How do Semantic Category Effects, Words of Different Levels of Concreteness and Syntactic Class Affect Vocabulary Retention in Young Asian EFL Learners?

KEITH TAYNTON

SEPTEMBER 2009

THE WESTMINSTER INSTITUTE OF EDUCATION AT OXFORD BROOKES UNIVERSITY

This dissertation is submitted in part fulfilment of the requirements governing the award of the Master of Arts in Education (TESOL).

Statement of originality

This Dissertation is an original piece of work which is made available for photocopying and for inter library loan, with permission of the Head of the Westminster Institute of Education.

Signed *Keith Taynton*

Acknowledgements

Many people have contributed to the production of this study and I would like to offer my heartfelt thanks for their time, support and assistance. First, my biggest thanks go to Rie Tanimoto for her late nights translating, and also to Yukie Tatsumi for her translations and advice. I must of course also thank the staff and first year students at Osaka International Owada Junior High School for their time and assistance while conducting this study and without whom I would not have had the opportunity to learn and write this dissertation. Special thanks must go to Kazuko Okuno and Utako Nagasako for their assistance with The Administration. Thanks must also go to Mayumi Fujita and Matt Lucas for moral support and understanding during the numerous crises. Lastly, but certainly not least, the teaching and administration staff at Oxford Brookes for their excellent tuition and especially my supervisor Paul Wickens for going the extra mile.

Abstract

Research into vocabulary learning has found that Semantic Category Effects (SCE) (Hakki Erten and Tekin, 2008; Finkbeiner and Nicol, 2003) hinder retention when words are learnt in semantically related categories. Other variables such as a word's concreteness and word class have also been found to change learnability. In this study young Japanese learners (*n*82, average age 12) were tested to see what effect manipulating Concreteness, Syntactic Class and Semantic Relationship had on learnability. Statistical analysis showed that previous studies' predictions about the effect of each variable had on learnability were generally confirmed.

However one discrepancy in the Qualitatively Differential Representational Framework (Duñabeitia's et. al. 2009; Crutch, 2006; Crutch and Warrington, 2006) led to a reconsideration of how the lexical network may be constructed and used. A Network Relationship Model proposes that words are stored in semantic categories which are affected by SCE. But an associative mode of input during learning may negate any SCE thus increasing learning efficiency. Implications of the analysis, teaching methods and approaches, and paths for future research are discussed.

1 Table of Contents

List of Tables and Figures

2 Introduction

Within the field of second language acquisition (SLA) one research focus has been the investigation of vocabulary learning (Hall, 1992:3). Newer ideas such as Lewis' lexical approach (LA) (Moudraia, 2001) emphasize the learning of vocabulary like single words, collocations and idioms, over a more traditional focus on grammar. Language, according to this approach "… consists of grammaticalized lexis, not lexicalized grammar" (ibid:2).

Harwood's review (2002) encompasses many of the problems facing practitioners wishing to implement LA, from lack of course books to access to corpuses for the raw materials to build lessons. One particular problem is that of teaching grammar and vocabulary:

> "Chomskyan generative paradigm cannot be claimed to describe language adequately and that the realms of grammar and lexis are neither readily definable nor even necessarily discrete (and hence in teaching lexis one can simultaneously be teaching grammar)" (2002:9).

Thus traditional boundaries which separate these aspects can be removed which changes how teachers need to approach grammar and vocabulary learning.

There may be value in learning set chunks of high frequency language by rote. This would rapidly increase a store of language and also give confidence to students when faced with communicative situations. Boers et al. (2006) found in their study that a ready store of formulaic phrases can contribute to oral proficiency.

Following from Krashen's input hypothesis (Brown et al., 2008:136-139) is the "extensive" approach where the student chooses material that is intrinsically interesting to read and gains vocabulary incidentally rather than through explicit instruction. Although there is controversy around the efficiency to this approach it does seem that new words and other improvements, such as better spelling and increased grammatical knowledge, can be attained from extensive input, and enhanced even more with post reading vocabulary exercises (Min, 2008).

While extensive and lexical approaches may be suitable for students with an existing L2 lexis, and much research has been done in this area in adults at intermediate level and above, a question remains on how to effectively and efficiently teach beginners, especially children, a basis of vocabulary from which they can develop into more independent learners (Comesana et al., 2009). This study investigates the reported phenomenon of interference effects, word class and concreteness on retention of vocabulary in young Asian learners with a view to increasing understanding of how best to present vocabulary to enable better retention.

3 Literature review

One of the contentious areas of investigation related to vocabulary acquisition is the best way to teach vocabulary and which methods enhance retention and use, for example implicit or explicit tuition (Broady, 2008). It has been argued (Nation, 2001:297-301; Oxford and Crookall, 1990 in Webb, 2007:64) that teaching words in a decontextualised way such as L1-L2 word pair translation, does not provide as much information about word use for the learner as when learned in context. Webb (2007) investigated the effect of teaching words in context and decontextualised conditions, and found no significant difference in understanding across a wide range of metaknowledge such as orthography and syntax between glossed contextualized sentences and translation pairs.

Another aspect which is being debated is the way L2 relates to L1 and primary conceptual knowledge. Three basic positions by Potter et al. (1984, in de Groot and Hoeks, 1995:685) are: A "word association hypothesis" where L1 and L2 have direct connections to each other and L1 mediates L2 conceptual access; "Concept mediation" where L2 is directly connected to conceptual knowledge; and a "developmental hypothesis" where L2 is initially mediated through L1 to concepts but eventually direct connections to the conceptual store are made as the learner becomes more proficient.

3.1 The Revised Hierarchical Model and the L2 Mental Lexicon

Two models related to this are Revised Hierarchical Model (RHM) by Kroll and Stewart (in Comesana et al., 2009:23; Li et al.'s 2009) and the L2 Mental Lexicon model by Hall (1992) that postulate that initially L2 learners access concepts through L1 lexis, while later on developing direct connections to concepts as proficiency increases. However several studies, among them Comesana et al. (2009) and Ferre et al. (2006), seem to show that L2 lexis can directly access the conceptual store of knowledge, even in beginners, which opposes the RHM's prediction (Comesana, 2009:23), but is allowed for in Hall's providing no original frame of L1 reference can be found by the learner to match the concept and L2 word.

Hall's model posits that initially, at least, L2 is parasitic on L1 form, frame and concept. He theorizes that a new L2 word form is initially, and inevitably, compared to a L1 "translational equivalent" (TE) by the learner (Hall, 1992:46), which then passes through a frame (a set of metaknowledge about the word's position, use and associations in a lexicon) before being connected to the conceptual knowledge which is the meaning of the word (ibid:11-25). If a new word is discovered that has no TE then the lexical processor builds a new frame for the new word which then connects to the concept (ibid:41-42). Gradually, once more evidence is accumulated about meaning, the strength of the connection between L2 word and concept becomes firm. Hall states that the implications of this model are that when learning new L2 vocabulary words should be given in a simple L1-L2 vocabulary list with the translation of L1 word being selected by the teacher as the most common form in use for the L2 word (ibid:47).

Presumably the direct connection of L2 to concept when no TE exists requires a lot of cognitive effort as the learner seeks to negotiate meaning of the L2 word, and concept, with existing conceptual knowledge. For example, for students of Japanese the Japanese concept of *wabsi sabi* (the idea of a certain kind of balance, peace and harmony) has no direct translation into English which requires a lot more explanation of the idea and *experience* of it, for the development of the concept for the foreign learner. According to Hall's model this would build a direct connection from L2 (Japanese) to the conceptual store, bypassing L1 (Mother Tongue) even in beginners. However the negotiation of meaning would presumably take place in L1 for the beginner which begs the question of what relationship the concept would have between L1-L2 connections during the learning process – a point not elaborated on in Hall's model.

3.2 Semantic Category Effects or Spread of Activation

One effect that has been found to hinder retention of words are semantic interference conditions known as the Semantic Category Effect (SCE) or Spread of Activation (SOA). Interference models propose that activating words that are semantically related at a conceptual level cause interference, manifested as longer recall times and difficultly in remembering the correct meaning or form, for subsequent recall of semantically related words (Damien, 2005:1381).

As Howard asserts

> "These [studies'] results definitively demonstrate the occurrence of cumulative interference for word retrieval by prior retrieval of other exemplars of the same semantic category—cumulative semantic inhibition" (2006:464).

There is a solid base of research into these phenomenon in other fields (e.g. Arevalo et al., 2007; Damian and Als, 2005; Howard, 2006) and these researchers approach the problem by using sophisticated MRI scanning techniques, latency tests and so forth. But their interest lies in brain function and the effects of conditions like dyslexia, and so they usually use only L1 participants. However, this research also has implications for second language teaching and learning as some studies seem to show a similar effect for L2 learners, which has been suggested to have implications for the organisation and delivery of vocabulary teaching in EFL/ESL text books and classrooms (Hakki Erten and Tekin, 2008).

Evidence that beginner L2 learners access concepts directly, and are thus susceptible to semantic interference effects, comes from the results of Comesana et al.'s experiment (2009) which tested response times in a translation exercise for beginners and more proficient L2 learners. According to the RHM beginners should access concepts through L1 and so a semantic interference effect should not be at work as much as in more proficient L2 participants who have stronger bonds to the conceptual store. The study showed that even beginner learners show some evidence of a SOA effect which means that they could be accessing conceptual information directly. If there was no such effect it would mean learners access concepts through L1 lexis.

Two other studies on interference effects on novice L2 learners (Hakki Erten and Tekin, 2008; Finkbeiner and Nicol, 2003) seem to support the hypothesis that an SCE hinders acquisition of second language vocabulary. However there are some potential weaknesses in these two studies which the present study hopes to address.

Hakki Erten and Tekin's experiment, building on the work of Finkbeiner and Nicol, investigated the SCE in schoolchildren in a real teaching situation using English words. They taught two sets of related and unrelated words and measured the results of retention over a period of time, concluding that there was a significant interference effect which impeded word acquisition when vocabulary was presented in semantically related sets. However their results have one quirk (2008:415) which isn't adequately explained and which may be the result of a systematic error in the study.

The quirk is that while both semantically related sets seemed to show a significant failure in retention, one unrelated set also showed a negative difference between initial learning scores and subsequent recall after 10 days. This is contrary to the hypothesis which states that unrelated sets should be easier to learn. One possible explanation is that the words used in the unrelated set were of different lexical classes (adjectives and nouns), were ambiguous (e.g. students could have confused the meaning of swim as either a verb or noun during learning), or had different levels of concreteness: the adjective *happy* which has a concreteness rating of 355 which is abstract, compared to *goat* which scores 636, very concrete (Wilson, 1988). This may have confounded the learning of the unrelated set but not through an interference effect. This error leads to a consideration not mentioned in either study: that of the effect of the constructs Word Class and Concreteness on retention.

Finkbeiner and Nicol's more sophisticated study into learning and production of L2 when trained in semantically related or unrelated sets offers a more robust set of results. However one issue with this study is that the words used for both related and unrelated sets seems to have come from the same sets (2003:373) so the unrelated condition cannot truly be said to have unrelated words, rather words from the four selected categories jumbled up and presented randomly.

There was no consideration of how, even though the items were presented randomly and not in sets, this effectively pseudo unrelated condition affected the results. As Damien and Als noted (2005:1373), the spread of activation may last from a few seconds up to several minutes, and also seems to have influence over a wide range of different activated nodes which means that perhaps even at a subconscious level activation in the four related categories chosen by Finkeiner and Nicol modified the outcome.

