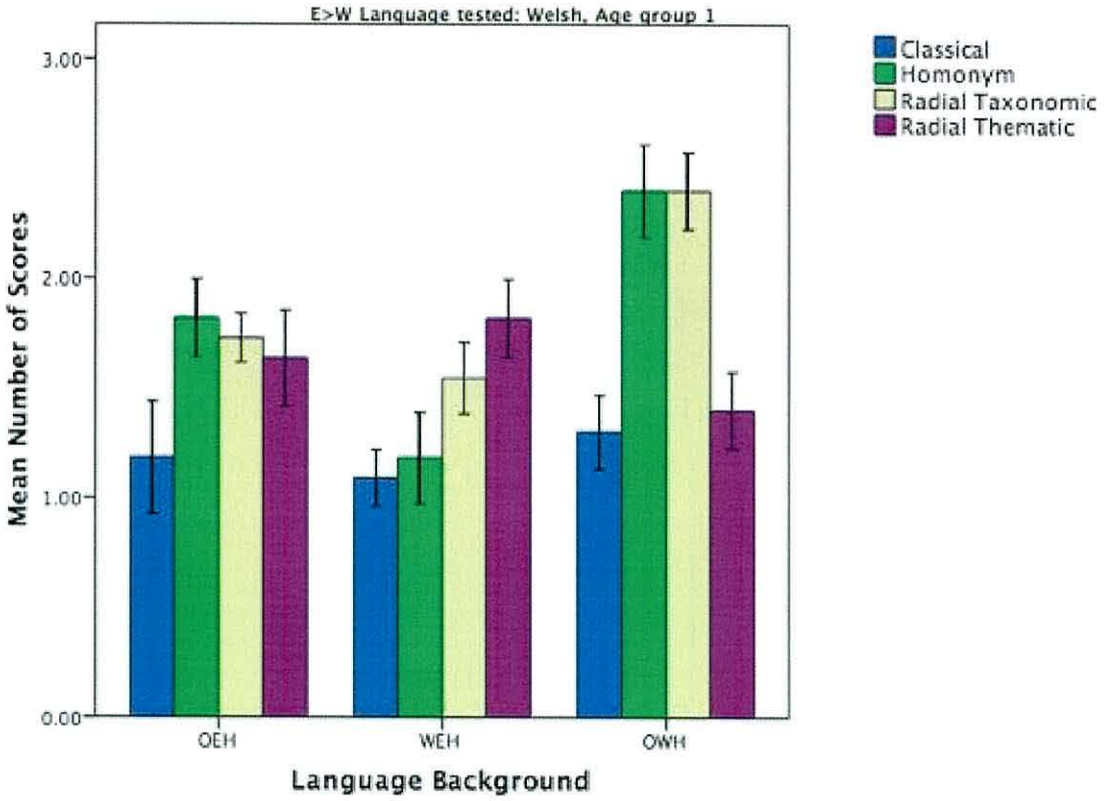


Figure 6: Word Group x Home Language where English is wider than Welsh by children aged 5 to 7.

Following up on the significant interactions, simple effects analyses were made into age group and Age Group x Home Language. For children aged 5 to 7, significant effects were found for word group, F (3,27)= 4.88, $p<.008$ and width, F (1,29) =219.11, $p<.001$. The

two way interaction between Word Group x Width, $F(2,27)=4.76$, $p<.016$ was also significant. Differences between home language groups $F(2,29)= 3.55$, $p<.042$, were significant with pairwise comparisons showing significant differences with children from the WEH group performing worse than children in the OWH group *MD* -.459, $p<.013$. Pairwise comparisons of the word groups found significant differences, with children performing worse on classical word stimuli compared with stimuli from other word groups: homonym *MD* -.356, $p<.009$; radial taxonomic *MD* -.365, $p<.001$; and radial thematic *MD* -.262, $p<.047$ There was also a significant difference between narrower and wider language, with children performing better on items in the narrower category *MD* 1.088, $p<.001$. The performance by children aged 5 to 7 when English is wider than Welsh is shown in Figure 6, and where Welsh is wider than English is shown in Figure 7.

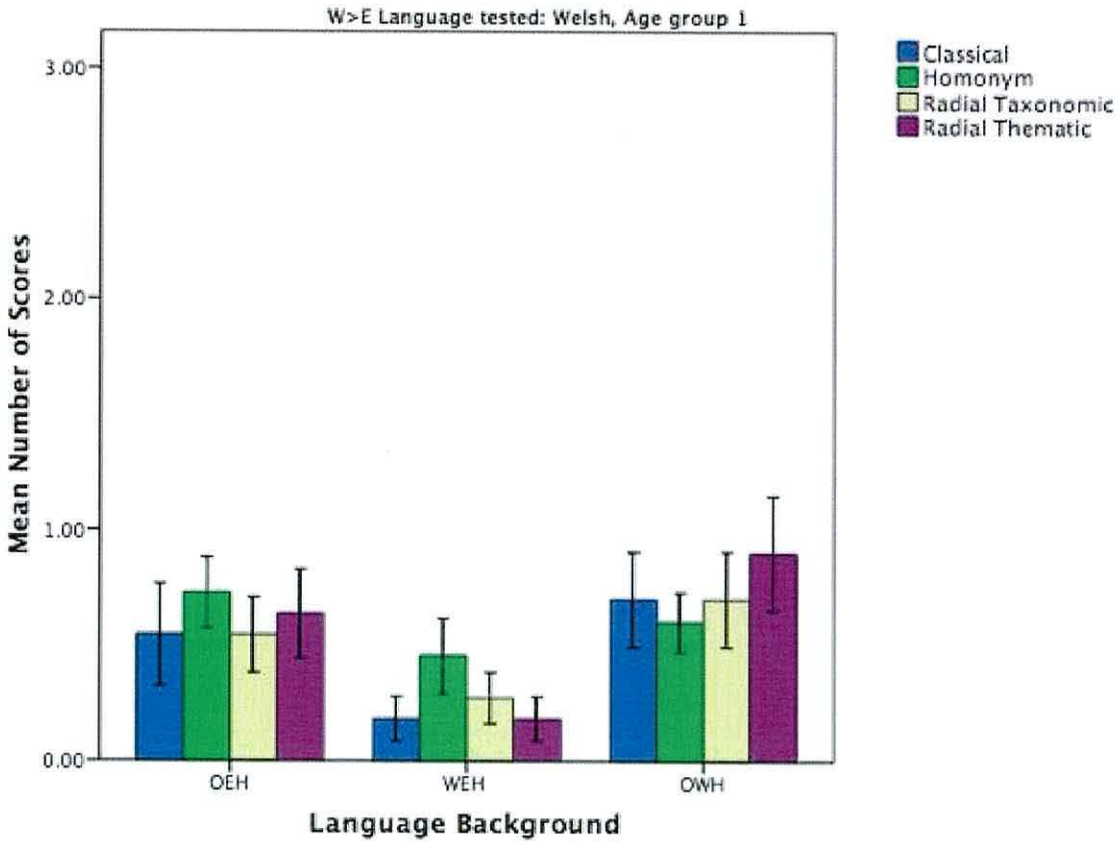


Figure 7: Word Group x Home Language where Welsh is wider than English for children aged 5 to 7.

Simple effects analyses of children aged 7 to 9 showed that word group, F (3,31)= 9.74, p<.001 and width, F (1,33) =51.42, p<.001 were significant. A two way interaction was present between Word Group x Width, F (3,31)=16.96, p<.001. Pairwise comparisons of the word groups found significant differences between the classical category with the homonym category *MD* -.431, p<.003; and also with the radial taxonomic category *MD* -.545, p<.001. Children performed better on radial taxonomic items compared with radial thematic items: *MD* .377, p<.001. With regards to width, similar results to those of children aged 5 to 7 were

seen with children performing better when width was narrower than wider *MD* 1.077, $p<.001$. Home Language F (2,33)= 3.41, $p<.045$. Pairwise comparisons of home language showed a significant difference between the performance of children in the OEH and WEH groups *MD* .380, $p<.015$. Children's performance when English is wider than Welsh is shown in Figure 8, and when Welsh is wider than English is shown in Figure 9.

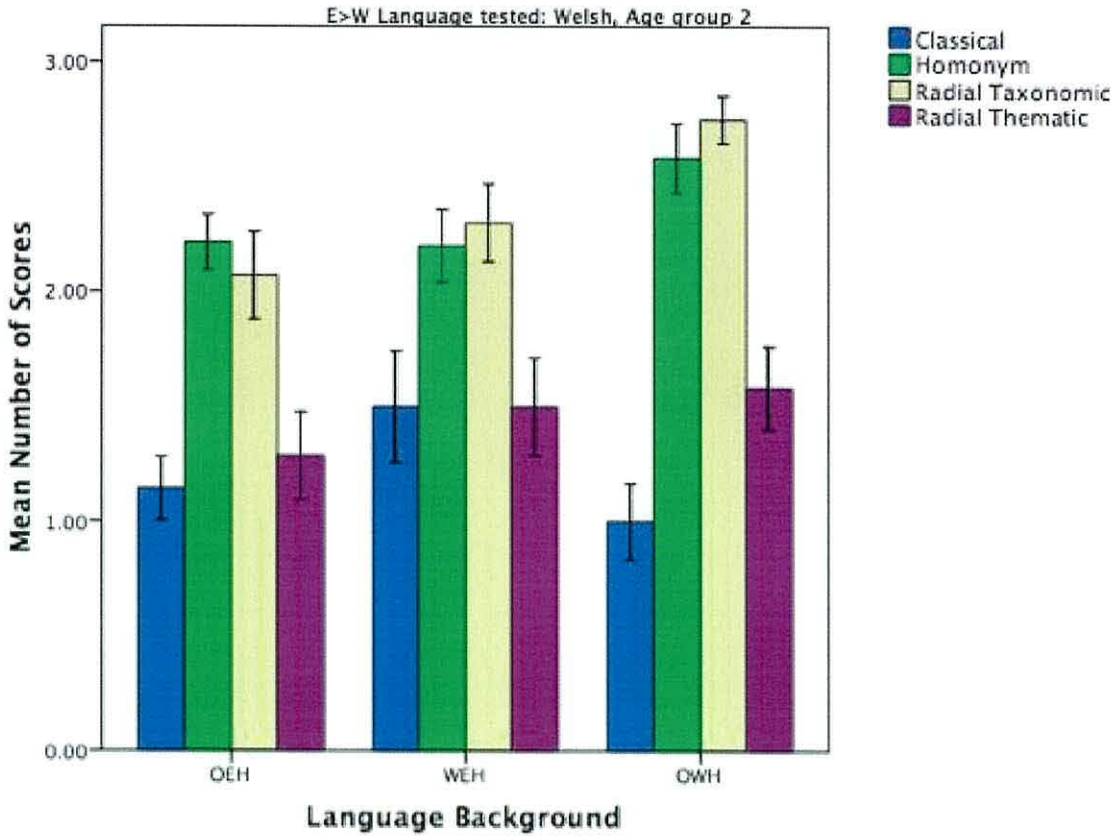


Figure 8: Word Group x Home Language where English is wider than Welsh for children aged 7 to 9.

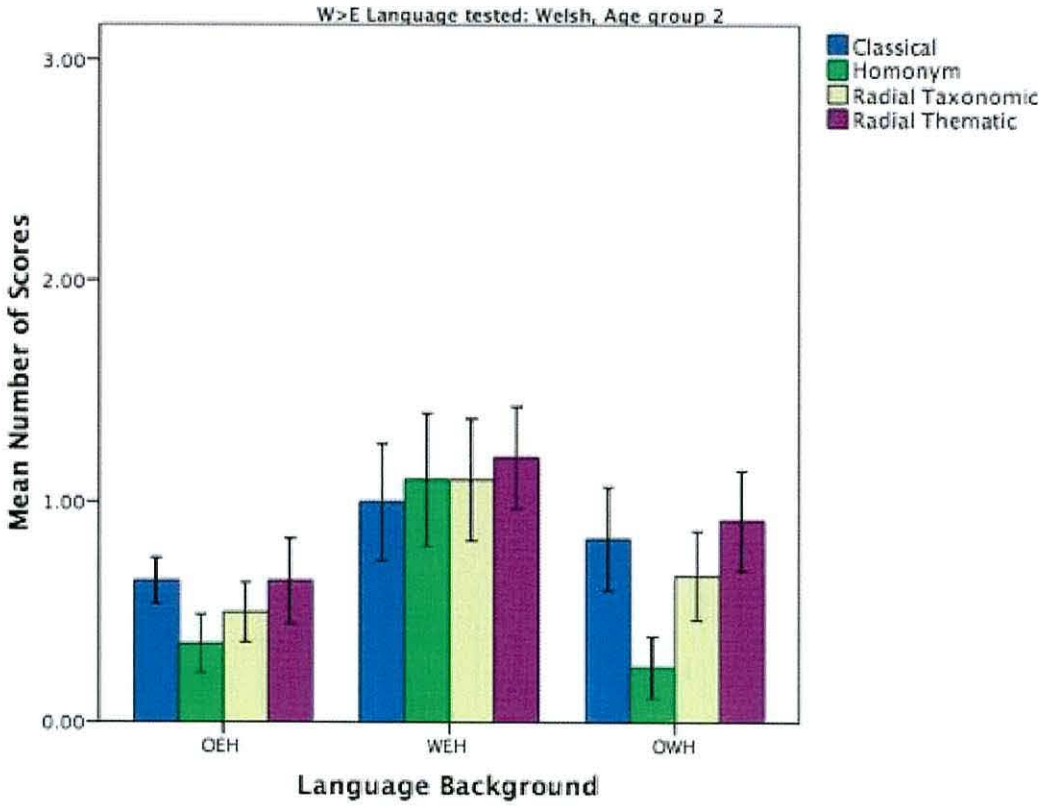


Figure 9: Word Group x Home Language where Welsh is wider than English for children aged 7 to 9.

Simple effects analyses of the performance of children aged 9 to 11 demonstrated that width was significant F (1,34) =22.76, p<.001, and a two way interaction of Word Group x Width was significant F (3,32)= 12.78 , p<.001. A three way interaction was also present between Word Group x Width x Home Language F (6,66)= 2.34, p<.041. Children also performed significantly differently on word groups with pairwise comparisons showing children performing worse on classical word items compared to radial taxonomic items *MD* -.290, p<.015. When English is wider than Welsh is shown in Figure 10, and when Welsh is

wider than English is shown in Figure 11.

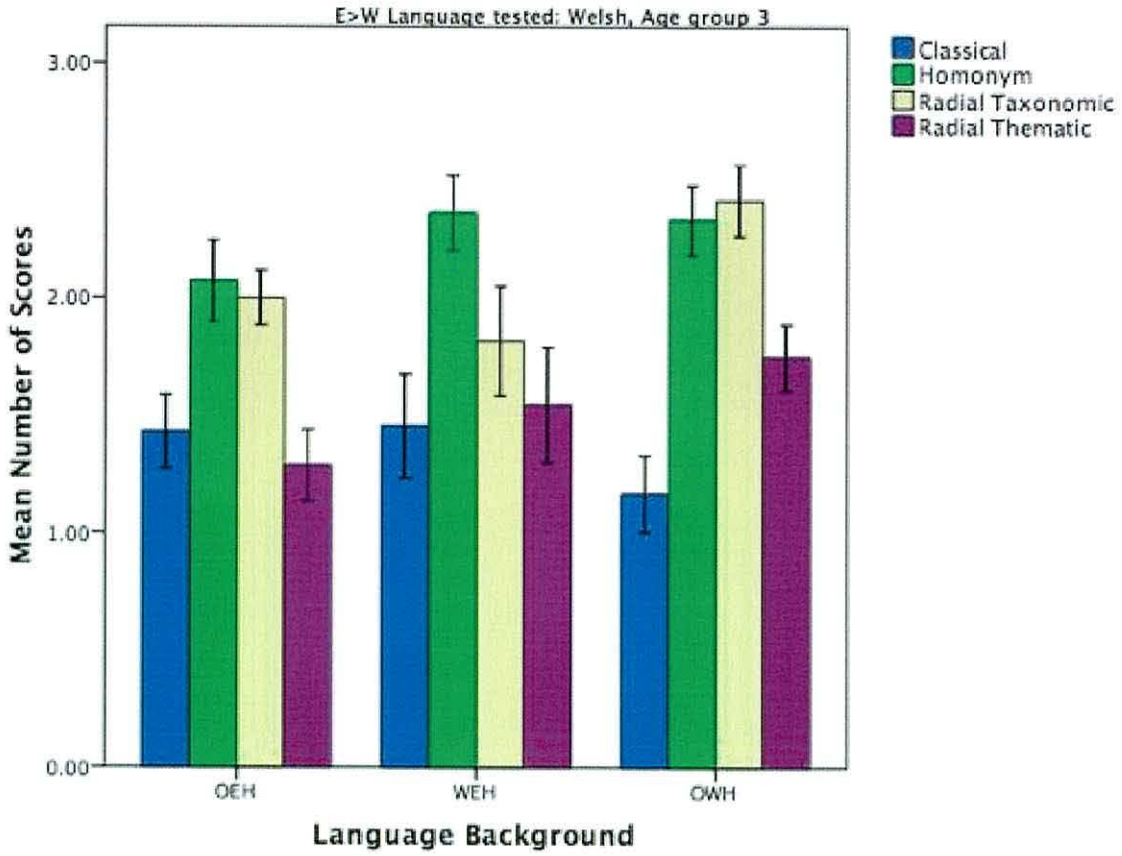


Figure 10: Word Group x Home Language where English is wider than Welsh by children aged 9 to 11

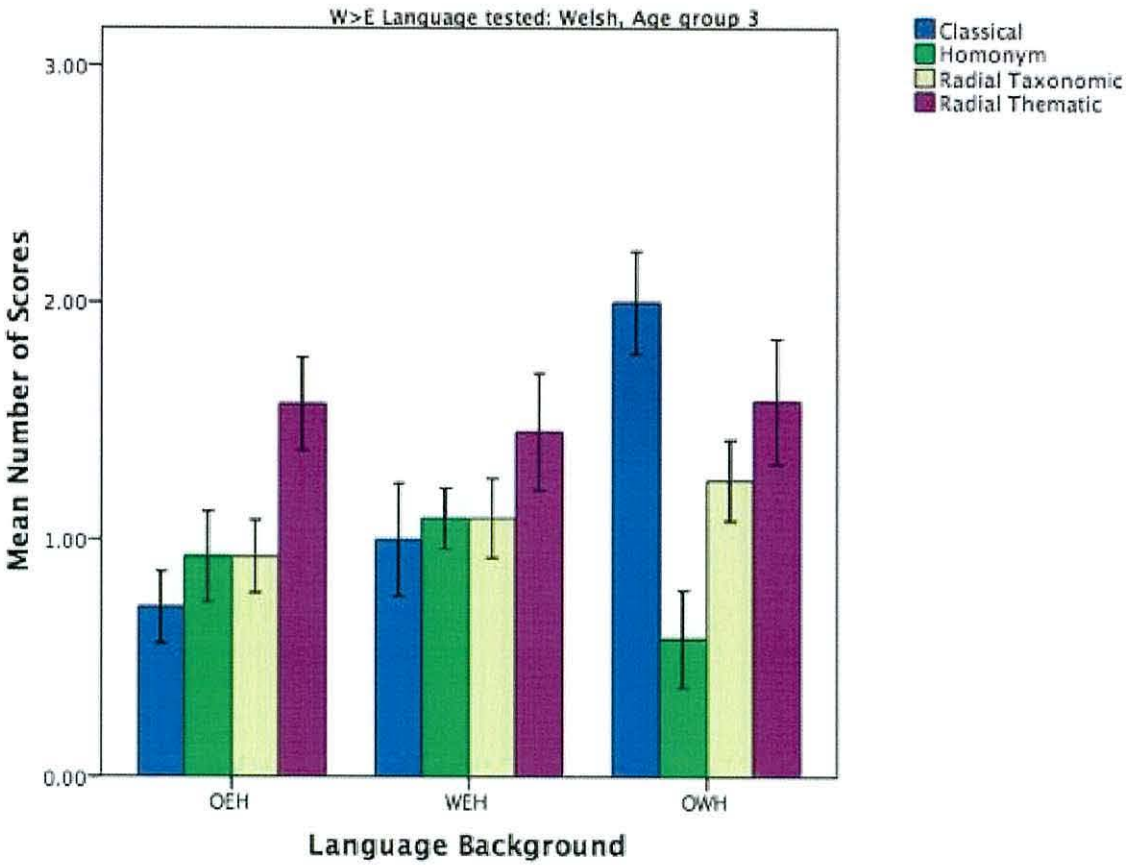


Figure 11: Word Group x Home Language where Welsh is wider than English by children aged 9 to 11.

The performance of the older children aged 9 to 11 in the OEH group demonstrated that width was significant F (1,13) =13.08, $p<$. 003, (narrower 1.70 vs wider 1.04). Word group was not found to be significant. However, the interaction between Word Group x Width F (3,39) =6.95, $p<$. 003 was significant. The WEH children in the showed a near significant difference by width F (1,10) =4.73, $p<$. 055, and for Word Group x Width F (3,30) =2.70, $p<$.

063. In the OWH group, width was significant F (1,11) =7.57, $p<$. 019. The two way interaction between Word Group x Width was also significant F (3,27) =11.34, $p<$. 001.

7.5.2 English

Analyses of variance were again performed on the data in which width (Welsh wider than English, and English wider than Welsh) and word category group (classical, homonym, radial taxonomic and radial thematic) were within subject variables, and age group (children aged 5 to 7, 7 to 9, 9 to 11) and home language group (Monolingual English, OEH, WEH, and OWH) were treated as between group variables.

Results showed main effects and word group, F (3,363)= 23.051, $p<$.001 and width, F (1,121) =185.45, $p<$. 001, age group F (2,121)= 9.71, $p<$.001; and home language group F (3,121) = 2.69, $p<$.049.

In general, children from the Monolingual English group performed significantly better than those from children in the OEH group *MD* .209, $p<$.007. For word groups there were significant differences between the classical word group (word group 1) and the radial taxonomic word group (word group 3) *MD* .260, $p<$.001; and the radial thematic group (word group 4) *MD* -.347, $p<$.001. There were significant difference between the performance of children on items in the homonym word group and the radial taxonomic word group *MD* .284, $p<$.001; and between the homonym word group and the radial thematic word group *MD* .371, $p<$.001. Children tested in English performed better on items when Welsh was the wider language *MD* 1.178, $p<$.001. Children in the youngest age group performed the worst with pairwise comparisons of the age groups finding significant differences between children aged

5 to 7 and 7 to 9: *MD* -.233, $p<.001$; and between children aged 5 to 7 and those aged 9 to 11: *MD* -.296, $p<.001$.

Theses results were modified however, with two way interactions which were found to be significant for Width x Word Group, F (3,363)= 26.71, $p<.001$; Word Group x Age Group F (6,363) = 2.65, $p<.015$, and Width x Age Group, F (2,121)= 5.93, $p<.003$. The two way interaction between Age Group x Home Language Group F (6,121)= 2.93, $p<.043$ was also significant.

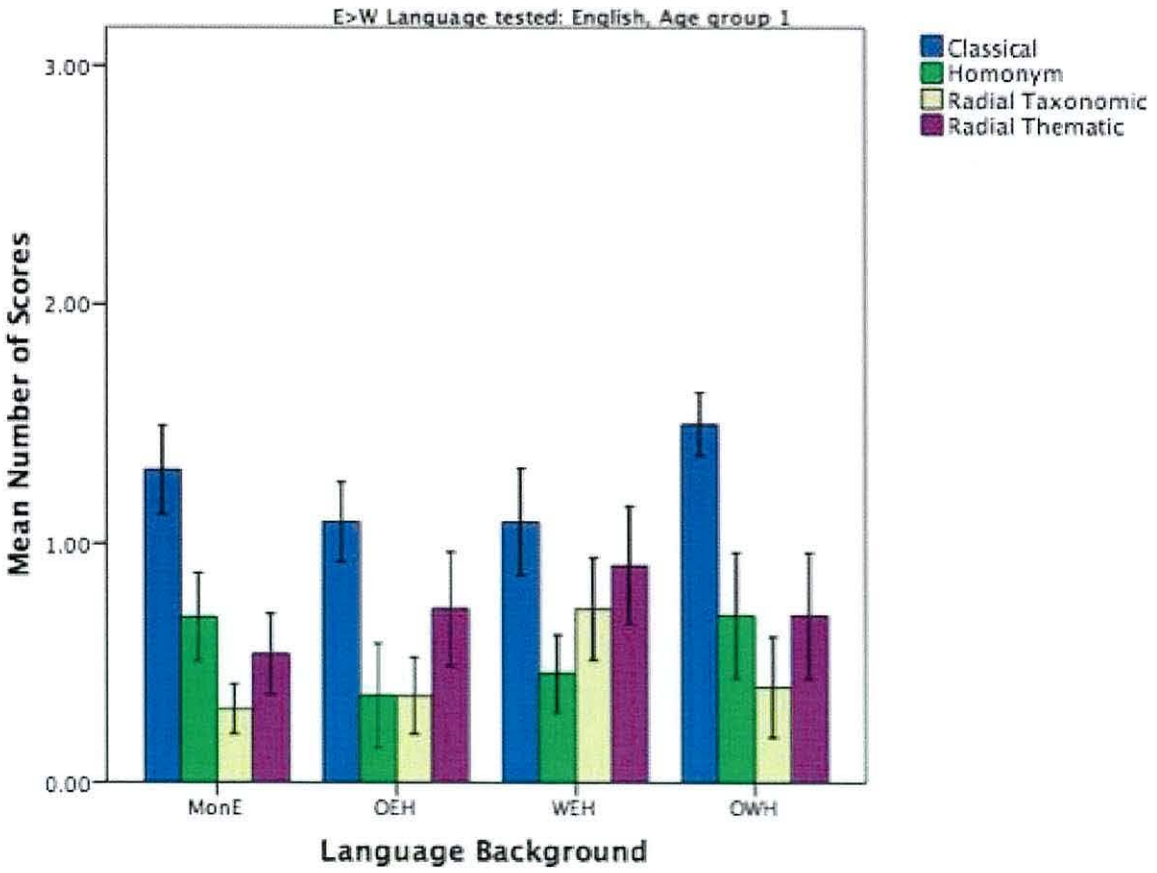


Figure 12: Word Group x Home Language where English is wider than Welsh by children aged 5 to 7.

To explain these interactions, follow up analyses were conducted for each age group. Children aged 5 to 7 performed significantly differently on word groups, F (3,39)= 6.14, p<.002, by width, F (1,41) =91.15, p<. 001, and for Word Group x Width, F(3,39)=10.38, p<.001. Pairwise comparisons of the word groups found in general participants performed worse on the classical group items when compared with radial taxonomic items *MD* -.296, p<.005; radial thematic *MD* -.411, p<.001. Children in general also did better on narrower items than on wider items *MD* 1.388, p<.001.

