UNIVERSITY OF GOTHENBURG
DEPT OF LANGUAGES AND LITERATURES

ENGLISH

Semantic, thematic or unrelated?

A study about how English textbooks for Upper Secondary School present new vocabulary

Charlotte Åberg, 891006
BA Degree paper, 15 hec
Interdisciplinary Degree Project
Teacher Education Programme LP01

Supervisor: Pia Khölmyr
Examiner: Mats Mobärg

Report number:

Abstract

This study is an analysis of how four textbooks, directed at students in upper secondary school, present new vocabulary: in semantic sets, in themes or in unrelated groups. It aims to show how vocabulary is presented, whether the way it is organized is in agreement with the pedagogical implications found in previous research. It also aims to somewhat facilitate teachers' evaluation of textbooks. The material consists of four books published after 2012, since there was a new Swedish curriculum in 2011 (Gy11). The possible effects of this change should therefore not affect the results in this study. For analyzing these books, I have used content analysis with mixed modes, that is, with traits from both quantitative and qualitative usage of content analysis.

Many researchers favor the thematic organization of new vocabulary. It has shown to help learners retrieve words from memory more easily. Presenting new vocabulary in unrelated sets neither helps nor impedes learning. The semantic organization has shown to impede learning. Despite this, researchers claim that textbooks primarily present new vocabulary in semantic sets. However, this study shows different results than those previously conducted. This could depend on, for example, the possibly different traditions in Swedish schools or that the text types might be different from those observed in previous research.

This study clearly shows that textbooks do not follow the directions provided by research. Even though Skolverket clearly states that teachers have to be up to date with research, many teachers do not have enough time to evaluate textbooks. Textbook writers however, have to keep in mind what sells and what does not. The significant differences between the results in this study and others lead to the conclusion that more research is needed in the field.

Key words: Vocabulary, Semantic, Thematic, Unrelated, Textbook analysis

Table of contents

1. Introduction

As English is a widespread language, learning it has become fundamentally important, which is why it is one of the core subjects in Sweden.

> The English language surrounds us in our daily lives and is used in such diverse areas as politics, education and economics. Knowledge of English increases the individual's opportunities to participate in different social and cultural contexts, as well as in global studies and working life. (Skolverket, 2011a).

The most important aspect entailed in learning a language is vocabulary, since ”…accurate communication depends largely on an extensive knowledge of vocabulary” (Folse, 2004, p 10). This is why teachers need to be able to tell whether a textbook presents vocabulary in efficient ways, as teachers should be updated in pedagogical research and research in their subjects (Skolverket, 2011b). Sheldon (1988) states that textbooks ”…represent […] the visible heart of any ELT [English Language Teaching] programme. The selection of a particular core volume signals an executive educational decision in which there is considerable professional, financial and even political investment” (p 237). Teachers need a lot of time to evaluate textbooks, especially since the writers might not always have only pedagogical implications as a motivation for how to structure the contents, but also a big interest in what sells or not. Today, however, eight out of ten teachers do not have the time to evaluate their course material (Stridsman, 2014). This means that a school might buy textbooks without anyone knowing if they follow the pedagogical implications in research. Hence, the textbooks might not provide help for teachers or learners if they are used. The textbooks might also have cost a lot of money. It is important this money is spent on a book that the teacher will want to use.

This study describes how English textbooks for upper secondary school present vocabulary, by analyzing wordlists. The words in each wordlist are categorized as *semantic*, *thematic* or *unrelated words*. *Semantic* means presenting a list of words that have a semantic relationship, for example topic-related words from a set such as *animals*. *Thematic* means presenting words together which belong to a superordinate topic (*theme*). The connection requires a cognitive process. Words that are not related to other words in the wordlists belong to the third category: *unrelated words*.

The results from the investigation are compared to what present research claims is a preferable way of presenting new vocabulary. The study could fill a gap in recently published

textbooks' ways of presenting new vocabulary and their accuracy compared to present research. Hopefully, the study will somewhat facilitate teachers' choice of material, or at least the evaluation of the textbooks analyzed here from a vocabulary learning point of view.

1.1 Aims and research questions

The study aims to show if the textbooks present vocabulary in *semantic*, *thematic* or *unrelated* sets and whether this is in accordance with what research has shown to be preferable. I wish to provide a means for teachers to efficiently evaluate the vocabulary part of these textbooks.

My research questions are

- In what different ways is vocabulary presented in different textbooks: in *semantic sets*, in *themes* or in *unrelated sets*?
- Does research claim *semantic*, *thematic* or *unrelated* presentation as preferable?

1.2 Material

In 2011, there was a new Swedish curriculum for upper secondary school (Gy11). Even though this is difficult to tell with any certainty, the new curriculum might have had some effect on the content of and organization in textbooks (books that are used in school and specifically written for educational purposes).

Even if the possible effect of a new curriculum on textbooks would have been interesting to observe, this study aims to show *new* textbooks' presentation of vocabulary. The limitation is therefore set at the year of the most recent curriculum. However, since the curriculum was implemented in 2011, it is unlikely that any textbooks published during that year show any sign of the changes in the curriculum. Thus, the time limit is set at 2012 and forward.

I have chosen to focus on English textbooks for Swedish students in upper secondary school. The first year in upper secondary school, when the students read English 5 (i.e. the course most students take in their first year in upper secondary school), might often be focused around gathering the class, that is, summing up what the students have learnt in compulsory school in order to make sure they start on the same level of proficiency in year two. This could mean that the focus is not on new material. For this reason, I have chosen to analyze books from English 6 (i.e the course most students take in their second year in upper secondary school).

The range of this study necessitates a limited number of books, which is why I have chosen to analyze four textbooks from the mentioned period of time. It would be preferable if the material consisted of the books most used in practice today. However, another study would have to be done in order to tell which these are, since publishing houses do not officially present how many copies are sold of each book series, that is which book is the most popular one (Nordlund, 2013). The material represents some of the major publishing houses in the context of Swedish school literature: *Liber*, *Natur och kultur*, *Gleerups*, and *Sanoma utbildning* (R. Vinde, personal communication, November 11, 2014) and *Studentlitteratur*, which many would probably also consider as one of the major publishing houses when it comes to textbooks.

1.2.1 Pioneer

The first book in this study is *Pioneer 2* written by Christer Lundfall, Eva Österberg and Jeremy Taylor, published in 2013 by *Liber*. Like all the textbooks in this study, *Pioneer 2* (henceforth *Pioneer*) is part of a series that provides books for both English 5 and 6 for upper secondary students. *Pioneer* is divided into six chapters, each with texts and/or song lyrics for reading and listening. To each of these texts, there is a wordlist and texts for introduction purposes (e.g. instruction for a theme in a chapter, p 8). These wordlists appear after each text. There are exercises in each of the chapters; one type of exercises is *listening*. Here too, there are wordlists. Since they do not belong to a text, which is required in order to establish the meaning of words that might vary depending on the context, they are not included in this study. Throughout the books, different text types are presented, such as news reports, argumentative texts and poetry. The topics for the chapters are *Under the surface* (Ch 1), *Hearts* (Ch 2), *Style* (Ch 3), *Survivors* (Ch 4), *Turning points* (Ch 5) and *Time* (Ch 6). These are followed by two texts, which make out a chapter called *Extra reading*. One of the texts is non-fiction and the other one fiction (*One flew over the cuckoo's nest*). These two texts have wordlists that are included in this study, given that they follow the same structure as those in the numbered chapters. In *Pioneer*, there is a mixture of fiction and non-fiction.