When dealing with lexical relationships there are several interrelated aspects which should be considered when studying interference effects. As already mentioned in Hakki Erten and Tekin's study, the effect of Word Class was not controlled for which may have confounded the results. Several studies have indicated that different types of words seem to be harder or easier to learn. Laufer's review suggests that "certain grammatical categories are more difficult to learn than others. Nouns seem to be the easiest; adverbs --the most difficult; verbs and adjectives -- somewhere in between" (1990: section 2.2). Dockrell's review of research also seems to support this notion commenting that nouns are perhaps easier to acquire than adjectives and verbs (2007:579), and Hall mentions that connections between L1 and L2 words are very complex (1992:49) which may explain why "simple" concrete nouns are easier because of the relative ease with which they can be related directly to real word objects sensed directly by sight, touch etc.

3.3 Organization of Language –Qualitatively Different Representational Framework

Several researchers have investigated word class and linguistic organisation in the brain. For example Bedney and Schill (2006), Druks (2002) and Gentner (2006) show that different word classes are located in different areas of the brain, and that some word classes are perhaps learnt more easily than others. This concurs with the results of Dockrell and Laufer, with concrete nouns being the first to be acquired in infants and abstract verbs later on, and this seems to be common to both L1 and L2 learners of all ages (Gentner, 2006:553-554).

Some recent evidence points to an organizational system where concrete and abstract lexis are stored in qualitatively different ways, with concrete items being stored in categories and abstract items in an associative network. Two studies (Crutch 2006, Crutch and Warrington, 2006) analyzed the way dyslexic patients identified and defined abstract and concrete items and concluded that there could be two such organization systems at work: “abstract concepts are represented in an associative network while concrete concepts are represented in a categorical framework” (Crutch, 2006:91).

These results seem to be verified by a Duñabeitia et al. study (2009) which determined how subjects categorized pictures of abstract or concrete items. The Qualitatively Different Representational Framework (QDR) theorizes that “abstract words are assumed to be organized mainly by semantic association and concrete words mainly by semantic similarity” (Duñabeitia et al., 2009:285), although they caveat it with the possibility that there is just one single network with different classes having different connection strengths (ibid).

The QDR model predicts that interference effects for abstract and concrete words are different. While concrete terms presented in semantic categories will be influenced by SCE, abstract terms presented in *associated* patterns will be less subject to such an effect.

> “If the representations supporting abstract concepts are stored in an associative network, it would be predicted that deep dyslexic semantic error responses to abstract words should be more likely to bear an associative than categorical relationship to the target. Furthermore, if the representations supporting concrete concepts are stored in a categorical framework, then semantic error responses to concrete words should be more likely to have a categorical rather than associative relationship to the target” (Crutch, 2006:92).

This means that Semantic Category Effects should be less visible for abstract words, because they are stored in an *associative* pattern, than concrete words in *semantically related categories*. This associative pattern would presumably be the result of a process like free association where personal meaning is the drive behind the relationship of the words selected and which can perhaps seem rather random compared to the semantic relationships in categories. Because of this apparent randomness we can predict that abstract sets have *less* interference than concrete words when presented in *semantically related category* sets, and so should be retained more easily in comparison.

This study compares semantically related and unrelated category sets, but not associative patterns which might affect abstract words more. If the results show less interference in general for abstract than concrete sets then it may point to the validity of the QDR model and future studies would need to compare semantically and associatively related sets to see how SOA affects each condition. However if the results show that both abstract and concrete sets are affected by SCE then it may point to a need to revise the QDR.

3.4 Summary

In summary, it seems that learning L2 vocabulary is a complicated process which has to take into account many factors, among them: Word Class, semantic Relationship, and level of Concreteness. While several separate studies have shown how interference effects can hinder semantically related words' retention, no study could be found that combined different syntactic words classes and different levels of abstraction to investigate the validity of interference or word class learnability model predictions.

By combining related and unrelated semantic category sets with different types of word class and levels of abstraction hopefully this study will provide a better idea of the relationship between words and any effects which increase or decrease learnability. This understanding could provide educators with a more informed choice about how and what to present to learners for beginners, as well as provide evidence for the validity of models which explain lexical organisation.

4 A Description of the Study

There now follows a description of the study, participants, research questions, design and validity issues.

4.1 Students' Profiles and Class Conditions

The study used a repeated measures two factor design (Hinton, 2004:193-205) with 82 Junior High School first grade Japanese children with an average age of 12 over a period of four weeks. Full ethical considerations were given and consent of the school and parents was obtained for the study (appendix 11.6). The school is private, set in suburban Osaka, and informal sampling of students' parents' occupations suggest that they generally come from a middle class background. All participants were treated during regular class times.

The students are streamed into three classes: 1-1 (n27) 1-2 (n27) 1-3 (n28) according to results of the school entrance exam which means that they form roughly homogenous groups of ability with 1-1 being the "lowest" and 1-3 the "highest". For the English classes given during the treatment each group is split into two, with one half (n13 or n14) taught by a Japanese English teacher and the other by a Native Speaker (NS, the researcher) in adjacent classrooms. After 25 minutes each half changes classrooms and teachers.

This raises the possibility of a systematic confounding effect that either half could be influenced as they either start the treatment with the NS fresh but not "warmed up" for English, or come to the treatment after 25 minutes of English "warmed up" but tired or distracted. However this may have been mitigated to some extent as the halves changed the starting classes (either Japanese or NS teacher) each week.

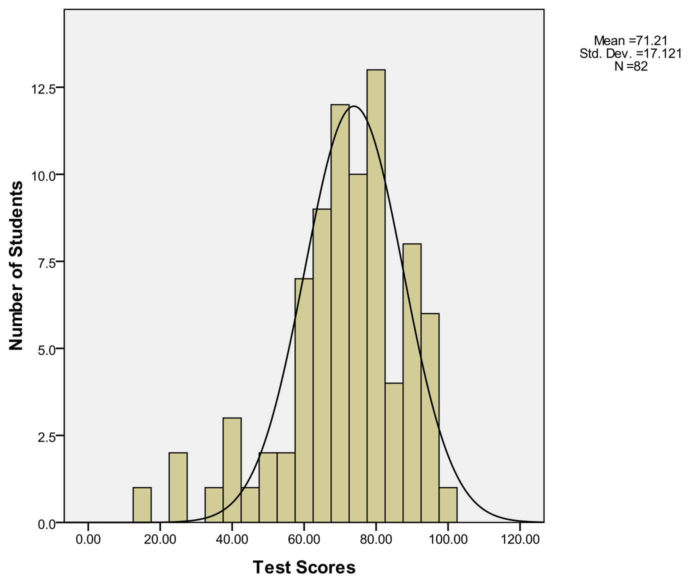


Figure 4-1 Histogram of students' English test scores and normal distribution

The Japanese English teacher's examinations carried out just before the treatment provided the following results (Fig 4.1) which give some idea of the levels of English competence in a standardised test after 6 weeks of formal tuition at Junior High School. These results show a fairly normal distribution of English knowledge and ability on the school test, from which it is assumed that, within the constraints of sample selection (Hinton, 2004: 48-57), this could be a representative sample from the population.

4.2 Research Questions and Predictions

A pilot study was conducted (Taynton, 2009) which showed that the conclusions of other studies seemed to be valid but there was scope for further investigation due to some anomalies in the results, and that the area was worth investigating in more detail.

The study seeks to refine previous research into interference effects and learnability difficulty by combining three factors together, and comparing which types of words and in what arrangement are retained more easily.

The constructs examined are:

- Word Class (WC) – either nouns or verbs
- Concreteness (C) – the level of abstraction or concreteness as defined by ratings from the MRC database
- Relationship (R) – semantically related categories of words; or randomly chosen unrelated sets of words but selected according to WC and C.

The following questions were investigated:

1) What is the effect of grouping words in semantic and unrelated sets on retention of word form and meaning?
2) What are the effects of different levels of Concreteness and Word Class on retention of word form and meaning?
3) Is there an interaction between Relationship, Word Class and Concreteness on learnability of word form and meaning?

Several predictions about the results can be made according to the theory and evidence already discussed. It is predicted that concrete words will be easier than abstract words, nouns will be easier than verbs, and unrelated sets easier than related sets.

However, when arranged according to the sets being studied here some differences emerge. The SCE predicts that semantically related concrete sets will be harder to learn, while the QDR shows that abstract sets presented in semantically related groups will not be affected by SCE. However, what is clear from previous studies is that there are certain factors which influence the ability to learn different types of words.

It is assumed that each factor has a weighting on the level of difficulty. Table 4-1 shows which factors are harder or easier, and this can predict how learnable different sets of words will be.

Hard	Easy
Related	Unrelated
Abstract	Concrete
Verbs	Nouns

Table 4-1 - Prediction of Learnability

For example Related Concrete Nouns (one difficult factor, two easy ones) could be easier or present similar levels of difficulty as Unrelated Abstract Verbs (one easy, two hard factors). SCE states Related sets are harder then Unrelated sets, but Concrete Nouns are easier than Abstract Verbs. It is these other factors which may provide contrary evidence to data from previous studies, and show how other factors influence learning. Table 4-2 shows all of the different combinations and their acronyms being tested in this study.

4.3 Variables and Test Design

The three independent variables are semantic Relationship (related or unrelated), Word Class (noun or verb) and Concreteness (concrete or abstract). The dependent variable is the retention of form and meaning between L1-L2 over time.

There were eight conditions being observed:

UCV - Unrelated Concrete Verbs	**UCN** – Unrelated Concrete Nouns	**UAV** – Unrelated Abstract Verbs	**UAN** – Unrelated Abstract Nouns
RCV – Related Concrete Verbs	**RCN** – Related Concrete Nouns	**RAV** – Related Abstract Verbs	**RAN** – Related Abstract Nouns

Table 4-2 Abbreviations for Each Condition

The Semantic Category Effect (SCE) is a result of a spread of activation in conceptual memory. Vocabulary storage can be conceived of as similar to a network of lexical items stored and joined together by connections like a spider's web. Each word, or node, in the network, which is connected to the conceptual store of knowledge, has developed semantic connections to other nodes (by frames of metaknowledge in Hall's model, 1992).

The effect of activation is that when a node is accessed a form of energy passes to nearby related nodes just as a pebble causes ripples in a pond. This spread of energy makes the nearby nodes active and, perhaps, more readily brought to conscious attention. The problem is when learning or retrieving vocabulary, forming connections, or getting meanings and form at the L1 or L2 to conceptual knowledge level, cognitive processes like recall and memorisation become confounded by other related activated nodes. This causes weak or unclear connections to be formed, or slower and more incorrect responses to meanings and form (Hakki Erten and Tekin, 2008; Finkbeiner and Nicol, 2004; Comesana et al., 2009).

To test SCE, groups of semantically related by category or unrelated by semantic category words were carefully chosen in English according to the assumed level of knowledge and understanding of the students (the full list is in appendix 11.2). They were then translated into Japanese, and the translator then produced three more choices, two of which were closely semantically related, and one which was opposite to the primary meaning. It was assumed that the opposite meaning has just a strong connection to the semantically related similar meanings and was therefore just as confounding as the two related foils. For example the primary word being taught is Motorcycle. Two semantically related words are scooter and trike, while the opposite is car.

The second research question investigates the effect of Word Class and Concreteness on retainability; that is, whether nouns or verbs, which are either more or less concrete as measured by the MRC psycholinguistic database (Wilson, 1988), are easier or harder to learn. This condition is unrelated to the SCE as the words are measured according to different criteria – the level of Concreteness and its Word Class.

This means comparisons can be made between abstract nouns and concrete nouns, or concrete verbs and abstract nouns to find which have been retained better. There are several possible permutations for analysis which will hopefully give an idea of variances between different word classes, concreteness and learnability.

While the first two questions look at essentially the same type of thing, the third question looks to see if there are any interactions between Word Class, Concreteness and Relationship. For example, are semantically related concrete verbs more easily learnt than unrelated, (i.e. randomly chosen), abstract nouns.

