

**AN INVESTIGATION INTO THE POTENTIAL OF
A CORPUS-INFLUENCED SYLLABUS FOR
PRIMARY ENGLISH LITERACY EDUCATION IN JAPAN**

by

ERI HIRATA

A thesis submitted to
The University of Birmingham
for the degree of
DOCTOR OF PHILOSOPHY

Department of English
College of Arts and Law
The University of Birmingham
March 2011

UNIVERSITY OF
BIRMINGHAM

University of Birmingham Research Archive

e-theses repository

This unpublished thesis/dissertation is copyright of the author and/or third parties. The intellectual property rights of the author or third parties in respect of this work are as defined by The Copyright Designs and Patents Act 1988 or as modified by any successor legislation.

Any use made of information contained in this thesis/dissertation must be in accordance with that legislation and must be properly acknowledged. Further distribution or reproduction in any format is prohibited without the permission of the copyright holder.

ABSTRACT

The research presented in this thesis investigates the feasibility of a corpus-influenced syllabus for primary literacy education in Japan. It achieves this with reference to two aspects of the context within which such an initiative might be developed. One is the cultural context; that is, the demands of primary ELT in Japan. Therefore this research explores policy makers' and teachers' views, the texts frequently used in primary ELT classrooms, and some aspects of teacher training. The other focus is from a linguistic viewpoint, concerned with the identification of linguistic features which pupils need to learn for the development of their English literacy. This thesis describes an innovative method for identifying such features. The cultural context was investigated by means of three surveys, the first of which was used to inform the choice of texts to include in the corpus. The surveys reveal a lack of attention to literacy teaching and teacher education in primary ELT in Japan, but also point to some potential for syllabus development. The research offers support for a corpus-influenced syllabus for teaching English literacy, while concluding that there is a need for incorporating it into teacher education and developing teaching methodologies which suit the pedagogic context of the Japanese primary school classroom.

DEDICATION

This thesis is dedicated to my beloved grandfather, Ryuichi Hirata (29th May, 1918 - 2nd July, 2009), who always loved his family, and encouraged my study in the UK and supported me in times of difficulty. “一生勉強 - life is an endless learning” as he always said to me, therefore I continue to learn. Thank you, *Ojiichan*.

ACKNOWLEDGEMENTS

First and most importantly, I would like to express my deepest gratitude to my supervisor, Dr Alison Sealey, for all her valuable guidance, inspiration, and kindest support throughout the period of my study. Without her thorough constructive criticism and moral support in my difficult times, this thesis could not have been completed. I owe her greatly.

I would like to show my deepest appreciation to all the members of academic and administrative staff in Birmingham who provided me warm, kind and valuable support throughout my study. Especially, Dr Pernilla Danielsson, for the surgery sessions patiently and kindly provided; Professor Susan Hunston for her support throughout my study at Birmingham - it was the meeting with her at Westmere in March 2004, that I decided to study my MA at Birmingham. I am also truly grateful for her original idea of the topic in this study and her guidance when I was lost. I am also greatly in debt to Dr Crayton Walker, who was a lead supervisor for MA study at Birmingham in 2004-5, and who always had time when I had questions. Dr Ann Darnton, Dr Suganthi John and Dr Nicolas Groom for their friendliness and warm support.

I am most indebted to publishers and copyright holders who gave me permission to study the copyrighted materials, and also teachers and students who participated in surveys. Without their understanding and cooperation, this study could not have been completed. I am also grateful for the scholarship granted to me in 2008-9 from the Faculty of Humanities.

I am also truly grateful for the support from administrative staff, especially Ms Marilyn Washbrook, Ms Michelle Cullen, Ms Sheila Bradley and friends I met in Birmingham - Yoko (a special friend who was also like a mother to me, and who is now peacefully in heaven), Tess, Caroline, Anne, Rachel, Juliet and Gary who were so helpful and supportive when I was acting as postgraduate representative in 2007-2008. Now that I am far away, I miss you all and the time we had at and outside the PG-tips meetings.

I must also thank my colleagues in Japan, who were always understanding about my situation and encouraged my continuation of this study.