1.2.2 Viewpoints

Second is *Viewpoints 2* (henceforth *Viewpoints*), written by Linda Gustafsson and Uno Wivast, published in 2012 by *Gleerups*. This series includes books for English 5, 6 and 7. *Viewpoints* consists of texts that are extracts from novels and language exercises that follows the theme and/or vocabulary that appeared in the specific text. Following each text is also a

writing exercise dealing with different text types, such as *Informal letter* or *Review*. For each text, the vocabulary appears on the page where the word appears in the text. At first glance, it may therefore appear as if there is a wordlist on every page where there is text. I have chosen to sum up these chunks of words into one item for analysis, since they belong to the same text and can therefore be analyzed from the point of view of that particular text's *topic*. The book is divided into five main parts/themes with four texts each. Though these parts/themes are not numbered chapters in the book, I have chosen to give them numbers in order to facilitate the analysis as well as the possibility to use this study to further research the same material. Each chapter has a theme that each text in that specific chapter is somehow connected to: *Young perspectives* (Ch1), *Angels of love* (Ch 2), *Moments in time* (Ch 3), *Viewpoints* (Ch 4) and *Stories in history* (Ch 5). There is a sixth part of the book that provides sample texts. There are no wordlists in this part, and so it is not included in this study.

1.2.3 Read and log on

Third is *Read and log on* which is part of a series where each book has a different title with the same theme: *Read and proceed* and *Read and take off* are two of the other books. The series has books for English 5, 6, and 7 and also for other levels of English. *Read and log* on is for English 5. It is written by Cecilia Augutis, Dave Draper, John Whitlam and Eva Österberg, published in 2014 by *Studentlitteratur*. It is the only book in this study for English 5, the reason being that it is the only one from the relevant time span. Even though the different years might show some difference in the content, the new curriculum might have an effect the books' structures and their presentation of, for example, vocabulary. The wordlists appear in the margin on the relevant pages next to the texts.

1.2.4 World Wide English

The fourth book is *World Wide English 6* (henceforth *World Wide English*) written by Christer Johansson, Kerstin Tuthill and Ulf Hörmander, published in 2014 by *Sanoma utbildning*. The book is divided into twenty chapters, each represented by a text: either a short authentic text, an extract from a longer authentic text or poems. Except for the twenty chapters there are some unnumbered parts called *English worldwide*. The purpose is to introduce learners to writers from the English-speaking world which the learners might not be familiar with (www.sanomautbildning.se). In these parts, there are no wordlists. There are only two explicit themes or subjects, except for the *English worldwide*. For chapter eleven to fifteen, the common theme is *Classics*. There is also a theme for a separate part, which is not a

chapter in itself, namely *Poetry through the centuries*. To each poem there is a separate wordlist. There are two versions of *World Wide English*, one for the Social Sciences Programme (SP) and one for the Natural Sciences Programme (NP). I have chosen the version that is adapted to suit SP students, since this is a programme where students can study languages in depth. It is also the most popular programme and has a larger number of students with 59,367 students, compared to NP with 41,309 students (school year 2013/2014) (www.skolverket.se). This might have had an effect on the contents, hence a possible effect on the chosen vocabulary. This should be taken into account when reflecting the generalizability of this study.

1.2.5 Echo

Natur och kultur is one of the major publishers of textbooks in Sweden. Their *Echo* is a series for English 5 and 6 written by Kevin Frato, Henry Alastair, Åke Persson, Marie Kabala-Rejment and Bodil Jonsson. It was published in 2013 and it is the only book by *Natur och kultur* published during the relevant time span. Hence, *Echo* qualifies to be included in this study on a number of criteria. In spite of this, it cannot be included as there are no wordlists in the book. For each text, there are exercises that concern different areas. One of these areas is vocabulary. These exercises are similar to those in the other textbooks, and would therefore allow for comparison in a study where these exercises are included as part of the material. Even though it would be interesting to analyze why the writers of this book have chosen not to sum up a traditional wordlist, there is no space for such an analysis in this essay. And so, even though *Natur och Kultur* is one of the major publishing houses, *Echo* has been excluded from this study.

1.3 Method

In this study, I analyze the vocabulary presented in wordlists in textbooks from three categories: *semantically related* words, *thematically related* words and *unrelated* words. The method for this study is content analysis (henceforth CA), a method with roots in the study of mass communication (Domas White & Marsh, 2006). CA is a flexible method, which has come into wide use in a lot of fields (Hsieh & Shannon, 2005; Domas White & Marsh, 2006). It allows for studies that include both quantitative and qualitative features; that is a *mixed mode* or a *hybrid approach* (Domas White & Marsh, 2006). CA can be adapted to suit the unique needs of every researcher and/or field, as well as the specific research question and

strategy. Also, CA allows for replication and does not rely on the authority of the researcher. At the same time, it involves specialized procedures: "[t]he important criterion is not numeric correspondence between coders but conceptual consistency between observation and conclusion" (p 38).

In this study the main focus is on the wordlists that appear with each text, using *semantic*, *thematic* and *unrelated*. The categories were established a priori. However, I have considered possible trends in the material, for example if there would have been patterns of presenting verbs with the same structure of inflection together (e.g. *run*, *ran*, *run* together with *become*, *became*, *become*), to determine whether these categories suffice. Researchers have used other categories than *semantic*, *thematic* and *unrelated*, to look at how new vocabulary is presented, such as a grammatical point of view (i.e. word class) or a perspective of high frequency words (Nordlund, 2013; Ljung, 1990). However using *semantic*, *thematic* and *unrelated* allows for the possibility to compare with most previous research (Domas White & Marsh, 2006) both regarding categories of words and pedagogical implications by, among others, Tinkham (1997), Nation (2000; 2001), Folse (2004), Papathanasiou (2009) and López-Jiménez (2014). These categories also suit the traditional structure of textbooks; namely short texts created or chosen for learning purposes with new vocabulary in mind.

Within *semantic* the words are divided into *sets*. In one set all the words are related to one another. If, for example *apple*, *peach*, *yellow* and *purple* were presented in a wordlist, they would all fall into the semantic category. They would also fall into two separate sets, *apple* and *peach* being the first set (*fruits*) and *yellow* and *purple* the second (*colours*), as they are not all four related to each other. The same pattern follows in the thematic category, the words are divided into different *themes*. For example, the words *hair* and *blow dryer* make out one theme, while *tree* and *climb* make out a second theme. All four words belong to the thematic category.

In order to be able to show how the vocabulary is presented in each book individually, I have first counted the words in each wordlist, dividing them into *semantic*, *thematic* or *unrelated*. I have also counted the number of *sets* (*semantic*) or *themes* (*thematic*) in which the words from the first two categories appear. I have then compared the results from each book with the others, in order to be able to generalize how vocabulary is presented in the overall material.

1.4 Definitions

Words that are presented in semantic sets are linguistically related in meaning: they are part of a common superordinate concept (Tinkham, 1997). López-Jiménez (2014) and Folse (2004) use the example of *family members* to show what a semantic set is and Nordlund (2013) refers to semantic categories such as *food, clothes* and *animals*. *Family members* could include words such as *brother*, *father* and *cousin*, while *food* could include *sandwich*, *pizza*, *mash* and *meatballs*. Another example of a semantic set comes from Tinkham (1997). The set is *parts of the face* and includes the words *eye, nose, ear, mouth* and *chin*. Papathanasiou (2009) calls a set like the ones above *topic-related words*. One example is the topic *crime*, including words as *mugging*, *terrorism* and *forgery*. Further, Tinkham (1997) describes that words in a semantic set "…directly descend as co-ordinates under a common superordinate concept." (p 143). The words in a semantic set belong to the same part of speech (López-Jiménez, 2014; Tinkham, 1997). Together with topic-related words (i.e. words from the same semantic set) synonyms, antonyms or homonyms constitute the semantic category.