4.4 Study Design

A repeated measures factorial design, as opposed to random assignment to control and treatment groups, was the only practical solution to the constraints of the scheduling situation. Although this may threaten generalisability, it offers the benefit that individual differences are reduced (Hinton, 2004:74-80).

A receptive test was designed (Nation, 2001:24-32) using L1-L2 translation word pairs and each student group was treated in as similar manner as possible to avoid confounding variables. The teaching took place in a regular classroom using a projector and a PowerPoint presentation, and students sat individually to prevent copying. An example of the presentations slides can be seen in the appendix 11.1.

All students had visual access to the screen and none reported an inability to read or see the words. Observation of the class during testing found no substantiated instances of copying or cheating.

The schedule of the teaching of different conditions could not use configurations such as Latin square (Pezzullo, 2009) to ensure that it was balanced for carryover effects or influences such as test fatigue or time of day so each condition was scheduled randomly.

4.5 Teaching and testing

The teaching and testing of each set of words was conducted in three stages.

4.5.1 Pretest

A pretest “*before”* immediately before the teaching of the words. This was to check for previous knowledge of the words and used as part of the ANOVA to test for teaching effects by comparing the results of each test over time.

4.5.2 Teaching and first post test

The teaching of each word and its L1 meaning had two parts. Each word was projected first in English on a big screen in a darkened room. Students looked at it and repeated the word out loud three times by repeating the researcher’s speech. After 10 seconds a second slide, which had the English with Japanese translation, appeared and the students repeated the English word one more time.

In total the English word was visible for 20 seconds, the Japanese translation for 10 seconds, and the word had been practiced verbally four times in English. There were only a few instances where there was a distraction or not enough time to complete all the recitations. When students had viewed all the slides in a group they took an immediate post test *"after 1"*. There was no time limit and timings were not kept.

4.5.3 Second post test

After seven days a delayed post test, *"after 2"*, was given to see how much had been retained. This period, also used in the Comesana et al. study (2009:31), was presumed to give enough time for any effects of the independent variables to take place.

An example of the three tests can be seen in appendix 11.3. To prevent memorisation affecting results each word and each choice order were randomly mixed up so that they appeared in different positions on every *before, after 1* and *after 2* answer sheets.

4.6 Tests and teaching schedule

The delayed post test and the teaching of the next set of words were in the same lesson period. The *after 2* delayed test was given before the next set of words were taught. Mindful of the possibility that the SOA can last for several minutes a distractor activity, lasting three to four minutes, was held between *after 2* and teaching the next set of words.

This was a short video clip of a popular British TV talent show, and students were asked to give their opinion about it by answering yes or no to two questions. This distractor activity was presumed to give time for any carry over effects to diminish so that the next set of words would not be influenced by the immediately given test.

4.7 Participation

Due to the length of the study and various institutional scheduling demands it was not possible for complete sets of data to be obtained with all students. Participation ranges from 52 students to 81. However, the number of students in every test is comfortably above the 30 recommended for statistical analysis (Hinton, 2004:55). But this variance may contribute some imbalances to the results which is a factor that needs to be considered during analysis and discussion.

4.8 Validity and Reliability

Can it be claimed that the study is actually measuring the constructs of SOA and word class/concreteness on learnability? The three previous studies on SOA for L2 learners (Hakki Erten and Tekin, 2008; Finkbeiner and Nicol 2004; and Comesana et al., 2009) all used L1-L2 word translation pairs to measure the effects of interference so it is assumed that it is a valid method for such measurements. There follows a brief discussion of the construction of the word sets which were made to measure the variables. The complete lists can be seen in appendix 11.2.

Groups of nine words were chosen for either related or unrelated semantic sets. Semantically related categories such as animals, personality characteristics and ways of writing were constructed. Unrelated sets were based on picking random words, but constrained by the criteria of Word Class and Concreteness. These criteria guided the selection of words as either verbs or nouns with varying degrees of Concreteness as defined, where possible, by the MRC Psycholinguistic Database (Wilson, 1988).

According to this database, and various other studies measuring participants' perceptions of abstractness and concreteness (for example Altarriba et al., 1999), scores toward 700 are more concrete and those towards 100 are more abstract. Chosen words were checked for these scores on the database and inappropriate words were discarded. As much as possible words towards abstract or concrete were chosen, with each group matched as closely as possible together. For example the scores for two sets are as follows:

Related Concrete Verbs (RCV) and MRC Concreteness rating		**Related Abstract Nouns (RAN) and MRC Concreteness rating**	
Write	446	Intelligence	275
Draw	442	Courage	277
Paint	577	Cowardice	263
Sketch	535	Kindness	261
Trace	371	Loyalty	261
Spray	514	Patience	266
Note	525	Pride	270
		Honesty	278

Table 4-3 Examples of Concreteness Ratings from the MRC Database

Although none of the words were at extreme ends of the scale it was felt that there was enough distance to justify its classification as abstract or concrete, with the halfway score of 350 being the deciding factor. Crutch uses a dividing line of 550 but notes that other studies have used arbitrary dividing lines, with 400 being used in several studies. As Crutch says:

> "Concreteness constitutes a continuous variable. As a consequence, the method for classifying concepts into 'abstract' and 'concrete' classes on the basis of normative data rather than qualitative definition or intuition is not clear-cut " (2006:93).

The words were then translated into Japanese and then two semantically similar and one opposite words were chosen giving a set of four choices for each answer. This was to provide equally difficult conditions for both related and unrelated sets as semantically related networks would be activated for both conditions. An example of an answer sheet can be seen in appendix 11.3.

5 Analysis of Results

After all the tests had been collected and marked by the researcher they were entered into SPSS for analysis. Random checks were made to test for accuracy in marking but no errors were found. However it is possible that some mistakes were made but there were not enough resources for double marking the entire collection of sheets.

A 2(Relationship: related, unrelated) x2 (Concreteness: concrete, abstract) x2 (Word Class: noun, verb) ANOVA (appendix 11.4) showed that individually Relationship, Concreteness and Word Class were all statistically significant factors in the learning of vocabulary. When cross referenced with each other, Concreteness *Word Class, and Relationship*Concreteness*Word Class were significant factors, but Relationship*Concreteness and Relationship*Word Class were not.

Repeated measures one way ANOVA tests were run on all the *before*, *after 1* and *after 2* results (section 4.5) to check if the teachings had had an effect on knowledge and the results were not due to random chance (Motulsky, 1999). This showed that all the students were affected by the teaching and that there was a very unlikely chance that the results were due to random chance (p=0 in all cases). This shows that there was an effect of the teaching on the mean scores of students and that they had retained something from the teaching therefore any results would probably have been influenced by SCE and/or Word Class and Concreteness.

Multiple "*t* tests" were performed on various combinations and these are presented in table 5-1 (page 23, full results in appendix 11.5). Statistically significant results are sorted according to level of significance. SPSS results for these tests are two tailed but for these study a one tailed result is needed (Hinton, 2004:42-44) so the figures were converted according to the rule presented by Elvers:

> "6. Determine if we can reject the null hypothesis or not. The decision rule is: if the one-tailed critical t value is less than the observed t AND the means are in the right order, then we can reject H_0." (Elvers, 2009).

Therefore although at a two tailed level of significance some results are greater than α0.05, according to the results of the procedure above for one tailed significance 12 pairs were statistically significantly different at α0.05.

	Pair	Sig. (2-tailed)	LF result	LFR	LFR2
1	UCN-RAN	.000	2,0	3,0	2,0
2	UCN-RAV	.001	2,0	3,0	2,0
3	UCN-UCV	.003	2,0	3,1	2,0
4	RCN-RAN	.009	2,0	2,0	2,0
5	UCN-RCV	.010	2,0	3,0	2,0
6	UAV-RAN	.016	0,0	1,0	2,0
7	RCN-RAV	.033	2,0	2,0	2,0
8	RCN-UCV	.048	2,0	2,1	2,0
9	UAN-RAN	.058*	0,0	1,0	2,0
10	UAV-RAV	.066*	0,0	1,0	2,0
11	RCN-RCV	.072*	2,0	2,0	2,0
12	UAV-RCV	.078*	0,0	1,0	2,0
13	UCN-UAN	.122	2,0	3,1	2,2
14	UCN-RCN	.312	2,2	3,2	2,2
15	RCV-RAN	.329	0,0	0,0	0,0
16	UAN-RAV	.401	0,0	1,0	2,0
17	RAV-UCV	.432	0,0	0,0	0,0
18	UAV-UCV	.436	0,0	1,1	2,0
19	UAN-RCV	.451	0,0	1,0	2,0
20	RCV-UCV	.503	0,0	0,0	0,0
21	UAN-UAV	.504	0,0	1,1	2,2
22	UAV-RCN	.672	0,2	1,2	2,2
23	RCN-UAN	.714	2,0	2,1	2,2
24	UCN-UAV	.716	2,0	3,1	2,2
25	RCV-RAV	.749	0,0	0,0	0,0
26	RAN-UCV	.857	0,0	0,1	0,0
27	UAN-UCV	.883	0,0	1,1	2,0
28	RAN-RAV	.885	0,0	0,0	0,0

Table 5-1 *t* test results on pairs of conditions, and LF scores.

Red highlights incorrect predictions. * significant at 1 tail

The study investigated three things:

1) What is the effect of grouping words in semantically related and unrelated sets on retention of word form and meaning?
2) What are the effects of different levels of Concreteness and Word Class on retention of word form and meaning?
3) Is there an interaction between Relationship, Word Class and Concreteness on learnability of word form and meaning?

The data will now be discussed with a breakdown of the statistical results, followed by a discussion of the implications.

5.1 Review of predictions

This study investigated the role of three factors on the learnability of words when placed in different combinations of those factors. A brief review of these factors follows.

Concreteness (C) – this is the level of imaginability – real world items tend to be easier to remember than abstract concepts. Further, each type may be stored in a different way according to the QDR: abstract words in an associative network and concrete in a category network. (Crutch 2006:91; section 3.3).

Relationship (R) – Previous research (section 3.2) indicates that semantically related sets of words (which have not been controlled for Word Class or Concreteness) are harder to learn than unrelated sets which are chosen randomly. The effect of manipulating Word Class and Concreteness is unknown, but according to QDR abstract words should not be affected as much by semantically related groupings.

Word Class (WC) – It seems that different syntactic classes of words are stored in different areas of the brain (section 3.3), and each class seems to have a different level of learnability with nouns being easier to learn than verbs.

Each of these factors has a effect on the learnability of a word. Concreteness and Word Class combine to make a word more or less memorable, while putting words together in semantically related should cause SCE to hinder learning while or unrelated groups should be free from such effects. However, according to the QDR, abstract words should not be so affected when presented in semantically related groups as they are possibly stored in a qualitatively different way – *associatively* – and so less subject to *semantically* related category effects.

The ANOVA (appendix 11.4) showed that two combinations of factors were significant in influencing learnability, and two were not. One surprising finding was that Relationship was not an influencing factor when combined with Word Class or Concreteness. However Relationship was a factor when combined with both Word Class and Relationship.

Some of these factors when combined could affect the learnability in a way that is different if one factor was considered on its own. For example Concrete Nouns in an Unrelated set should be the most learnable: U+C+N. However one factor may have a negative effect on the set: Related Concrete Nouns should be less learnable due to being in a related set so perhaps the equation would be R-(C+N). See table 4-1 for reference.

The results of the *t* test on pairs (table 5-1, appendix 11.5) provides an idea of the weighting and combination of factors by showing which sets are more or less learnable and which factors seem to be influencing these results. These pairs will be used to see how these factors combine to determine learnability.

If each factor has a weighting of 1 then combinations of factors can be calculated to see how the ANOVA was derived with reference to the data, and how this can be related to theory. If each set of three factors is broken down into pair combinations then a formula, named the Learnability Formula (LF), can show how each pair interacts.