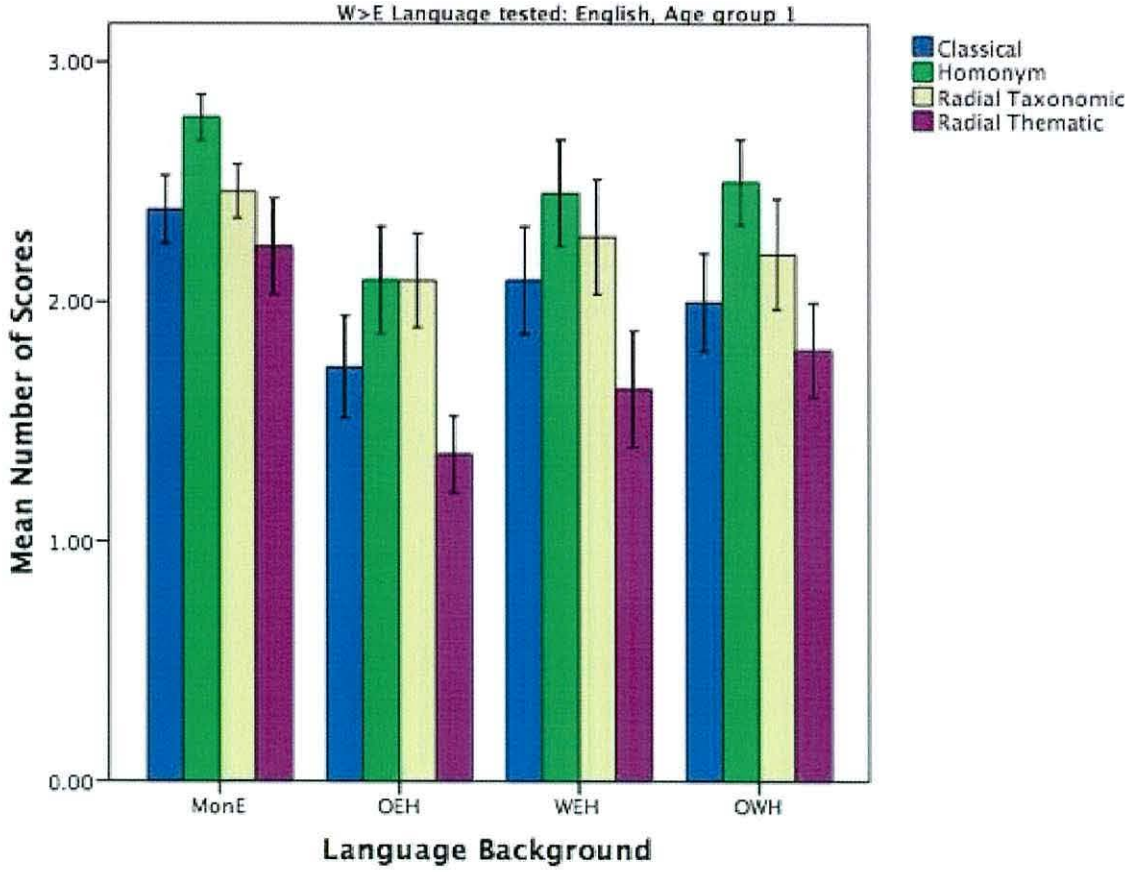


Figure 13: Word Group x Home Language where Welsh is wider than English by children aged 5 to 7.

Children aged 7 to 9 showed significant effects for word group, F (3,37)= 9.77, p<.001 and Width, F (1,39) =70.12, p<.001, and for Word Group x Width, F (3,37)= 3.95, p<.001. Pairwise comparisons of the performance of children for the word groups found significant differences between the classical category with the radial thematic *MD* -.258, p<.004; and with radial non thematic items *MD* -.334, p<.001(participants performing worse on classical items than the other two word groups). There was also a significant difference in the performance of children on narrower and wider language test items *MD* 1.38, p<.001.

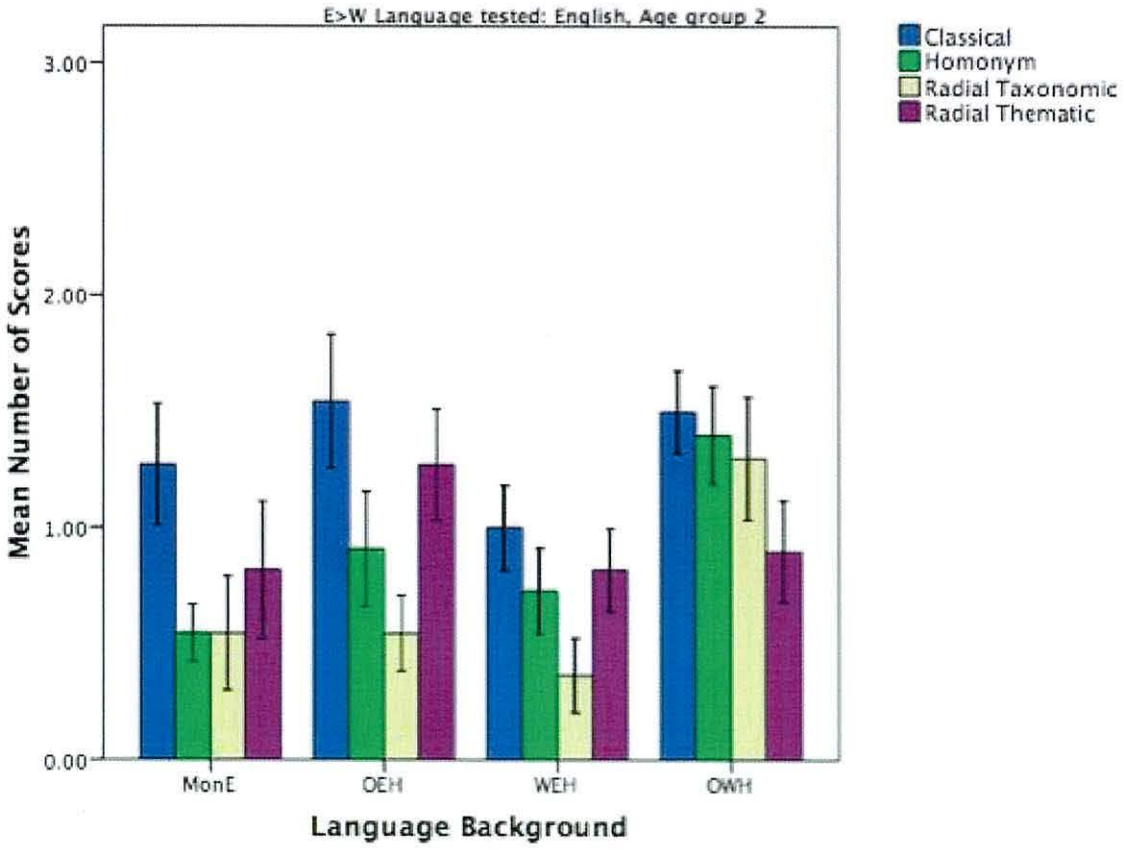


Figure 14: Word Group x Home Language where English is wider than Welsh by children aged 7 to 9.

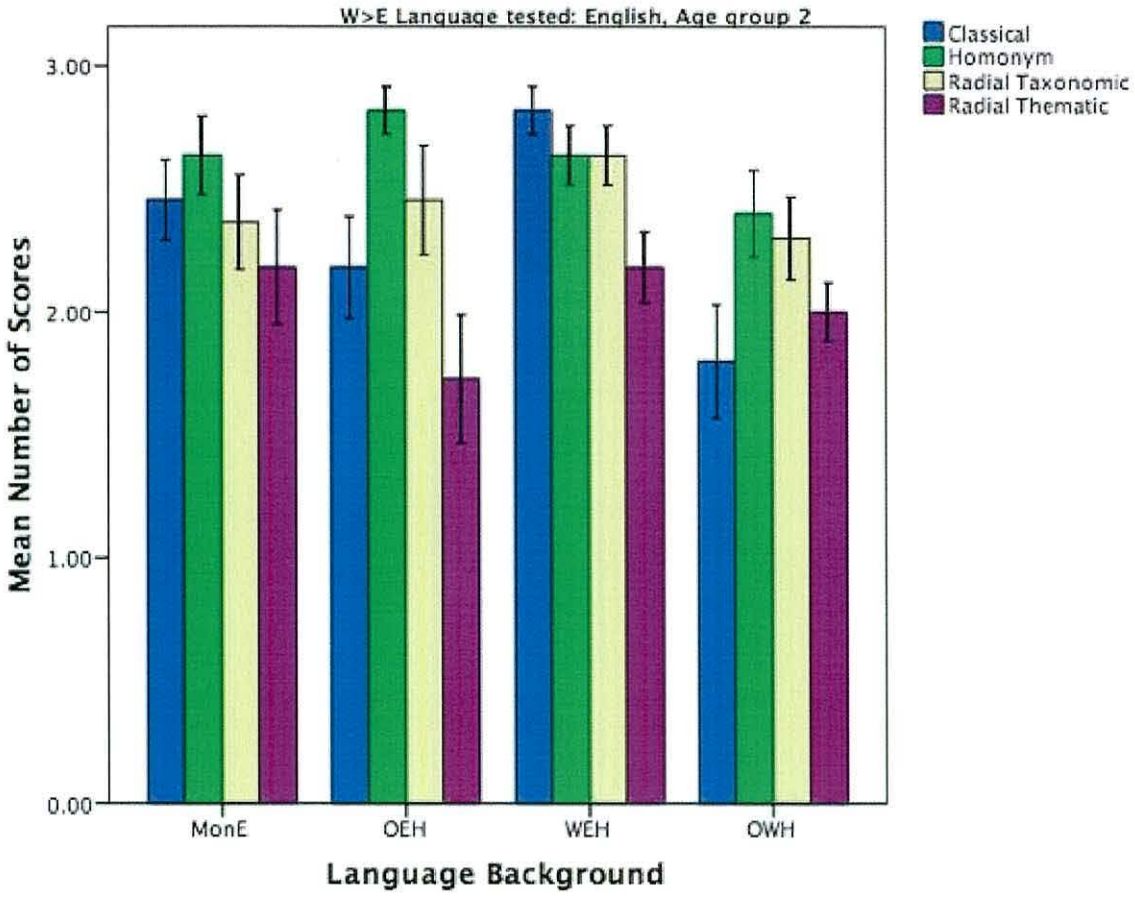


Figure 15: Word Group x Home Language where Welsh is wider than English by children aged 7 to 9.

Children aged 9 to 11 followed a pattern similar to that of other age groups with width, F (1,41) =30.06, p<.001 and word group, F (3,39) =12,15 p<.001 both significant. and a two way interactions between Word Group x Width, F(3,39)= 11.58, p<.001. But they also showed Width x Home Language interaction, F(3,41)= 3.17, p<.001. Children did not perform as well on classical categories when compared to the other categories, homonym *MD* -.307, p<.002; radial taxonomic *MD* .232, p<.020; and radial thematic *MD* .301, p<.012.

Overall children performed better on homonym category items, doing significantly better than radial taxonomic *MD* .539, $p<.001$; radial thematic *MD* .608, $p<.001$. Children also performed better on the test items where Welsh was wider than English *MD* .756, $p<.001$.

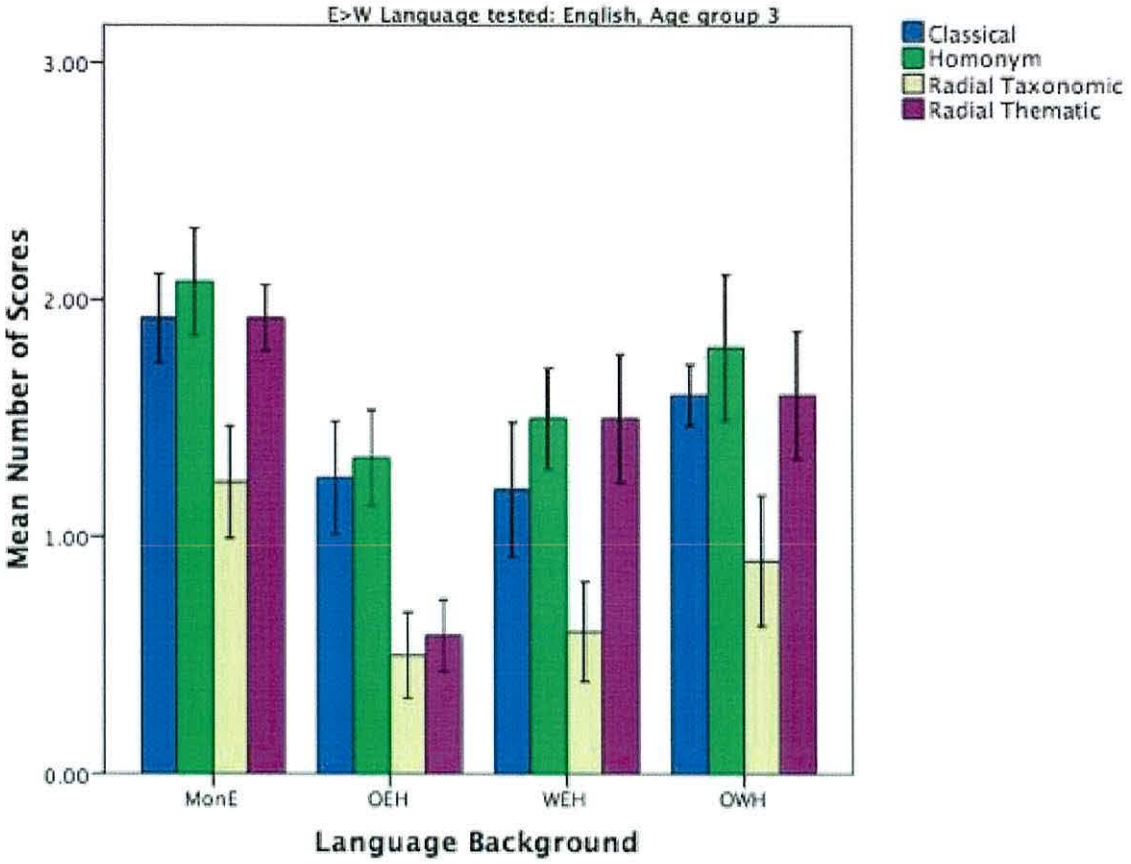


Figure 16: Word Group x Home Language where English is wider than Welsh by children aged 9 to 11

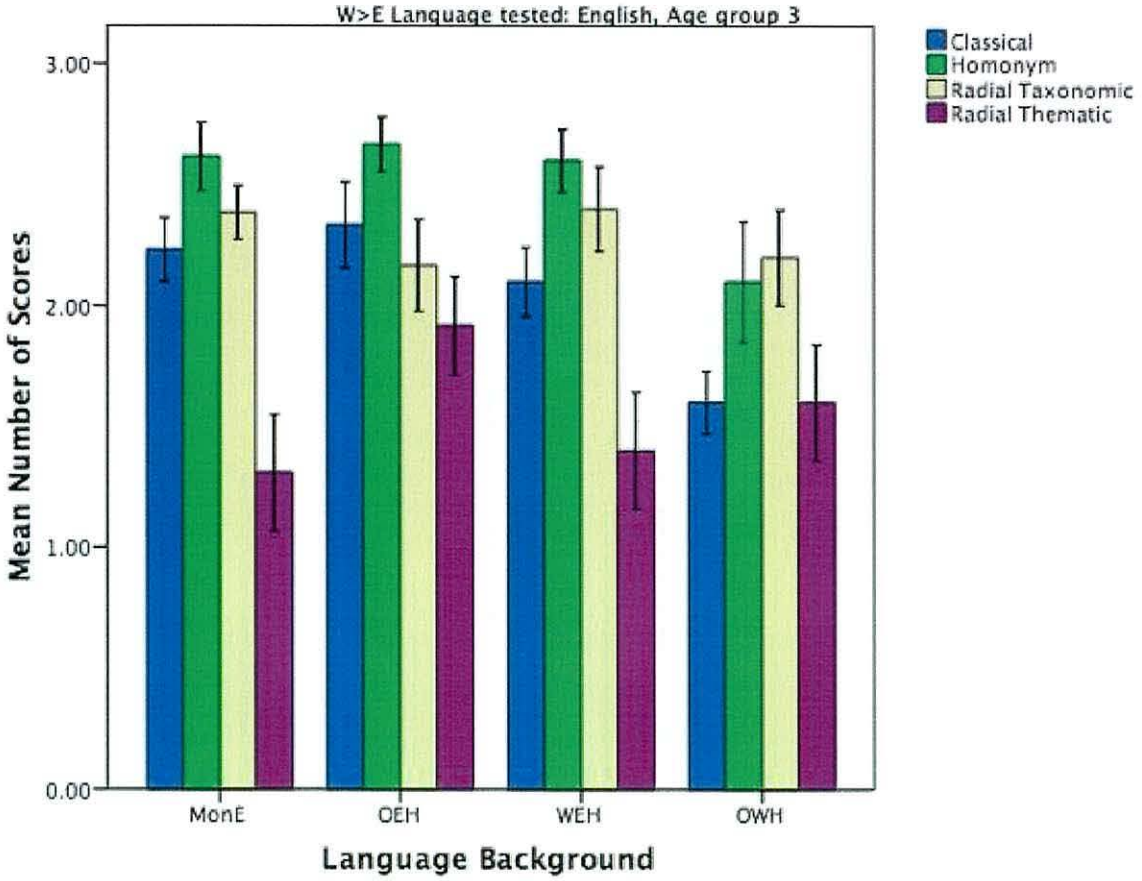


Figure 17: Word Group x Home Language where Welsh is wider than English by children aged 9 to 11

7.6 Categorisation task with reaction times

This element of the categorisation task study aimed to identify whether there were any differences in the time taken to respond to the choices presented during the test. The inclusion of reaction times in the study design was to gauge whether there could be differences across groups in the time taken required to make a decision. Differences in reaction times of choices observed in participants of different groups may indicate that participants from some language backgrounds were less certain that a particular referent was an appropriate choice in the test

language. The collection of data in this manner also showed the frequency of T1 being selected before T2 and conversely how often T2 was selected before T1.

The reaction time data were recorded along with the collection of the categorisation task choice data. As each child participating in the test pressed the 'go' button the reaction time function began, the position of each selection was noted and the time in milliseconds when each picture was selected (i.e. when a child pressed the image of the word on the screen).

In the English test condition, with items therefore wider in English than in Welsh, children selected T1 directly before T2 on 85% on items selected from the homonym group. Children were also quicker in the homonym group to select the second choice (T2) when compared to other groups. In instances when T2 was selected before T1, children were again more likely to select both target sequentially on items in the homonym group. There was also a contrast noted in the difference in reaction times in performance by children on items the homonym group when T1 was selected before T2 (1971ms), and when T2 was selected before T1 (3460).

When children were tested in Welsh, with items therefore wider in Welsh than in English, children selected T1 directly before T2 predominantly in the radial thematic group (49%) with homonym group items selected less frequently (7%). Children were much quicker to select items from the radial taxonomic group than the radial thematic group. Children only picked T1 directly after T2 in the radial categories (approximately the same amount of times). There was also a marked difference in reaction times in performance by children on items in both radial groups when T1 was selected before T2 (radial taxonomic 2622ms, radial thematic

3465ms), and when T2 was selected before T1 (radial taxonomic 2439, radial thematic 2998ms).

7.7 Discussion

The basic premise of the constructivist model is that a dual approach of semantic organisation and cognitive processing is required for children to manage two languages. Linguistic organisation by itself is insufficient to enable the child to navigate through the complexity of actions required to make sense of his/ her world. The results were predicted to show that the performance of all of the bilinguals would possibly differ from each due to differing language dominance, but all were likely to perform differently to the monolingual group in the English test condition. In the Welsh condition, the OWH group had been predicted to do best.

The investigation of errors, i.e. in which direction (taxonomically or thematically) the children extended their choices, is examined in Chapter 8. Although the extended choices data was not intended to be an error analysis for all the selections beyond T1 (or T1 and T2 when an item was wider) for each item. A child received a score of one for an item if they correctly choose 1 or 2 pictures. If a child scored zero for an item, the data did not include the reason why the participant's choice was incorrect. For example, if an item was narrower (only 1 correct picture should be selected) and a participant scored zero; they may have selected both target items, only target 2, or only selected some of the other taxonomic or thematic distractors. The possibilities for the inclusion of this supplementary analysis is further discussed in chapter 10.

7.7.1 Children tested in Welsh

In interpreting the results the initial point of reference was to look at any differences between language background groups, specifically on performance on the classical category. Performance on the homonym words was expected to be better in all language background groups as there was no conceptual link between the two words representing the category, only a homophonic relationship. In the homonym group, when the item was narrower in the test language, there should be minimal interference from the dominant language, therefore there should be no reason to expect a carry over from one language to the other.

All children performed better on items which were narrower (English wider than Welsh), with performance on items when Welsh was wider tending to be low in all groups. Children in the WEH group tended to perform similarly across all word groups on items which were wider in Welsh, while the other two home language groups do not perform well on homonym and radial taxonomic groups (which is the opposite of performance on items where English is wider). The relatively lower performance on homonym groups, especially in the Only Welsh at home group, was not expected in predicting the outcomes of the study.

The study results showed that on items where Welsh was wider than English, children performed better on homonym and radial taxonomic categories. This result was true for all participants from all three age groups. This result in line with the prediction made previously that children would have the greatest difficulty on the classical items and would probably do worse on radial thematic than radial taxonomic.

Children in the older age group performed better on items where Welsh was wider than English when compared to the other two age groups. When Welsh was wider, children

appear to be more conservative with their choice in all language background groups only selecting the first target on the majority of items. Older children from the OWH group performed the best on items in the classical word group which is in line with the prediction that classical word categories would be the most difficult for children. As predicted, where Welsh was wider than English, the OEH group (least like the OWH group) performed worst on the classical categories. A possible explanation for the performance of children from the OEH group is that they are underextending the wider category as a result of the influence of the narrower category in the dominant language. This is particularly the case for items in the classical group (although it is important to note that performance was generally low on classical groups).

7.7.2 Children tested in English

When English was wider than Welsh there were generally no differences found between the performance of the different home language groups. Generally all children did much better in this condition compared to when Welsh was the wider language. It had been expected that children in the OWH group language may have difficulty with stimuli which were wider in English than Welsh as a result of over extending the word in the narrower language due to the influence of their dominant language (Welsh). This was not confirmed by the results of the study. Children were possibly cautious in their selection. Although all language background groups tended to do worse on the classical word group and performance did not tend to be any better in the older age group.

When children were tested in English, similarities were noted to the results obtained

for children tested in Welsh. Regarding English wider than Welsh items, the performance in relation to Width demonstrated that children find it easier to correctly select items which were in the Welsh wider than English (the test item in English being narrower.). When English was wider, all children (including monolingual English) the 5-7 year old children and the 7-9 year old children performed much better in the classical group than for other word groups.

Children generally performed less well on radial items, especially on radial taxonomic items, when an item was wider in the English test condition compared to the other word groups. The finding that the performance on radial thematic items was better than that for radial taxonomic items was also observed in the older children tested in Welsh (although the performance of the radial taxonomic group was not markedly worse than the other word groups between children tested in Welsh).

When Welsh was wider, children again performed generally better for the taxonomic and homonym (homonym best) items although they also generally performed fairly well on classical items (with monolingual constantly performing in a similar way for classical items and radial taxonomic items).

Also when Welsh was wider, there seemed to be an improvement in the performance of children in the OWH group on classical items by children aged 9-11 compared to children aged 7-9 from the same home language. Children from the other home language groups in both age groups (7-9 year olds and 9-11 year olds) perform better on classical items when compared to children from the OWH group. This is again in line with the prediction that the OWH group would have the greatest difficulty for items in the classical word group, possibly as a result of overextending the word category to in narrower language (English) as a result of

a broader category in the dominant language (Welsh).

The reaction time observations also provided some supplementary information, showing that the possible effects of width may be differ across different word groups in English and in Welsh (e.g. the tendency for children tested in English to choose T1 and T2 directly after each other for items in the homonym group, whilst children tested in Welsh tended to choose T1 and T2 directly after each other for items in the radial thematic group). This finding merits further investigation. Extrapolating from the results in both English and Welsh, it appears that the level of exposure, and therefore input, does have an effect on the ability to correctly categorise items. This is particularly the case when the structure of a category is more opaque for a child to learn its boundaries.

Chapter 8. Taxonomic and Thematic Choices

8.1 Introduction

A third set of analyses was conducted on the same data to examine the choices of distractor items in order to identify whether children extended their choices beyond the targets to either taxonomically or thematically related referents. The study was designed to identify influential factors involved in the categorisation of specific word categories – taxonomic and thematic and to compare the performance of children from different language backgrounds. The study used non-verbal stimuli test which was performed by individual children in either English or Welsh. The study adopted Markman's (1991) theory of taxonomic bias, which proposed that children will prefer to associate new words with objects that are taxonomically related rather than thematically related. Markman argued that young children naturally label new concepts with a mutually exclusive label i.e. each concept has its own classification. They only gradually lose this tendency as they learn about more complex relationships.

8.2 Participants

As outlined in Chapter 7, participants were recruited from 14 primary schools in North and West Wales, 1 school in Cheshire and 1 school in Berkshire. Participants in the study were placed into one of four different groups depending on their home language background, monolingual English, OEH, OWH and WEH.

8.3 Non-linguistic stimuli

As outlined in 7.3.3, the picture stimuli consisted of 6 pictures for each test item. The selection of pictures for each test item was chosen according to a clearly defined criteria. This consisted of two of the six pictures representing two target referents, two taxonomically linked (to the targets) distractors, and two thematically linked (to the targets) distractors. The study was designed to include two language conditions, with the Welsh condition involving only Welsh instructions and the English condition only involving English instructions. The words identified for the categorisation task were wider in one language than the other (W>E for Welsh wider, E>W for English wider) i.e. the two targets for a test item could be named by a single word in the wider language, however would be named by two words in the narrower language. For example in the Welsh condition, the word *codi* is wider in Welsh than in English. In Welsh *codi* can represent lifting a box or getting out of bed. Therefore for the two target stimuli for the *codi* test item, the word *codi* could be used to select (or name) both target pictures. Whereas in the English condition (narrower language) the word *lifting* should only be used to name the first target (T1), as lifting a box, but it would not be appropriate to name the second target (T2), *lifting* in the sense of getting out of bed (see Figure 18). The task was designed so that there were four different word categories with two nouns and one verb for each word category in both languages.