Another term that is frequently used to express this semantic relationship between words is *lexical.* It is used in Nation (2001) and Nation (2000), in which he provides an example of a lexical set, *fruits* by giving examples of which words could be included; *banana*, *orange*, *apple*, *plum*. When compared to Tinkham's (1997), López-Jiménez's (2014) and Papathanasiou's (2009) definitions and examples above, it becomes clear that Nation (2000; 2001) uses *lexical set* to describe what is here referred to as a *semantic* set. Even though Nation uses *lexical* in the two books previously mentioned, the word *semantic* is used in Nation and Newton (1997), as well as in Folse (2004), Tinkham (1997), Nordlund (2013) and in López-Jiménez (2014). Hence, using *semantic* facilitates comparison.

The second category is *thematic*. Tinkham (1997) describes thematically related words to be part of the same *frame* or *schema*. The thematic category includes words from different parts of speech, connected by the fact that they are cognitively, rather than linguistically, associated with a thematic concept (López-Jiménez, 2014; Tinkham, 1997). A thematic concept is thus established through cognitive associations between the words.

It is important to point out that there is a difference between *topic-related* words (i.e. *semantically* related words) and cognitively (i.e. *thematically*) related words (Papathanasiou, 2009). The first requires lexical knowledge, while the second requires a psychological, cognitive process. To tell the difference between a *semantic set* and a *theme*, the cognitive

association is key. An example of a theme (*school*) is that between the words *student*, *backpack* and *geography book*. In order to make out an association between these words a cognitive process is required. This process could consist of these steps: students often have backpacks to carry their school equipment in. A geography book is part of the equipment used by a student and is likely to be put in a backpack, which is carried by that student. One could also consider the cognitive process in questions: who carries a backpack (*student*) and what other things or processes concerns such a person (*school*, *geography book*). Another example is the theme *frog* which includes the words *pond*, *hop*, *frog*, *green*, *swim* and *slippery* (Tinkham, 1997). These processes leading to a cognitive association are not necessary to discover semantic relationships between words. This is why thematic themes fit in language learning situations, since these are not focused on linguistic analysis (López-Jiménez, 2014).

Despite this difference the categories are not mutually exclusive. Tinkham (1997) claims that some words in a thematic cluster could be semantically related but that this ambiguity is rarely a problem, "…many clusters are easily perceived as clear examples of one or the other type" (p 142). However, a set can be part of a theme. If we look at the words *skirt* and *pants*, they are part of a semantic set (*clothes*). At the same time, these words can be part of the theme *shopping* along with *buy* and *saleswoman*.

In this study, the ambiguities that do appear have been placed in the semantic category if they are considered likely to cause confusion in a learning situation. Supposing that the words *skirt* and *pants* in the previous example are completely new to a learner, they could probably cause confusion. They will then be placed in the semantic category. If the words were counted in both *semantic* and *thematic* they would add to the total in both categories. However, neither the total number of words nor the number of words in *unrelated* have changed. Counting them twice (once in each category) would therefore give a false percentage in the semantic and thematic category. Other examples of ambiguities such as this one are presented in the results analysis.

The definition of the unrelated sets is all the words that appear in a wordlist but do not relate to any of the other words in that wordlist in either a thematic or semantic way. There might be categories into which these words could be considered to belong. For example they could be the same word class or they might be high-frequent words. In fact, most words are probably possible to connect with one another if one searches carefully enough.

2. Previous research

There has been a lot of research in the field of language learning. Since learning vocabulary is a big part of learning a language, it is no wonder the area of vocabulary is a big field of research in itself. Some of the researchers in this field focus on the advantages and disadvantages of presenting words in semantic, thematic or unrelated ways. However, it seems there is no previous research on vocabulary divided into these categories in English textbooks for Swedish upper secondary students.

According to Tinkham (1997), the most frequent way to present vocabulary to students is in semantic sets. He has investigated how 48 learners learn and remember English words by comparing them with artificial words in different language learning activities. The English words were put in either semantic sets or thematic themes and the artificial words worked as a control group. He explains that even though the semantic organization facilitates the making of vocabulary activities, "…such clustering of similar items impedes rather than enhances learning" (p. 138); something which might come as a surprise for textbook writers who often rely on the semantic organization of new vocabulary. While the semantic organization impedes learning, thematic organization facilitates it. Psychological research predicts that a thematic group of words (i.e. one theme) would be even easier to learn than a set of unrelated words, according to Tinkham. He continues that curriculum writers and programme planners should consider the fact that they facilitate vocabulary learning by organizing vocabulary in thematic clusters in their texts.

Nation (2000; 2001), a language teacher and PhD supervisor in the field of vocabulary learning, clearly states that presenting new words that are semantically related causes confusion for the learner. According to Nation, the only time to mention semantic relationships between words is when most of them are already established. This issue also applies to a situation where the related words are similar in form too, e.g *Tuesday* and *Thursday* (Nation, 2000). However, there is no danger in presenting *cold* as the opposite of *hot* provided that the learner is already familiar with *hot* (Nation & Newton, 1997). These similar words should not only be presented at different times, but in "[w]idely differing contexts" (Nation 2000, p 9).

Nation (2000) summarizes what researchers have been able to show by performing experiments such as Tinkham's and problematizes that there are researchers who still claim that the semantic relationships between the words facilitates learning since it "requires less

learning" and since related words are more easily retrieved from memory. Also, they claim, the presentation of semantically related words shows how knowledge can be organized, it "reflects the way such information is stored in the brain" (p 6) and the meaning of the words are even clearer when a learner can appreciate how they are related to other words in that particular set. In spite of these claims, Nation states that research shows it is more time-consuming to learn semantically related words, rather than unrelated words or words that are related in a storyline (i.e. *thematic*).

Not only do some researchers continue to claim that there are positive aspects to the semantic organization of new vocabulary, but textbook writers continue to structure textbooks from this point of view, according to Nation (2000). Nation explains that despite the knowledge of the dangers with presenting related words together, "[w]e do not have to look very far in textbooks to see that opposites (…), free associates (…) and lexical sets (…) are often presented together" (6). One of the reasons the semantic organization continues is that "[i]t is difficult for course designers, as well as teachers and learners, to appreciate that items such as months, days of the week, and numbers are best learnt, initially, when not learnt together" (p 7). However, reorganizing words from a semantic set helps learning even when changed into a mixed group; that is, the semantically related words do still appear but they are mixed together with some unrelated words. This means it would be easier to learn *red*, *chair*, *run* and *black* rather than *red*, *black*, *yellow* and *green*. Nation (2000) promotes that words be organized so that the new words can create a sentence. Teachers and textbook writers should "[d]raw on the findings of research…" (p 10) and "[a]void inference from related words" (Nation, 2001, p 92).