5.2 Analysis of results according to the Learnability Formula

5.2.1 Example 1 Pair 1 UCN – RAN

The Learnability Formula (LF) is based on giving each combination of factors a weighting (see table 4-1 page 17) and then seeing how the factors in each pair either combines to make it more or less learnable. To illustrate this, examples from table 5-1 will be discussed. Firstly, UCN and RAN (pair 1, table 5-1).

Unrelated Concrete Nouns (UCN):

UCN ((U-C)=0 (C+N)=2 (U-N)=0) = 2

Relationship and Concreteness have no interaction so U-C (1-1) = 0

However Concreteness and Word Class do interact, and both concrete words and nouns are easy to learn so C+N (1+1) = 2

Again, Relationship and Word Class have no interaction so U-WC=0 (1-1) = 0. So UCN has a LF score of 2.

Related Abstract Nouns (RAN):

In the RAN set according to the above formula the total is 0:

RAN ((R-A)=0 (A-N)=0 (R-N)=0) =0

R-A: Relationship and Concreteness have no interaction = 0

A-N: Concreteness and Word Class do interact, but abstract words are harder to learn so any benefit of easier nouns is cancelled out by their abstract nature, thus A-N = 0

R-N: Relationship and Word Class have no interaction = 0. The total score for RAN = 0.

The total result of the LF for UCN (2) and RAN (0) = LF:2,0 which means that UCN is easier to learn than RAN.

A paired samples *t* test revealed a statistically reliable difference between the mean of UCN and RAN: $t(81) = 3.65$, $p = .000$, $\alpha = .05$. That is, UCN is significantly easier to learn than RAN due to Word Class (noun) and Concreteness (concrete). This shows that the LF has predicted correctly.

Table 5-1, column LF, shows how the Learnability Formula correlates to the results of the *t* tests. The LF predicated 71% of results correctly so clearly there are some problems with the LF and further analysis of the anomalies follows.

There were nine pairs that were not correctly predicted by the LF: Pairs 6, 9, 10, 12, 13, 22, 23, 24.

One common factor that each of these pairs have is that when the LF and the results of the *t* test do not match, meaning that the pairs are significantly different or not, the other factor that seems to influence the learnability is Relationship.

Three examples show how this factor seems to play a part, which is in line with the ANOVA analysis that shows R*C*WC together is a significant combination.

5.2.2 Example 1 Pair 13 UCN - UAN

UCN – UAN (table 5-1 pair 13) are statistically not different in terms of learnability, $t(67)=1.56$, $p=0.122$, but have a LF of 2,0 which predicts that UCN is more learnable. In this case the power of an unrelated noun in the UAN set could reduce the effect that Concreteness has on learnability. Abstract nouns become as easy to learn as concrete nouns when presented in unrelated pairs. Compare pair 1 UCN – RAN, $t(81)=3.65$, $p=.000$, where there is a big difference in learnability, and this shows Relationship plays a part in learnability.

5.2.3 Example 2 Pair 23 RCN - UAN

This is another anomaly and should be significantly different according the LF (2,0) but, according to the *t* test: $t(66)=0.368$, $p=.714$, is not. The factor that is working to make these pairs similar seems to be Relationship. A related set of concrete nouns is equal to an unrelated abstract noun group so the deciding factor, as in 9.2.2, could be Relationship. An easy set of concrete nouns (RCN) is disturbed by the SCE, while a difficult set of abstract nouns is helped by being in an unrelated set so less hindered by SCE. In these two cases it seems that Relationship can offer a clue to why these pairs are similar.

5.3 Reformulation of the LF

From the above analysis it seems that Relationship influences the results as predicted by SCE. Therefore a reformulation of the LF is needed and can be shown as:

When a set contains an unrelated factor, add 1 to the total LF.

For example if this rule is applied to the anomalies noted above then the results are:

Pair 6: UAV – RAN

(U-A)=0 (A-V)=0 (U-V)=0 U=+1 = LF=1

(R-A)=0 (A-N)=0 (R-A)=0 R=0 = LF=0

The results of this reformulation can be seen in table 5-1 column LFR

This reformulation does not however increase the reliability of the LF, which has the same accuracy of 72% due to other pairs becoming unbalanced (e.g. 16, 19, 26).

It seems that LFR cannot account for these discrepancies and explain the triple combination of factors (Relationship*Concreteness*Word Class) shown in the ANOVA.

5.3.1 A special case – Unrelated Abstract sets

One of the most obvious patterns that can be seen in the results is that UAx (x = nouns or verbs) seems to have a special power. It seems that unrelated abstract sets increase learnability when compared to either related or unrelated sets of other factors. According to the QDR, abstract words should not be as affected by either semantically related or unrelated groups, but these results seem to indicate that when compared to RAx sets, UAx is more powerful. However this seems to be an exception to the general rule that the LF gives with regards to weightings and combinations.

If the original LF (5.2) is adjusted (to become LFR2) to include the rule: *If a set contains a combination of UA, then add 2 points to that set* the LF produces results, seen in table 5-1 column LFR2, with an accuracy of 85%.

One explanation for this may be in the data. The UAx sets generally had fewer participants than other sets which may have caused a statistical error. One way to test the accuracy of this is to redo the studies with random samples and the same number of participants in each group.

Another explanation is that Relationship may have some impact on Concreteness despite it being rejected in the ANOVA. In the significantly different results (table 5-1)

UAx sets were more learnable than RAx sets in all three anomalous cases (pairs 6,9,10). This indicates perhaps that related groups of abstract words are affected somewhat by a SCE even though according to the QDR (section 3.3) they are stored in an associative network which is not meant to be so affected by being grouped together in semantic categories.

If SCE occurred in related abstract groups it would make them harder to learn just as related concrete groups are more difficult. For example RAx groups should not be so hindered by SCE so they should be more learnable when compared to RCx groups. However no significant cases of this pattern can be seen. In fact, pair 7 in table 5-1 shows the reverse – RCN is more learnable than RAV, which is predicted by LFR2 (Concreteness and Word Class are stronger predictors than Relationship).

By looking at the reverse pattern to see if UAx has an advantage over UCx there are no such significant pairs. A comparison of RAx and UCx, and UAx and RCx gives only one result: pair 12 UAV – RCV. This indicates that there seems to be some SCE effect in related abstract groups, and less so in unrelated groups which mimics the patterns of concrete groups.

UCx is more learnable than RAx or UAx in four cases (pairs 1, 2, 3, 5) where it would be expected to be similar due to no SCE in RAx or UAx. This shows that there does not seem to be any advantage for abstract words in semantically related or unrelated groups which is different to the predictions of the QDR.

Overall it seems that LFR2, which accounts for a SCE in abstract groups, can account for some of these anomalies and points to a theoretical weakness in the QDR. However, a true comparison cannot be made as no sets were made to test the QDR theory of interference effects in associative groups of abstract words. Future studies would need to compare associative and semantically related abstract and concrete words to get a clearer picture of how these two variables interact.

5.3.2 Remaining Anomalies

After modification LFR2 can predict results at 85%. However four pairs remain that it does not correctly predict: Pairs 16, 18, 19 and 27.

5.3.2.1 Pairs 16: UAN – RAV, 18: UAV – UCV, 19: UAN – RCN, and 27: UAN - UCV

LFR2 results for these pairs are 2,0 which predicts a significant difference in learnability but the *t* test for each shows quite conclusively no significance. The common factor in all of these pairs is UAx. Theory would predict that unrelated abstract sets would be more learnable and indeed LF correctly predicted the results as being equally difficult to learn. This result is somewhat puzzling. It may actually be offering evidence that abstract sets are, as predicted by QDR, not affected by Relationship. This is counter to the proposal of LFR2 that they are subject to SCE. Clearly there is a grey area around this which requires more research.

5.3.3 SCE Not Found in Likeliest Pair

Another very interesting result, which seems to refute the results of previous SCE studies, is pair 14 UCN – RCN. Essentially this pair shows the results of related and unrelated sets similar to those used by Hakki Erten and Tekin (2008) and Finkbeiner and Nicol (2003). These studies seemed to show a clear SCE but here the results, $t(80)=1.018$, $p=.312$, show no significant difference in learnability: LFR2=2,2 due to both sets containing concrete nouns.

This is a very puzzling finding which may be explained as follows: This study measured more constructs than the others (section 3.2), so while it does seem that SCE could not be replicated they do have some effect on learnability when combined with other factors. This is because it seems that the factors of Concreteness and Word Class in combination play a more important role in learnability, with Relationship having an effect in some combinations of words. Previous SCE studies did not consider such factors and so while valid for the measures they took, they are not comparable here. Also, as previously mentioned, there were some potential problems with the previous studies' methodologies which may have influenced the results.

5.4 What is Missing from the Significant Results

One group that is missing from the results that are easier to learn is RAx. However UAx appear on the more learnable side four times. The QDR proposes that abstract words are not so hindered by semantic categories, but there is some evidence that abstract words are boosted by being in unrelated sets, while related abstract sets are harder to learn.

It seems that UCx and UAx sets are almost comparable in terms of learnability when compared to RAx sets. For example pairs 1, 2, 3, 5 are UCN which LFR2 predicts will be easy to learn. However, pairs 6, 9, 10 and 12 are UAx sets which, like UCN, is more learnable when compared to both RAx and RCx.

But, when UAx and UCx pairs are compared (pairs 13, 18, 24, 27) then no significant differences arise. Perhaps the boost given to abstract words in an unrelated set equalizes the advantage given to concrete words.

The main point to be made from this finding is that it seems abstract words are indeed affected by SCE when learnt in a semantically related set, and that an unrelated set of abstract words is easier to learn. This is similar to the way concrete words behave and so the proposition of the QDR that abstract words are stored differently seems to be contradicted and points to the need for an alternative hypothesis of how language is stored and related within a network. This is explored in more depth in section 6.1.

5.5 Summary

In summary, the models the LF concept is based on, such as SCE and QDR, do not adequately explain the results of this study. In general LF can only predict with 72% accuracy, and if only significant results are included then it predicts at only 66%. However, there are some results that do follow previous models expectations and so the task is to find out what adjustments need to be made to previous models to account for these discrepancies.

It seems that there is some grey area around the QDR proposal that abstract words are not as affected when put in semantically related categories. This study shows that when SCE in abstract related sets occur the LFR2 is a more accurate predictor. However, there are also some anomalies which seem to support the QDR model. In short this study has produced some mixed findings with regards to abstract sets, but in general supports previous research into SCE and learnability of Word Class and Concreteness.

As this study is possibly the first time that such diverse factors have been brought together to be tested in combination, there will most likely be some gaps where the models do not adequately overlap and so some revision of their assumptions may resolve these issues.

6 Discussion

The study showed that models, like QDR and SCE, when combined in the form of a Learnability Formula (LF) have some success in predicting results of paired *t* tests (72% for all results, or 66% for only significant results). This means that the premises upon which the models are founded must have some truth. However in this study several results arose which could not be adequately explained by the LFR2 which may point to some of the theories being in need of adjustment.

The three research questions can now be discussed according to the results.

6.1 Research Question 1

What is the effect of grouping words in semantically related and unrelated sets on retention of word form and meaning?

Previous studies into SCE (section 3.2) showed that grouping words of (mostly) concrete nouns in semantically related or unrelated categories initiated an interference effect that caused semantically related groups to be more difficult to recall in a later test. This idea is given theoretical support by the QDR which states that abstract and concrete nouns are stored differently, with concrete words in a categorical framework and abstract ones in an associative network, therefore concrete words should be affected more by SCE and abstract semantically related sets less.

The ANOVA analysis from this study (appendix 11.4) showed that Relationship*Word Class and Relationship*Concreteness had no significant relationship. In general this was found to be true when the paired *t* test results were examined (table 5-1, page 23); however there was one anomaly that could be partially explained if Relationship did affect Concreteness.