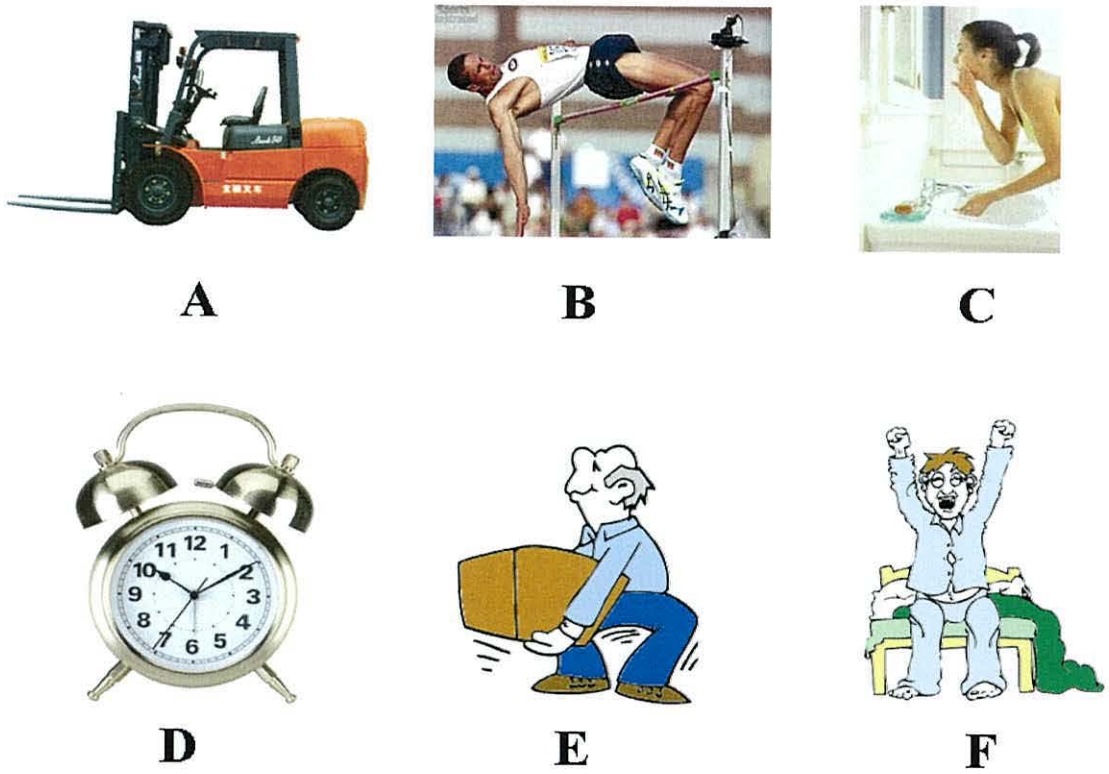


Figure 18: An example of non-linguistic stimulus from the categorisation task item *codi* used in the Taxonomic and Thematic Study

8.4 Results

8.4.1 Welsh

Analyses of variance were performed on the data in which width (Welsh wider than English, and English wider than Welsh), word category group (classical, homonym, radial taxonomic and radial thematic) and extended choice (i.e whether extended choice beyond target was taxonomic or thematic) were within subject variables, and age group (children aged 5 to 7, those aged 7 to 9 and 9 to 11) and home language group (OEH, WEH, and OWH) were treated as between group variables. The number of times participants selected taxonomic and thematic extensions for items in each of the four word groups was recorded. The results

showed significant main effects for word group, F (3,243)= 165.43, p<.001; width, F (1,81) =15.84, p<. 001; and extended choice group F (1,81)= 167.35, p<.001. Results of between-subject factors did not show significant main effects for age group p<.524; nor for home language p<.232.

When tested on items from the classical category, in general children chose significantly fewer extensions than with the homonym word group *MD* -.111, p<.024; with the radial taxonomic word group *MD* -.564, p<.001; and the radial thematic group *MD* -.922, p<.001. There was a significant difference between performance for homonym and radial taxonomic groups *MD* .453, p<.001, and with radial thematic groups *MD* .811, p<.001. (Children performed better on homonym group items than the other two word category groups) There was also a significant difference between children's performance for radial taxonomic and radial thematic groups: *MD* .358, p<.001. In addition, there was a significant difference between the amount of thematic and taxonomic choices made *MD* .121, p<.001.

The main effects were modified however, with two and three way interactions for Word Group x Age Group, F (6, 243)= 3.53, p<.002; Word Group x Home Language group, F (6, 243)= 3.80, p<.001; Word Group x Age Group x Home Language group F (6, 243) = 3.80, p<.001, Word Group x Width F (3, 243) = 44.98, p<.001; Word Group x Width x Age, F (6, 243) = 2.61, p<.018; Word Group x Extended Choice group, F (3, 243) = 45.11, p<.001; Word Group x Width x Extended Choice group F (3, 243) = 78.96, p<.001; and extended choice x Language Background group F (2,81) = 3.56, p<.033. The two way interaction: Age x Home Language was also not significant p<.482.

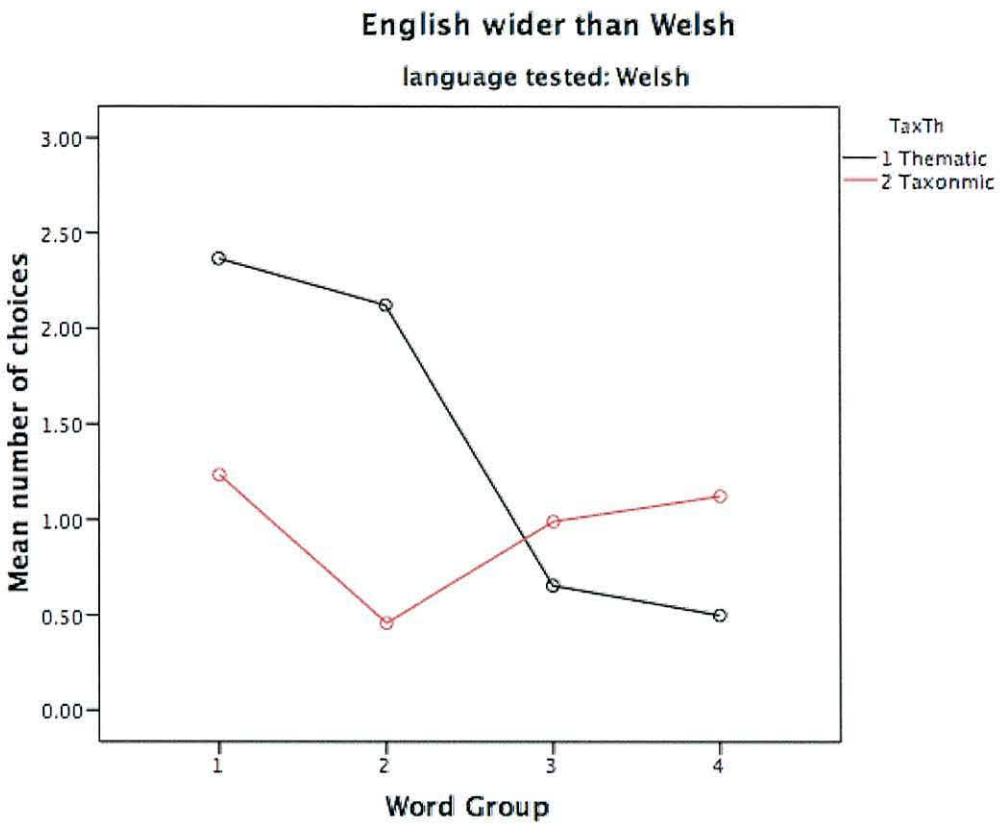


Figure 19: Extended Choices x Word Group of children tested in Welsh on items wider in English than in Welsh

Follow-up analyses indicated that for the choice of classical categories, width was significant, F (1,81) =181.74, p<. 001, and the extended choice group was also significant F (1,81) =20.55, p<. 001. The three way interaction between Width x Age group x Home Language group was significant F (4,81) =2.73, p<. 001. A significant interaction was also found in Width x Extended Choice group F (1,81) =65.18, p<. 001. When Welsh was wider than English there were no significant differences between the performance of children from

different age groups or between language background groups, although there was a significant difference in the number of taxonomic (.989) and thematic choices (.656) made by children p<.001. The difference between taxonomic and thematic extensions by children was significant with children selecting more thematic (2.37) pictures than taxonomic (1.23) pictures.

Performance relating to the homonym categories indicated that width was significant, F (1,81) =5.76, p<. 019, extended choice group was also significant F (1,81) =34.19, p<. 001. There were also two way interactions between width x extended choice, F (1,81) =74.66, p<. 001. Children made significantly less taxonomic (.533) than thematic choices (1.644) p<.001.

For the radial taxonomic words width was significant, F (1,81) =52.58, p<. 001, as was the Width x Age Group F (2,81) =3.53, p<. 034, and the Width x Home Language interaction. Extended choices was also significant F (1,81) =34.19, p<. 001. The Width x Extended Choice interaction was significant, F (1,81) =204.68, p<. 001. There were also significant differences to be seen on items where English was wider than Welsh between children aged 9 to 11 and 7 to 9 (*MD* .433), p<.006. In the case of home language groups, children from the OWH group extended their selection less than children from the OEH group (*MD* -.483,) p<.002 and from children from the WEH home group (*MD* -.383,) p<.014. Children also made significantly more thematic choices (2.122) than taxonomic (.456) when English was wider. There was a significant difference within the extended choice group F (1,81) =32.29, p<. 001, with children selecting significantly more taxonomic pictures (1.122) than thematic choices (.500) p<.001.

For the radial thematic words width was significant, F (1,81) =271.64, p<. 001, as was

the Width x Age Group F (2,81) =9.34, $p<.$ 001. The number of extensions made by children was also significant F (1,81) =70.48, $p<.$ 001. The results for the home language groups was close to being significant F (2,81) =2.56, $p<.$ 083.

When English was wider than Welsh there were significant differences between performance of children from the different age groups F (2,81) =3.94, $p<.$ 0823. Children aged 9 to 11 (1.78) performed significantly better than those aged 5 to 7 (1.28) $p<.006$. There was a significant difference within the extended choice group F (1,81) =49.80, $p<.$ 001, with children selecting significantly less taxonomic pictures (1.133) than thematic choices (1.922) $p<.001$.

The difference between the number of choices made by children in the OWH group and the Welsh and English at home groups were nearing significance $p<.053$. Children differed in their extended choices F (1,81) =27.81, $p<.$ 001, with children selecting significantly less taxonomic pictures (.122) than thematic choices (.644) $p<.001$.

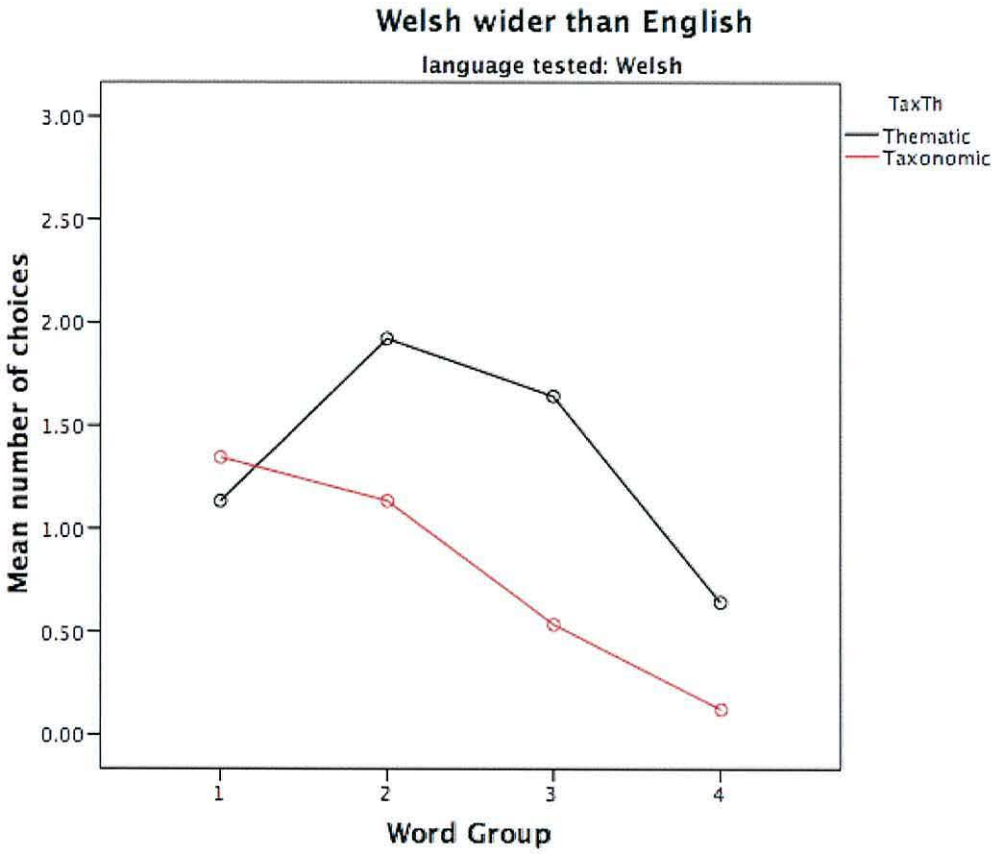


Figure 20: Extended Choices by Word Group of children tested in Welsh on items wider in Welsh than in English

To examine these effects further analyses were conducted on each age group separately. Children aged 5 to 7 showed differences by word groups F (3,81) 52.65, $p<$. 001, by width F (1,27) =4.81, $p<$. 037, and by extended choice group F (1,27) =40.07, $p<$. 001. The following two and three way interactions were also significant, Word Group x Width F (3,81) =14.52, $p<$. 001. Word Group x Width x Home Language F (6,81) = 2.23, $p<$. 048, Word Group x Extended Choice Group F (3,81) =14.51, $p<$. 001, and Word Group x Width x Extended Choice Group F (3,81) =20.89, $p<$. 001.

For children aged 5 to 7 in the OEH group, word group was found to be significant F (3,27) =23.38, p<. 001 although width was not found to be significant. These children selected the most distractor items on items from classical and homonym groups when compared to the radial groups ps<.050. Children aged 5 to 7 also selected more thematic distractors than taxonomic alternatives (MD.513, p<. 002. Two and three way interactions were also found between Word Group x Width F (3,27) =8.12, p<. 001, Word Group x Extended Choice Group F (3,27) =5.56, p<. 004, and Word Group x Width x Extended Choice Group F (3,27) =4.44, p<. 012.

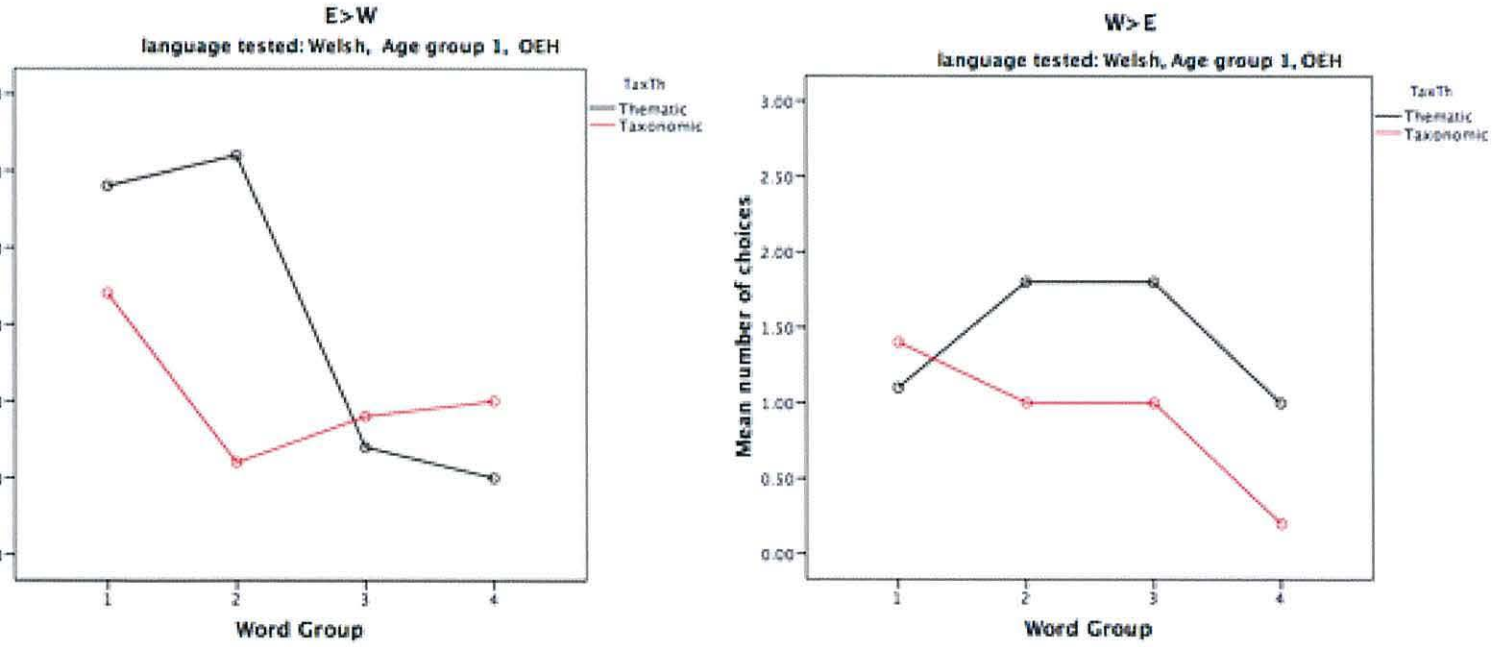


Figure 21. Word Group by Extended Choice in the OEH Group
for children aged 5 to 7.

For children aged 5 to 7 from the WEH group word group was significant F (3,27) =11.52, p<. 001 although width was again not found to be significant. Children aged 5 to 7 did not make many extensions in the radial categories (especially in the case of radial thematic group items) and the majority of distractor choices were made in the homonym and classical group items (pairwise comparisons homonym vs radial taxonomic p<.006, homonym vs radial taxonomic).The extended choice group was significant F (1,9) =20.76, with children aged 5 to 7 selecting markedly more thematic than taxonomic distractors (*MD*.675) p<. 001. Two and three way interactions were also found between Word Group x Extended Choice group F (3,27) =8.43, p<. 001, and Word Group x Width x Extended Choice group F (3,27) =10.16, p<. 001.

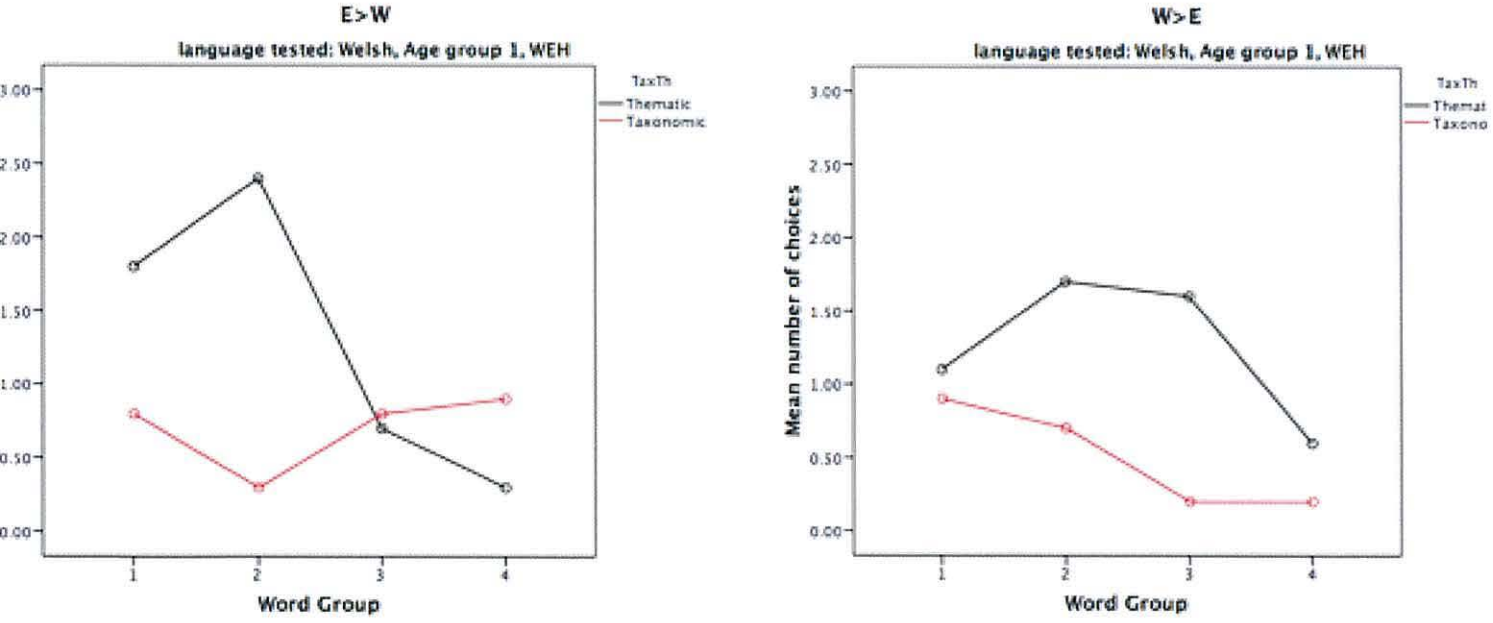


Figure 22. Word Group x Extended Choice in the WEH group by children aged 5 to 7

For children aged 5 to 7 from the OWH group, word group was significant F (3,27)

=25.13, $p<.001$ although width, as with the other home language groups, was not found to be significant. Children aged 5 to 7 selected significantly more items in the classical group than the other three word groups $ps<.002$, and conversely selected significantly less in the radial thematic group than the other three word groups $ps<.001$. Children aged 5 to 7 from the OWH group tended to select the same number of taxonomic and thematic distractors. Two and three way interactions were found between Word Group x Width F (3,27) =8.61, $p<.001$, and Word Group x Width x Extended Choice Group F (3,27) =9.74, $p<.001$.

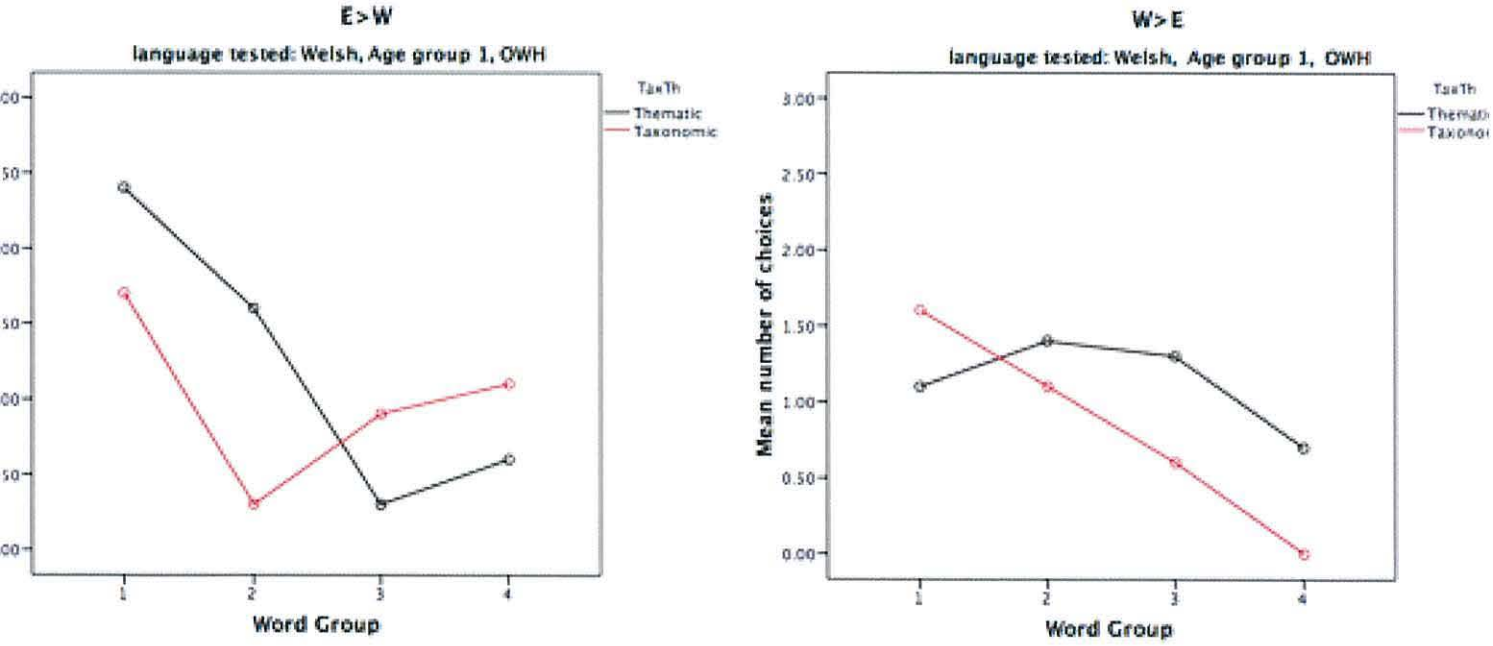


Figure 23. Word Group x Extended Choice in the OWH group
by children aged 5 to 7

Children aged 7 to 9 showed main effects for word group F (3,81) =58.89, $p<.001$. width was not significant The extended choice group was significant F (1,27) =67.14, $p<.001$, as was Word Group x Width F (3,81) =16.52, $p<.001$, Word Group x Home Language F

(6,81) = 3.76, $p<.048$, Word Group x Extended Choice Group F (3,81) =20.27, $p<.001$, and Word Group x Width x Extended Choice Group F (3,81) =22.61, $p<.001$. Children aged 7 to 9 from the OEH group selected more distractors in the classical group than on radial thematic items $p<.004$, and more homonym distractors than both radial taxonomic ($p<.004$) and radial thematic items ($p<.001$). Children also selected more distractors on radial taxonomic items than radial thematic items ($p<.007$). The extended choice group was significant F (1,9) =17.80, $p<.002$. Two and three way interactions were also found between, Word Group x Extended Choice Group F (3,27) =6.04, $p<.003$, and Word Group x Width x Extended Choice Group F (3,27) =4.79, $p<.008$.