Folse (2004) takes a clear stance against organizing new vocabulary in semantic sets. He summarizes what researchers in the field have shown. He presents eight myths that are common about second language vocabulary learning. One of these myths is the idea of semantic sets as a good way of presenting new words to a learner. Folse states that research certainly shows that organization in semantic groups is a bad way of presenting new vocabulary. Despite the fact that the semantic organization could confuse the learners, it continues to exist in teaching and textbooks "…because it is much easier for textbook writers and teachers to present vocabulary in semantic sets such as *family members*, *animals*, or *days of the week* [my italicization] than design creative vignettes to accommodate all of the words in a vocabulary list" (p 4). Even though Folse admits to the fact that there is no one perfect way to teach vocabulary, he states that research shows that thematic word groups help the students remember vocabulary.

Papathanasiou (2009) problematizes the different views on how new vocabulary should be presented; in semantic sets or not. She has compared the results from two of her own intermediate classes, one of which learned semantically related words and the other unrelated words. She also refers to the experimental evidence, among others from Tinkham (1997), that shows the disadvantages of the semantic set. Although this research exists, Papathanasiou explains that researchers continue to support presenting related words since this means learning the distinction between similar words. It also means that learners can cover whole areas of meaning. She claims that "[c]ourse book writers are driven to present semantically related vocabulary items mostly because of their own perceptions of the communicative needs of their students. As a consequence, these course books are divided into various units responding to any situation in which students might find it necessary to communicate in the L2 [second language] (for example, visiting a doctor)" (p 314). Papathanasiou's research shows that unrelated vocabulary is preferred when it comes to beginners but that with intermediate learners there is no significant difference. However, she explains that this might depend on lack of motivation with the younger, intermediate learners, since they presented low scores with both related and unrelated sets. This leads to the conclusion that there is a need for re-evaluation of pedagogical practice and that course book writers also need to consider facilitating vocabulary learning by not using semantic sets. Papathanasiou, like Nation (2000; 2001) mentions the fact that learners who are already familiar with certain words might not be confused by vocabulary being presented in semantic sets. This leads to Papathanasiou's (2009) conclusion that the semantic sets could be used at a secondary stage, thus adding cumulatively to what is already known and thereby avoiding interference.

López-Jiménez (2014) explores how vocabulary is treated in second language (L2) textbooks for students in Spanish schools. Her material consisted of 12 textbooks directed to young adult learners, plus their teacher's editions. One of her research questions is "[w]hich type of lexical organization (semantic/thematic) is used more frequently in the presentation of the new vocabulary?" (p 169). She states that "…the results of research into L2 vocabulary learning have not been taken into consideration. This can be seen in the overuse of semantic sets and the insufficient recycling found in the teaching units" (p 163). Like Tinkham (1997), López-Jiménez (2014) claims that "[p]sychological research generated by interference theory has shown that organizing the new vocabulary in semantic sets actually impedes rather than facilitates L2 learning" (p 165).

As was previously mentioned, there are no studies of vocabulary in English textbooks for Swedish students that investigate the presentation of vocabulary from the point of view used here. However, some studies have been conducted on vocabulary distribution in English textbooks for Swedish learners. In a study focused on current trends in vocabulary teaching in practice and implications of good techniques, Sökmen (1997) problematizes the difficulty of learning random material versus the risk of cross association (i.e. that the meaning of the first word is applied to the second word and vice versa) that follows learning semantically related words. Nordlund (2013) has explored the ways of introducing new vocabulary in Swedish textbooks in the subject of English for students in 4th to 6th grade. She claims that there is no comprehensive study focused on the presentation of how new vocabulary is presented in textbooks in English for young learners in Sweden. Her study is mainly focused around word classes and frequency of words in textbooks. However she does also look at different semantic domains presented in these books, concluding that these occur in the textbooks and that the domain *clothes* reappear in all books. She states that research in the area is particularly important since teachers are not given the proper amount of time needed to thoroughly evaluate textbooks. The most comprehensive study of vocabulary in English textbooks for Swedish upper secondary students, Ljung (1990), investigates only how frequent certain words and word classes are.

It appears that the most preferable way of presenting vocabulary is in themes. Unrelated sets seem to neither facilitate nor impede learning. However, researchers in the field speak against presenting words in semantic sets. A semantic relationship is not dangerous in itself. It might be confusing if a learner is confronted with many semantically related words at once.

3. Results

In this chapter, the results from each book are first presented independently; followed by tables showing a comparison between the books. After this presentation, I present and problematize some of the ambiguities that appear in the material.

3.1 The books

The first book presented is *Pioneer*. In *Pioneer*, there is a clear dominance of unrelated words (80.7%), followed by the semantic category (10.9%). The most rare way to present vocabulary is in themes (8.3%). There are, however some lists in which more words are presented in thematic ways than in semantic ways, (see Appendix 1). The lists with 100% of the words belonging to the same category (*unrelated*) are from short lists belonging to either the lyrics of a song or to a poem. The highest percentage, when excluding *unrelated*, is found in the thematic category (66.7%), even though this is the overall smallest category. However, this list consists of only three words, meaning that the largest percentage in *Pioneer* (excluding *unrelated*) is represented by only two words.

One wordlist shows a large number of semantically related words (21 words) and, like many other wordlists, a large number of sets (7 sets). However, this wordlist has one set with a lot of words in it (9 words), such as *separates*, *item* and *garment*, all related to the semantic topic *clothing*.

Second are the results from *Viewpoints*. Most words are unrelated (85.0%) in the wordlists and the semantic category (8.2%) is larger than the thematic (6.8%). However, one list in *Viewpoints* (see Appendix 2) differs from the other lists. 24.9% of the words in this list belong to the thematic category, which is a high percentage compared to the overall thematic category (i.e 6.8%). The 24.9% percentages are represented by 33 words, which makes it the greatest number of words when excluding *unrelated*. Still excluding *unrelated*, the second greatest number (25 words, 16.5%) is also in the thematic category rather than the semantic category. Even though the semantic category is generally larger than the thematic, the largest numbers (excluding *unrelated*) thus appear in the thematic category.

The theme with the largest number of words appears in the largest group of words, when excluding *unrelated*, that was previously mentioned, namely 29 out of 33 words. The theme (*warfare*) includes words such as *ambush*, *artillery*, *take cover*, *strike*, *mined* and *gunfire*.

One example of a list with a lot of semantically related words contains 24 semantically related words. However, there are 11 different sets that these words are divided into. Among the 24 words, 18 appear in pairs and the last six are in groups of three; meaning that only two or three words are actually related to one another.

Third are the results from *Read and log on* (see Appendix 3). In *Read and log on*, the unrelated category is the largest one (81.6%). However, the thematic category (9.5%) is slightly larger than the semantic category (9.0%). *Read and log on* is a shorter book, the texts are fewer and shorter and so, the wordlists are shorter. The highest percentage, when excluding *unrelated*, appears in the thematic category (17.2%). The largest number of words also belongs to the thematic category (13 words). 11 out of the 13 words in the wordlist in Ch 7 belong to one theme; *battle*, including words such as *battlement*, *combat*, *reload* and *arrow*.

Last are the results from *World Wide English* (Appendix 4). In *World Wide English*, the second largest category is *thematic* (11.3%). This is also the category in which the highest percentage (41.5%), when excluding *unrelated*, appears. The shortest lists in *World Wide English* consist of words from poems and these are the only lists where all words appear in one single category: *unrelated*.

The largest number of words (24 words) appears in the thematic category, as well as the second largest (22 words). In the list with 24 words, the words are distributed over four themes (*advertising*, *job searching*, *at the office* and *words related to economy*), three of them have seven words each, for example *job searching* includes words such as *résumé*, *hire*, *make ends meet* and *skilled*.