A Learnability Formula (LF) model (section 5.2) was proposed that gave a formula for predicting the results according to the theories presented from other research. Its accuracy, initially 72%, increased to 85% once a modification was made that took account of a relationship between Relationship*Concreteness (LFR2, table 5-1).

However, one surprising result seemed to contradict the other studies into SCE by not showing a significant difference between semantically unrelated and related sets of concrete nouns (section 5.3.3). However the LFR2 model correctly predicted the result that no difference would be found.

In general it was found that Relationship had an effect on learnability and did seem to fit the SCE model, but perhaps in a different way when other factors like WC or C were introduced. By itself it was not able to overturn or hinder a set's learnability as the other factors of Concreteness and Word Class were found to be a more powerful force affecting learnability. This is supported by other SCE studies (section 3.2) that found semantic category (i.e. Relationship) had a clear interference effect on learning. But, when combined with other variables, Relationship is not the most powerful factor.

When this data is analysed in relation to the results predicted by the QDR some differences emerge which can be accounted for in two ways. The QDR (Crutch, 2006; Crutch et. al., 2006; Duñabeitia et. al. 2009; section 3.3), supposes that abstract words are represented in a semantically *associative* pattern, while concrete words are organized according to *categorical* semantic similarity. Therefore semantically related sets of words should not activate interference effects for abstract words as much as, or at all, compared to semantically related concrete sets.

In general this was the case, however, there were several sets (table 5-1 pairs: 6, 9, 10) that compared semantically related and unrelated abstract sets which showed a significant difference in learnability in favour of the unrelated condition where there probably should not have been.

6.1.1 A problem with related abstract sets and QDR

This data shows that interference effects among semantically related abstract sets of words (RAx) can occur, which seems contrary to the assertion that semantic category relationship is not an organizing principle, and therefore not so affected by SCE, for abstract word networks.

There are two possible reasons for this discrepancy. The first is a threat to validity: semantic sets in this study were not subject to extensive testing to ensure that they were actually part of semantic sets, instead they were chosen a priori by the researcher with some reference to previous research. By misjudging this a type 1 error could be made.

Secondly, the QDR does not consider how lexical network access might be dynamic depending on the stimulus provided. The Duñabeitia et. al. study (2009), which they claim is the first to confirm the QDR in healthy patients, offered concrete pictures for abstract associations, for example a picture of a nose, which was found to be associated with the abstract word *smell*. This study seems to show that Semantic Category Effects are present in semantically related categories of abstract words.

The difference between the Duñabeitia et al. study and this is that of focus of research and presentation. They focused on trying to find evidence for the QDR, and presented single words with pictures in sequence and measured saccadic eye movements to ascertain which pictures were fixated on for the longest period of time when hearing a keyword. This study presented nine words in semantically related or unrelated sets translated into L2 to see which groups were more easily remembered with any difference being attributed to SCE.

According to QDR abstract and concrete words may be organized differently, but it seems they are still affected when tested for SCE which could mean that the way that words are activated determines their relationship to each other. One feature of language that may help explain this is that words can be seen in two ways: either in a semantic category (e.g. colours, cognitive functions, or a subspecies of ant) or by generating associations through activities like free association (green -> grass -> cow -> MacDonalds->health->happiness…). The important thing to note here is that both semantic categories and free associations might contain words that differ in Concreteness and Word Class, meaning that Concreteness is not the main factor which organises vocabulary.

It seems reasonable to conclude that a level of concreteness is one point on a continuum, not an either/or, (cf Crutch, 2006:93) so the point made by the QDR that abstract or concrete words are organised in *qualitatively* different *networks* is perhaps counterintuitive. But, as Duñabeitia et al., (2009:285) hedge, perhaps there is just *one network* with different connection strengths. If Concreteness is not the prime organizer of lexis then another factor must be at work.

6.1.2 Semantically Related and Associative Networks – a Model

The data that is missing that would provide more evidence to support the QDR is research into the way associative networks behave in relation to interference effects. Figure 6-1 (page 34) shows the gaps in the knowledge of the QDR. While data from this and other studies can be found to more or less confirm the process on the right side of the model, no research could be found that shed light on the way associative networks behaved.

Such research would have to investigate how words behave in associatively connected groups. The difference can be seen in table 6-1 and 6-2 where the choices are different: either associative or semantic relationship.

Target word	Associative words			
classify	順序を決 order	定する censor	検閲する determine	分類する classify

Table 6-1 Example of Related Abstract Verbs with **associative** words

Target word	Semantically related words			
classify	分ける divide	仕切る separates	統一する unified	分類する classify

Table 6-2 Example of **semantically** Related Abstract Verb set from this study

This unknown means that it is possible that associated groupings may not be affected by SCE. To account for this the Network Relationship Model (NRM) in figure 6-1 shows at what point the difference may occur.

6.1.3 Explanation of the Network Relationship Model

Starting at the bottom, learning is affected directly by the mode of presentation. For groups of words in a semantic category the network relationship will be categorical and the right path of processing will be taken, which is subject to SCE.

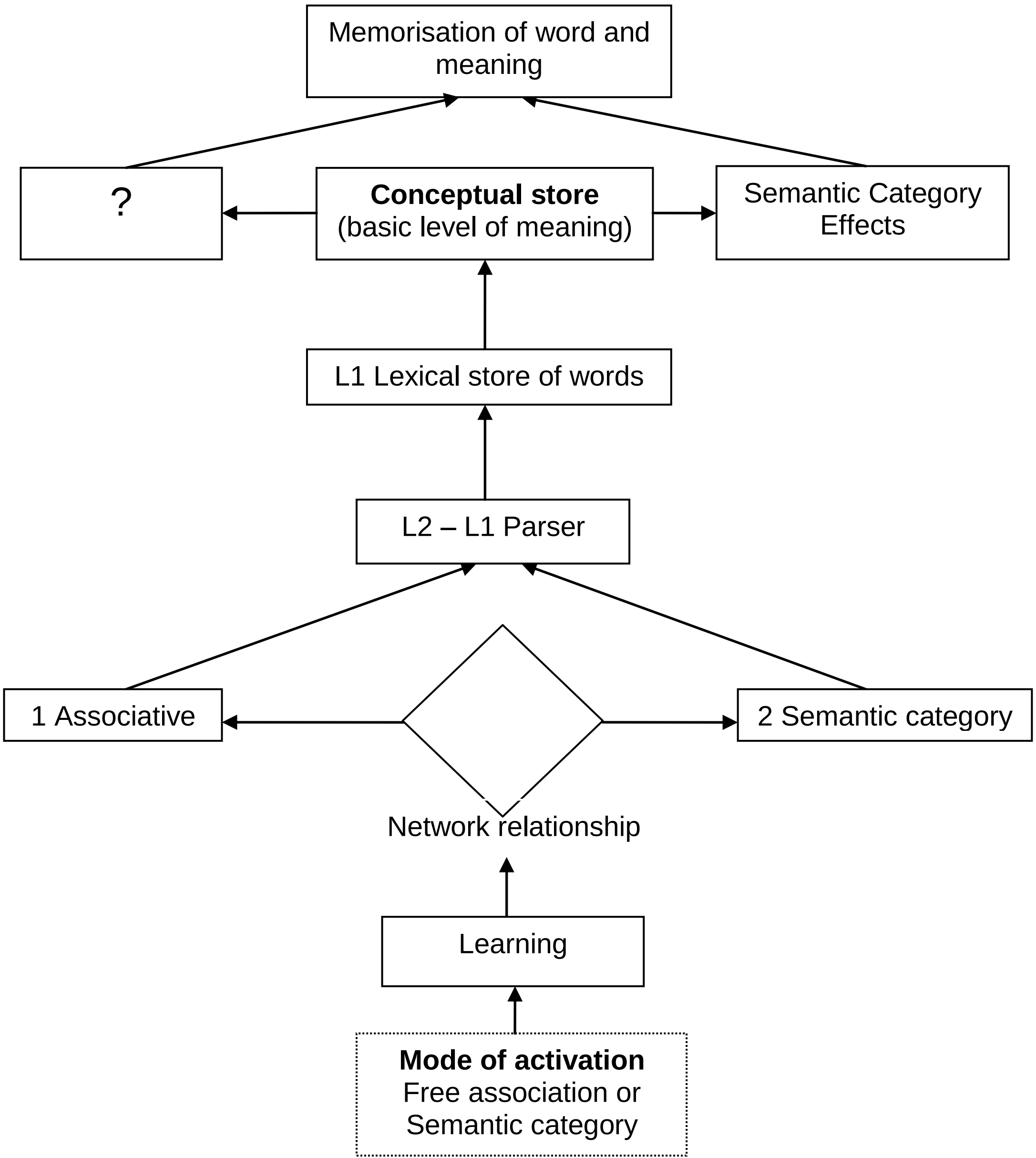


Figure 6-1 The "Network Relationship Model" of how input format can affect lexical processing

However for sets of words grouped associatively the left path will be taken, and it is here that there is no data to confirm the hypothesis that associated groups of words will not be affected in the same way as categorical groups. This is represented by a "?".

Interference effects occur after the negotiation of meaning in L1 at the conceptual level which means that SCE occurs in L1, for beginners at least. Students with higher levels of vocabulary would have a larger interlanguage and possible this may in itself become an independent network which is subject to SCE directly instead of through an intermediary L1-L2 translation process (cf section 3.1).

The L1-L2 negotiation of meaning process is conducted by a parser (such as Halls Mental Lexicon Model or the HRM – section 3.1). This gets the L1 translation and then delivers the match of concept and word in L1 and L2 from the conceptual store. This is then affected by SCE after which it is memorised.

While this model proposes that most of the processes used during learning L2 are the same (the parser and L2-L1 conceptual store look up), the way the words are accessed are qualitatively different according to an associative or semantic pattern. This is indicated by the left or right paths: boxes 1, Associative, and 2 Semantic Category. At this point *different types of connections* are initiated when accessing the lexical store although the same *processes* to find meaning are still used (e.g. Hall's TE frames – section 3.1). In other words, these different types of connections are activated according to the mode of input and affect the way the lexical store is accessed.

The results of this study show that related abstract words, which should be stored more in an *associative* array, and thus not so affected by semantic categorical Relationship effects, are actually interfered with when being learnt in semantic categories. The Relationship of words determines the interference by SCE, with related groups being interfered with more than unrelated groups regardless of abstraction.

While it can be said that there is some evidence to confirm theories like SCE and QDR, there is a hole left by lack of data into associative relationships. In light of this it is interesting to speculate what may be happening on the left side and this may have implications for vocabulary learning.

If the lexical store is visualised as a three dimensional spider's web, with words being individual nodes, then an extra dimension can explain the problem in the QDR that there is a qualitative difference between concrete and abstract networks which the data here seems to partially contradict.

If a semantic category in the lexical store, which contains all the items in that category both abstract and concrete, is thought of as a concentric ring in the web then Spread of Activation effects occur due to the close proximity of nodes in the connected circle (Crutch, 2006:91-92). An associative relationship is quite different in nature – instead of a concentric ring, associations travel *vertically* and horizontally in the web accessing different nodes in different categories/rings in a much more dynamic and *unrelated* way. This pattern mimics the unrelated condition found in this study which in general tended to be less affected by SCE perhaps due to larger distances between nodes.

Figure 6-2 (page 37) shows how this relationship works. The blue concentric circles are how the words (represented as blue dots) are stored in a semantic category – their relationship is close and accessed along the same axis – horizontal. Many different layers of categories store the words in those categories without any direct *semantic* connections.

Associative relationships are vertical, represented as green connecting lines, as well as horizontal. They can access nodes from within the same category and also jump vertically to other nodes in different categories. Their relationship is qualitatively different. It is an associative pattern which depends on a participant's experience. This network also arises regardless of the level of concreteness – that is abstraction is not a deciding factor in choice of the next node.