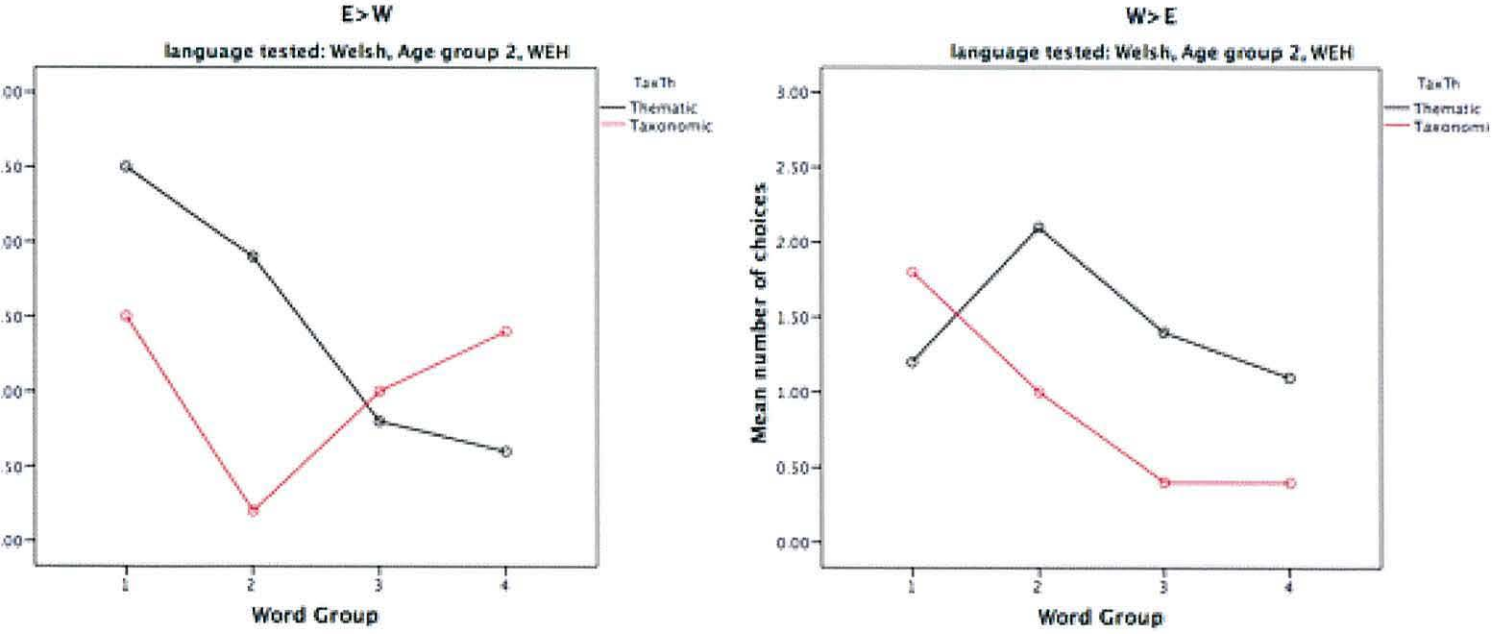


Figure 24. Word Group x Extended Choice in the OEH group
by children aged 7 to 9

For children aged 7 to 9 from the WEH group, word group was significant F (3,27)

=16.45, $p<.001$,. Width was not found to be significant. Children aged 7 to 9 also selected significantly more thematic than taxonomic choices F (1,9) =27.99, $p<.001$. Children of this age group again selected more distractors on classical word items than the other three word groups $ps<.001$. Two and three way interactions were also found between Word Group x Extended Choice F (3,27) =7.24, $p<.001$, and Word Group x Width x Extended Choice Group F (3,27) =7.16, $p<.001$.

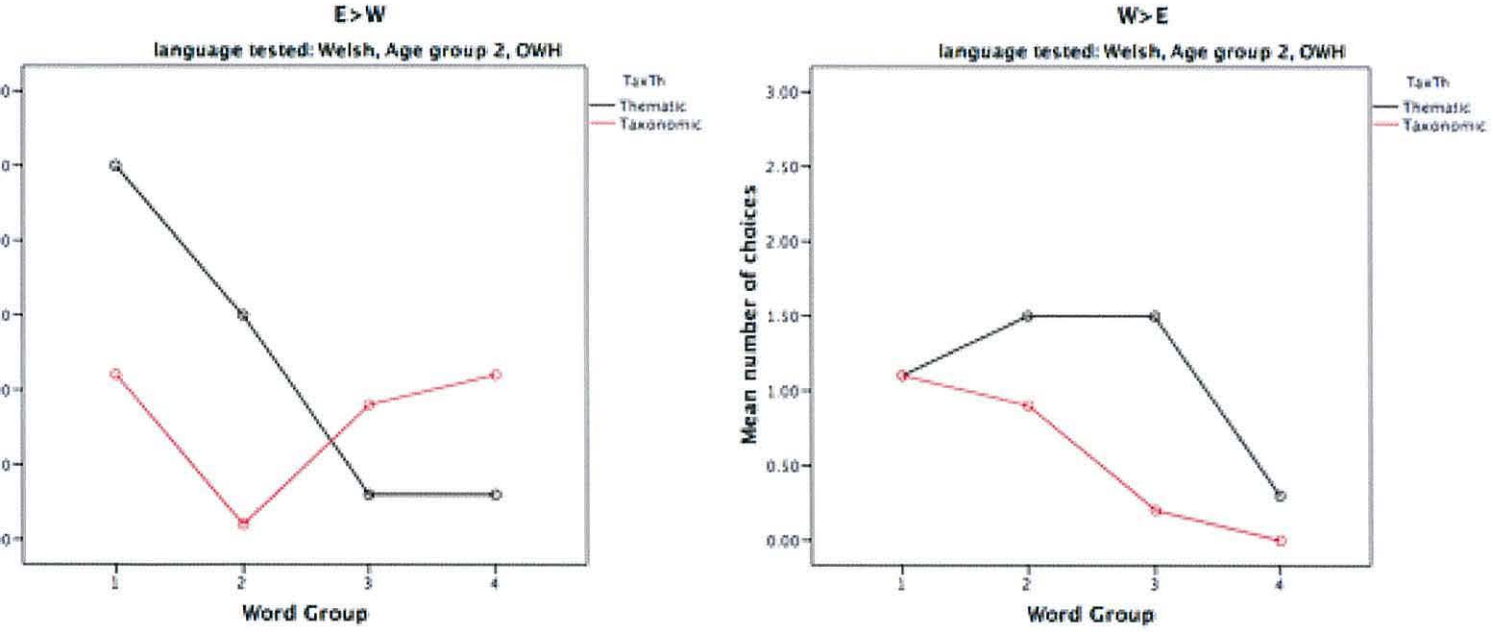


Figure 25. Word Group x Extended Choice in the WEH group by children aged 7 to 9.

Children aged 7 to 9 from the OWH group performed similarly to the other home language groups selecting distractors on the classical items than on the other word groups $p<.001$. Width was not found to be significant. The extended choice group was again significant, with more thematic distractors selected F (1,9) =24.10, $p<.001$. Two and three way

interactions were also found between Word Group x Width F (3,27) =11.68, p<. 001, Word Group x Width F (3,27) =13.60, p<. 001, and Word Group x Width x Extended Choice Word Group F (3,27) =14.55, p<. 001.

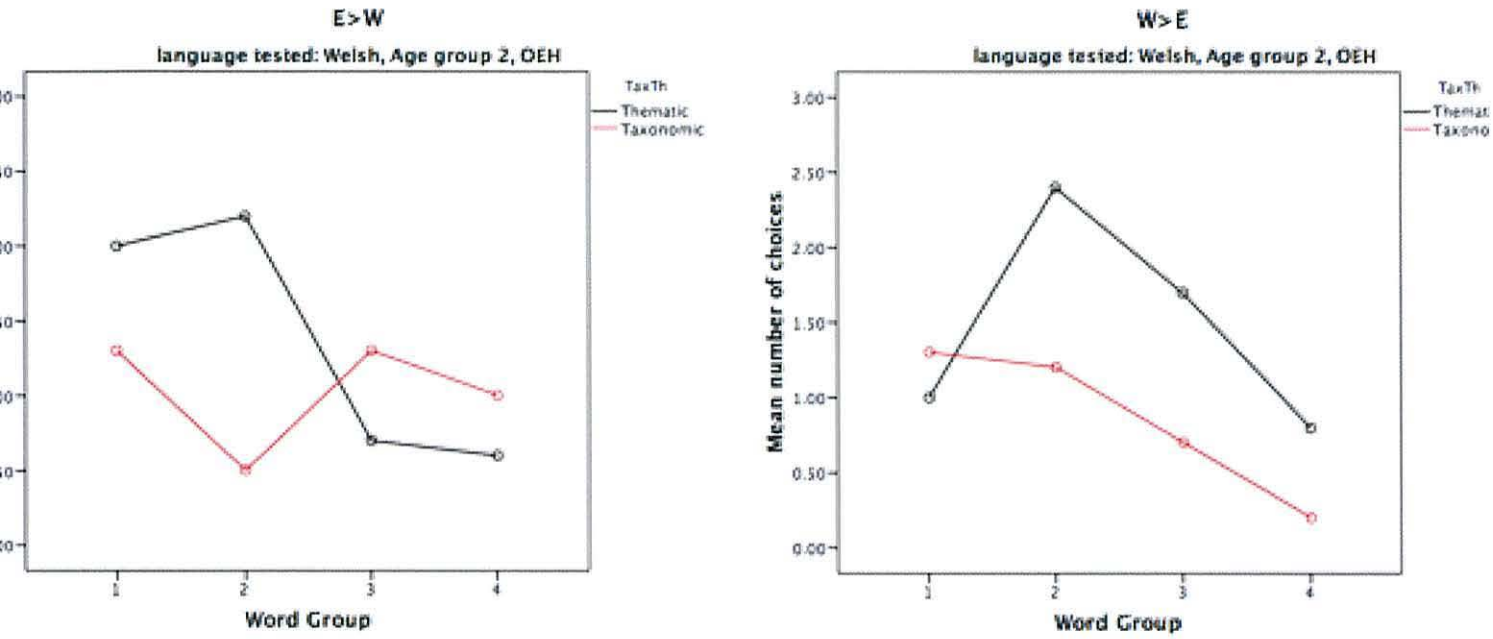
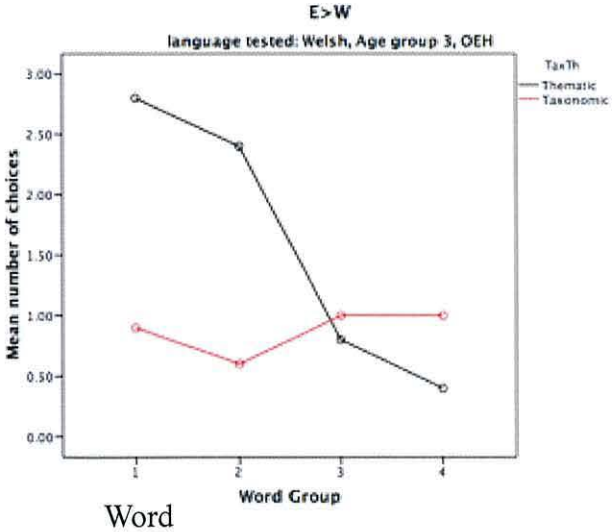


Figure 26: Word Group x Extended Choice in the OWH group
by children aged 7 to 9

Children aged 9 to 11 showed main effects for word group F (3,81) =60.28, p<. 001. Width was also found to be significant F (1,27) =13.76, p<. 001. The extended choice group was again significant F (1,27) =68.32, p<. 001. Two and three way interactions were also found between Word Group x Width F (3,81) =18.90, p<. 001, Word Group x Extended Choice Group F (3,81) =13.51, p<. 001, and Word Group x Width x Extended Choice Group F (3,81) =41.91, p<. 001.

Children aged 9 to 11 from the OEH group selected more distractors in the classical

group than both radial groups *ps*<005, as was the case on homonym test items *ps*<.001, with word group being significant *F* (3,27) =26.29, *p*<. 001. Children of this age also selected fewer distractors on radial taxonomic items than on radial thematic items *p*<.003. There was once again no difference in the number of distractors selected between the narrower and wider items. These children again selected more thematic distractors (*MD*.575) *F* (1,9) =49.59, *p*<. 001. Two and three way interactions were also found between, Word Group x Width *F* (3,27) =6.54, *p*<. 002, Word Group x Extended Choice Group *F* (3,27) =5.01, *p*<. 007, and Word Group x Width x Extended Choice Group *F* (3,27) =19.35, *p*<. 001.



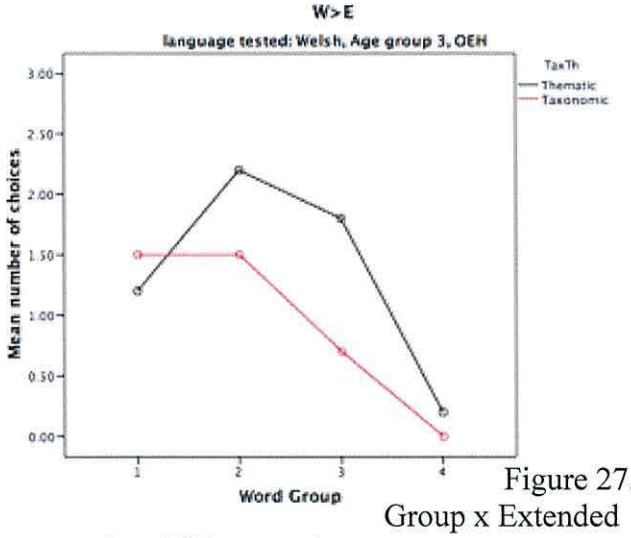


Figure 27. Word Group x Extended

Figure 27: Choice in the OEH group by children aged 9 to 11.

For children aged 9 to 11 from the WEH group word group was significant *F* (3,27) =19.75, *p*<. 001, with pairwise comparison showing more distractors selected in the classical category than both radial groups *ps*<.005, more in homonym group items than both radial groups *ps*<.010, and between radial taxonomic and thematic *p*<.016. Width was not found to

be significant. Thematic distractor choices were again prevalent (*MD*.713) F (1,9) =20.29, p<. 001. Two and three way interactions were also found between Word Group x Width Group F (3,27) =6.49, p<. 002, Word Group x Extended Choice Group F (3,27) =3.73, p<. 023, and Word Group x Width x Extended Choice Group F (3,27) =13.81, p<. 001.

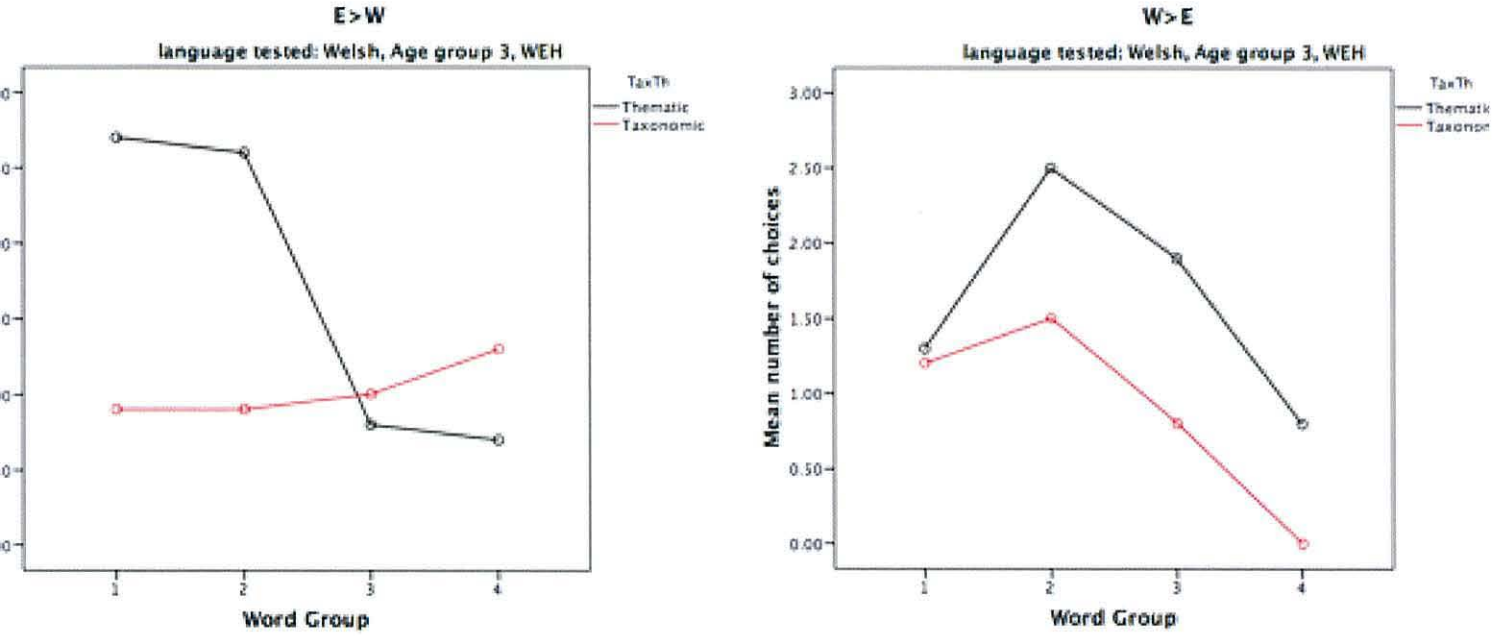


Figure 28: Word Group x Extended Choice in the WEH group
by children aged 9 to 11

Children aged 9 to 11 from the OWH group performed similarly to the other two home language groups by selecting more distractors in the classical and homonym groups compared to both radial groups *ps*<.050 (word group being significant F (3,27) =15.97, p<. 001). Comparisons between the number of items selected between widths was also found to be significant F (1,9) =18.69, p<. 002. Children from this age group again selected more thematic distractor choices than taxonomic distractors F (1,9) =16.00, p<. 003. Two and three way interactions were also found between Word Group x Width F (3,27) =6.40, p<. 002,

Word Group x Extended Choice F (3,27) =7.59, p<. 001, and Word Group x Width x Extended Choice Word Group F (3,27) =11.17, p<. 001.

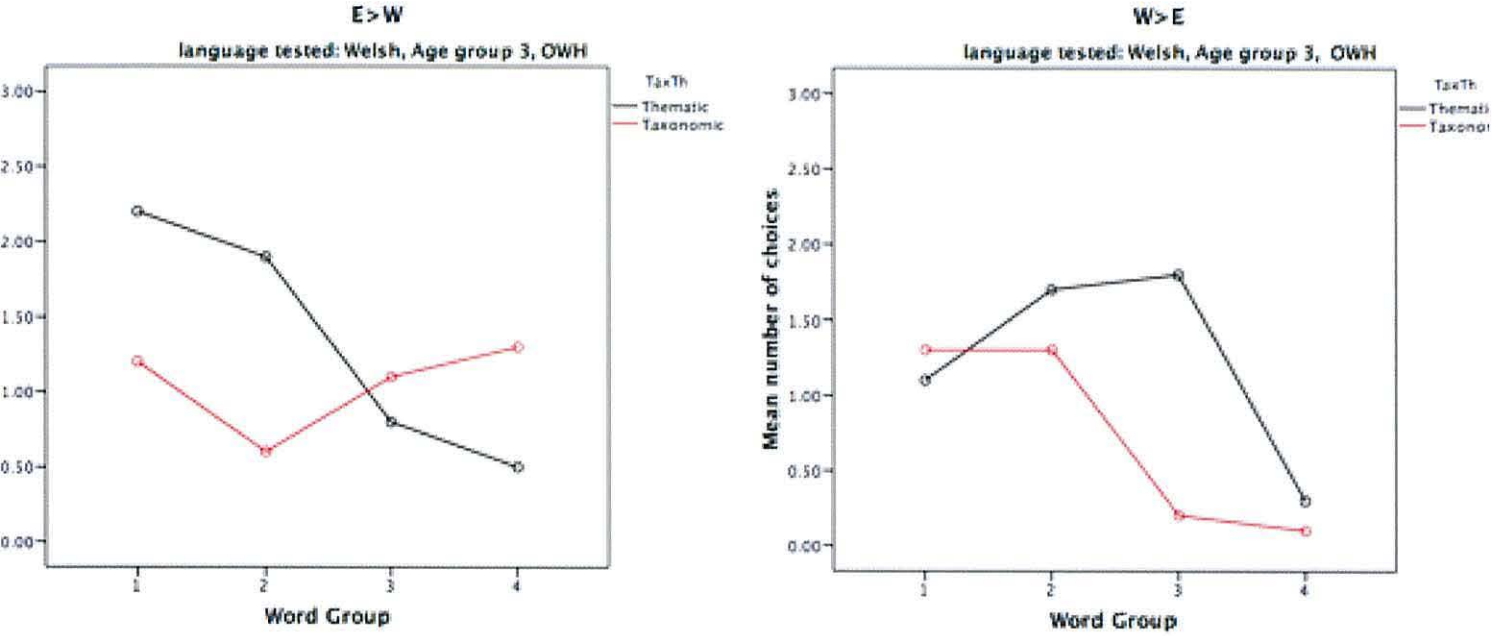


Figure 29: Word Group x Extended Choice in the OWH group by children aged 9 to 11

8.4.2 English

Results of the main effects showed that word group, F (3,324)= 179.05, p<.001, width, F (1,108) =10.42, p<. 001, and for extended choice F (1,108)= 267.29, p<.001. Results of between-subject factors did not show main effects for age group or and home language group. Two way and three way interactions were found to be significant for Width x Word Group, F (3,324)= 134.47, p<.001; Extended Choice Group x Age Group F (3,324) = 9.61, p<.001, and Word Group x Width x Home Language Group, F (6,324)= 2.16, p<.004.

Follow-up analyses investigating each word category indicated that on classical categories, width was significant, F (1,108) =667.56, p<. 001, extended choice group was also

significant F (1,108) =13.27, $p<.$ 001 with pairwise comparisons showing that children chose more thematic choices (1.721) than taxonomic (1.438) choices. There was a significant difference between the number of thematic (2.14) and taxonomic choices (2.41) made by children p<.046. The difference between taxonomic and thematic extensions by children was again significant when Welsh was wider than English, F (1,108) =69.23, with children selecting less taxonomic (.842) objects than thematic (1.68) referents $p<.001$.

Performance on the homonym categories indicated that width was significant, F (1,108) =8.36, $p<.$ 005, extended choice was also significant F (1,108) = 252.94, $p<.$ 001. There were also two way interactions between Width x Extended Choice, F (1,108) =4.55, $p<.$ 035, Extended Choice x Age Group, F (2,108) =13.32, $p<.$ 001, and a three way interaction between Width x Extended Choice x Age Group, F (1,108) =3.87, $p<.$ 024. There were no significant differences between age group or home language group. When English was wider than Welsh there was no significant differences found between the performance across age groups and between home language groups. There was a significant difference found between the number of taxonomic and thematic extensions made by children with pairwise comparison showing more thematic (2.15) than taxonomic (.733) choices. There was also a significant interaction between Extended Choice Group x Age Group F (1,108) =11.41, $p<.001$. When Welsh was wider than English there were no significant differences between age groups and between home language groups, although there was a significant difference between the number of taxonomic (1.34) and thematic choices (2.41) made by children $p<.001$. There was also a significant interaction between Extended Choice x Age group F (1,108) =3.29, $p<.041$.

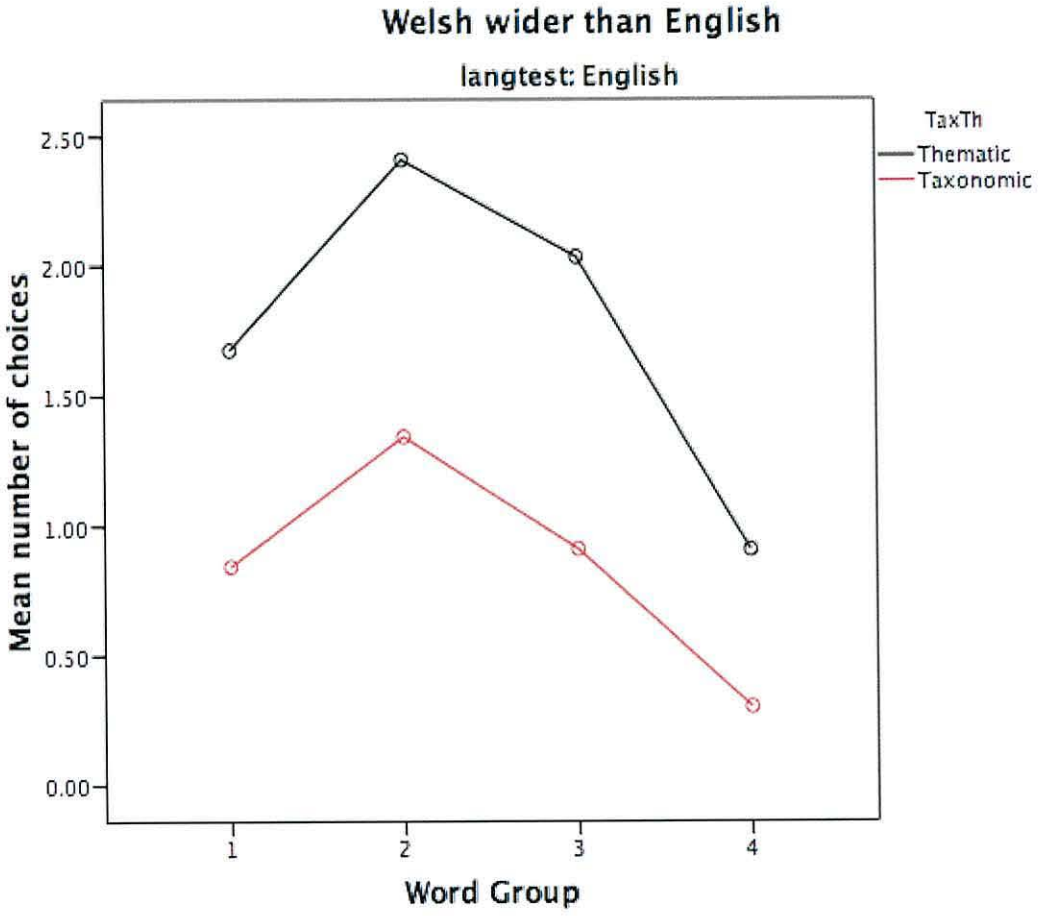


Figure 30: Extended Choice x Word Group by children tested in English on items Welsh wider than English

For the radial taxonomic words width was significant, F (1,108) =47.48, p<. 001. The extended choice group was also significant F (1,108) =81.75, p<. 001, with a pairwise comparison showing children selecting more thematic choices than taxonomic (*MD*.592) p<.001. The Width x extended choice interaction was significant, F (1,108) =119.20, p<. 001. There were no significant differences found between age groups or home language groups. Where English was wider than Welsh, children selected significantly less thematic choices

(.733) than taxonomic (1.13), $p<.003$. When Welsh was wider than English there were no significant differences between age groups and between home language groups, although there was a significant difference with extended choice F (1,108) =5.29, $p<$. 001, with children selecting significantly more thematic referents (2.03) than taxonomic referents (.908) $p<.001$.

For the radial thematic words width was significant, F (1,108) =332.63, $p<$. 001, as was the Width x Age Group F (2,108) =5.43, $p<$. 006. the extended choice group was also significant F (1,108) =158.57, $p<$. 001, with children selecting more thematic (1.658) than taxonomic (.821) referents $p<.001$. Width x Extended Choice Group F (2,108) =7.72, $p<$. 006. was also significant. There were no significant differences between age group nor home language group, although home language groups were close to being significant F (2,81) =2.56, $p<$. 083. When English was wider than Welsh there were no significant differences found between language background groups or age groups. Within the extended choice group F (1,108) =82.75, $p<$. 001, children selected significantly less thematc referents (.900) than taxonomic choices (1.29) $p<.015$ When Welsh was wider, children aged 5 to 7 made significantly more extensions than children aged 7 to 9 2 (*MD* .450) $p<.009$, and children aged 9 to 11 (*MD*.363) p.035. Differences were not found between home language groups. There was a significant difference between the extended choices F (1,108) =42.48, $p<$. 001, with children selecting significantly taxonomic referents (.300) than thematic choices (.908), p<001..