The large number of words does not, however, guarantee one large theme or set. One example is a list with 16 words in the thematic category. At first glance, this could be considered as a high number of words in a theme. However, not all 16 words are related to each other. For instance the words *manger*, *scent*, and *sentence* are all in the thematic category, but they are not related to each other. *Manger* is related to *the Babe*, *scent* to *chop* and *stand trial* to *sentence* (the themes are *the bible, cooking* and *crime/justice*).

Below is a presentation of the texts' total number of words and how they are distributed over the categories. In *Semantic*, the number of words that appear together with other semantically related words are presented as well as their percentages of the total amount of words they make out. The same structure goes for *Thematic* and *Unrelated* and finally a presentation of the total number of words in the books.

Table 1. Comparison books – number of words

Book	Semantic	Thematic	Unrelated	Total
Pioneer	127 (10.9%)	97 (8.3%)	938 (80.7%)	1162
Viewpoints	241 (8.2%)	201 (6.8%)	2501 (85.0%)	2943
Read and log on	35 (9.0%)	37 (9.5%)	319 (81.6%)	391
World Wide English	225 (9.8%)	260 (11.3%)	1819 (78.9%)	2304
Total	628 (9.2%)	595 (8.8%)	5577 (82.0%)	6800

In the overall material, the unrelated words clearly make out the largest category with 82% of the total amount of words. Second are the semantically related words (9.2%) and third the thematically related words (8.8%). However, between the last two categories the difference is only marginal (0.4 percentage points). *Pioneer* has the highest percentage of semantically related words (10.9%) and *Viewpoints* the lowest (8.2%). *World Wide English* has the highest percentage of thematically related words (11.3%) and *Viewpoints* the lowest (6.8%). The thematic category is the second largest category in two of the books, *Read and log on* and *World Wide English*.

World Wide English shows the same dominance of the unrelated words as the other books. Even though the unrelated words are the most common, the percentage (78.9) is lower than what is seen in the other books; *Pioneer*, 80.7%, *Viewpoints*, 80.1% and *Read and log on* 81.6%.

The second table shows an average of how many words belong to each set (W/S), or theme (W/T) in the semantic and thematic category respectively.

Table 2. Comparison books - words per set/theme

Book	Semantic	Set(s)	W/S	Thematic	Theme(s)	W/T
Pioneer	127	52	2.4	97	21	4.6
Viewpoints	241	107	2.3	201	35	5.7
Read and log on	35	15	2.3	37	6	6.2
World Wide English	225	83	2.7	260	47	5.5
Total	628	257	2.4	595	109	5.5

The average number of words per theme (5.5 words) is more than twice as high as that of words per semantic set (2.4 words). Even though there are more words that are related to each other in a semantic way, these words do not always make out a huge semantic set, but rather they appear in pairs or groups of three. In the thematic category, however, the average theme

incorporates five or six words.

3.2 Ambiguities

Even though the numbers show a trend in presenting words in unrelated sets, other parts of the investigation did provide difficulties and ambiguities. There are ambiguities when it comes to deciding whether a word belongs to a semantic set or a theme; or if the semantic set may be part of a theme. If this is the case, is it always obvious whether the word should be placed in the semantic or the thematic category? There are cases where it is uncertain whether a word belongs to any of the related words or to the *unrelated* category. The decision depends on various factors exemplified here.

First there is a variable of cultural differences, meaning traditions connected with religion, certain patterns that are usual in a particular geographic area, or habits in a family, that needs to be taken into account. As for the semantic category, there is one example of a semantic topic, namely *kitchen items*, relying on culture in *Pioneer*. Some might consider *iron* to be an item belonging to the kitchen, similar to *oven* and *stove*. However, the book is directed to upper secondary students in Sweden. Since I doubt that learners in Sweden would consider an iron as a kitchen tool, I have chosen not to include this word in the group of semantically related words. Looking at the thematic category, there is a theme in *Viewpoints*; *Islam*, including words such as *Eid-Al-Fitr* and *prayer*. The word *Um*, meaning *the mother of someone*, in Arabic appears in the same wordlist as *Eid-Al-Fitr* and *prayer*. Some might consider *Um*, being an Arabic word of courtesy, closely related to the religion of Islam. However, there are Muslims that do not speak Arabic. There are also Arabic speaking people who are not Muslims. Even though it might be the case that ethnicity, language and religion can sometimes be considered as a whole, I have therefore chosen not to include *Um* in the theme. There is another problem here as well. One might question the purpose of presenting Arabic words in an English textbook, as well as including them in this study. I have included them since they are part of the new vocabulary presented the same way as *prayer*, with translations (both to English and Swedish) in a wordlist.

Second, regarding words as part of a semantic set or a theme might depend on the learners' knowledge of a particular subject. An example of such words is *X-men (serie)* and *comic (serietidning)* (*Pioneer*). They are translated with two different words, but the first Swedish word *serie* is sometimes used as a short version of the second *serietidning*. If the learners know that *X-men* is an example of a comic, this should not cause any confusion. It is

likely that even though the learners do not know this, they can figure out that *X-men* is the name of something. However, looking at only the information given in the wordlist, there is no information provided to facilitate learning and impede confusion. Another example is the theme *church* in *World Wide English*. Here, *confession* is presented as well as *confessional box* and *confessional door*. Should these three words be considered as semantically or thematically related to one another? They belong to one semantic set (part-whole relationship of a confessional box). However, it might be considered as highly probable that *door* and *box* are words that are already known to a learner at this level (Thorén, 1967). The combinations with *confessional* may be presented in the wordlist since they appear in what a learner might find as a new context. Thus, they are probably words from one theme rather than from one semantic set in the view of the learner. Another case where words are related in a semantic as well as a thematic fashion is also in *World Wide English*. The synonyms *valiant* and *bold* are presented in the wordlist, hence semantically related as well as part of the theme: *fight*/*battle*. The case here is different to that of *confessional door*/*box*, since the words are likely put in the wordlist because they are new. Since they are translated with synonyms in Swedish: *tapper*/*modig* for *valiant* and *djärv*/*oförskräckt* for *bold*. These minor differences are likely to be lost and/or confused with one another. This pair has therefore been considered as semantically related rather than thematically.

A large amount of ambiguities can be found in the previously mentioned list with 24 words in *World Wide English*. There are four different semantic sets and four different themes. In one of the themes, *advertising*, the words *advertising campaign*, *advertising agency* and *advertising* appear. Much like the case with *confession/-al*, I have considered it likely that learners at this level already know the words *campaign* and *agency*. Hence, they are considered as thematically related. Another theme in this text is *looking for a job* where one could consider including the semantically related words *employ*, *employment* and *employee*. However, a learner must then be familiar with both the words and the effect of suffixes such as *–ment* and *–ee* in order to not confuse these words with one another. Opposite to the example with *advertising*, I find it probable that learners at this level do not know these suffixes. Hence, learners might confuse the words. I have therefore chosen to put these three words in the category *semantically related words*.

Other than previously mentioned issues depending on learners' vocabulary proficiency is the variable of the learners' knowledge of certain authentic texts and text types. One example where this could have an effect on the results is in *World Wide English* (p 216). Here, the text is an extract from Shakespeare's *Twelfth night*. If we were to consider the themes in

the wordlist, there is one, *nicknames for someone you love* (*mistress*, *pretty sweeting* and *sweet-and-twenty*). However, if we were to consider *Twelfth night* as our theme, all the words in the list would be part of that theme. Even though it could be claimed to be a text that most people might know of, it is unlikely that learners know it well enough to consider *journey* and *mirth* to be connected in that they appear in the theme of the text. *Twelfth night* has therefore been regarded as a text like all others in the books. The only theme is therefore *nicknames*.