As associations can travel up and down as well as left and right, the gaps between nodes is larger which may be big enough for Spread of Activation effects not to affect learning. In this way an associative network could have a built in protection against interference effects which makes learning more efficient, just as learning unrelated groups of words seem to be.

The qualitative difference mentioned in the QDR could be just a matter of direction of travel within the network. Different rings of abstract and concrete nodes are accessed dynamically during association, but primarily organized according to semantic category, not Concreteness. Therefore there are not separate networks but one giant three dimensional network accessed in different ways. This access is switched on at the time when vocabulary is presented, either box 1 or 2 in figure 6-1.

One other issue which relates to this concept and which could influence learnability is the evidence presented in section 3.3 that different word classes are stored in different areas of the brain. If such physical neural networks exist then the actual distance between nodes becomes a physical reality and the Spread of Activation becomes potentially vastly weaker. These two factors could have implications for how vocabulary is taught, and these issues are discussed further in section 8.

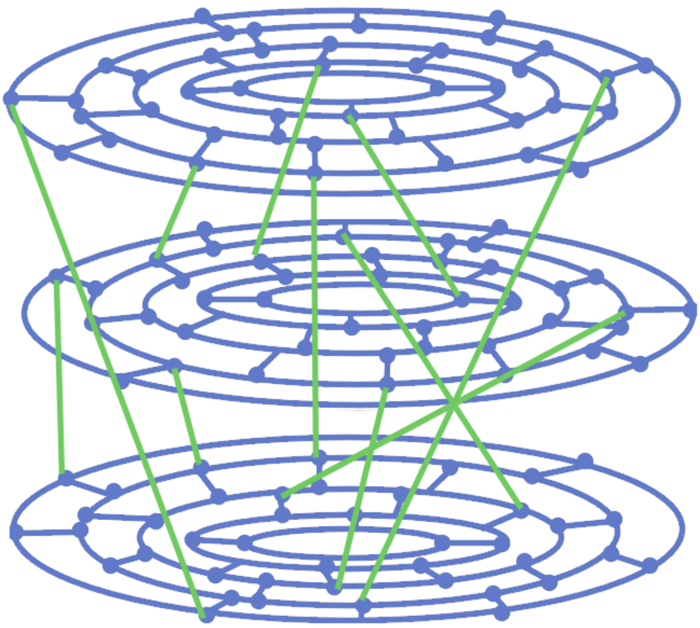


Figure 6-2 Representation of semantic categories and associative connections in lexical memory

6.1.4 Summary of Research Question One

It was found that Semantic Category Effects that hinder learning by confusing memorisation of a word by a process of Spread of Activation were found to have an effect in this study. This supports the conclusions from previous SCE studies. However it was found this factor, Relationship, was not necessarily the most powerful predictor of interference of learning.

Furthermore, these results point to a need for extra research in the area that the QDR relates to. While this study used only semantically associated groups of words, future studies need to test to see if, and to what extent, abstract and concrete sets of words are affected when grouped according to associative patterns compared with semantically related or unrelated groups of abstract and concrete words. The results of this data would give a better idea of if, as the QDR supposes, the two different types of words are stored in different systems, or as seems to be the case here, share a system which is affected by SCE effects.

6.2 Research Question 2

What are the effects of different levels of Concreteness and Word Class on retention of word form and meaning?

This study found clear evidence that a word's syntactic classification and level of concreteness affected its learnability. Using the Learnability Formula model (section 5.2) words determined as "easy" by previous research, such as concrete nouns, were found to be statistically significantly easier to learn than words considered "harder" such as abstract verbs.

These results confirm the conclusions of other research into this area (section 3.3) that different syntactic categories and concreteness affect learnability.

6.3 Research Question 3

Is there an interaction between Relationship, Word Class and Concreteness on learnability of word form and meaning?

This has largely already been discussed in 6.1. Although at face value it seems that these three factors do not interact to influence learnability (section 5.2) the reformulated LFR2 (section 5.3) shows that Relationship combined with the other factors C and WC do cause a significant difference in learnability.

6.4 Summary of Research

Overall it seems that general predictions of previous research were replicated here and that the phenomenon of Semantic Category Effects does seem to hinder learning. It also seems that words belonging to different semantic classes and levels of abstractness contribute to a word's learnability. Positive relationships were identified between the variables and a model, the LFR2, was formulated that was 85% accurate in predicting learnability.

One inconsistency was found with the QDR model and a new model that seeks to explain how words are connected in a network was developed based upon the data and analysis from this study. It is also highly theoretical as no data could be found to test its claims and this would need to be tested in future.

7 Weaknesses in this Study: Threats to Reliability and Validity

No study can claim to be perfect and this study makes no such assertions especially given some potentially serious issues with validity.

7.1 Testing of Translation of Word Groups

There are three pairs which do not fit into the data (table 5-1: 16, 19, 27) and no reasonable explanation can be given which makes sense according to theory. Apart from possible confounding factors in the study design, the way that that the mind works to relate L2 information to L1 information could be the cause of the problem. Hall's model and the RHM both show how mental mapping of L2 occurs to L1 or directly to the conceptual store.

One explanation for the incongruity of these results is that different languages may not have direct equivalents and that the equivalent, or substitute that is used instead, may actually have a different L1 relationship in terms of Concreteness, Word Class or Relationship. What this problem brings to light is the way that SCE affects L1-L2 interlanguage.

This problem could have caused some of the results seen here and, due to the problems inherent in such translation situations, cannot be easily accounted for and controlled for bias or confoundability. This study tried as much as possible to consider these factors but, as always, no experiment in this field can be perfect and future experiments would need to carefully consider how processes described by the HRM can affect results.

Another major threat is the lack of testing of the selected words to see if they actually are part of semantically related categories. The words were selected by the researcher after looking for results of previous research or tables of data which offered such lists, but were surprisingly hard to find. Therefore due to time and resource constraints best efforts were made to develop the lists according to a priori assumptions of what constitutes a semantic category and the words that should be included or excluded. This fundamental issue would need to be carefully considered in future research to ensure that no errors can be found in the lists.

7.2 Controlling Groups to Equalize Words' Factors

A related issue is lack of control of factors such as word length, number of syllables etc. that other studies usually consider. In this study this factor was not considered until after the word lists had been created and again due to time constraints was not rectified so the lists were left untested for this possible confounding factor.

However for concreteness and abstraction ratings best efforts were made to acquire reasonably grouped sets of words in accordance with data from the MRC (Wilson, 1988). But, a rating could not be found for all words so some room exists for error which could have affected results.

In terms of methodology, the length of the study, over four weeks, may have contributed to fatigue although this was partially controlled for by randomizing combinations of lessons so that different sets were taught to different groups of students at different times so one external factor should not have systematically confounded the results.

7.3 Generalisability

This study is quite narrow in its focus in terms of participants and, although quite different in relation to other similar studies which use adults (usually of University age), it may not be generalisable to other demographics. Comesana (2009:25) comments that children bilinguals may have different brain patterns to adults and so, combined with the evidence that different types of words are housed in different areas of the brain (section 3.3) it could be that even at this young age with relatively little L2 education difference in brain patterns are emerging that may not apply to adults without such experiences at a similar young age. This would presumably affect older differently educated participants in a different way.

7.4 Pretesting for Prior Knowledge

In contrast to the Hakki Erten and Tekin study, a pretrial of words was not conducted to test for prior knowledge. This reduces the possibility of memorisation, recognition or learning before the treatment. However this meant that all the selected words were assumed to be unknown by the students which was not the case in one of the conditions (RCV) which showed an above average mean of correct responses compared to the other conditions' results (281 correct answers opposed to an average of 187).

Pre-existing knowledge reduces the ability to measure the effect for a condition as the variability would not be as great between prior and after treatment measurements. However, as this was only noted for one condition it is assumed that the other conditions with similar pre-treatment averages can be treated as valid. A future study would need to more rigorously test for prior knowledge.

Another issue is that students' knowledge of the translated words into Chinese characters (Kanji) was unreadable or unknown due to their level of education (Kanji are taught in fairly standard levels throughout the school going period). This was informally reported on several occasions and could have affected results as the other known foils would simply be ignored and the “novel” unknown character simply remembered and chosen bypassing any effect being tested. In future this would need to be considered more carefully, but here was unfortunately beyond the resources of the researcher to adequately manage.

8 Implications and Recommendations

The study’s results points to a need to examine how to best teach and learn vocabulary in the classroom. It was found that Semantic Category Effects do hinder learning and so teaching words in such sets should be avoided.

With regards to Concreteness and Word Class, it seems that these factors affect both L1 and L2 acquisition (section 3.3). Ellis (1994:chapter 3) has an extensive discussion on first and second language acquisition, and Nation (2001:chapter 3) gives many examples of vocabulary activities. Among these, extensive input and Lexical Approach activities, which mimic more naturalistic language acquisition, may be beneficial when combined with learning traditional vocabulary lists. This would give a top down and bottom up method where students learn not only meanings of words but quickly put them to use in whole language activities.

For beginners, learning vocabulary, initially at least, through word lists in L1-L2 pairs seems to be the most obvious and efficient method (Webb, 2007). This study has provided some evidence about the learnability of different types of words and therefore teachers can be more aware of needing to spend more time in different activities on particular types of words than others. In short – all words are not the same and should be treated differently depending on WC and C. Further recommendations about how best to teach words that vary in these two factors is beyond the reach of this study which seeks only to show that learnability differences exist.

The Network Relationship Model (NRM – figure 6-1) supposes that Concreteness is not the determining factor in how a lexical network is organised. What does seem to influence how the lexical store is accessed is the relationship among words that are presented. If given in semantically related sets, SCE occur and vice versa. Unrelated sets seem to mimic a more associative, random, pattern and are less prone to SCE, regardless of Concreteness. Therefore the mode of presentation seems to determine how the lexical store is accessed.

This hypothesis of vertical/horizontal access has implications for vocabulary learning. If associative patterns of access in the lexical spider’s web (up/down/left/right) are free from category effects due to the bigger distances between nodes it seems reasonable to conclude that learning could be enhanced by building up patterns of words through a free association exercise such as the classic “mind map”.

However if spread of activation effects also occur when nodes are activated in free association it could be that SCE like effects will also hinder retention in an associative pattern as well as a semantic category pattern. This could be the case when particularly strong associations are built within the network in which case the Relationship becomes like a semantically related category condition and causes SCE like effects due to a strong *associative*, rather than a *category,* bond.

It is not clear what the effect of topic would have on SCE. It is assumed that a topic is quite general and during topical discourse more associative rather than semantically related vocabulary will be generated as a topic has the potential for more personal, thus associative, meanings to be brought to mind. Nevertheless, topics can also be semantically related as well so it could be that generating words around a topic would cause SCE. This is clearly also an area for future research.

8.1 An NRM Vocabulary Exercise

In an exercise developed according to NRM, the student or teacher could choose the initial word and free associate in L1. After a set of 10 or so words has been generated they are translated into L2. These lists would need regular review and also some other activity such as producing sentences or paragraphs would be necessary to enable meaningful use of the language.

In this way a highly personal vocabulary list is built up which builds on existing connections in the network but is hopefully free from SCE. It also has the benefit of being intrinsically interesting for the student as the meanings are personal, and motivating as patterns of interest are related directly to their own experience and so much more relevant to what they want to communicate in L2 (cf Krapp & Schiefle, 1986; Krapp, 2002, 2005; Deci & Ryan, 1985; Hidi & Renninger, 2006).

But is this practice enough to develop a stable interlanguage. How would the knowledge learned be put to practical uses in communicative situations, for example. A more structured approach, especially for beginners, is necessary that balances input and building of vocabulary networks. One approach would be to combine extensive input programmes of reading and listening as well as a lexical approach with vocabulary building activities.