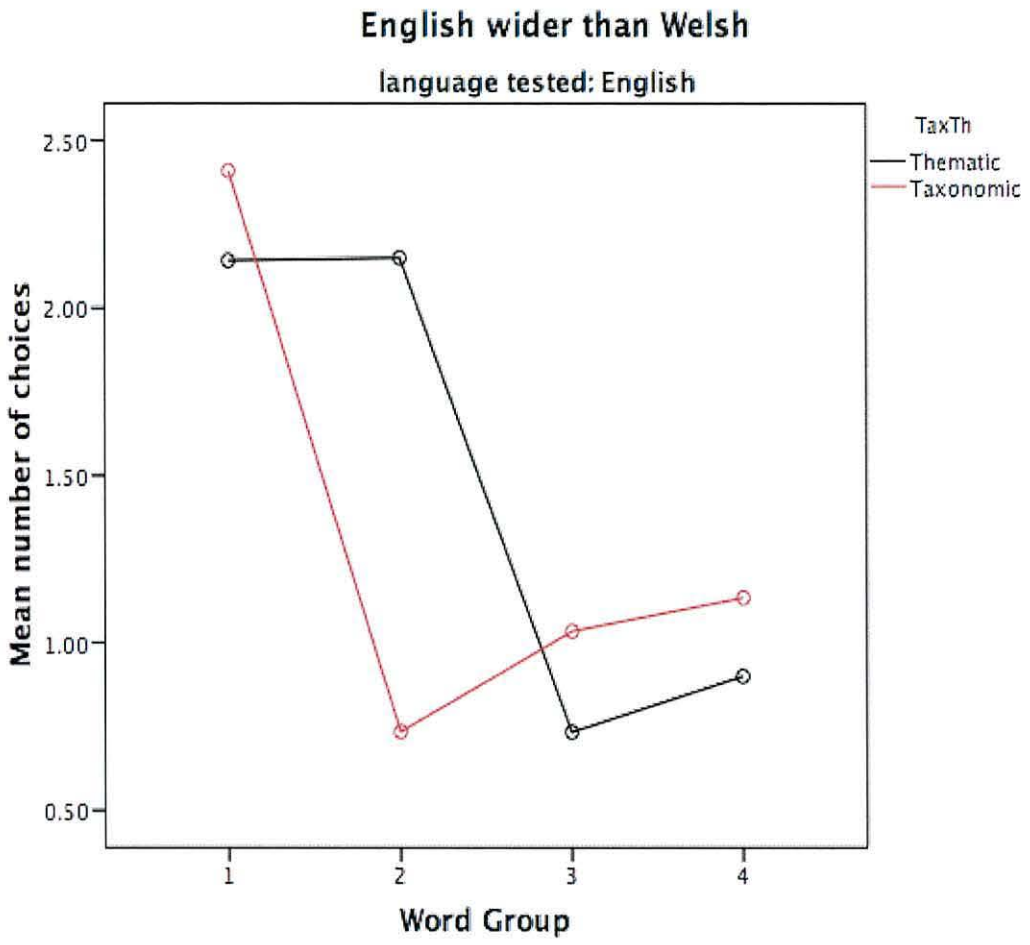


Figure 31: Extended Choice by Word Group by children tested in English on items English wider than Welsh

8.5 Discussion

8.5.1 The performance of children tested in Welsh

When children were tested in Welsh on Welsh wider items, children from all language background groups selected thematic distractors more often than taxonomic in all word groups. This did not follow the prediction that children would predominantly select more taxonomic distractors. Although the study did not make any specific predictions regarding

differences between word groups, children selected significantly fewer taxonomic and thematic distractors in the radial thematic category in comparisons to the other word groups. Children from all age groups selected the most taxonomic distractors in the classical word category. Children aged between 5 and 7 years across all home language groups, had a preference for selecting thematic distractors in the homonym and radial taxonomic word categories, with number of selections being higher than the number of thematic choices in the classical and radial thematic word groups, and between the number of taxonomic distractor selections in the homonym and radial taxonomic groups. There tended to be a significant difference in the number of thematic choices made by children from all age groups compared to the number of taxonomic choices they made in the radial taxonomic word category. Although there was no main effect of width, children did perform in the same manner on word groups across both widths.

When English was wider than Welsh there was a tendency for children not to choose thematic distractors in either radial groups. Children tended to make more taxonomically extended choices than thematic over radial categories. In the homonym group children tend to pick fewer taxonomic distractors compared to other groups and there was generally a large difference across all groups in the choice of more thematic than taxonomic in this word group. The number of taxonomic distractors made by children when English was wider than Welsh tended to be similar across all word groups in all participants (the only exception being fewer choices made of items in the homonym group).

8.5.2 The performance of children tested in English

Children tested in English on stimuli where Welsh was wider than English consistently selected more thematic choices than taxonomic items in all word groups, with this effect being apparent across all age groups and home language groups. This was similar to the performance of children who were tested in Welsh and was again not in line with the study's predictions that more taxonomic distractors would be selected if children extended their choice beyond the targets. For the radial thematic group, children across all age groups tended to select a high number of taxonomic distractor choices. This was also the case for the number of selected thematic distractor choices, which was markedly fewer than in the radial thematic group compared to the other word groups. When test items were wider in English than Welsh, children tended to select significantly more thematic distractors selected compared to taxonomic distractors in the homonym word group. This was the case for children across all language background groups. Children consistently made the most taxonomic choices from the classical word group with a similar number of taxonomic distractors and thematic distractors chosen.

Chapter 9. Memory Task

9.1 Introduction

This chapter reports the memory task performed by the same monolingual and bilingual children as those who performed the categorisation task. The study was designed to provide additional evidence as a follow up to the main categorisation test to better understand how children categorised items from word groups (classical, homonym, radial taxonomic, and radial thematic) and to compare the performance of children from different language backgrounds. The study used non-verbal stimuli test which was performed by individual children in either English or Welsh. The task examined whether having two words for two concepts or one word encompassing both items would be most beneficial for recall by children. The task was also intended to investigate whether the structure of inclusion within a word group would also affect the ability to recall both pictorial representation of the test word (i.e. the test words for memory task were target 1 and target 2 from the categorisation task).

9.2 Procedure

Two hundred and ten children took part in the task and were allocated to individual language groups as outlined for the main categorisation task in Chapter 7. Some of the participants from the total 238 who took part in the categorisation task were omitted from the memory task. This was as a result of the initial intention for the analysis to include group totals which would have required equal participant group numbers. A memory task was performed with both representations of the noun targets (target 1 and target 2) items from the categorisation task included. This consisted of 32 target items in total. A PowerPoint

presentation was shown to the individual child with each slide displaying one digital image of the target items on the screen. Participants viewed all 32 slides once, and were then requested to count backwards from five to zero, before being instructed to say out loud, in any order, as many of the items they recalled seeing in the slide show. The request to count backwards was inserted to prevent any effects of rehearsal. Each child completed the memory task and the main categorisation task in the same language and in the same session, with the memory task completed after the main categorisation task. The items in the memory task are outlined in Table 2 below:

	One category	**Homonyms**	**Radial Taxonomic**	**Radial Thematic**
E>W				
1 2	STICK/ FFON **Walking stick** Stick	NAIL/HOELEN **Nail** Fingernail	KEY/GORIAD **Door key** Computer key	DISH/DYSGL **Plate** Food
1 2	WALL/MUR **Castle wall** House wall	CHEST/BREST **Anatomy** Chest box	HAND/LLAW **Hand of body** Hand on clock	EARTH/DAEAR **Earth (planet)** Earth/Ground
W>E				
1 2	CYSGOD/SHADOW **Shadow** Shade	YSGOL/SCHOOL **School** Ladder	PEN/HEAD **Head** Top of mountain	UCHEL/HIGH **High** Loud
1 2	BAWD/THUMB **Thumb** Big toe	HWYL/FUN **Fun** Sail	PWRS/PURSE **Purse** Udder	FFORDD/ROAD **Road** Way

Table 2: Noun items used in the Memory Task

9.3 Results

9.3.1 Tested in Welsh

Analyses of variance were performed on the data in which width (Welsh wider than English, and English wider than Welsh), word category group (classical, homonym, radial

taxonomic and radial thematic) and memory target (target 1 or target 2) were within subject variables, and age group (children aged 5 to 7, 7 to 9 and 9 to 11) and home language group (OEH, WEH, and OWH) were treated as between group variables. The results showed significant main effects for word group, F (3,243)= 10.27, p<.001; and memory target F (1,81)= 138.37, p<.001.

Children performed significantly different on items in the classical category with pairwise comparisons showing differences between the classical word group with the homonym word group *MD* .081, p<.045; and the radial thematic group *MD* -.139, p<.001 (classical group performed better than the homonym group, worse than radial thematic) There was a significant difference between the performance of children for homonym and radial taxonomic word groups *MD*.-133, p<.001. There was also a significant difference between their performance for radial taxonomic and radial thematic: *MD* .192, p<.001. In addition, there was a significant difference between the amount of target 1 and target 2 choices *MD* .289, p<.001. The main effects were modified however, with two and three way interactions found to be significant for Word Group x Width x Target Choice, F (3, 243)= 7.89, p<.002; Target Choice x Age Group F (2, 81) = 5.11, p<.008,

Follow-up analyses examining each category type indicated that on classical categories, width was not significant, but the target choice was significant F (1,81) =28.10, p<. 001. There were also two and three way interactions between Target Choice x Age Group F (4,81) =4.50, p<. 014, between Width x Target Choice F (1,81) =6.96, p<. 010, between Width x Age Group x Home Language group was significant F (4,81) =2.68, p<. 037.

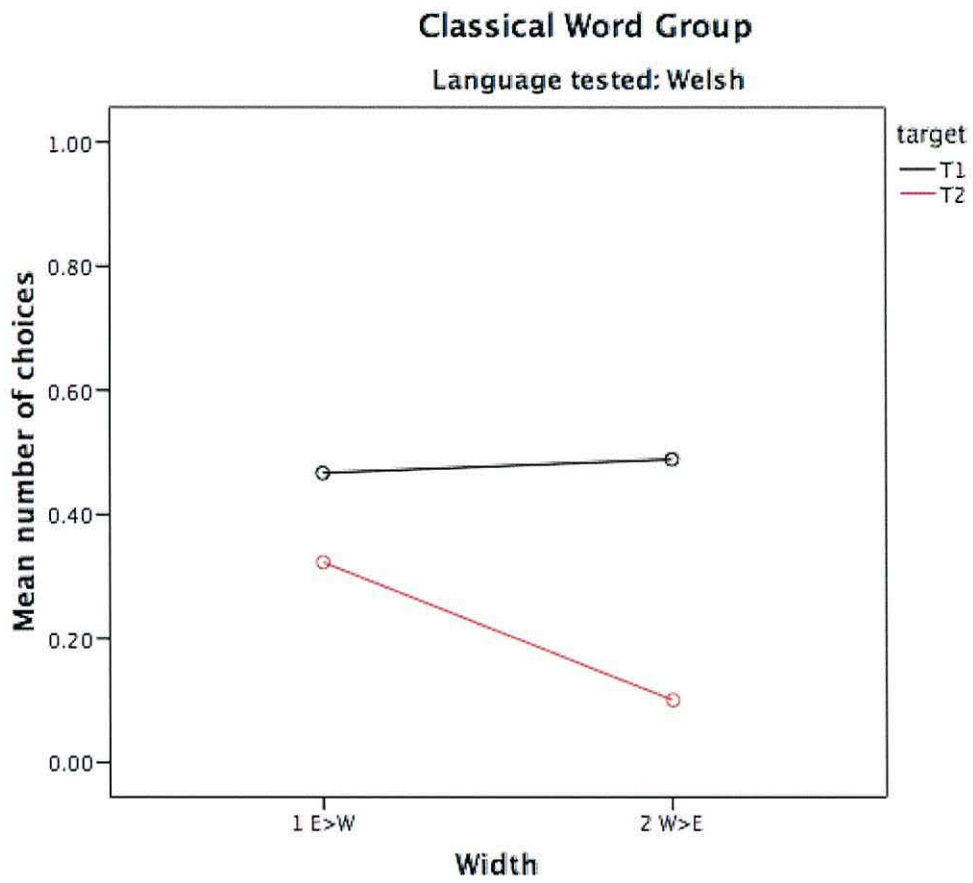


Figure 32: Mean number of target 1 and target 2 selected in the Classical Word Group x Width when tested in Welsh

The children's performance relating to the homonym categories indicated that target choice (selecting the picture which was central to the target group, i.e. item which T1 in categorisation task) was significant, F (1,81) =32.29, p<. 001,with more target 1 selections made than target 2. There were also two way interactions between Width x Extended Choice, F (1,81) =74.66, p<. 001. There were no significant differences between age groups, although significant differences were found in home language groups F (2,81) =3.80, p<. 026. Pairwise comparisons showing that children in the OEH group performed significantly different to

children in both the WEH group *MD* -.183 *p*<. 011, and the OWH group MD -.150 *p*<.037.

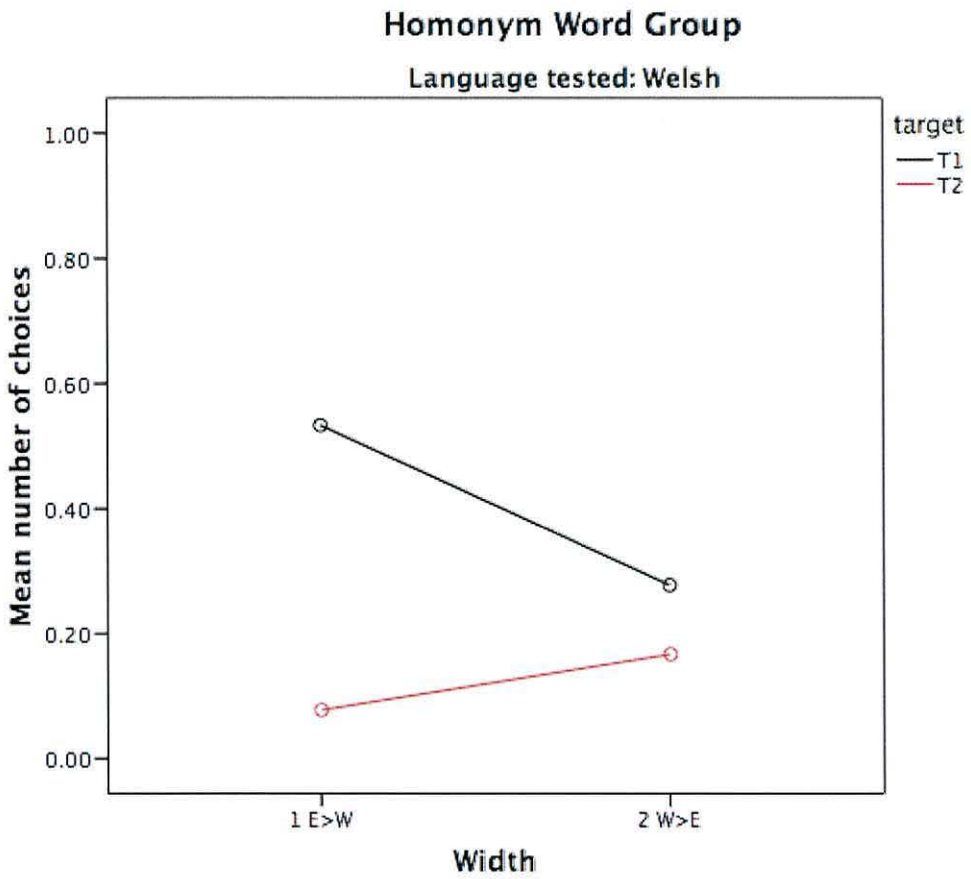


Figure 33: Mean number of target 1 and target 2 choices in the Homonym Word Group x Width when tested in Welsh

For the radial taxonomic words width was significant, *F* (1,81) =6.15, *p*<. 015, as was target choice *F* (4,81) =68.76, *p*<. 014, together with Width x Target Choice *F* (2,81) =4.33, *p*<. 041.. Where English was wider than Welsh children chose target 1 significantly more than target 2 *MD* .456 *p*<.001. On stimuli where Welsh was wider than English there was a

significant difference for target choice F (1,81) =24.92, p<. 001, with children remembering significantly more target 1 words (.322) than target 2 words (.056). .

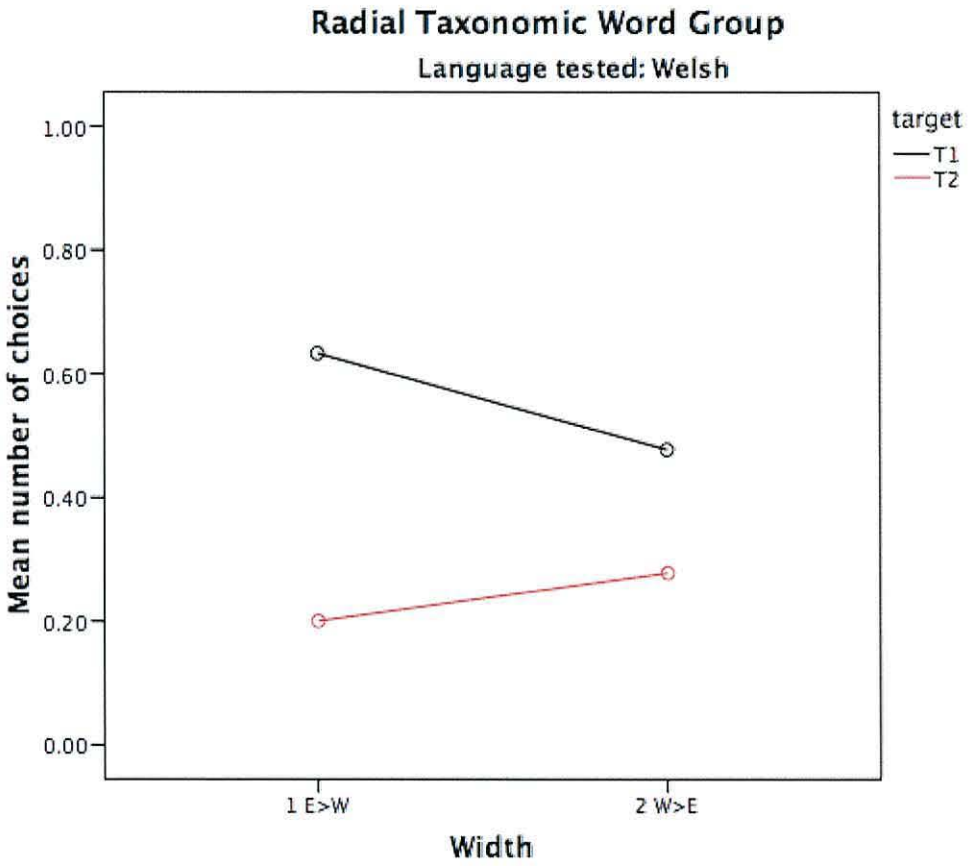


Figure 34: Mean number of target 1 and target 2 choices in the Radial Taxonomic Word Group x Width when tested in Welsh

For the radial thematic words children's target choice was significant, F (1,81) =26.31, p<. 001. Width x Target Choice was also significant F (1,81) =4.91, p<. 030. as was Target x Age Group x Home Language F (4,81) =2.78, p<. 032. There were no significant differences between age group nor between home language group.

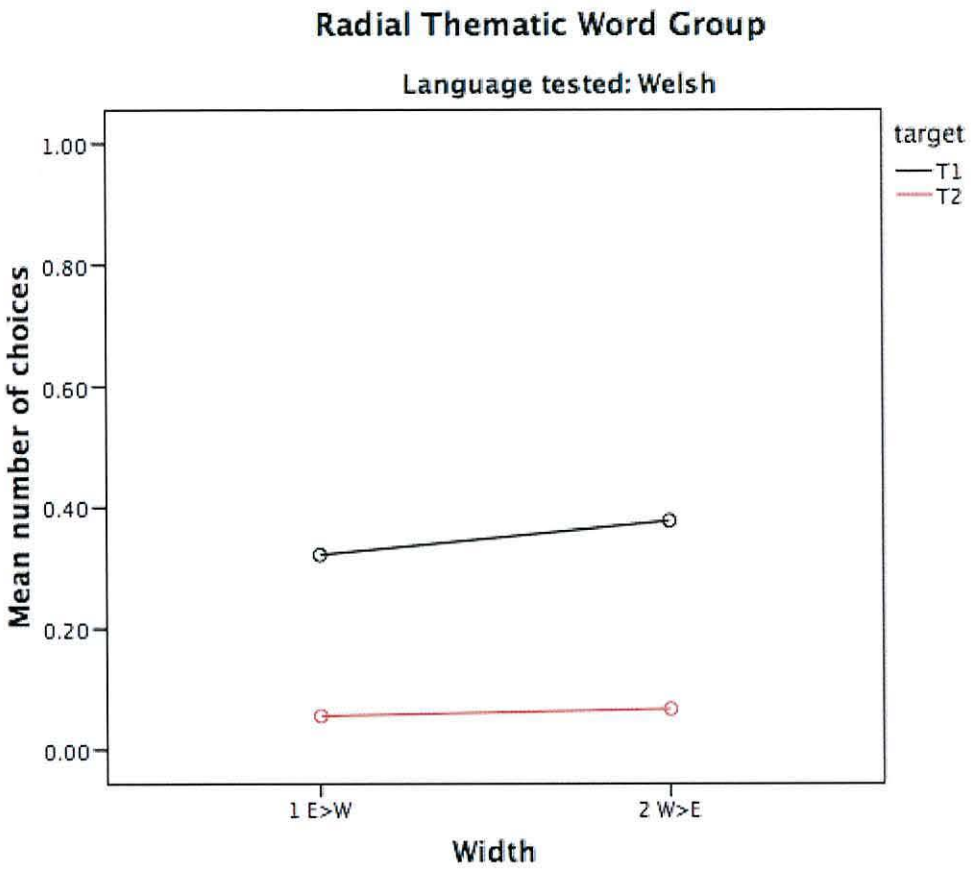


Figure 35: Mean number of target 1 and target 2 choices in the Radial Thematic Group x Width when tested in Welsh

9.3.2 Tested in English

Analyses of variance were performed on the data in which width (Welsh wider than English, and English wider than Welsh), word category group (classical, homonym, radial taxonomic and radial thematic) and memory target (target 1 or target 2) were within subject variables, and age group (children aged 5 to 7, 7 to 9 and 9 to 11) and home language group (Monolingual, OEH, WEH, and OWH) were treated as between group variables.

The results showed significant main effects for word group, F (3,324)= 30.77, p<.001;

and target choice target F (1,108)= 210.09, p<.001, Width was close to being significant, F (1,108) =3.76, p<. 055. Differences between the age groups were also found to be significant F (2,108) p<.001.

Children performed significantly differently on items in the classical category with pairwise comparisons showing differences between their performance with classical word group and the homonym word group *MD*.258 p<.001; the radial taxonomic group *MD* -.165 p<.001, and the radial thematic group *MD* .358 p<.001. There was a significant difference between the children's performance for homonym and radial taxonomic word groups *MD*.-094, p<.001, and the radial thematic word group *MD* .100 p<.001. There was also a significant difference between their performance for radial taxonomic and radial thematic word groups: *MD* .194, p<.001. In addition, there was a significant difference between the number of target 1 and target 2 choices children made *MD* .318, p<.001. Pairwise comparisons of age groups found significant differences between children aged 9 to 11 and those aged 7 to 9. *MD* .80 p<.001, and aged 5 to 7. *MD* .147 p<.001.

The main effects were modified however, with two and three way interactions found to be significant for Word Group x Age Group F (6, 324)= 2.13, p<.049; Target Choice x Age Group x Home Language F (6, 108) = 2.21, p<.008; Word Group x Width x Home Language F (9, 324)= 2.36, p<.013; Word Group x Target Choce x Home Language F (9, 324)= 2.09, p<.030; Word Group x Width x Target Choice F (9, 324)= 5.79, p<.001.

Follow-up analyses examined each word group individually. For the classical category, target choice was significant F (1,108) =42.86, p<. 001, with children remembering significantly more target 1 picture representations than target 2 (*MD*.308), as was width F

(1,108) = 10.86, p<.001. Age group was found to be significant F (1,108) =42.86, p<. 001, with pairwise comparisons showing significant differences between children aged 5 to 7, those aged 7 to 9 and 9 to 11. Where English was wider than Welsh, children chose target 1 significantly more than target 2, *MD* .308 p<.001 but there was no differences in performance between language groups nor in age groups. Where Welsh was wider than English there was again significant difference in target choice F (1,108) =21.87, p<. 001, with children remembering significantly more target 1 words (.617) than target 2 words (.308).

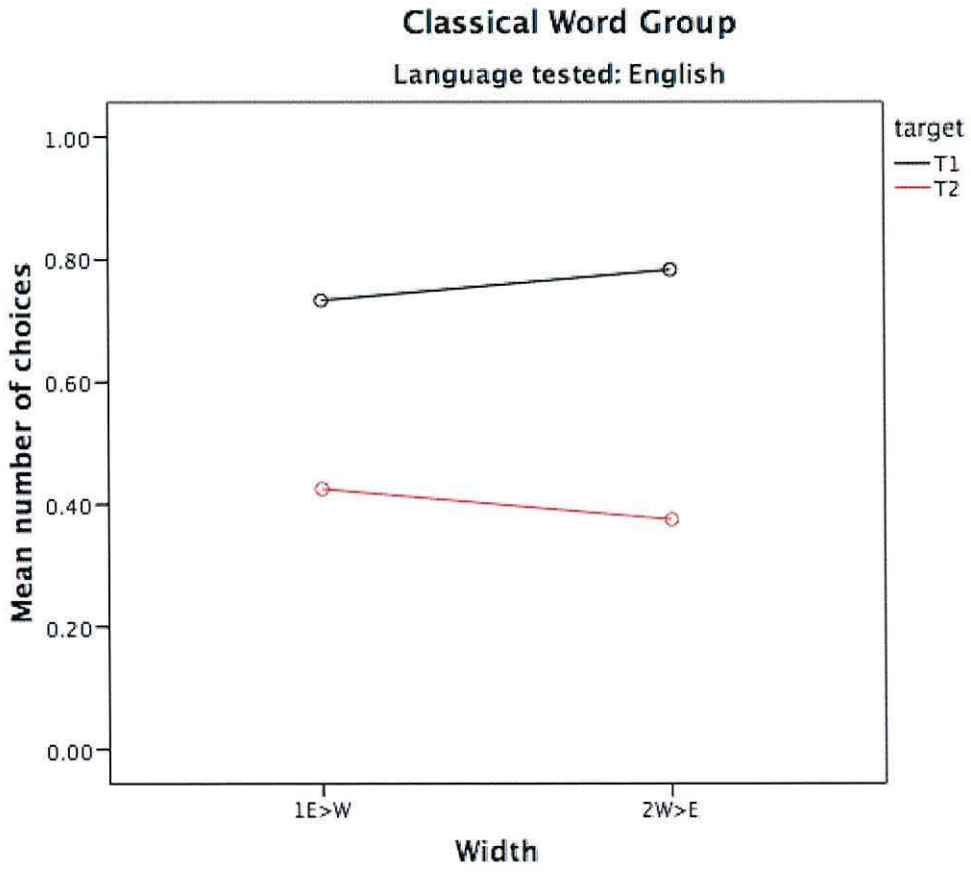


Figure 36: Mean number of target 1 and target 2 choices in the Classical x Width when tested in English

With performance on homonym categories target choice was significant, F (1,108) =38.70, $p<$. 001. with more target 1 selections made than target 2, as was width F (1,108) =11.95, $p<$. 001. There were also two way interactions between Width x Age Group, F (2,108) =3.76, $p<$. 026, Target Choice x Age Group F (2,108) =3.72, $p<$. 028, Target Choice x Home Language group F (2,108) =3.72, $p<$. 011. There were significant differences between age groups F (2,108) =5.09, p<. 008. Pairwise comparisons showing that children aged 5 to 7 performed significantly worse than children aged 9 to 11, *MD* -.219 $p<$. 002. There were no significant differences for language background groups.