Another problematic factor is how to count certain words. However, only one problematic case has been found in this study. This case, where one word is related to two other words in two different semantic ways, is found in *World Wide English* in the list on page 203: *society* and *companion* are translated with the same word: *sällskap*. From the text, the learner cannot know the difference. Hence, a learner will probably view them as synonyms with the exact same meaning. *Society* is also related to *high society*, translated into *högre kretsar*, as a variation or hyponym. *Society* is here counted as an item in *semantically related words*, even though it is part of two sets, each with two words. It would seem that two sets would require a minimum of four words. However, *society* cannot be counted twice in its category, since it adds one word only to the total number of words. And so, there are two sets but only three words in this example.

The results do not only depend on deciding where to put words and how to count them, but also the writers' choice of words. These may, in turn, depend on the text type. Synonyms such as *scuffed* and *worn* (*Viewpoints*, p 9) may appear simply because they provide a variation in a text, in this example a novel. In other texts, a writer may wish to repeat words or phrases for an effect (however any repeated word would not be marked more than once in a wordlist). Choosing between different words may also depend on their form. One possible example of this is in the poems in *World Wide English* (p 218) where one criterion for the choice of words could be the fact that they are similar in form and sound when pronounced, such as *swallow* and *sparrow*. This is of course open for interpretation by the reader. To summarize, if the textbooks had had other texts from which to choose vocabulary, the results might have been different. This could also have been the case if the textbook writers would have chosen to change, for instance, the use of synonyms.

Other than the previously mentioned variables, there is one more: the effect of a text's subject. The content of a text can possibly have an effect, not only which words are put in a wordlist and how they are related, but also the number of words necessary to put in a wordlist. In two of the books, the largest groups of words appear in similar themes: *warfare* (*Viewpoints* 3.4) and *battle* (*Read and log on*, Ch 7). These themes are perhaps of the nature

where a lot of words appear that are new to a learner at this level, hence the comparatively great number of words.

And so, the results clearly state that the most usual way of presenting words is in unrelated sets, followed by semantic sets and then thematic themes. Even though the semantic set is the second largest category in total numbers and percentages, each set has fewer words than each theme. Variables such as ambiguities between theme and set/-s, cultural differences and what could be considered as already known to a learner at this level, have been taken into consideration in the study.

4. Discussion

Nation (2000), Folse (2004) and Lopèz-Jiminèz (2014) claim that there is still a trend to present new vocabulary in semantic sets. However, this study, with this particular material, seems to show different trends in the results than previous studies have shown. Semantically related words do appear, but they are rarely presented with more than two words from the same set. Also, there are no wordlists that consist of only semantically related words. And so, even if a learner might confuse the words within each group, this might not be the case with all the words in the category of semantically related words. Nation (2000) states that mixing semantically related words with unrelated words makes learning easier than it would have been learning only words from a semantic set. Hence, the semantic relationships between the words in the material might not impede learning in the way that research expresses as highly problematic. It is therefore impossible to tell whether the number of semantically related words shown in this study actually impedes learning.

The thematic relationship claimed to facilitate learning (Folse, 2004) might, much like the case with the semantic category and the impediment of learning, not be as obvious as it appears at first glance in the results. In order to draw any conclusions, we need to know whether the number of themes affect learning. If it does, then the themes that do facilitate learning might be those with a large number of words. This means that even the lists with a large number of words in the thematic category shown in this study, cannot be assumed to facilitate learning.

The fact that the presence of semantically related words might not cause problems if only one of the presented words in the wordlists is new to learners is also important to take into consideration. It could, however, be considered unlikely that the textbook writers presume that all but one of the semantically related words in a list are already established. The words that are put in a wordlist seem to be what is considered as the difficult words, likely to be unknown by the learner. This could actually explain the fact that there is no clear pattern that shows which words appear in the lists, neither semantic nor thematic sets/themes dominate the list. It could therefore be reasonable to presume that the research claims speaking in favor of a thematic organization and against a semantic organization, have not been taken into account when choosing words for the wordlists. This could mean that the functional way of presenting semantically related words, namely in a cumulative way that benefits mental organization of words, might not occur.

Despite the research claims that favor thematic organization of new vocabulary, for example by Tinkham (1997), there are also those who claim that the semantic relationships help learning, since it facilitates storing words in the memory. Also, learners acquire knowledge about the structures, differences and patterns among these closely related words. However, according to Nation (2000), if semantically related words are presented at the same time there is a risk the learner will not learn any of them and/or mix them up. Hence, the learner will not learn the words and also they lose the knowledge of the structures, patterns and differences. If it is desirable that learners learn how to distinguish between the different meanings or structures and to easier store words in the memory, this should be an argument against presenting words belonging to the same semantic set in the same wordlist. The arguments for semantic organization of new vocabulary can thereby not be used to promote putting semantically related words in the same wordlist.

From the results it seems that the text type might have some effect on what words are put in a wordlist. The differences between my results and that of previous research might then depend on what texts are used in different textbooks. Other factors that might contribute to the fact that my results differ from what previous research has shown, could perhaps be differences depending on teaching traditions and whether or not it is considered to be important to use authentic texts. Also, the proficiency level of the learners as well as their presumed knowledge of different texts, text types and themes could have an impact. Hopefully, the textbook writers have also considered what might interest the students, thus including certain texts and excluding others. There might also be a trend of choosing unrelated words into a wordlist. If this is the case, teachers might assume that this way of presenting vocabulary is the best way. Teachers could therefore look for this when choosing a textbook. They might expect that textbook writers are familiar with what helps and what impedes learning, thus concluding that the unrelated sets are preferable since they dominate textbooks. This could in turn lead to textbook writers continuing to present vocabulary in unrelated groups; it sells more, since the teachers now look for it in textbooks.

5. Conclusions and pedagogical implications

This study investigates how words are presented in four English textbooks for Swedish students in upper secondary school: in semantic sets, thematic themes or unrelated sets. My purpose has been to provide a means for teachers to evaluate the books. My purpose has also been to show whether the textbooks present vocabulary in a way that is preferable or not according to researchers in the field. The material, *Pioneer*, *Viewpoints*, *Read and log on*, and *World Wide English*, are published by the major textbook publishing houses in Sweden. The method, content analysis, has made possible to show both quantitative and qualitative aspects.

This study contradicts what has previously been shown by research about how new vocabulary is presented in textbooks, namely that vocabulary is organized in semantic sets. In this study, words are mostly presented in unrelated sets. Even if the semantic category is the second largest in the overall material, two out of four books studied here have a larger amount of words in the thematic category than in the semantic one. From this, I draw the conclusion that it is likely that the textbook writers have not taken into consideration the fact that many researchers consider it easier to learn vocabulary in themes than in unrelated sets and also that introducing new words in semantic sets could impede learning. It would seem that the words that are marked, that is put in a wordlist, are chosen for other reasons. One probable reason could perhaps be that certain words are considered as completely unknown for a learner at this level, hence supposedly *difficult* words.

The results indicate that the text type might have an effect on the vocabulary presentation. Also, the subject in the text could be important for the amount of new words and which category these words belong to. The texts used in a textbook and their subjects might depend on various factors such as proficiency levels, authenticity, students' presumed interests and traditions.