However one thing is clear from this research: a vocabulary list should not be put in semantic category in order to avoid SCE. Also, more time and effort has to be spent on learning more abstract words and words in different syntactic classes. As Harwood (2002) notes this means much more recycling of language than is commonly given in modern textbooks.

If the NRM is correct, and Spread of Activation effects are not active when vocabulary is accessed by association, some time can also be given to a learner centric activity of building a personal dictionary of words using free association.

In conclusion, vocabulary learning is a complicated process that can be hindered by all three factors studies here. A different perspective on how lexical networks are built and methods that take advantage of that structure could provide a more efficient way to build vocabulary and learn language.

9 Conclusion

This study investigated what effects Semantic Category Effects have on groups of words that are different in Concreteness and Word Class as well as Relationship. It found that there were significant differences in learnability when these variables were manipulated and confirmed the results of previous studies that SCE does affect learnability. However, it was found that Word Class and Concreteness also affect learnability and to a greater degree than semantic grouping alone. When combined R, WC and C also had a powerful effect on learnability.

A potential weakness was found in the QDR model which states that abstract words are stored in associative networks, but this study showed that abstract groups of words were subject to semantic category effects. Therefore words may be stored in semantic networks and be affected by SCE regardless of level of concreteness. However, semantic categories can be accessed associatively which may, according to the predictions of the NRM, not be affected by SCE due to distance between nodes.

The implication of this network organisation is that vocabulary should be learnt in semantically unrelated sets, and possibly the most efficient method would be student centred free association combined with extensive input and Lexical Approach activities. Future studies would need to confirm the validity of the NRM predictions and also test the effectiveness of free association as a vocabulary learning method.

10 References

Altarriba, J., Bauer, L., Benvenuto, C., 1999, Concreteness, context availability, and imageability ratings and word associations for abstract, concrete, and emotion words, *Behavior Research Methods, Instruments, & Computers* v31 n4, p578-602.

Arevalo, A., Perani, D., Cappa, S. F., Butler, A., Bates, E., Dronkers, N., Action and Object Processing in Aphasia: From Nouns and Verbs to the Effect of Manipulability, *Brain and Language*, v100 n1 p79-94.

Bedny, M., Thompson-Schill, S., 2006, Neuroanatomically Separable Effects of Imageability and Grammatical Class during Single-Word Comprehension, *Brain and Language*, v98 n2 p127-139.

Benelli, B., Belachhi, C., Gini, G., Lucangeli, D., To define means to say what you know about things' : the development of definitional skills as metalinguistic acquisition, *Journal of Child Language* v33 p71–97

Boers, F., Lindstromberg, S., 2005, Finding Ways to Make Phrase-Learning Feasible: The Mnemonic Effect of Alliteration, *System: An International Journal of Educational Technology and Applied Linguistics,* v33 n2 p225-238.

Boers, F., Eyckmans, J., Kappel, J., Stengers, H., Demecheleer, M., 2006, Formulaic Sequences and Perceived Oral Proficiency: Putting a Lexical Approach to the Test, *Language Teaching Research,* v10 n3 p245-261.

Boers, F., Lindstromberg, S., 2008, Phonemic repetition and the learning of lexical chunks: The power of assonance, *System: An International Journal of Educational Technology and Applied Linguistics,* v36 p423-436.

Brown, R., Waring, R., Donkaewbua, S., 2008, Incidental vocabulary acquisition from reading, reading-while-listening, and listening to stories, *Reading in a Foreign Language*, v20 n2 p136-163.

Chun, K., 2008, What effect do mixed sensory mode instructional formats have on both novice and experienced learners of Chinese characters? *Learning and Instruction* v18 p96-108

Colosi, L., 1997, *The Layman's Guide to Social Research Methods*, Retrieved on 26 May 2009 from: http://www.socialresearchmethods.net/tutorial/Colosi/lcolosi1.htm .

Comesana, M., Perea, M., Pineiro, A., Fraga, I., 2009, Vocabulary Teaching Strategies and Conceptual Representations of Words in L2 in Children: Evidence with Novice Learners, *Journal of Experimental Child Psychology*, v104 n1 p22-33.

Crutch, S., Ridha, B., Warrington, E., 2006, The Different Frameworks Underlying Abstract and Concrete Knowledge: Evidence from a Bilingual Patient with a Semantic Refractory Access Dysphasia Bilingual Patient with Semantic Refractory Access Dysphasia *Neurocase* v12, 151–163.

Crutch, S., 2006, Qualitatively Different Semantic Representations for Abstract Further Evidence from the Semantic Reading Errors of Deep Dyslexic Patients, *Neurocase*, v12 n2, 91- 97.

Damian, M., Als, L., 2005, Long-Lasting Semantic Context Effects in the Spoken Production of Object Names, *Journal of Experimental Psychology: Learning, Memory, and Cognition*, v31 n6 p1372-1384.

Deci,E.L., Ryan,R.M., 1985, *Intrinsic Motivation and Self-Determination in Human Behavior*, New York, Plenum.

De Groot, M.B., Hoeks, J., 1995, The Development of Bilingual Memory: Evidence from Word Translation by Trilinguals, *Language Learning*, v45 n4 p683-734.

Dockrell, J., Braisby, N., Best, R. 2007 Children's Acquisition of Science Terms: Simple Exposure is Insufficient, *Learning and Instruction* v17 577-594.

Duñabeitia, J., Avilés A., Afonso, O., Scheepers, C., Carreiras, M., 2009, Qualitative Differences in the Representation of Abstract Versus Concrete Words: Evidence from the Visual-World Paradigm, *Cognition* v110 p284–292.

Duyck, W., Szmalec, A., Kemps, E., Vandierendonck, A., 2002, Verbal Working Memory is Involved in Associative Word Learning Unless Visual Codes are Available, *Journal of Memory and Language* v48 p527–541.

Ellis, R., 1994, *The Study of Second Language Acquisition*, OUP, Oxford.

Elvers, G., 2009, *Using SPSS for t Tests*, Retrieved on 31 August 2009 http://academic.udayton.edu/gregelvers/psy216/spss/ttests.htm.

Ferre, P., Sanchez-Casas, R., Guasch, M., 2006, Can a Horse Be a Donkey? Semantic and Form Interference Effects in Translation Recognition in Early and Late Proficient and Nonproficient Spanish-Catalan Bilinguals, *Language Learning*, v56 n4 p571-608.

Field, A., 2005*, Discovering Statistics Using SPSS, Second Edition (and sex, drugs and rock 'n' roll)*, SAGE Publications.

Finkbeiner, M., Nicol, J., 2003, Semantic Category Effects in Second Language Word Learning, *Applied Psycholinguistics* v24 p369–383.

Forster, K., 2004, Category Size Effects Revisited: Frequency and Masked Priming Effects in Semantic Categorization, *Brain and Language*, v90 n1-3 p276-286.

Garson, D., 2009, *Univariate GLM*, Retrieved on 5 June 2009 from: http://faculty.chass.ncsu.edu/garson/PA765/anova.htm#assume.

Hakki Erten, I., Tekin, M., 2008 Effects on vocabulary acquisition of presenting new words in semantic sets versus semantically unrelated sets, *System* v36 p407–422.

Harwood, N., 2002, Taking a Lexical Approach to Teaching: Principles and Problems, *International Journal of Applied Linguistics,* v12 n2 p139-55.

Hall, C., 1992, *Making the Right Connections: Vocabulary Learning and the Mental Lexicon*, Retrieved on 13 March 2009 from: http://www.eric.ed.gov/ERICWebPortal/contentdelivery/servlet/ERICServlet?accno=ED363128.

Hinton, P., 2004, Statistics Explained *2nd Ed*., Routledge.

Hidi, S., Renninger, A., 2006, The Four-Phase Model of Interest Development, *Educational Psychologist*, v41 n2 p111-127 2006.

Hidi, S., 1990, Interest and Its Contribution as a Mental Resource for Learning *Review of Educational Research*, v60 n4, p549-571 Winter 1990.

Howard, D., Nickels, L., Coltheart, M., Cole-Virtue, J., 2006, Cumulative Semantic Inhibition in Picture Naming: Experimental and Computational Studies, *Cognition*, v100 n3 p464-482.

Kaufman, A., 2003, *Practice Effects, Speech and Language forum*, Retrieved on 9 April 2009 from: http://www.speechandlanguage.com/cafe/13.asp, (originally in Encyclopedia of Human Intelligence, Vol. 2, pp. 828-833). Sternberg, R (ed), Macmillan Publishing Company).

Krapp, A., 2005, Basic Needs and the Development of Interest and Intrinsic Motivational Orientations, *Learning and Instruction*, v15 n5 p381-395 Oct 2005

Krapp, A., 2002, Structural and dynamic aspects of interest development: theoretical considerations from an ontogenetic perspective, *Learning and Instruction* v12 n4 383–409 2002

Krapp, A., Schiefle, U., 1986, The Development of Interests: Research Programs in the Federal Republic of Germany, *Speech at the Annual Meeting of the American Education Research Association* 1986, Retrieved on 27th April 2008 from: http://www.eric.ed.gov/ERICWebPortal/contentdelivery/servlet/ERICServlet?accno=ED 273653.

Laufer, B., and Girsail, N., 2008, Form-focused Instruction in Second Language Vocabulary Learning: A Case for Contrastive Analysis and Translation*, Applied Linguistics* v29 n4 p694–716.

Laufer, B., 1990, Why are some words more difficult than others?--Some Intralexical Factors that Affect the Learning of Words, *IRAL: International Review of Applied Linguistics in Language Teaching*, v28 n4.

Li, L., Mo, L., Wang, R., Luo, X., Chen, Z. 2009 Evidence for Long Term Cross-Language Repetition Priming in Low Fluency Chinese-English Bilinguals, *Bilingualism: Language and Cognition*, v12 n1 p13-21

Molinari, G., Tapiero, I., 2007*,* Integration of new domain-related states and events from texts and illustrations by subjects with high and low prior knowledge, *Learning and Instruction* v17 p304-321.

Motulsky, H., 1999, *Analyzing data with GraphPad Prism*, Graphpad software Inc., Retrieved on 10 July 2009 from: http://www.graphpad.com/articles/AnalyzingData.pdf.

Nation, I., 2001, *Learning Vocabulary in Another Language*, Cambridge.

Min, H-T., 2008. EFL vocabulary acquisition and retention: Reading plus vocabulary enhancement activities and narrow reading. *Language Learning* v58 n1 p73–115.

Moudraia, O., 2001, *Lexical Approach to Second Language Teaching*. ERIC Digest, Retrieved on 7 July 2009 from: http://www.eric.ed.gov/ERICWebPortal/contentdelivery/servlet/ERICServlet?accno=ED455698.

Pezzullo, J. C., 2009, *Latin Squares for Constructing "Williams Designs", Balanced for First-order Carry-over (Residual) Effects*, Retrieved on 3 June 2009 from: http://statpages.org/latinsq.html.

Saito, S., Logie, R., Morita, A., Law, A., 2008, Visual and phonological similarity effects in verbal immediate serial recall: A test with kanji materials, Journal *of Memory and Language* v59 p1–17.

Schnur, T., Schwartz, M., Brecher, A, Hodgson, C., 2006, Semantic Interference during Blocked-Cyclic Naming: Evidence from Aphasia, *Journal of Memory and Language*, v54 n2 p199-227.

Schnotz, W., & Kürschner, C., 2007, A Reconsideration of Cognitive Load Theory, *Educational Psychological Review,* v19 p469–508.

Taynton, K., 2009, Does the semantic category effect influence learning abstract nouns and verbs in foreign language learning?, Submission to MA Education degree course, Oxford Brookes University, UK.

Wilson, M.D. (1988) The MRC Psycholinguistic Database: Machine Readable Dictionary, Version 2. *Behavioural Research Methods, Instruments and Computers*, v20 n1 p6-11.