These past few years have been the most challenging years of my life, for losing close friends through sudden death, and seeing the most important person in my life suffering from illness. A very special thanks must go to my parents and my family for their unfailing love and belief, and especially to my fiancé, for his love, patience, support and constant encouragement (plus chocolate and coffee supplies) that enabled me to complete this study.

CONTENTS

Chapter One

Introduction	1
1.1 General aims of this study	1
1.2 Research context: Primary English education in Japan	1
1.2.1 The background of the study	1
1.2.2 The benefits of corpus linguistics to syllabus design	3
1.2.3. The benefits of storybooks in learning to read	5
1.2.4. Aims and scope of the present study	6
1.3. Outline of the thesis	8

Chapter Two

Context and Culture	11
2.1 Introduction	11
2.2 Primary English language teaching in Japan.....	11
2.3 Action Plan 2003 and Course Guidance 2008 in primary education	12
2.4 Conflict of ideologies in the Action Plan and its implementation.....	15
2.4.1 International understanding or simply English education?	15
2.4.2 Promoting egalitarianism versus individualised education	17
2.4.3 Classroom orientation	19
2.5 Current teaching context and additional issues in primary ELT.....	22
2.5.1 In relation to teaching syllabus and materials	22
2.5.2 Cultural contexts and Communicative Language Teaching (CLT).....	26
2.5.3 Teacher Education	28
2.6 Cultural education and literature.....	32
2.7 Summary.....	37

Chapter Three

Corpora in the Classroom.....	39
3.1 Introduction.....	39
3.2 Syllabus design	40
3.2.1 Definition of syllabus - syllabus as a part of materials design	40
3.2.2 Type A syllabuses.....	41
3.2.3 Type B syllabus	52
3.2.4 Summary: Integrated syllabus and teacher training	56
3.3 Learning and teaching of vocabulary	58
3.3.1 Learning of vocabulary.....	58
3.3.2 Teaching of vocabulary	65
3.3.3 Selection of vocabulary items for a syllabus	69
3.3.4 Lexical syllabus	72
3.4 Teaching and learning grammar	81

3.4.1 Noticing	82
3.4.2 Focus on form	84
3.5 Corpora and English language teaching.....	86
3.5.1 Historical perspectives on vocabulary teaching	86
3.5.2 Corpora, literature and language teaching	89
3.6 Identification of pedagogically useful items	95
3.6.1 A historical overview of methods of extracting vocabulary items for teaching.....	95
3.6.2 Identification of PUI from SLA (=English taught) research perspectives	99
3.7 Summary and direction of this research.....	104

Chapter Four

Study design	107
4.1 Introduction.....	107
4.2 Overall design of this study	107
4.3 Specific objectives of this study: research questions	108
4.4 Methods - 1: cultural challenges.....	110
4.4.1 Survey I: Survey on Japanese primary teachers' views	111
4.4.2 Survey II & Survey III: Surveys on teacher education.....	115
4.4.3. Methods of analysis of surveys	120
4.5 Methods 2: linguistic challenges	122
4.5.1 The corpus design: a 'pedagogically informed corpus'	123
4.5.2 Corpus building procedure: decisions and challenges.....	129
4.5.3 The corpus: basic statistical information	132
4.5.4 Corpus analysis: basic concepts and procedure.....	134
4.6 Summary.....	140

Chapter Five

English literacy teaching and Teacher education in Japan	141
5.1 Introduction.....	141
5.2 English literacy teaching in primary ELT: based on Survey I.....	142
5.2.1 Characteristics of respondents and their attitudes towards literacy work	142
5.2.2 The use of storybooks in any form of teaching	142
5.2.3 Reasons for the choice of books/storybooks for the classroom: summary and key-points	146
5.2.4 Books in use or familiar to teachers	149
5.2.5 Conclusions and limitations of the findings of Survey I	150
5.3. Teacher education for TEYL in Japan: Survey II and III.....	151
5.3.1 Views from current teachers: based on Survey II.....	152
5.3.2 Views from prospective teachers: based on Survey III	160
5.3.3 In relation to literacy teaching	164
5.3.4 Conclusions and limitation of the findings of Survey II and III.....	167
5.4 Local and contextual considerations for literacy syllabus design	169
5.5 Conclusion	172