Interestingly, I have discovered, but not further investigated, the fact that *Echo*, the recently published English textbook series by *Natur och Kultur* does not present vocabulary in wordlists. This could perhaps depend on the fact that there is a tradition that considers guessing from context a good pedagogical technique to learn vocabulary (Sökmen, 1997). The book could be included in a study similar to this if the scope would have sufficed to explore the exercises for vocabulary learning. This would be a topic for future research. Another interesting topic to look at would be similarity in word forms. It might be the case that learners easily confuse or mix up words that are similar in form (Nation, 2000), especially if

learners are unfamiliar with fixes and their effects. For example suffixes' connection with different word classes in such cases as *employ* (verb) and *employment* (*-ment*; noun) (*World Wide English*, p 231). However, similar words (in form), belonging to the same word class do occur, such as *response* and *responsibility* (*World Wide English,* p 114), meaning that the knowledge of suffixes previously mentioned might not always prevent mix-ups. For further studies, it would be interesting to investigate whether these similarities have any effect on learning and if so whether they help or impede learning. It would also be interesting to investigate whether teachers actually use textbooks for vocabulary learning purposes in practice. If they do it would be interesting to see how they use them.

Wordlists with new vocabulary to be learnt at the same time should, according to many researchers (for example Tinkham, 1997; Nation, 2000; 2001, Folse, 2004; Papathanasiou, 2009; Lopéz-Jiminèz, 2014), not include words that are related in a semantic way. It is therefore clear that textbooks need to be revised so that they help rather than impede learning. However, it is not obvious whose responsibility it is to make this change: whether it is up to the teachers or textbook writers. As Sheldon (1998) claims, the textbooks are part of a commercial industry. Hence, the writers have more than pedagogical implications to take into consideration. Also, Skolverket (2011b) does state that teachers should be updated with research in their fields. However, teachers rarely have enough time to thoroughly evaluate textbooks (Nordlund, 2013), something that might lead to the fact that certain teaching techniques are still employed, even if they have been proven by research to impede learning. And so, if it is a teacher's responsibility to evaluate a textbook thoroughly enough to decide whether it agrees with the pedagogical research claims, there has to be a realistic amount of time set aside for evaluating textbooks. This is the only way we can be sure that learners are provided with the best possible conditions to learn.

References

Primary sources

Lundfall, C., Österberg, E. & Taylor, J., (2013). *Pioneer 2*. Stockholm: Liber

Gustafsson, L. & Wivast, U., (2012). *Viewpoints 2*. Malmö: Gleerups

Augutis, C., Draper, D., Whitlam, J. & Österberg, E. (2014). *Read and log on*. Lund: Studentlitteratur

Johansson, C., Tuthill, K. & Hörmander, U. (2012). *World Wide English*. Stockholm: Sanoma Utbildning

References

Domas White, M. & Marsh, E.E. (2006). Content analysis: a flexible methodology. *Library trends, 55 (1)*, 22-45

Folse, K. S. (2004) Myths about Teaching and Learning Second Language Vocabulary: What Recent Research Says. *TESL Reporter 37* (2), 1-13

Hsieh, H-F. & Shannon, S. H. (2005). Three approaches to content analysis. *Qual Health Res 15:1277*. Sage. Retrieved 2014-10-29 from http://qhr.sagepub.com/content/15/9/1277

Ljung, M. (1990). *A study of TEFL vocabulary*. Stockholm: Almqvist & Wiksell International

López-Jiménez, M. (2014) A critical analysis of the vocabulary in L2 spanish textbooks. *Porta Linguarum 21*, 163-181. Retrieved 2014-11-01 from http://hdl.handle.net/10481/30489

Nation, I. S. P. & Newton, J. (1997). Teaching vocabulary. In J. Coady, & T. Huckin (Ed.) *Second language vocabulary acquisition*, 238-254. Cambridge: Cambridge University Press

Nation, I.S.P. (2000). Learning vocabulary in lexical sets: dangers and guidelines. *TESOL Journal, 9 (2),* 6-10

Nation, I. S. P. (2001). *Learning vocabulary in another language.* Cambridge University Press.

Nordlund, M. (2013) *Ordförrådet och textboken – En analys av läromedlet Good Stuff för skolår 4-6*. Examensarbete magisterexamen språkdidaktik. Umeå Universitet

Papathanasiou, E. (2009). An investigation of two ways of presenting vocabulary. *ELT Journal, 42 (4)*, 313-322. Oxford: Oxford University Press

Sanoma Utbildning. Retrieved 2014-11-21 from http://www.sanomautbildning.se/Laromedel/Gymnasie--vuxenutbildning/Engelska/Baslaromedel/Worldwide-English/Fordjupad-information/

Sheldon, L. E. (1988). Evaluating ELT textbooks and materials. *ELT Journal, 42* (4), 237-246. Oxford: Oxford University Press

Skolverket (2011a). *English subject Curriculum 2011*. Retrieved 2014-10-04, from http://www.skolverket.se/polopoly_fs/1.174543!/Menu/article/attachment/English%20120912.pdf

Skolverket (2011b). *Curriculum for the Upper Secondary School 2011*. Retrieved 2014-10-11 from http://www.skolverket.se/om-skolverket/publikationer/visa-enskild-publikation?_xurl_=http%3A%2F%2Fwww5.skolverket.se%2Fwtpub%2Fws%2Fskolbok%2Fwpubext%2Ftrycksak%2FRecord%3Fk%3D2975

Skolverket, statistics. Retrieved 2014-11-22 from http://www.skolverket.se/statistik-och-utvardering/statistik-i-tabeller/gymnasieskola/skolor-och-elever/skolor-och-elever-i-gymnasieskolan-lasar-2013-14-1.215752

Sökmen, A.J. (1997). Current trends in teaching second language vocabulary. In N. Schmitt & M. McCarthy (Ed.) *Vocabulary: Decsription, Acquisition and* Pedagogy, 152-161. Cambridge: Cambridge University Press

Stridsman, S. (2014, November). Läromedelslotteriet. *Skolvärlden*. November, 29-34

Thorén, B. (1967). *10000 ord för tio års engelska: ordlista*. Lund: Liber Läromedel

Tinkham, T. (1997). The effects of semantic and thematic clustering on the learning of second language vocabulary. *Second Language Research*, 13 (2), 138–163. DOI: 10.1191/026765897672376469