Where English was wider than Welsh children chose target 1 significantly more than target 2, *MD* .408 $p<$.001. There was a significant difference in the performance for age groups F(2,108) =6.39. Pairwise comparisons showing that children aged 9 to 11 performed significantly different to both children aged 5 to 7, *MD* .387 $p<$. 002, and those aged 7 to 9, *MD* -219 $p<$. 312. Pairwise comparisons also showed that children in the OWH group performed significantly different to all three of the language background group (Monolingual English *MD*-.267, $p<$.047; OEH, *MD*-.267, $p<$.035; and OWH, *MD*-.267, $p<$.026).

Where Welsh was wider than English there was again significant difference in children's target choice F (1,108) =4.82, $p<$. 030, with children remembering significantly more target 1 words (.425) than target 2 words (.308). Target Choice x Age Group was also significant F (3,108) =6.59, $p<$. 001.

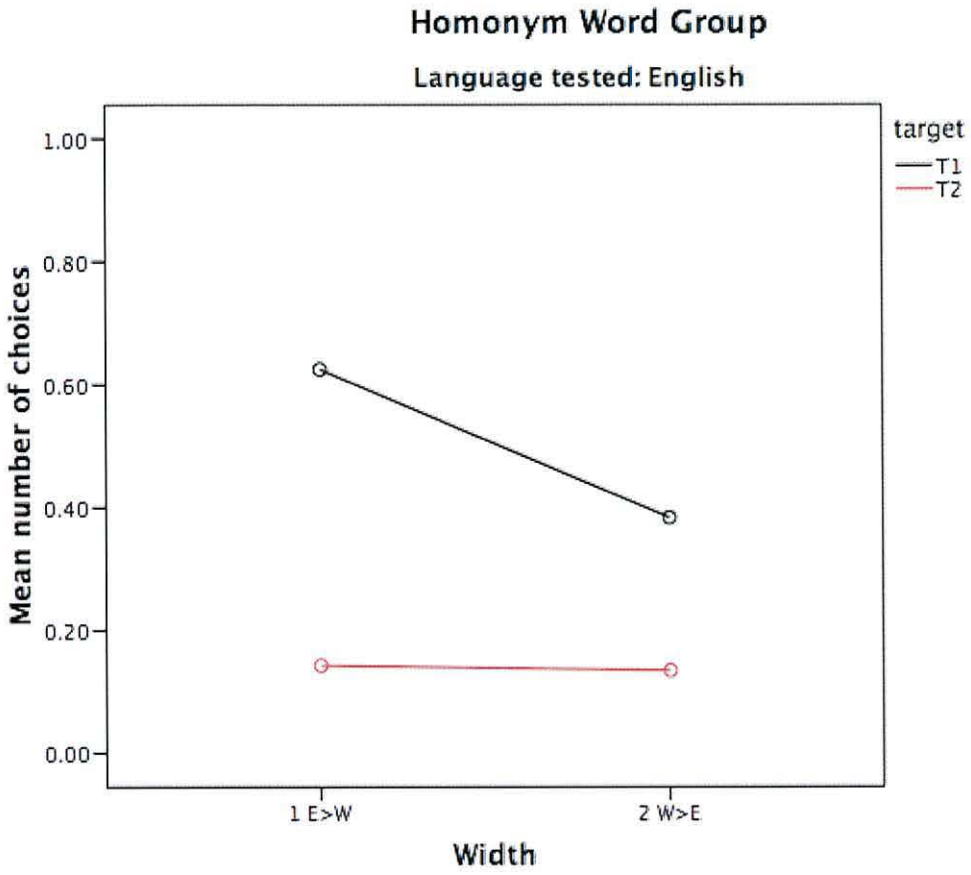


Figure 37: Mean number of target 1 and target 2 choices in the Homonym Group x Width when tested in English

For the radial taxonomic words width was significant, F (1,108) =13.87, p<. 001, as was target choice F (1,108) =69.87, p<. 001, with Width x Target Choice interaction F (2,108) =7.17, p<. 009 also showing significance. Children from different age groups and home language performed similarly with no significant difference found. Where English was wider than Welsh, children chose target 1 significantly more than target 2, *MD* .483 p<.001. There were also differences in performance between language groups that children from the monolingual English group performed significantly different to those in the WEH group

MD.183 $p<.006$. On stimuli where Welsh was wider than English there was a significant difference target choice F (1,108) =18.60, $p<.001$, with children remembering significantly more target 1 words (.333) than target 2 words (.092). Target x Home Language groups was also significant F (3,108) =2.91, $p<.038$.

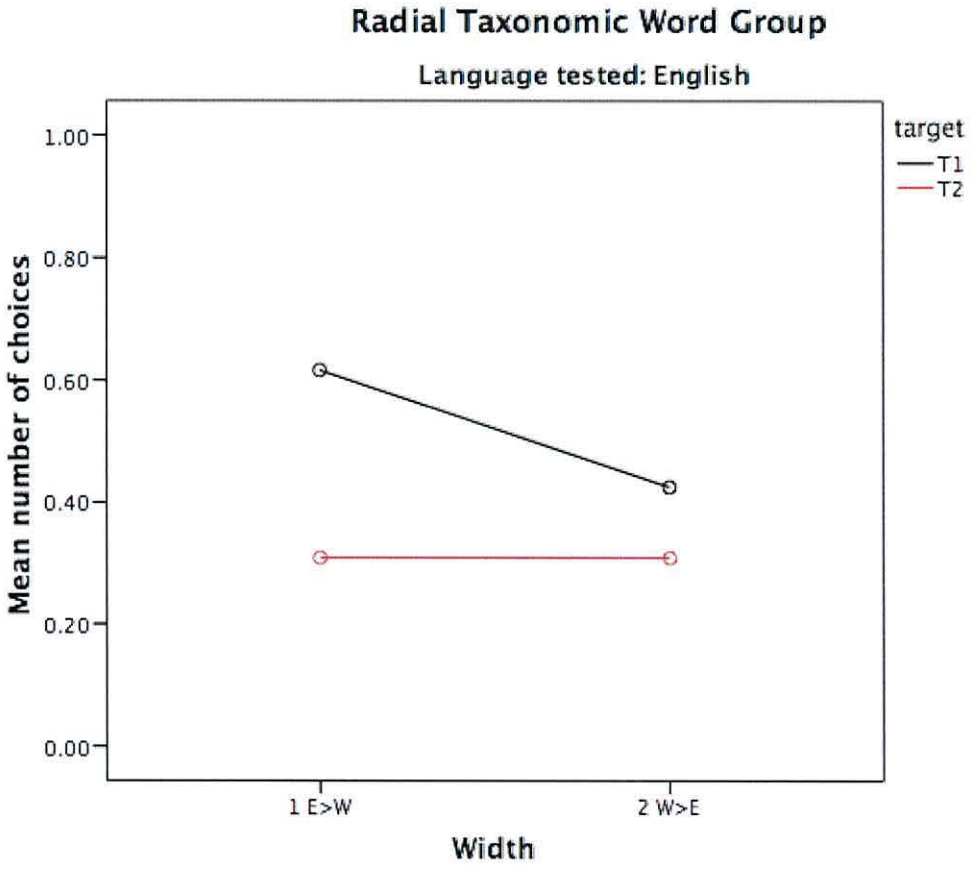


Figure 38: Mean number of target 1 and target 2 choices in the Radial Taxonomic Word Group x Width when tested in English

For the radial thematic words target choice was significant, F (1,108) =92.41, $p<.001$, but width was not significant. Width x Target Choice was significant F (1,81) =4.91, $p<.19$. There were no significant differences between age group or home language group. Children

once again recalled significantly more items target 1 and target 2, *MD*.337. Where English was wider than Welsh, children chose target 1 significantly more than target 2 F (1,108) =22.25, $p<$. 001. There was no difference in performance found between language groups or age groups. On stimuli where Welsh was wider than English, there was a significant difference in target choice F (1,108) =76.75, $p<$. 001, with children remembering significantly more target 1 words than target 2 words, *MD* .425 $p<$001.

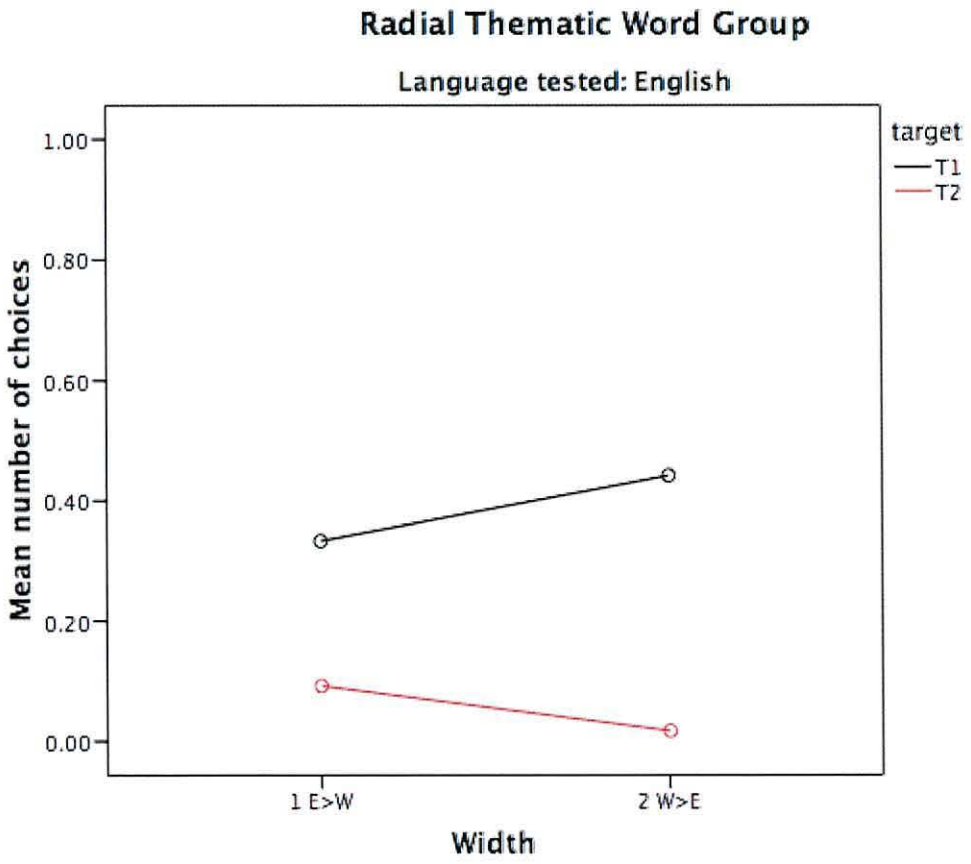


Figure 39: Mean number of target 1 and target 2 choices in the Radial Thematic Group x Width when tested in English

Following up on the significant interactions, simple effects analyses were made into Age and Age x Home Language. For the monolingual English children aged 5 to 7, neither word group nor width was found to be significant. The memory of target (greater recall for target 1) was significant F (1,9) =6.83, $p<$. 028. For children aged 5 to 7 in the OEH group, word group was significant F (3,27) =4.15, $p<$. 015, however no significant effect for width was found. Pairwise comparisons showed significant differences between the items children recalled between radial taxonomic and homonym word group, *MD*.250 $p<$.032, and with the radial thematic group, *MD*.275 $p>$.012. Children also selected target 1 on more occasions than target 2 $p<$. 001. Children aged 5 to 7 in the WEH group performed similarly across word groups although width was found to be significant F (1,9) =6.43, $p<$. 032. Children aged 5 to 7 in both the WEH and OWH groups 1 recalled more target 1 items than target 2 items.

Children in the Monolingual English children group aged 7 to 9 recalled a similar amount of items across words groups and widths but were constant in remembering target 1 more than target 2, F (1,9) =16.99, $p<$. 002. For the performance of children aged 7 to 9 in the OEH group, word group was significant F (3,9) =4.84, $p<$. 029, with pairwise comparisons showing significant differences between the radial categories with children remembering more radial taxonomic items than radial thematic group, *MD*.350 $p>$.010. They were also better at remembering target 1 items <. 001. For children aged 7 to 9 in the WEH group, word group was significant F (3,9) =4.33, $p<$. 013 and width was also found to be significant F (1,9) =6.43, $p<$. 018. The target choice group was significant F (1,9) =12.96, $p<$. 006. pairwise comparisons showing that children remembered more items from the classical group than homonym (*MD*.375) and with radial thematic items, *MD*.425 $p<$.050. This was also the

case for radial taxonomic items compared to homonym and radial thematic items *ps*<.050. For children aged 7 to 9 in the OWH group, word group was significant F (3,9) =4.33, p<. 013, however the target choice group was not significant. Similar to the performance of the WEH group, children recalled more items in the classical group than and homonym word groups, *MD*.325 p<.039, and radial thematic word groups *MD*.475 p<.018.

For the monolingual English children aged 9 to 11, word group was significant F (3,9) =6.79, p<. 028. Pairwise comparisons showing that children recalled significantly more items in the classical word group than the other three word groups p<.050. However there was no significant difference in the amount of target 1 and target 2 choices. For children in the OEH group (p<.038) and WEH (p<.001) only target choice (more target 1 than target 2 recalled) was significant with similar amount of items remembered between word groups and widths. For children aged 9 to 11 in the OWH group, word group was significant F (3,9) =6.00, p<. 001, as was target choice group F (1,9) =14.35, p<. 001. Pairwise comparisons showing that children selected target 1 more than target 2 (*MD*.325), and that children recalled classical word group items than homonym word groups, *MD*.475 p<.012, radial taxonomic items, *MD*.500 p>.034, and radial thematic items, *MD*.625 p>.006.

9.4 Discussion

The memory task was designed to provide additional evidence to the findings of the main categorisation task. The prediction was that children would recall fewer items for those items that were named by a single word than for those that had two labels for each referent. This would hold true across word groups although there was a possibility that there would be

more items remembered in the homonym group, due to there being no conceptual link in place between the referents of the two homonyms used. The memory task was administered directly after the categorisation task therefore it is possible that children would potentially be influenced by their performance in the categorisation test.

Results for children tested in Welsh show that children generally performed in a similar way across word groups, with the most notable result being that all children found it easier to recall target 1 items than target 2. There was no real difference between widths with target 1 being recalled more than target 2 on narrower and wider items. Children in general performed better on classical word category items. However, they did not select target 2 as often on items where English was wider than Welsh in the classical word category. Children performed next best (this was also true in English) on items from the radial taxonomic word group items. However, possibly as a result of width (although this was not true in all age and language background groups), children selected fewer items Welsh was wider than English.

A similar pattern was seen in the children tested in English, with target 1 always being remembered more than target 2. Children remembered generally more items from the classical word group (which was not the same as the results from the categorisation task) and fewest from the radial thematic group. Some differences between groups were observed with older children in age group 3 correctly remembering more items than children in age group 1. It appears that children performed better on English wider than Welsh items (this is fairly difficult to interpret and is not constant in the interaction term and does not appear to be the case for the monolingual English children).

Why do children appear to perform better on target 1 rather than target 2 items

constantly across all word groups? It could be argued that the performance of children on test items in the categorisation task which were narrower could be associated with the higher level of target 1 choices in the memory task. Children tended to perform better in the categorisation task on category items which were narrower in the test language. When an item was narrower the children needed to only select target 1 to be correct on the item and to ensure that they did not select target 2. The position of target 1 as more central to being a better exemplar for a category may have also been influential in children recalling more target 1 than target 2 items. The hypothesis for this task expected children to also remember more target 2 items correctly in the narrower condition compared to items which were part of a broader category. Although this was true in some instances (e.g. items in the classical group tested in English), it did not generally appear to be the case. Children did not generally perform as well in the wider category of the categorisation task, usually underextending the word category membership and selecting target 2. Although children on the whole did select target 1 which may also possibly explain why the performance of target 1 is not generally affected by width.

Chapter 10 General Discussion

10.1 Overview

The study provides a contribution to the available evidence in the little-studied area of word categorisation by bilingual Welsh/ English speaking children. It was designed to test how bilingual Welsh/ English speaking children categorise nouns and action verbs and to assess from the findings how participants develop the knowledge needed for managing both languages. Both English and Welsh differ in the categorisation of word meanings and these differences provide significant challenges to early bilinguals. The analyses used the findings from sets of analyses to compare results with other models and theories examined in early chapters of the thesis. The discussion will initially present the main findings and conclusion of the study before then considering them in greater detail with reference to the goals of the study.

10.2 General Findings

The study achieved its aim of identifying evidence which supported the view that bilingual Welsh/ English children exhibit behaviour which supports the constructivist model. This was defined as the occurrence of two separate sets of categorisation, a semantic system and a conceptual space in which the child develops cognitive abilities to make sense of the world. The lack of taxonomic bias across age groups and category group was a notable finding as this supports the view that bilingual children rely more on cognitive skills to categorise rather than structured linguistic processes. In addition, there was strong evidence that children had greater difficulty with processing items from the classical word group. This

was seen in the poorer performance of children from the OEH group when tested in Welsh, and in the similar performance, when tested in English, of the OWH group across the three age groups (with children aged 5 to 7, 7 to 9 and 9 to 11 not performing as well as the other three home language groups on items from the classical words group).

This study identified unexpected results for differences between word groups with the differing amount of taxonomic and thematic distractors which are worthy of further investigation. This study marked an initial attempt to investigate semantic categorisation by bilingual Welsh/English children. The work has demonstrated that this research area is one which would merit far greater attention from researchers in the future. These conclusions are further discussed in the next sections.

10.3 Categorisation processing

The basic premise of the constructivist model is that a dual approach of semantic organisation and cognitive processing is required for bilingual children to be able to categorise concepts and apply the correct label to novel words. Linguistic organisation by itself is insufficient to enable the child to navigate through the complexity of actions required to make sense of his/ her world. The wide variety of potential difficulties from polysemy to a possible lack of co-equivalence between the two languages requires a dynamic and cognitive approach (Pavlenko, 2009). The findings from the study focused on what lessons can the experience of Welsh/ English bilinguals can tell us about the distinction between semantic organisation and the formation of conceptual ideas. The study, in addition, considered the nature of different word categories and how the bilingual child develops two sets of categories

simultaneously using both his/ her cognitive abilities and semantic systems. Many of the study's findings are in line with predictions made by the constructivist model which provide explanations on how both languages are represented in bilinguals, and how they manage complex and differing amount of input when participating in categorisation activities.

The study results showed that children almost always performed better on homonym and radial taxonomic categories when Welsh was wider than English. This result was true for all participants from all three age groups. This result was in line with the study prediction made previously that children would have the greatest difficulty on the classical items and would probably do worse on radial thematic than radial taxonomic items. This complies with the constructivist model in that the membership for categories in the classical group are less transparent and would require greater language input to be categorised correctly (Gathercole, 2007). This was demonstrated by the results where children in the older age group performed better on items where Welsh was wider than English when compared to the other two age groups.

It was predicted that children in the older group would perform better as a result of greater input. This was supported by the fact that older children from the OWH group performed best on items in the classical word group compared to the other two home language groups. This indicates that they were more confident in understanding category membership for a less clear category group. There was improvement as well in the performance of children in the OEH group and those in the WEH group in English wider than Welsh items for children aged 5 to 7 compared to children aged 7 to 9 and those aged 9 to 11. This further supports the view that there is a tendency for children not to specifically

understand the word category at a younger age, rather than because they are overextending the word category to include the wider choice as well.

Performance with the homonym words was expected to be better as there was no conceptual link between the two words representing the category, only a homophonic relationship. This would be especially true when the item was narrower in the test language as there should be less interference from the dominant language. As a consequence therefore, there should be no reason to expect a carry over, nor a need to use the structure of the dominant language to guide performance from one language to the other. As predicted, where Welsh was wider than English, the OEH group (least like the OWH group) performed worst on the classical categories. A possible explanation for the performance of children from the OEH group is that they were underextending the wider category as a result of the influence of the narrower category in the narrower language. This being particularly the case for items in the classical group, although it is important to note that performance was generally low on classical groups. The probable explanation for the performance of children from the bilingual groups is that they were underextending the wider category as a result of the influence of the narrower category in the narrower language. Again as predicted, where Welsh was wider than English, the OEH group (least like the OWH group) performed worst on the classical categories.

All children who took part in the categorisation test performed better on items which were narrower (when English was wider than Welsh) and their performance on items when Welsh was wider tended to be poor for all groups. The result may indicate that children did better in identifying a central item (target 1) as being part of a group than an item which was

more peripheral (target 2). This finding appears to support the family resemblance and prototype view that children find it easier to identify the best exemplar compared to one which is not as central to the word category meaning (Lakoff, 1987; Markman, 2003). This finding appears to show that children were conservative in their choices. The possibility of children performing better on narrower items as a result of their adopting a more conservative approach to the task must also be considered.

Some studies have indicated that if transfer is to take place, it is more likely to occur in bilinguals who receive a balanced equal input from both languages (Gathercole, 2007). When tested in Welsh, children in the WEH group tended to perform in a similar way across all word groups on items which were wider, whilst the other two background groups did not perform as well on homonym and radial taxonomic groups (this was an opposite finding to the performance on items where English is wider). The lack of difference in performance in WEH between different word category groups is a finding which should be investigated further. The relatively poor performance on homonym groups, especially in the OWH group, was interesting as it did not meet the expectations of the early study predictions. It would be useful to carry out an error analysis on all items to obtain a clearer view as to why participants were incorrect to ascertain if the poor performance of participants from different home language groups was due to similar reasons.

The study has provided an interesting starting point to speculate how differences in the children's performance on tasks can be used to explain the manner in which bilingual children develop and use different categories depending on the type of category which is most appropriate for the individual language; and to what degree, if any, does the child's dominant

language influence the categorisation of word labels in the second language. There is a growing interest in better understanding how bilingual children manage to develop two sets of categorisation within their conceptual space. The debate between a shared space or two separate areas continues although the findings of this study support the view that the bilingual child needs to consider the requirements of each language in categorising and labelling new concepts. There is some evidence to suggest a shared storage for the conceptual representations of the bilingual's two vocabularies with links between concepts and lexical names in the two languages (Dong, Gui & McWhinney, 2005). The difficulties of mapping conceptual meaning across languages as different as Welsh is from English should not be underestimated. It has been suggested that bilinguals need to re-structure boundaries rather than map concepts across to their dominant language (Pavlenko, 2009).

10.4 Extended choice

It was noted in earlier chapters that Vygotsky (1962) proposed that children have a preference for thematic categorisation whereas older children and adults have a preference for taxonomic categorisation. This approach was challenged by Markman (1989) and others who suggested that children may prefer to associate new words with objects that are taxonomically related rather than thematically related. The constraints on children in acquiring meanings for new labels include Markman's (1991) theory of mutual exclusivity which highlighted the paradox that although it would be natural for children to create new categories for every novel word, i.e. the categories would be mutually exclusive, this is likely to quickly become unmanageable.

Using this approach, according to Markman, children assess how categories are organised by being able to link novel terms with collective labels, for example, a *tree* is part of the collective label *forest*. This association maintains the simplicity of the relationship associated with the idea of mutual exclusivity, with novel terms being linked to another label but not being defined by the label i.e. the *tree* is not the same label as *forest*. There is no consensus amongst researchers whether bilinguals are less likely to use mutual exclusivity on the basis that they are already familiar with two terms for an object. Davidson and Tell (2005) found that both monolingual and bilinguals retained mutual exclusivity for the naming of whole objects. Healey and Scarabella (2008) found no differences between the performance of monolinguals and bilinguals on a picture naming task although bilinguals performed differently when receiving socio-cognitive cues.

The findings of the extended choices task for this study did not follow the prediction and support Markman's (1991) theory of taxonomic bias. When children were tested in Welsh on Welsh wider items, children from all language background groups selected generally more thematic distractors than taxonomic in all word groups. Although the study did not make any specific predictions regarding differences between word groups, children selected significantly fewer taxonomic and thematic distractors in the radial thematic category in comparisons to the other word groups.

In terms of which category was most affected, children from all age groups selected the most taxonomic distractors in the classical word category. Children aged between 5 and 7 years across all language background groups, had a preference for selecting thematic distractors in the homonym and radial taxonomic word categories, with the number of

selections being higher than the number of thematic choices in the classical and radial thematic word groups; and between the number of taxonomic distractor selections in the homonym and radial taxonomic groups.

There was a significant difference in the number of thematic choices made by children from all age groups compared to the number of taxonomic choices they made in the radial taxonomic word category. Although there was no main effect of width, children did perform in the same manner on word groups across both widths.

When English was wider than Welsh there was a tendency for children not to choose thematic distractors in both radial groups. Children tended to make more taxonomic extended choices than thematic on radial categories. In the homonym group children tended to pick fewer taxonomic distractors compared to other groups and there was generally a large difference across all groups in the choice of more thematic than taxonomic in this word group. The number of taxonomic distractors made by children when English was wider than Welsh tended to be similar across all word groups in all participants (the only exception being fewer choices of items in the homonym group).

Children tested in English on stimuli where Welsh was wider than English consistently selected more thematic choices than taxonomic items in all words groups with this effect being achieved across all age groups and language background groups. This was similar to the performance of children who were tested in Welsh and was again not in line with the study's prediction that more taxonomic distractors would be selected if children extended their choice beyond the targets.