Wiederman, M., & Whitley Jr., B., (Eds) 2001, Handbook for Conducting Research on Human Sexuality, Lawrence Erlbaum.

Quinn, W., Kinoshita, S., 2008, Congruence Effect in Semantic Categorization with Masked Primes with Narrow and Broad Categories, *Journal of Memory and Language*, v58 n2 p286-306.

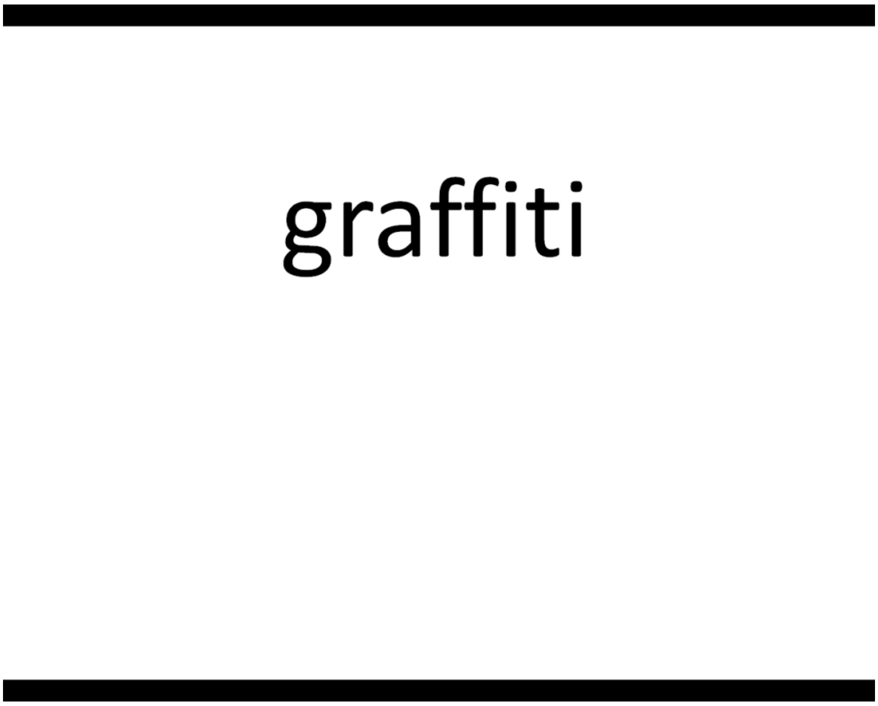
Yoshii, M., 2006, L1 and L2 Glosses: Their Effects on Incidental Vocabulary Learning, *Language Learning & Technology,* v10 n3 p85-101.

11 Appendices

11.1 Example of Powerpoint slides

Two Powerpoint slides from the teaching presentations. The slides were set on a 10 second timer.

First slide, English only - 10 seconds.



Second slide with translation – 10 seconds.

11.2 Word lists used

Word lists used in the study and their concreteness rating (Conc).

Uxx indicates semantically Unrelated sets, Rxx are semantically related sets. See table 6-1 for more explanation.

* data not available from the MRC (Wilson, 1988).

UCV	Conc
swallow	547
ring	593
nail	598
stick	604
light	550
wind	552
pick	502
photograph	590
pound	*

UCN	Conc
saw	532
translator	*
gravel	587
motorcycle	*
needle	608
skull	570
hurricane	576
petrol	*
dishcloth	*

UAV	Conc
plan	357
pray	372
punish	344
approve	*
divide	*
expect	*
impress	*
belong	*
share	*

UAN	Conc
crime	387
employment	424
poverty	343
memory	284
slavery	393
faith	*
freedom	277
rumor	311
relaxation	285

RCV	Conc
etch	*
trace	371
sketch	535
spray	514
carve	*
print	*
paint	577
note	525
graffiti	*

RCN	Conc
eagle	616
antelope	*
toad	658
badger	*
python	580
wolf	595
mule	592
parakeet	*
caterpillar	586

RAV	Conc
classify	*
match	535
remember	*
translate	*
tell	306
explain	*
define	*
understand	*
think	*

RAN	Conc
loyalty	261
pride	270
dedication	*
meanness	*
intelligence	275
cowardice	263
patience	266
honesty	278
courage	277

11.3 Test sheets for before, after1 and after2 tests

Example set of answer sheets: *before*, *after1* and *after2*.

Name　　　　　　　　　　　　　　　　Class 1 - 1-26　　　　　ravb4

例 dog	自転車	犬	勉強	先生
classify	分類する	統一する	仕切る	分ける
match	合わす	反発する	結ぶ	調和する
remember	覚える	思い出す	戻す	忘れる
translate	翻訳する	変換する	解読する	省略する
tell	言う	話す	黙秘する	伝える
explain	演説する	説明する	解明する	述べる
define	決める	曖昧にする	明確にする	確実にする
understand	納得する	理解する	誤解する	把握する
think	考える	思う	ぼんやりする	想像する

Name　　　　　　　　　　　　　　　　Class 1 - 1 - 14　　　　　ravaf1

explain	述べる	説明する	演説する	解明する
remember	思い出す	戻す	忘れる	覚える
classify	仕切る	分ける	分類する	統一する
match	調和する	合わす	結ぶ	反発する
tell	言う	伝える	黙秘する	話す
define	決める	曖昧にする	明確にする	確実にする
understand	誤解する	納得する	把握する	理解する
think	想像する	考える	思う	ぼんやりする
translate	解読する	翻訳する	省略する	変換する

Name　　　　　　　　　　　　　　　　Class 1 - 1 - 22　　　　　ravaf2

remember	戻す	思い出す	覚える	忘れる
classify	仕切る	統一する	分ける	分類する
match	反発する	合わす	調和する	結ぶ
tell	話す	伝える	黙秘する	言う
think	ぼんやりする	想像する	思う	考える
explain	解明する	説明する	述べる	演説する
understand	理解する	誤解する	納得する	把握する
define	曖昧にする	確実にする	明確にする	決める
translate	解読する	変換する	翻訳する	省略する

11.4 SPSS results of ANOVA of Test Scores

Repeated Measures Design 2*2*2 General Linear Model Tests of Within-Subjects Effects

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[a]
relationship * concretness	Sphericity	3.563	1	3.563	1.774	.191	.046	1.774	.254
	Greenhouse	3.563	1.000	3.563	1.774	.191	.046	1.774	.254
	Huynh-Feldt	3.563	1.000	3.563	1.774	.191	.046	1.774	.254
	Lower-	3.563	1.000	3.563	1.774	.191	.046	1.774	.254
Error(relationship *concretness)	Sphericity	74.313	37	2.008					
	Greenhouse	74.313	37.000	2.008					
	Huynh-Feldt	74.313	37.000	2.008					
	Lower-	74.313	37.000	2.008					
relationship * wordclass	Sphericity	.089	1	.089	.041	.841	.001	.041	.054
	Greenhouse	.089	1.000	.089	.041	.841	.001	.041	.054
	Huynh-Feldt	.089	1.000	.089	.041	.841	.001	.041	.054
	Lower-	.089	1.000	.089	.041	.841	.001	.041	.054
Error(relationship *wordclass)	Sphericity	80.620	37	2.179					
	Greenhouse	80.620	37.000	2.179					
	Huynh-Feldt	80.620	37.000	2.179					
	Lower-	80.620	37.000	2.179					
concretness * wordclass	Sphericity	23.054	1	23.054	14.501	.001	.282	14.501	.960
	Greenhouse	23.054	1.000	23.054	14.501	.001	.282	14.501	.960
	Huynh-Feldt	23.054	1.000	23.054	14.501	.001	.282	14.501	.960
	Lower-	23.054	1.000	23.054	14.501	.001	.282	14.501	.960

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[a]
Error(concretness*wordclass)	Sphericity	58.821	37	1.590					
	Greenhouse	58.821	37.000	1.590					
	Huynh-Feldt	58.821	37.000	1.590					
	Lower-	58.821	37.000	1.590					
relationship * concretness * wordclass	Sphericity	31.317	1	31.317	14.719	.000	.285	14.719	.962
	Greenhouse	31.317	1.000	31.317	14.719	.000	.285	14.719	.962
	Huynh-Feldt	31.317	1.000	31.317	14.719	.000	.285	14.719	.962
	Lower-	31.317	1.000	31.317	14.719	.000	.285	14.719	.962
Error(relationship*concretness*wordclass)	Sphericity	78.725	37	2.128					
	Greenhouse	78.725	37.000	2.128					
	Huynh-Feldt	78.725	37.000	2.128					
	Lower-	78.725	37.000	2.128					
	Lower-	145.616	37.000	3.936					

a. Computed using alpha = .05

11.5 Paired Samples Statistics

Results of the *t* test paired samples. See table 4-1 for details of abbreviations.

Pair No.	Pairs Compared	Paired Differences					t	df	Sig. (2-tailed)
					95% Confidence Interval of the				
		Mean	Std. Deviation	Std. Error Mean	Lower	Upper			
1	UCN-RAN	.96341	2.38537	.26342	.43929	1.48754	3.657	81	.000
2	UCN-RAV	1.0555	2.61032	.30763	.44216	1.66895	3.431	71	.001
3	UCN-UCV	.98387	2.47275	.31404	.35591	1.61183	3.133	61	.003
4	RCN-RAN	.72840	2.44955	.27217	.18675	1.27004	2.676	80	.009
5	UCN-RCV	.69512	2.38663	.26356	.17072	1.21952	2.637	81	.010
6	UAV-RAN	.81013	2.93553	.33027	.15260	1.46765	2.453	78	.016
7	RCN-RAV	.56944	2.22550	.26228	.04648	1.09241	2.171	71	.033
8	RCN-UCV	.68852	2.66171	.34080	.00683	1.37022	2.020	60	.048
9	UAN-RAN	.64706	2.76311	.33508	-.02176	1.31587	1.931	67	.058*
10	UAV-RAV	.65217	2.89942	.34905	-.04434	1.34869	1.868	68	.066*
11	RCN-RCV	.46914	2.31347	.25705	-.04242	.98069	1.825	80	.072*
12	UAV-RCV	.53165	2.64495	.29758	-1.12408	.06079	1.787	78	.078*
13	UCN-UAN	.51471	2.70705	.32828	-.14054	1.16995	1.568	67	.122
14	UCN-RCN	.28395	2.51115	.27902	-.27131	.83921	1.018	80	.312
15	RCV-RAN	.26829	2.47481	.27330	-.27548	.81207	.982	81	.329
16	UAN-RAV	.23729	2.15230	.28021	-.32360	.79818	.847	58	.401
17	RAV-UCV	.28302	2.60456	.35776	-.43489	1.00093	.791	52	.432
18	UAV-UCV	.28814	2.82264	.36748	-.44745	1.02372	.784	58	.436
19	UAN-RCV	.22059	2.39929	.29096	-.36016	.80134	.758	67	.451
20	RCV-UCV	.19355	2.26041	.28707	-.38049	.76759	.674	61	.503
21	UAN-UAV	.23077	2.77133	.34374	-.45593	.91747	.671	64	.504
22	RCN-UAV	-.12821	2.66463	.30171	-.72899	.47258	-.425	77	.672
23	RCN-UAN	.10448	2.32336	.28384	-.46223	.67119	.368	66	.714
24	UCN-UAV	.12658	3.08581	.34718	-.56460	.81777	.365	78	.716
25	RCV-RAV	.08333	2.20595	.25997	-.43504	.60171	.321	71	.749
26	RAN-UCV	.06452	2.81606	.35764	-.65063	.77966	.180	61	.857
27	UAN-UCV	.05769	2.80345	.38877	-.72279	.83818	.148	51	.883
28	RAN-RAV	-.04167	2.42892	.28625	-.61243	.52910	-.146	71	.885

Results of t tests on combinations of variables (*significant at 1 tail)