Appendix 1. Pioneer – calculation

Page(s)	Semantic	Set(s)	Thematic	Theme(s)	Unrelated	Total
13	6 (12.0%)	3	0 (0.0%)	-	44 (88.0%)	50
20	6 (7.6%)	3	14 (17.8%)	1	59 (74.7%)	79
29	6 (11.5%)	3	2 (3.8%)	1	44 (84.6%)	52
41	9 (11.0%)	1	0 (0%)	-	73 (89.0%)	82
53	15 (34.9%)	6	9 20.9%)	2	19 (44.2%)	43
63	0 (0.0%)	-	0 (0.0%)	-	3 (100%)	3
64	0 (0.0%)	-	0 (0.0%)	-	7 (100%)	7
70	7 (11.1%)	3	0 (0.0%)	-	56 (88,9%)	63
80	7 (8.8%)	3	2 (2.5%)	1	71 (8.8%)	80
94	21 (33.3%)	7	2 (3.2%)	1	40 (63.5%)	63
101	8 (14.5%)	4	6 (10.9%)	1	41 (74.5%)	55
109	9 (19.1%)	4	0 (0.0%)	-	38 (80.9%)	47
111	2 (40.0%)	1	0 (0.0%)	-	3 (60.0%)	5
117	2 (4.1%)	1	15 (30.6%)	2	32 (65.3%)	49
132	4 (5.0%)	2	19 (23.8%)	3	57 (71.2%)	80
142	0 (0.0%)	-	0 (0.0%)	-	17 (100%)	17
146	0 (0.0%)	-	0 (0.0%)	-	7 (100%)	7
147	0 (0.0%)	-	0 (0.0%)	-	3 (100%)	3
154	0 (0.0%)	-	7 (10.9%)	2	57 (89.0%)	64
163	5 (14.3%)	2	3 (8.6%)	1	27 (77.1%)	35
171	4 (19.0%)	2	0 (0.0%)	-	17 (81.0%)	21
176	3 (14.3%)	1	5 (23.8%)	1	13 (61.9%)	21
183	0 (0.0%)	-	5 (21.7%)	1	18 (78.3%)	23
192	0 (0.0%)	-	0 (0.0%)	-	7 (100%)	7
193.1	0 (0.0%)	-	2 (66.7%)	1	1 (33.3%)	3
193.2	0 (0.0%)	-	0 (0.0%)	-	4 (100%)	4
198	2 (7.4%)	1	0 (0.0%)	-	25 (92.6%)	27
207	0 (0.0%)	-	2 (10.5%)	1	17 (89.5%)	19
220	2 (3.1%)	1	2 (3.1%)	1	60 (93.8%)	64
229	7 (16.7%)	3	2 (4.6%)	1	33 (78.6%)	42
234	2 (4.3%)	1	0 (0.0%)	-	45 (95.7%)	47
Total	127 (10.9%)	52	97 (8.3%)	21	938 (80.7%)	1162

Appendix 2. Viewpoints - calculation

Ch	Semantic	Set(s)	Thematic	Theme(s)	Unrelated	Total
1.1	9 (8.3%)	4	8 (7.4%)	2	91 (84.2%)	108
1.2	4 (2.8%)	2	7 (4.0%)	1	164 (93.7%)	175
1.3	6 (5.4%)	2	5 (4.5%)	1	101 (91.0%)	112
1.4	3 (2.7%)	1	0 (0.0%)	-	110 (97.3%)	113
2.1	7 (5.8%)	3	6 (5.0%)	1	107 (89.2%)	120
2.2	6 (7.0%)	3	0 (0.0%)	-	80 (93.0%)	86
2.3	3 (2.4%)	1	13 (10.5%)	2	108 (87.1%)	124
2.4	22 (12.3%)	10	4 (2.2%)	1	153 (85.5%)	179
3.1	12 (5.6%)	6	21 (9.7%)	3	183 (84.7%)	216
3.2	12 (8.8%)	6	0 (0.0%)	-	124 (91.2%)	136
3.3	24 (12.2%)	11	4 (2.0%)	1	169 (85.8%)	197
3.4	8 (6.0%)	4	33 (24.9%)	3	92 (69.2%)	133
4.1	10 (14.9%)	5	0 (0.0%)	-	57 (85.0%)	67
4.2	10 (6.6%)	5	25 (16.5%)	4	116 (76.8%)	151
4.3	22 (10.0%)	8	2 (1.0%)	1	197 (89.1%)	221
4.4	22 (13.8%)	9	10 (6.3%)	2	127 (79.9%)	159
5.1	23 (12.6%)	10	21 (11.5%)	6	138 (75.8%)	182
5.2	12 (8.1%)	5	13 (8.8%)	3	123 (83.1%)	148
5.3	11 (10.1%)	5	10 (9.1%)	1	88 (80.1%)	109
5.4	15 (7.2%)	7	19 (14.0%)	3	173 (83.5%)	207
Total	241 (8.2%)	107	201 (6.8%)	35	2501 (85.0%)	2943

Appendix 3. Read and log on – calculation

Ch	Semantic	Set(s)	Thematic	Theme(s)	Unrelated	Total
1	8 (11.0%)	3	7 (9.6%)	1	58 (80.1%)	73
2	4 (6.9%)	2	10 (17.2%)	1	44 (75.8%)	58
3	7 (15.6%)	3	0 (0.0%)	-	38 (84.4%)	45
5	4 (8.2%)	2	0 (0.0%)	-	43 (93.4%)	47
6	7 (9.2%)	3	7 (9.2%)	2	62 (81.6%)	76
7	5 (5.4%)	2	13 (14.1%)	2	74 (80.4%)	92
Total	35 (9.0%)	15	37 (9.5%)	6	319 (81.6%)	391

Appendix 4. World Wide English – calculation

Page(s)	Semantic	Set(s)	Thematic	Theme(s)	Unrelated	Total
10	7 (14.0%)	2	10 (20.0%)	1	33 (66.0%)	50
22f	16 (17.0%)	3	2 (2.1%)	1	76 (80.9%)	94
36	5 (7.5%)	2	3 (4.5%)	1	59 (88.1%)	67
48	6 (7.1%)	3	4 (4.7%)	2	75 (88.2%)	85
64f	22 (18.0%)	8	17 (13.9%)	2	83 (68.0%)	122
80f	4 (2.7%)	2	17 (11.3%)	1	129 (86.0%)	150
102f	6 (4.6%)	3	9 (6.9%)	1	115 (88.5%)	130
114	2 (2.2%)	1	8 (9.0%)	1	79 (88.8%)	89
132	7 (8.6%)	4	6 (7.4%)	1	68 (85.2%)	81
138	2 (4.8%)	1	4 (9.5%)	1	36 (85.7%)	42
150	6 (6.4%)	3	15 (16.0%)	2	73 (77.7%)	94
176f	15 (6.9%)	6	30 (13.9%)	4	171 (79.2)	216
194f	14 (10.1%)	4	7 (5.1%)	1	117 (84.8%)	138
203	11 (10.0%)	5	4 (3.6%)	1	95 (86.4%)	110
213f	15 (8.7%)	4	16 (9.3%)	5	141 (82.0%)	172
216	0 (0.0%)	-	3 (15.0%)	1	17 (85.0%)	20
217	2 (9.1%)	1	7 (31.8%)	1	13 (59.1%)	22
218	2 (18.2%)	1	2 (18.2%)	1	7 (63.6%)	11
219	0 (0.0%)	-	0 (0.0%)	-	6 (100%)	6
220	4 (23.5%)	2	0 (0.0%)	-	13 (76.5%)	17
221	0 (0.0%)	-	0 (0.0%)	-	5 (100%)	5
222	4 (26.7%)	2	0 (0.0%)	-	11 (73.3%)	15
223	3 (20.0%)	1	2 (13.3%)	1	10 (66.7%)	15
231	14 (15.1%)	4	24 (25.9%)	4	55 (59.1%)	93
244f	2 (3.8%)	1	22 (41.5%)	3	29 (54.7%)	53
255f	18 (18.8%)	6	13 (13.5%)	3	65 (67.8%)	96
263	13 (16.3%)	4	9 (11.3%)	2	58 (72.5%)	80
273f	21 (14.3%)	8	15 (10.2%)	4	111 (75.6%)	147
286	4 (4.8%)	2	11 (13.1%)	2	69 (82.1%)	84
Total	225 (9.8%)	83	260 (11.3%)	47	1819 (78.9%)	2304