For the radial thematic group, children across all age groups tended to select a high

number of taxonomic distractor choices. This was also the case for the number of selected thematic distractor choices, which was markedly less than in the radial thematic group compared to the other word groups. When test items were wider in English than Welsh, children tended to select significantly more thematic distractors selected compared to taxonomic distractors in the homonym word group. This was the case for children across all language background groups. Children consistently made the most taxonomic choices from the classical word group with a similar number of taxonomic distractors and thematic distractors chosen.

In the absence of a taxonomic bias, other explanations are therefore needed to address this finding. An alternative approach is available with the concept of individual cognitive abilities which emphasises the role of the environment where language is used instead of the development of a theoretical hierarchy (Croft & Cruse, 2004). As thematic choices relate to labels which probably are associated with a practical world view from the child's perspective, these choices would be attractive. In some ways this is nearer to linguistic relativity where the world view shapes utterance. It is also closer to Fillmore's (1976) Frames Theory which proposed that a frame could represent a concept which requires a knowledge of a larger concept 'domain'. This finding also call to mind Rosch's (1975) work on prototype theory with the idea of categorisation of novel words using a 'best fit' approach .

The relationship between categories in radial categories revolves around related meanings of words form categories which bear family resemblances as 'meaning chains' (Lakoff,1987). The objects in the meaning chain have a shared attribute which link them. A category can then be structured radially through a number of sub-categories which are

extended in an ordered way. For example the representation *window* can have three meanings: 'an opening in the wall', 'a frame fitting into the wall' and 'the glass filling the frame fitting into the wall' with the three representations linked through the chain of meanings (Gathercole, 2007).

10.5 Recall and memory

The organisation of conceptual space to manage cognitive activity in the categorisation process for two languages is the subject of much debate as has been outlined in earlier chapters. (Kroll & Stewart, 1994). The interference of one language on the other and the method of accessing existing knowledge are two factors which require explanation.

The memory task was designed to provide additional evidence to the findings of the main categorisation task. The prediction was that children would recall fewer items which were wider in the test language compared to items which were narrower because there were two labels for each referent. This would hold true across word groups although there is a possibility that there would be more items remembered in the homonym group. This would be due to no conceptual link in place between the homonym category with only the same label used (i.e. homonyms name different categories). The memory task was administered directly after the categorisation task therefore it is possible that children would potentially be influenced by their performance in the categorisation test.

Results for children tested in Welsh show that children generally performed in a similar way across word groups, with the most notable result being that all children found it easier to recall target 1 items than target 2. There was no real difference between widths with

target 1 being recalled more than target 2 on narrower and wider items. Children in general performed better on classical word category items. However, they did not remember target 2 as often on items where English was wider than Welsh in the classical word category. Children performed next best (this was also true in English) on items from the radial taxonomic word group items. However, possibly as a result of width (although this was not true in all age and home language groups), children remembered fewer items when Welsh was wider than English.

A similar pattern was seen in the children tested in English, with target 1 always being selected more than target 2. Children remembered generally more items form the classical word group with children performing worst in the recall of items in the radial thematic group. Some differences between groups were observed with older children in age group 3 correctly remembering more items than children in age group 1. If there was any effect of width in children tested in English it appears that children performed better on English wider than Welsh items (this is fairly difficult to interpret and is not constant in the interaction term and does not appear to be the case for the monolingual English children).

Why do children appear to perform better on target 1 than on target 2 items constantly across all word groups? It could be argued that the performance of children on test items in the categorisation task which were narrower could be associated with the higher level of target 1 choices in the memory task (Paivio, & Desrochers, 1980). Children tended to perform better in the categorisation task on category items which were narrower in the test language. When an item was narrower the children needed to only select target 1 to be correct on the item and to ensure that they did not select target 2. The hypothesis for this task expected

children to also remember more target 2 items correctly in the narrower condition compared to items which were part of a broader category. Although this was true in some instances (e.g. items in the classical group tested in English), it did not generally appear to be the case. Children did not generally perform as well in the wider category of the categorisation task, usually underextending the word category membership and selecting target 2. Although children on the whole did select target 1 which may also possibly explain why the performance of target 1 is not generally affected by width.

10.6 Vocabulary tests and home language questionnaires

The performance of children in both the BPVS and the Prawf Geirfa, when considering raw scores, improved with age (this would be an expected outcome due to an increase in cognitive ability and an increased level of input for both languages). However, once scores had been standardised by age, performance between age groups was similar with the OWH group performing better than the other two bilingual groups across all age groups. This indicates the importance of input on vocabulary, with children from the OWH receiving much greater exposure to Welsh. There was no major difference between the bilingual groups in the BPVS. However, this seems to be at a cost to the vocabulary ability in English as was seen in the results of the bilingual groups' scores compared with monolingual English participants' scores. This could be due to a reduced input as a result of the need for dual language input, with a large difference between the scores of the monolingual group in age groups 2 and 3 as compared with the scores of the three bilingual groups in the same age groups. It is important to note that the scores of monolingual children were well above

average and the scores of the bilingual groups being a little below average. In line with findings from previous studies, both BPVS and Prawf Geirfa Cymraeg tests did not correlate with the performance in the main categorisation task (with the scores collapsed for word group) and did not provide significant results (there was only one correlation as described in Chapter 6).

The analysis of the evidence from the completed language background questionnaire was an opportunity to identify whether parental background had any impact on the bilingual child's ability to manage categorisation tasks in two languages. Little comparable evidence has been collected, however an examination of parental educational attainment and the performance of individual children did not suggest that this element was an influential factor. These results agreed with the findings of Gathercole (ed.,2007) that parental educational or occupational attainment levels were not major influences on the performance of children on the categorisation task. This finding was consistent across home language background groups. An interesting contrast was found between the educational attainment levels of both mothers and fathers between monolingual households and Welsh only households where educational attainment levels for advanced qualifications in the former were twice the level in the latter.

10.7 Limitations and further study

The study findings and conclusions provided valuable insights into the semantic categorisation by Welsh/English bilingual children as both a developmental and a categorisation process. However this research area is at an early stage and further work would provide opportunities to take advantage of advances in global research studies in this field.

The semantic categorisation of word categories has moved from an historical view which emphasised the classical theoretical construct to a more complex cognitive approach which relies on the child making comparisons and choice decisions in both classifying and interpreting novel concepts. Boundaries for these decisions do not appear to be absolute however and models such as Gentner & Markman's (1997) structure mapping engine provide a basis for investigating further the relationship between semantic systems and cognitive organisation.

The interpretation of the study findings highlighted considerations for further work. The nature of the categorisation task may be such that children had an advantage on test items when the test was narrower, as they would only have to select one target to be correct for that item. Although, as previously mentioned, if a child selected both items, he/she would receive a score of zero for that item. The inclusion of a full error analysis on the responses made by children would be a useful addition in further research.

For the memory task, there may have been an issue with the design, with children not providing the target responses which the study investigated. It was important that no orthographic labels were included in the memory task PowerPoint display, but for further research this issue should be addressed.

The use of reaction times for the study could be utilised in further research to investigate the differences between participants groups. The use of reaction times can show any differences in the time responses even if the selections made were the same. For example, children who were tested in their dominant language may have found it easier to correctly categorise objects, and will therefore have responded more quickly in selecting the correct

response. On items which were wider in the test language, there may also have been a difference in the times between selecting both targets. This would be useful supplementary data for considering the effect on categorisation behaviour between groups.

An additional investigatory technique which would be beneficial for use in further research would be the adoption of eye tracking techniques. It was observed during the testing procedure that children would sometimes fixate on certain pictures with their fingers placed above the picture, but would not ultimately select the picture. The use of eye tracking would allow the investigator to gain insight into how children process the items in the task, even if they do not ultimately select certain pictures. For example, whether they fixate on thematic or taxonomic distractors, or when an item is narrower, whether they also deliberate over selecting target 2.

Further research into the specific circumstances of the acquisition of Welsh and English by children will hopefully benefit from the experience of this initial study.

References

Aitchison, J, (2003), *Words in the mind: An introduction to the mental lexicon.* 3rd edition. London: Wiley Blackwell, 2003.

Altarriba, J., & Mathis, K. M. (1997). Conceptual and lexical development in second language acquisition. *Journal of Memory and Language,* **36**, 550-568.

Ameel, E., Malt, B., Storm, G. & Sloman, S. (2005). How bilinguals solve the naming problem. *Journal of Memory and Language*, **52**, 209-329.

Andersen, E.S. (1975). Cups and glasses: Learning that boundaries are vague. *Child Language* **2**, 79-103.

Andrews, D. R. (1994). The Russian color categories *sinij* and *goluboj*: An experimental analysis of their interpretation in the standard and emigre languages. *Journal of Slavic Linguistics*, **2**, 9-28.

Barsalou, L. W. (1989). Intraconcept similarity and its implications for interconcept similarity. In S. Vosniadou & A. Ortony *Similarity and analogical reasoning*. New York: Cambridge University Press, 76- 121.

Barsalou, L.W. (1996). Empirical evidence for perceptual simulation in conceptual processing. *International School on Downward Processing in the Perception Representation Mechanisms*. International School of Biocybernetics, Isle of Ischia, Italy.

Bates, E. & McWhinney, B. (1982). Functionalist approaches to grammar. In E. Wanner & L.Gleitman (eds.). *Language acquisition: The state of the art.* Cambridge: Cambridge University Press, 173-218.

Blaye, A. & Bonthoux, F. (2001). Thematic and taxonomic relations in preschoolers: The development of flexibility in categorization choices. *British Journal of Developmental Psychology*, **19**, 395-412.

Bloomfield, L. (1935). *An introduction to the study of language*. New York: Henry Holt.

Bloomquist, J. (2007). Developmental trends in semantic acquisition: Evidence from overextensions in child language. *First Language*, **27**(4), 407-420.

Bosch, L. & Galles, S.N. (2001). Evidence of early language discrimination abilities in infants from bilingual environments, *Infancy*, **2**(1), 29-49.

Bowerman, M. & Choi, S (2003). Semantic factors in the acquisition of rules for word use and sentence. D. Genter (ed.) *Language in mind: Advances in the study of language and thought.* Cambridge, MA.: MIT Press, 387-422.

Bowerman, M. & Choi, S. (2001). Shaping meanings for language: Universal and language-specific in the acquisition of spatial semantic categories. In M.Bowerman & S.C. Levinson (eds.), *Language acquisition and conceptual development.* Cambridge: Cambridge University Press, 475-511.

Bowerman, M. & Pederson, E. (1992). *Cross linguistic perspectives on topological spatial relationships.* San Francisco, CA: American Anthropological Annual Meeting.

Campbell, M. (2007). Word learning in preschoolers: Are bilingual 3-year-olds less guided by mutual exclusivity than their monolingual counterparts? Edinburgh: University of Edinburgh.

Carroll, J. B. (ed.) (1956). *Language, thought and reality: selected writings of Benjamin L. Worf.* Cambridge, Ma.: MIT Press.

Casenhiser, D. & Goldberg, A. E. (2005). Fast mapping between a phrasal form and meaning. *Developmental Science,* **8** (6), 500-5008.

Choi, S. (2000). Caregiver input in English and Korean: use of nouns and verbs in book-reading and toy-play contexts. *Journal of Child Language*, **27**, 69-96.

Choi, S. (2006). Influence of language specific input on spatial cognition: Categories of containment. *First Language*, **26**, 207-232.

Choi, S. (2009). Typological differences in syntactic expressions of Path and Causation. In V.C.M. Gathercole (ed.) *Routes to language: Studies in honor of Melissa Bowerman.* Mahwah, NJ: Lawrence Erlbaum Associates, 169-194.

Clark, E. V(1971). On the acquisition of the meaning of 'Before' and After'. *Journal of Verbal Learning and Verbal Behavior*, **10** (3), 266-275.

Clark, E. V. (1979). What's in a word? On the child's acquisition of semantics in his first language in V. Lee (ed.) *The language reader*. Milton Keynes, Open University Press, 186-234.

Clark, E.V. (1983). Convention and contrast in acquiring the lexicon. In Th. B. Siler & W. Wannenmacher (eds.) *Concept development and the development of word meaning.* New York: Springerverlag, 67-89.

Clark, E.V. (1987). The principle of contrast: A constraint on language acquisition. In B. MacWhinney (ed.) *Mechanisms of language acquisition.* Hillsdale. NJ: Lawrence Erlbaum Associates, 1-33.

Clark, E.V. (1993). *The lexicon in acquisition.* Cambridge: Cambridge University Press.

Clark, E.V. (1997). Conceptual perspective and lexical choice in acquisition. *Cognition*, **64**(1),

1-37.

Clark, E. V. (2004). How language acquisition builds on cognitive development. *Trends in Cognitive Science,* **8**(10), 472-478.

Clark, E. V. (2006). Color reference, and expertise in language acquisition. *Journal of Experimental Child Psychology*, **94**(4), 339-343.

Clark, E. V. (2007). Conventionality and contrast in language and language acquisition. *New Directions for Child and Adolescent Development,* **115**, 12-23.

Clark, E.V. & Sengul, C.J. (1978). Strategies in the acquisition of deixis. *Journal of Child Language,* **5**, 457-475.

Clark, E.V. & Svaib, T. (1997). Speaker perspective and reference in young children. *First Language*, **17** (49), 57-74.

Clark, E. V. & Wong, A. (2002). Pragmatic directions about language use: Offers of words and relations. *Language in Society*, **31**, 181-212.

Colunga, E. & Smith, L.B. (2005). From the Lexicon to Expectations about Kinds: A role for associative learning. *Psychological Review,* **112**(2), 347-382.

Collins, A. M. & Loftus, E. F. (1975). A Spreading Activation Theory of Semantic Processing. *Psychological Review*, **82**, 407-428.

Collins, A. M. & Quillian, M. R. (1969). Retrieval Time from Semantic Memory. *Journal of Verbal Learning and Verbal Behavior*, **8**, 240-247.

Cook, V., (2002). Background to the L2 user. In V. Cook (ed.), *Portraits of the L2 user.* Clevedon, UK: Multilingual Matters, 1-28.

Costa, A. (2005). Lexical access in bilingual production. In J.F. Kroll & A.M.B, De Groot

(eds.) *Handbook of Bilingualism: Psycholinguistic approaches*. Oxford: Oxford University Press, 308-325..

Costa, A., Alario,F.X., Caramazza, A. (2005). On the categorical nature of the semantic interference effect in the picture-word interference paradigm. *Psychonomic Bulleting Review*, **12**(1), 125-131.

Costa, A. & Caramazza, A. (1999). Is lexical selection in bilingual speech production language specific? Further evidence from Spanish-English and English-Spanish bilinguals. *Bilingualism: Language and Cognition,* **2**(3), 231-244.

Costa, A. & Caramazza, A. (2002). The production of noun phrases in English and Spanish: implications for the scope of phonological encoding in speech production. *Journey of Memory and Language*, **46**(1), 178-198.

Costa, A., Miozzo, M. & Caramazza,A. (1999). Lexical selection in bilinguals : Do words in a bilingual's lexicón compete for selection? *Journal of Memory and Language,* **41**, 365-397.

Costa, A. & Santestban,M (2004). Lexical access in bilingual speech production: Evidence from highly proficient bilinguals and L2 learners. *Journey of Memory and Language*, **50**(4), 491-511.

Coupland, N. et al (2006). Imagining Wales and the Welsh Language. *Journal of Language and Social Psychology*, **25**(4), 351-376.

Crystal, D. (1997). *The Cambridge Encyclopedia of Language.* Cambridge: Cambridge University Press.

Dahan, D., Magnuson, J.S. & Tanenhaus, M.K. (2001). Time course of frequency effects in

spoken-word recognition: Evidence from eye movements. *Cognitive Psychology*, **42**, 317-367.

Dahan, D. & Tanenhaus, M.K. (2005). Looking at the rope when looking for the snake: Conceptually mediated eye movements during spoken-word recognition. *Psychonomic Bulletin & Review*, **12(3)**, 453-459.

Davidson, D. & Tell, D. (2005). Monolingual and bilingual children's use of mutual exclusivity in the naming of whole objects. *Journal of Experimental Child Psychology*, **92** (1), 25-45.

De Groot, A. M. B. (1993). Word type effect in bilingual processing tasks: Support for a mixed representational system. In R. Schreuder & B. Weltans (eds.) *The bilingual lexicon*. Amsterdam. Philadelphia: John Benjamins, 27-51.

De Groot, A. M .B. (1992). Bilingual lexical representation: A closer look at conceptual representations. *Orthography, Phonology, Morphology, and Meaning*, **94**(3), 389-412.

De Groot, A. M. B., & Nass, G. L. J. (1991). Lexical representation of cognates and noncognates in compound bilinguals. *Journal of Memory & Language*, **30**, 90-123.

Deuchar, M. (1992). Can government and binding theory account for language acquisition? In C. Vide (ed.) *Lenguajes naturales y lenguajes formales VIII*. Barcelona: Universitat de Barcelona, 273-279.

Deuchar, M & Quay, S. (2000). *Bilingual Acquisition: Theoretical Implications of a Case Study*. Oxford: Oxford University Press.

Dijkstra, T., & Van Hell, J. G. (2003). Testing the Language Mode hypothesis using trilinguals. *International Journal of Bilingual Education and Bilingualism*, **6**, 2-16.

Dong,Y., Giu, S. & McWhinney, B. (1995). Shared and separate meanings in the bilingual mental lexicon. *Bilingualism:Language and Cognition*, **8**(3), 221-238.

Dunn, L., Whetton, C. & Burley, J. (1997). *The British Picture Vocabulary Scale*. 2nd Ed. London: NFER-Nelson.

Ellis. N. C., O'Dochartaigh, C., Hicks, W., Morgan, M. & Laporte, N. (2001). Cronfa Electroneg o Gymraeg (CEG): A one million word lexical database and frequency count for Welsh. URL: http://bangor.ac.uk/ar/cb/ceg.php.en accessed on 10.09.2009.

Elman, J. L. (2004), An alternative view of the mental lexicon. *Trends in Cognitive Sciences*, **8**(7), 301-306.

Evans, V. & Green, M. (2006). *Cognitive linguistics*. Edinburgh: Edinburgh University Press.

Fillmore, C.J. (1976). Frame semantics and the nature of language. *Annals of the New York Academy of Sciences: Conference on the Origin and Development of Language and Speech,* **280**, 20-32.

Fillmore, C.J. & Atkins, B.T.S. (2000). Delivering polysemy: The case of 'crawl'. In Y. Ravin & C. Leacock (eds.) *Polysemy - theoretical and computational approaches*. Oxford: Oxford Linguistics, 91-110.

Fitzpatrick, T. (2010). Review of doctoral research in second language acquisition in Wales (2003-2008), *Language Teaching*, **43** (4), 496-521.

Gathercole, V. C. M. (2002). Monolingual and bilingual acquisition: Learning different treatments of that-trace phenomena in English and Spanish. In K. D. Oller & R. Eilers (eds.), *Language and literacy in bilingual children*. Clevendon, UK: Multilingual Matters, 220-254.

Gathercole, V.C.M. (ed.) (2007). *Language transmission in bilingual families in Wales.* Cardiff: Welsh Language Board.

Gathercole, V. C. M. (2007). Miami and North Wales, so far and yet so near: A constructivist account of morpho-syntactic development in bilingual children. *International Journal of Bilingual Education and Bilingualism*, **10**, 224-247.

Gathercole, V.C.M. (ed.) (2009). *Routes to language: Studies in honour of Melissa Bowerman.* New York: Psychology Press.

Gathercole, V.C.M. & Moawd, R. (2010). Semantic interaction in early and late bilinguals: all words are not created equally. *Bilingualism: Language and Cognition,* 13(4), 385-408.

Gathercole, V.C.M. & Thomas, E.M. (2005). Minority language survival: input factor influencing the acquisition of Welsh. In J. Cohen, K.T. McAlister, K. Rolstad, and J. MacSwan (eds.), *Proceedings of the 4th International Symposium on Bilingualism.* Somerville ,MA: Cascadilla Press, 852-874.

Gathercole, V.C.M. & Thomas, E.M. (2009). Bilingual first-language development: Dominant language takeover, threatened minority language take-up. *Bilingualism: Language and Cognition,* **12**(2), 213-237.

Gathercole, V.C.M. et al (2006). *Language transmission in bilingual families in Wales.* Cardiff: Welsh Language Board.

Gentner, D. (1982). Why nouns are learned before verbs: linguistic relativity versus natural partitioning. In S.A.Kuczaj (ed.), *Language development.* Hillsdale,N.J.:Lawrence Erlbaum, 301-334.

Gentner, D. (2005). The development of relational category knowledge. In L.Gershskoff-

Stowe & D.H. Rakiston (eds.) *Building object categories in developmental time.* Hillsdale, NJ.: Lawrence Erlbaum Associates,245-275.

Gentner, D. & Boroditsky, L. (2009), Early acquisition of nouns and verbs: Evidence from Navajo. In In V. C. M. Gathercole (ed.) *Routes to Language: Studies in honor of Melissa Bowerman*, New York: Taylor & Francis, 5-36.

Gentner, D. & Boroditsky, L. (2001). Indviduation, relativity and early word learning. In M. Bowerman and S. Levinson, (eds.), *Language Acquisition and Conceptual Development*. Cambridge: Cambridge University Press.

Gentner, D. & Markman, A.B. (1997). Structure mapping in analogy and similarity. *American Psychologist*, **42**, 45-56.

Gerard, L. & Scarborough, D. (1989). Language-specific lexical access of homographs by bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition.* **15**(2), 305-315.

Glaser, E.R. & Dungelhoff, F-J. (1984). The time course of picture-word interference. *Journal of Experimental Psychology: Human Perception and Performance*, **10**(5), 640-654.

Glaser, E.R. & Glaser, M.O. (1989). Context effects in Stroop-like word and picture processing. *Journal of Experimental Psychology*, **118**(1), 13-42.

Gopnik, A., & Choi, S. (1995). Names, relational words and cognitive development in English and Korean speakers: Nouns are not always learned before verbs. In M. Tomasello & W. Merriman (eds.) *Beyond names for things: Young children's acquisition of verbs.* Hillsdale, NJ: Lawrence Erlbaum, 63-80.

Gopnik, A. & Meltzoff, A. N. (1986). Relations between semantic and cognitive development

in the one-word stage: The specificity hypothesis. *Child Development*, **57**, 1040-1053.

Gopnik, A. & Sobel, D. (2000). Detecting blickets: How young children use information about novel causal powers in categorization and induction. *Child Development,* **71**(5), 1205-1222.

Grainger, J. & Beauvillain, C. (1987). Interlexual homographs: Some limitations of a language selective access. *Journal of Memory and Language*, **26**, 658-672.

Greenberg, J. & Kuczaj, S. A. (1982). Towards a theory of substantive word-meaning acquisition in Stan A. Kuczaj (ed.) *Language development: Syntax and semantics*. Mahwah, NJ: Lawrence Erlbaum, 275-312.

Grosjean, F. (1982). *Life with two languages: An introduction to bilingualism.* Cambridge, MA: Harvard University Press.

Grosjean, F. (1992). Another view of bilingualism. In R.H.Jackson (ed.) *Cognitive processing in bilinguals*. Amsterdam: North-Holland Publishing Co., 51-62.

Grosjean, F. (1997). Processing mixed language: Issues, findings, and models. In A.M.B. De Groot & J.F. Kroll (eds.) *Tutorials in Bilingualism: Psycholinguistic perspectives*. Mahwah, NJ: Erlbaum, 225-254.

Grosjean, F. (1998). Transfer and language mode. *Bilingualism: Language and Cognition,* **1**(3), 175-176.

Grosjean, F. (2001). *The Bilingual's Language Modes*. Unpublished Paper. http://www.unine.ch/ltlp/pub/langmode.pdf accessed 10.09.2009

Guerssel, M. et al (1985). *A cross-linguistic study of transitivity alternations. Papers on the Parasession on Causitives and Agentivity.* Chicago: Chicago Linguistics Society, 48-

63.

Hampton, J.A. (1979). Polymorphous concepts in semantic memory. *Journal of Verbal Learning and Verbal Behaviour,* **18**(4), 441-461.

Hampton, J. A. (1981). An investigation of the nature of abstract concepts. *Memory and Cognition,* **9** (2), 149-156.

Halliday, M.A.K. & Hasan, R. (1976). *Language cohesion*. London: Longmans.

Haritos C. & Nelson K. (2001). Bilingual memory: The interaction of language and thought. *Bilingual Research Journal,* ***25***(4), 605–626.

Hart, B. & Risley, T.R. (1995). *Meaningful differences in the everyday experience of young American children*. Baltimore, MD: Brookes Publishing Co.

Haugen, E. (1956). *Bilingualism in the Americas*. Birmingham AL: University of Alabama Press.

Healey, E. & Sarabella, B. (2008). Are children willing to accept two labels for a single object? A comparative study of mutual exclusivity in monolingual and bilingual children. *Journal of Child Language,* **19**, 633- 658.

Holowka, S., Brossea-Lapre, F. & Petitto, L. A. (2002). Semantic and conceptual knowledge underlying bilingual babies' first signs and words. *Language Learning*, **52**(2), 202-262.

Hulk, A. & Muller, N. (2000). Bilingual first language acquisition at the interface between syntax and pragmatics. *Bilingualism: Language and Cognition,***3**, 227-244.

Jones, H. (2003). Wales in the censuses of population: a longitudinal study. Welsh Language Board. URL http://www.byig-wlb.org.uk/English/publications/Publications/2619.pdf

accessed 10.5.2009.

Keatley, C.W. Spinks, J.A. & de Gelder, B.(1994). Asymmetrical cross-language priming effects. *Memory & Cognition*, **22**(1), 70-84.

Kemler-Nelson, D.G. (1984). The effect of intention on what concepts are acquired. *Journal of Verbal Learning and Verbal Behavior*, **23**, 734-759.

Kowalski, K. & Zimiles, H. (2006). The relation between children's conceptual functioning with color and color term acquisition. *Journal of Experimental Child Psychology*, 94 (4), 275-366.

Krascum, R. M. & Andrews, S. (1998). The effects of theories on children's acquisition of family-resemblance categories. *Child Development,* **69**(2), 333-346.

Kripke, S. (1971). Identity and necessity. In M. Munitz (ed.) *Identity and Individuation*, New York: New York University Press,135-164.

Kroll, J.F. (1993). Accessing conceptual representations for words in a second language. In R. Schreuder & B. Weltans (eds.) *The bilingual lexicon*. Amsterdam. Philadelphia: John Benjamins, 53-81.

Kroll, J.F. & Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language,* **33** (2), 149-174.

Kroll, J.F., Michael, E., Tokowicz, N. & Dufour, R. (2002). The development of lexical fluency in a second language. *Second Language Research,* **18**, 137-171.

Kronenfeld, D.B., Armstrong, J.D. & Wilmoth, S. (1985). Exploring the internal structure of linguistic categories: An extensionist semantic view. In J.W.D. Dougherty (ed.)

Directions in cognitive anthropology, Urbana: University of Illinois Press, 91–113.

Kuczaj, S.A. (1982). Young children's overextensions of object words in comprehension and/or production: Support for a prototype theory of early object word meaning. *First Language*, **3**(8), 93-105.

Labov, W. (1973). *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania.

Lakoff, G. (1987). *Women, fire, and dangerous things.* Chicago: University Of Chicago Press

Lin. E. L., & Murphy, G. L. (2001). Thematic relations in adults' concepts. *Journal of Experimental Psychology: General*, **130**, 3-28.

Lindemann, P.G. & Markman, A.B. (1996). Alignability and attribute importance in choice. In G. Cottrell (ed.) *Proceedings of the eighteenth annual meeting of the Cognitive Science Society*. Hillsdale, NJ: Erlbaum, 358-363.

Lopez, M. & Young, R. K. (1974). The linguistic interdependence of bilinguals. *Journal of Experimental Psychology,* **102**, 981-983.

McCloskey, M. & Glucksberg, S. (1979). Natural categories: Well defined or fuzzy sets? *Memory & Cognition*, **6**(4), 462-472.

McGregor, K.A., Friedman, R. M., Reilly, R. & Newman, R.M. (2002). Semantic representation and naming in young children. *Journal of Speech, Language, and Hearing Research*, **45**. 332-346.

Majid, A.,Van Staden, M., Boster,J. & Bowerman, M. (2004). *Event categorization: A cross-linguistic perspective.* Chicago: Cognitive Science Society Annual Meeting Proceedings, 885-890.

Malt, B.C. et al (1999). Knowing versus naming: similarity and the linguistic categorisation

of artefacts. *Journal of Memory and Language*, **40,** 230-262.

Malt, B.C. & Sloman,S.A. (2003). Linguistic diversity and object naming by non-native speakers of English. *Bilingualism: Language and Cognition*, **6**, 47-67.

Malt, B.C., Sloman, S.A. & Gennari, S. (2003). Speaking versus thinking about objects and actions. In D. Gentner & S. Goldin-Meadow (eds.) *Language in mind: Advances in the study of language and thought.* Cambridge, Mass.: MIT Press, 81-111.

Markman, E. M. (1990). Constraints children place on word meanings. *Cognition*, **14**, 57-77.

Markman, E.M. (1991). *Categorization and naming in children: problems of induction.* Cambridge, MA: MIT Press.

Markman, E. M. (1994). Constraints on word meaning in early language acquisition. *Lingua,* **92**, 199-227.

Markman, E.M. (2003). *Constraints on word learning: Speculations about their nature,origins and domain specificity*. Stanford, CA: Stanford University.

Medin, D.L. & Smith, E.E. (1981). Strategies and classification learning. J*ournal of Experimental Psychology: Learning, Memory and Cognition*, **7**, 241-253.

Medin, D.L.(1989). Concepts and conceptual structure. American Psychologist, **44**(12), 1469-1481.

Meisel, J. (1997). The acquisition of the syntax of negation in French and German contrasting first and second language development. *Second Language Research*, **13,** 227-63.

Mervis, C.B.(1987). Child basic object categories and early lexical development. In U. Neisser (ed.) *Concepts and Conceptual Development: Ecological and Intellectual Factors in Categorisation*. Cambridge: Cambridge University Press, 201-233.

Meyer, D.E., & Schvaneveldt, R.W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology,* **90**, 227-234.

Morris, J & Hurst, G. (1991). Lexical cohesion computed by thesaural relations as an indicator of the structure of text. *Computational Linguistics*, **17**(1), 21-43.

Murphy, G. L. (2002). *The Big Book of Concepts*. Cambridge, MA: MIT Press.

Murphy, G.L. (2001). Causes of taxonomic sorting by adults: a test of the taxonomic-thematic shift. *Psychonomic Bulletin & Review,* **8**(4), 834-839.

Murphy, G.L. & Medin, D. L. (1985). The role of theories in conceptual coherence. *Psychological Review* **92,** 289-316.

Nazzi, T. & Gopnik, A. (2001). Linguistic and cognitive abilities in infancy: when does language become a tool for categorization? *Cognition*, **80**(3),11-20.

Nicoladis, E. (2002). What's the difference between 'toilet paper' and the 'paper toilet ? French English bilingual children's crosslinguistic transfer in compound nouns. *Journal of Child Language,* **29**, 843-863.

Osherson, D. N. and Smith, E. E. (1981) On the adequacy of prototype theory as a theory of concepts. *Cognition* **12,** 299-318.

Paivio, A (1986). *Mental representations: a dual coding approach.* Oxford: Oxford University Press.

Paivio, A. & Desrochers, A. (1980). A dual-coding approach to bilingual memory. *Canadian Journal of Psychology*, **34**(4), 388-399.

Paivio, A. & Lambert,W. E. (1981), Dual Coding and bilingual memory. *Journal of Verbal*

Learning and Verbal Behavior, **20**(5), 532-539.

Paradis, J. (2001). Do bilingual 2-year olds have separate phonological systems? *International Journal of Bilingualism,* **5**, 19-38.

Paradis, J. & Navarro, S. (2003). Subject realization and crosslinguistic interference in the bilingual acquisition of Spanish and English: What is the role of the input? *Journal of Child Language,* 30, 371-393.

Paradis, J. & Genesee, F. (1996). Syntactic acquisition in bilingual children: Autonomous or interdependent? *Studies in Second Language Acquisition,* 18, 1-25.

Pavlenko, A. (ed.) (2009). *The bilingual mental lexicon: Interdisciplinary approaches*. Bristol: Multilingual Matters.

Pavlenko, A. & Jarvis,S. (2002). Bidirectional transfer. *Applied Linguistics*, **23**(2), 190-214.

Pavlenko, A. & Jarvis, S. (2002). Bidirectional transfer. *Applied Linguistics* 23(2), 190-214.

Pena, E., Bedore, L. & Rappazzo, C. (2003). Comparison of Spanish, English, and bilingual children's performance across semantic tasks. *Language, Speech, and Hearing Services in Schools,* **34**, 5-16.

Piaget, J. (1954). *Intelligence and affectivity: Their relationship during child development.* Palo Alto, CA: Annual Review Inc.

Potter, M.C., So, K.F., Von Eckhardt, B. & Feldman, L.B.(1984). Lexical and conceptual representation in beginning and more proficient bilinguals. *Journal of Verbal Learning and Verbal Behavior,* **23**, 23-38.

Pye, C. (1996). *K'iche' Maya verbs of breaking and cutting*. Linguistics Colloquy, University of Kansas, February 1996.

Ramscar, M. et al (1997). Is there a place for semantic similarity in the analogical mapping process? *Proceedings of the 19th Annual Conference of the Cognitive Science Society,* Stanford University.

Ramscar, M. & Yarlett, D. (2007). Linguistic self-correction in the absence of feedback: A new approach to the logical problem of language acquisition. *Cognitive Science*, **31**, 927-960.

Rescorla, L.A. (1980). Overextension in early language development. *Journal of Child Language*, **7**(2), 321-335.

Rips, L. J. (1975). Inductive judgements about natural categories. *Journal of Verbal Learning & Verbal Behaviour*, **14**(6), 665-681.

Rips, L. J. (1975). Quantification and semantic memory. *Cognitive Psychology,* **7**(3), 307-340.

Rips, L.J.& Medin, D.L. (2005). Concepts, categories and semantic memory. In K.Holyoak & R.Morrison (eds.) Cambridge handbook of thinking and reasoning. New York: Cambridge University Press, 37-72.

Rosch, E.H. (1973). Natural categories. *Cognitive Psychology*, **4**, 328-350.

Rosch, E.H. (1975). Cognitive representation of semantic categories. *Journal of Experimental Psychology,* **104**(3), 192-233.

Samuelson, L. & Smith, L.B. (2000). Attention to rigid and deformable shape in naming and non-naming tasks. *Child Development,* **71**, 1555-1570.

Schrauf, D.C. & Rubin, R.W. (2003). Belief and recollection of autobiographical memories. *Memory and Cognition*, **31**(6), 887-991.

Schwanenflugel, P. & Ray, M. (1986). The relationship between category typicality and

concept familiarity: Evidence from Spanish and English-speaking monolinguals. *Memory and Cognition*, **14**(2), 150-163.

Silverberg, S. & Samuel, A.G. (2004). The effect of age on second language acquisition on the representation and processing of second language words. *Journal of Memory and Language*, **51**, 381-398.

Sloman, S.A., Malt, B.C. & Fridman, A. (2001). Categorization versus similarity: The case of container names. In U.Hahn & M.Ramscar (eds.) *Similarity and Categorization.* Oxford: Oxford University Press, 33-86.

Smiley, S. & Brown, A.L. (1979). Conceptual preferences for thematic or taxonomic relations: A nonmonotonic age trend from preschool to old age. *Journal of Experimental Child Psychology*, **28**, 249-247.

Smith, E. E., Shoben, E.J. & Rips, L.J. (1974). Structure and process in semantic memory: A future model for semantic decisions. *Psychological Review,* **1**, 214-241.

Stepanova, O. & Coley, J.D.(2002). The green-eyed monster: Concepts of envy and jealousy in Russian and English. *Journal of Cognition and Culture*, **2**, 235-262.

Stephany, U. & Voeikova, M (2003). On the early development of aspect in Greek and Russian child language, a comparative analysis. *ZAS Papers in Linguistics,* **29,** 211-224.

Swain, M. (1995). Three functions of output in second language learning. In G. Cook & B. Seidlhofer (eds.) *Principle and practice in applied linguistics: Studies in honour of H.G. Widdowson*. Oxford: Oxford University Press, 125-144.

Taylor, I. & Taylor , M. (1990). *Psycholinguistics: Learning and using lang*uage. Englewood

Cliffs, NJ.: Prentice Hall International.

Vainikka, A. & Young-Scholten, M. (1996). Gradual development of L2 phrase structure. *Second Language Research,* **12**(1), 7-39.

Vihman, M. et al (2007). Onset of word form recognition in English, Welsh, and English-Welsh bilingual infants. *Applied Psycholinguistics*, **28**, 475-493.

Vygotsky, L. S. (1962). *Thought and language.* Edited and translated by Eugenia Hanfmann and Gertrude Vakar. Cambridge, MA: MIT Press.

Wattenmaker, W. D., Nakamura, G. V., & Medin, D. L. (1988). Relationships between similarity-based and explanation-based categorization. In D. Hilton (ed.) *Contemporary science and natural explanation: Common sense conceptions of causality*. Brighton: Harvester Press, 205-241.

Waxman, S.R. (1994). The development of an appreciation of specific linkages between linguistic and conceptual organization. *Lingua*, **92**, 229-57.

Waxman, S.R. & Braun, I. (2005). Consistent (but not variable) names as invitations to form object categories: New evidence from 12-month-old infants. *Cognition*, **95**(3), 59-63,

Waxman, S.R. & Markov (1995). Words as invitations to form categories: Evidence from 12-to-13 month old infants. *Cognitive Psychology*, 29(3), 257-302.

Yip, V. & Matthews, S.(2000). Syntactic transfer in Cantonese-English bilingual child. *Bilingualism:Language and Cognition*, **3**(3), 193-208.

Yip, V. & Matthews, S (2005). Dual input and learnability: Null objects in Cantonese-English bilingual children. In J. Cohen et al (eds.) *Proceedings of the 4th International Symposium on Bilingualism.* Somerville,MA: Cascadilla Press, 2421-2431.

Appendix 1: Documents sent to Schools and Parents/Guardians

A1.1 Letter to Schools
A1.2 Parental Consent Letter
A1.3 Bilingual Questionnaire
A1.4 Monolingual Questionnaire

School Session 2009

Dear Head of School,

We are a conducting a study concerning language categories in Welsh and English in children and are looking for schools willing to participate. The current study will form the basis for the PhD thesis of a student working under my supervision in the School of Psychology, University of Wales, Bangor.

The study will examine children's categorization and the possible variables thought to influence this. The study will involve children in years 6, 5, 4 ,3,2,1 taking part in approximately 2 sessions, each lasting no longer than 20 minutes. The children will be tested on a one-to-one basis with a researcher. The tasks will be simple, interesting, and non-competitive.

The procedure involves the children looking at a PowerPoint slide show and judging which pictures match the words said using a touch screen monitor. The children will also be given the British Picture Vocabulary Scales to gain an understanding of their fluency in English. These scores will not be used to compare schools or children. Confidentiality will be respected throughout the study – no names will be attached to the data, and all reports will respect both the school's and the child's anonymity.

No child will be forced into participating, and if at any stage he or she wishes to withdraw, he or she will be free to do so. Parents will be welcome to come and observe the sessions if they wish. If you give permission for your school to participate, we will send the parents a consent letter describing the study.

If you have any questions or would like more information please do not hesitate to contact Professor Virginia C. Mueller Gathercole_on 01248 382624 pss116@bangor.ac.uk or Mr. Hedd Tomos at psp25f@bangor.ac.uk. If you should have any complaints about the study, please direct these to Professor Oliver Turnbull o.turnbull@bangor.ac.uk.

Yours sincerely,

Professor Virginia C. Mueller Gathercole

PRIFYSGOL
BANGOR
UNIVERSITY

Dear Parent / Guardian,

We will be conducting a study in your child's school concerning language categories in Welsh and English in children. The current study will form the basis for the PhD thesis of a student working under my supervision in the School of Psychology, University of Wales, Bangor.

The study will investigate the way that children categorise words and the possible influences on this process. The children will be tested on a one to one basis with a researcher. The tasks are simple, interesting and non competitive.

We are seeking your permission for your child to take part in the study. No child will be forced into participating, and if at any stage he or she wishes to withdraw, he or she will be free to do so. Parents will be welcome to come and observe the sessions if they wish.

We anticipate that each child will be seen for about twenty minutes on two occasions. The sessions will take place in the school at a time convenient to the school.

Confidentiality will be respected throughout the study – no names will be attached to the data, and all reports will respect both the school's and the child's anonymity.

Also, if you give permission for your child to participate, we would ask that you complete a short questionnaire concerning the child's language background. We will be sending you a questionnaire if you agree for your child to participate.

After the completion of the study, we can make available to you a short summary of the findings. If you would like us to send you this summary, we would appreciate it if you could supply your address below.

If you have any questions or would like more information please do not hesitate to contact Professor Virginia C. Mueller Gathercole_on 01248 382624 pss116@bangor.ac.uk or Mr. Hedd Tomos at psp25f@bangor.ac.uk. If you should have any complaints about the study, please direct these to Professor Oliver Turnbull o.turnbull@bangor.ac.uk. Thank you once again. We look forward to your response.

Yours sincerely,

Professor Virginia C. Mueller Gathercole_____

Please complete this slip as appropriate and return as soon as possible to the school or to the address shown above.

Child's name:_____ Date of birth:_____

School Year [please tick one]:

Year 1 ❑ Year 2 ❑ Year 3 ❑ Year 4 ❑ Year 5 ❑ Year 6 ❑

Parent's / Guardian's name:_____

Telephone number:_____

Address (optional):_____

- ❑ I consent for my child to participate in the study
- ❑ I do not consent for my child to participate in the study
- ❑ I understand that I may withdraw my child from the study at any point

Parent's / Guardian's signature:_____Date:_____

If we have any further studies may we contact you further? Yes ❑ No ❑

PARENTAL QUESTIONNAIRE FOR CHILD PARTICIPANTS

Thank you for allowing your child to participate in this study. We would like to get some background information on the language background of your child and your family, as well as information on your child's activities outside of school. Please feel free to answer these questions in any way you feel is appropriate, and if there is any question you would rather not answer, that is fine too. Just leave it blank and pass on to the next question.

Child's name:_____
Child's date of birth:_____
Child's gender: Male ☐ Female ☐
Child's place of birth:_____
Child's school year and teacher's name:_____

1. Which languages does your child speak? (Please tick all that apply)

☐ English
He/she began speaking English at age:_____
Please state approximately what percentage of the time your child speaks English:
A: 100% ☐ B: 75% ☐ C: 50% ☐ D: 25% ☐ E: 0% ☐
Please state approximately what percentage of the time English is currently spoken in the home:
A: 100% ☐ B: 75% ☐ C: 50% ☐ D: 25% ☐ E: 0% ☐

☐ Welsh
He/she began speaking Welsh at age:_____
Please state approximately what percentage of the time your child speaks Welsh:
A: 100% ☐ B: 75% ☐ C: 50% ☐ D: 25% ☐ E: 0% ☐
Please state approximately what percentage of the time Welsh is currently spoken in the home:
A: 100% ☐ B: 75% ☐ C: 50% ☐ D: 25% ☐ E: 0% ☐

☐ Other language(s) – please specify:_____

He/she began speaking this language at age:_____
Please state approximately what percentage of the time your child speaks this language:
A: 100% ☐ B: 75% ☐ C: 50% ☐ D: 25% ☐ E: 0% ☐
Please state approximately what percentage of the time this language is currently spoken in the home:
A: 100% ☐ B: 75% ☐ C: 50% ☐ D: 25% ☐ E: 0% ☐

2. Which languages were spoken in your home when your child was the following ages? (Including the language used by grandparents and siblings in speaking to your child)

A: 100% English
B: about 80% English, 20% Welsh
C: about 60% English, 40% Welsh
D: about 50% English, 50% Welsh
E: about 40% English, 60% Welsh
F: about 20% English, 80% Welsh
G: 100% Welsh
H: other combination
I: does not apply – child is not this old

From the child's birth until he/she turned two years of age_____
When he/she was 3 and 4 years of age_____
When he/she was 5 and 6 years of age_____
When he/she was 7 and 8 years of age_____
When he/she was 9 and 10 years of age_____

3. Please indicate which of the following activities your child participates in (tick all that apply):

- ☐ Uniformed organisations (e.g. Brownies, Cubs)
- ☐ Sports clubs (e.g. netball, football)
- ☐ Music (e.g. orchestra, choir)
- ☐ Yr Urdd
- ☐ Eisteddfodau
- ☐ After-school club
- ☐ Other – please specify_____

4. What is the mother's occupation? Please specify:_____

5. What is the father's occupation? Please specify:_____

6. Please indicate the highest level of education obtained by each parent:

Mother:
R 1 2 3 4 5 6 7 8 9 10 11 12 13 1 2 3+ 1 2 3+
Father:
R 1 2 3 4 5 6 7 8 9 10 11 12 13 1 2 3+ 1 2 3+

Primary Secondary GCSE A-level University Post+

7. In general, what language(s) can you speak?

Welsh _____ English _____ Other L1: [specify] _____ Other L2: [specify] _____

8. Has your child ever undergone speech or language therapy?
Yes ☐ No ☐

9. Has your child ever been treated for a hearing problem?
Yes ☐ No ☐

If 'Yes', when?

_____

10. Has your child ever been diagnosed with dyslexia?
Yes ☐ No ☐

11. Has your child ever been diagnosed with a behavioural problem?
Yes ☐ No ☐
If yes, please specify:_____

12. Has any other member of the family ever undergone speech or language therapy?
Yes ☐ No ☐
If yes, please give their relationship to the child:
☐ Mother
☐ Father
☐ Brother / Sister
☐ Other – please specify:_____

13. Has any other member of the family ever been treated for a hearing problem?
Yes ☐ No ☐
If yes, please give their relationship to the child:
☐ Mother
☐ Father
☐ Brother / Sister
☐ Other – please specify:_____

14. Has any other member of the family ever been diagnosed with dyslexia?
Yes ☐ No ☐
If yes, please give their relationship to the child:
☐ Mother
☐ Father
☐ Brother / Sister
☐ Other – please specify:_____

THANK YOU FOR TAKING THE TIME TO COMPLETE THIS QUESTIONNAIRE

PARENTAL QUESTIONNAIRE FOR CHILD PARTICIPANTS

Thank you for allowing your child to participate in this study. We would like to get some background information on the language background of your child and your family, as well as information on your child's activities outside of school. Please feel free to answer these questions in any way you feel is appropriate, and if there is any question you would rather not answer, that is fine too. Just leave it blank and pass on to the next question.

Child's name:_____
Child's date of birth:_____
Child's gender: Male ☐ Female ☐
Child's place of birth:_____
Child's school year and teacher's name:_____

1. Which languages does your child speak? (Please tick all that apply)

☐ English
He/she began speaking English at age:_____
Please state approximately what percentage of the time your child speaks English:
A: 100% ☐ B: 75% ☐ C: 50% ☐ D: 25% ☐ E: 0% ☐
Please state approximately what percentage of the time English is currently spoken in the home:
A: 100% ☐ B: 75% ☐ C: 50% ☐ D: 25% ☐ E: 0% ☐

☐ Other language(s) – please specify:_____
He/she began speaking this language at age:_____
Please state approximately what percentage of the time your child speaks this language:
A: 100% ☐ B: 75% ☐ C: 50% ☐ D: 25% ☐ E: 0% ☐
Please state approximately what percentage of the time this language is currently spoken in the home:
A: 100% ☐ B: 75% ☐ C: 50% ☐ D: 25% ☐ E: 0% ☐

2. Please indicate which of the following activities your child participates in (tick all that apply):

☐ Uniformed organisations (e.g. Brownies, Cubs)
☐ Sports clubs (e.g. netball, football)
☐ Music (e.g. orchestra, choir)
☐ After-school club
☐ Other – please specify_____

3. What is the mother's occupation? Please specify:_____

4. What is the father's occupation? Please specify:_____

5. Please indicate the highest level of education obtained by each parent:

Mother:	R	1	2	3	4	5	6	7	8	9	10	11	12	13	1	2	3+	1	2	3+
Father:	R	1	2	3	4	5	6	7	8	9	10	11	12	13	1	2	3+	1	2	3+
	Primary							Secondary			GCSE		A-level		University			Post-		

6. In general, what language(s) can you speak?

English _____ Other L1: [specify] _____ Other L2: [specify] _____

English:

a. As long as I can remember: _____

b. Since before going to school: _____

c. Once I started primary school: _____

d. Not until I started secondary school: _____

e. I learnt it as an adult

Other language:

a. As long as I can remember: _____

b. Since before going to school: _____

c. Once I started primary school: _____

d. Not until I started secondary school: _____

e. I learnt it as an adult _____

7. In general, what language(s) can your partner speak [if relevant]?

English _____ Other L1: [specify] _____ Other L2: [specify] _____

8. Has your child ever undergone speech or language therapy?
Yes ☐ No ☐

9. Has your child ever been treated for a hearing problem?
Yes ☐ No ☐

If 'Yes', when?

_____

10. Has your child ever been diagnosed with dyslexia?
Yes ☐ No ☐

11. Has your child ever been diagnosed with a behavioural problem?
Yes ☐ No ☐
If yes, please specify:_____

12. Has any other member of the family ever undergone speech or language therapy?
Yes ☐ No ☐
If yes, please give their relationship to the child:
☐ Mother
☐ Father
☐ Brother / Sister
☐ Other – please specify:_____

13. Has any other member of the family ever been treated for a hearing problem?
Yes ☐ No ☐
If yes, please give their relationship to the child:
☐ Mother
☐ Father
☐ Brother / Sister
☐ Other – please specify:_____

14. Has any other member of the family ever been diagnosed with dyslexia?
Yes ☐ No ☐
If yes, please give their relationship to the child:
☐ Mother
☐ Father
☐ Brother / Sister
☐ Other – please specify:_____

Appendix 2:Examples of Non-linguistic Stimuli PowerPoint Slides

A2.1 Categorisation Task
A2.2 Memory Task

A

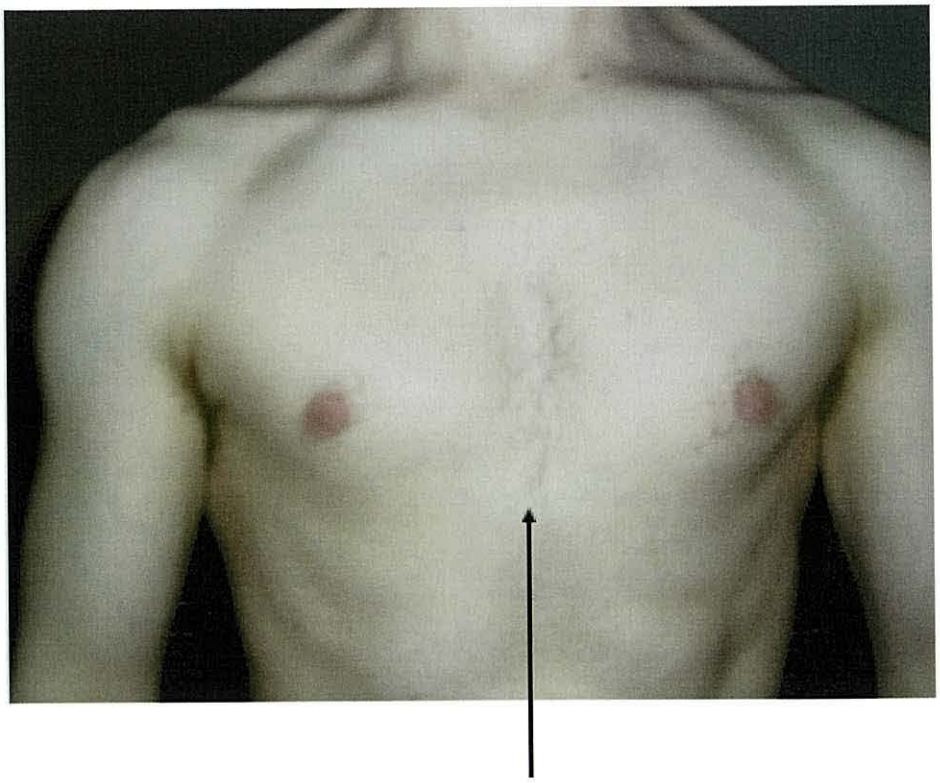
B

C

D

E

F



Appendix 3: An Example of a Scoring Sheet for the Categorisation Task Study

Name:_____
Date:_____

Year:_____

School:_____

TEST ITEM	WELSH	RESPONSES	ENGLISH	RESPONSES
1	ANIFAIL		ANIMAL	
2	COCH		RED	
3	MELYN		YELLOW	
4	EISTEDD		SITTING	
5	GLAS		BLUE	
6	GWYRDD		GREEN	
7	BAWD		THUMB	
8	PAFFIO		BOXING	
9	BWRW		RAINING	
10	BREST		CHEST	
11	CHWYRNU		SNORING	
12	CODI		LIFTING	
13	COLLI		LOSING	
14	PEN		HEAD	
15	DYSGL		DISH	
16	DAEAR		EARTH	
17	FFORDD		ROAD	
18	PLYGU		FOLDING	
19	PORI		GRAZING	
20	LLAW		HAND	
21	HWYL		FUN	
22	GORIAD		KEY	
23	HOELEN		NAIL	
24	PWRS		PURSE	
25	YSMYGU		SMOKING	
26	FFON		STICK	
27	MUR		WALL	
28	YSGOL		SCHOOL	
29	CYSGO		SHADOW	
30	SHADOW		HIGH	