Chapter Six

Comparison of Sub-corpora: Culture.....	174
6.1 Introduction.....	174
6.2 Cultural and linguistic differences between EOCL and JTCL	175
6.2.1 Methodology and procedure.....	175
6.2.2 Comparison of keywords and key-keywords	176
6.2.3. Results and Discussion: comparison of keywords	178
6.3 Summary and implication of cultural keywords.....	199

Chapter Seven

Using a corpus of children's literature to identify pedagogically useful items	203
7.1. Introduction.....	203
7.2 Methods of extracting 'pedagogically useful' lexical items.....	205
7.2.1 Frequency list	205
7.2.2 Phraseology	212
7.2.3 Colligational framework – automatic identification of patterns.....	223
7.3 Methods of extracting pedagogically useful items: Stage 2.....	231
7.3.1 Rationale for having Stage 2 analysis.....	231
7.3.2 Implementation of Method	232
7.3.3 Analysis and Discussion Stage 2: co-selection.....	234
7.4 Summary of integrated methods: Stage 1 and Stage 2	241
7.5 Limitations.....	243
7.6 Conclusion	243

Chapter Eight

Pedagogically Useful Items for Japanese Young Learners.....	246
8.1 Introduction.....	246
8.2 Implications of PPUIs for primary ELT in Japan	247
8.2.1 PUI-1: Contractions.....	248
8.2.2 PUI-2: Semantic grouping.....	251
8.2.3 PUI-3: Sequences containing articles and prepositions.....	255
8.2.4 PUI 4: Repeated sequences.....	257
8.2.5 Connection between oracy and literacy.....	260
8.3 Summary: criteria for deciding what is pedagogically useful for Japanese YLs....	261

Chapter Nine

Corpus-influenced syllabus design.....	263
9.1 Introduction.....	263
9.2 Corpus-influenced syllabus design for primary ELT in Japan	264
9.2.1 Implications of phraseological findings, SLA and reading pedagogy.....	264
9.2.2 Corpus-influenced syllabus design.....	268
9.3 Pedagogical considerations and recommendations.....	280
9.3.1 Teaching methodology	280
9.3.2 Practical examples	283
9.4 Implications of corpus-based research for TEYL teacher education	300
9.5 Limitations.....	305

9.6 Conclusion	307
-----------------------------	------------

Chapter Ten

Conclusion	309
10.1 Introduction.....	309
10.2 A prototype syllabus for Japanese primary ELT	310
10.3 Reflections on how the thesis has met its aims	317
10.3.1 The cultural context of Japanese primary ELT	317
10.3.2 Pedagogic corpus initiatives and aspects of culture	320
10.3.3 Pedagogic corpus analysis methods and implications.....	322
10.4 Contribution of this thesis	325
10.5 Limitations.....	327
10.6 Further research	329

References.....	332
------------------------	------------

Appendicies	354
--------------------------	------------

LIST OF FIGURES

Figure 3.1	Frequency of grammatical errors (Ano, 2005, p.15)	100
Figure 3.2	Acquisition order of children whose L1 was Spanish (Dulay and Burt, 1973)	101
Figure 3.3	Ranking of morpheme scores in the NICT JLE corpus (Izumi and Isahara, 2004 p.69)	102
Figure 5.1	Use of written materials and/or storybooks	143
Figure 5.2	The use of translated Japanese storybooks	145
Figure 5.3	Use Japanese translated texts, if available?	145
Figure 5.4	Reasons for choosing books for teaching young learners	147
Figure 5.5	Teachers' teaching experience in general and in teaching English	152
Figure 5.6	Teachers' perception of the skills they need to improve	153
Figure 6.1	Positional frequency of collocates for <i>woman</i> and <i>man</i> , span 5:5	182
Figure 7.1	Semantic grouping of <i>and X1 and X2 and X (Nouns)</i>	227
Figure 7.2	Semantic grouping of <i>and X1 and X2 and X3 (Gerund)</i>	228
Figure 7.3	Colligational framework <i>at the X1 of X2</i>	229
Figure 7.4	Colligational framework <i>on the X1 of X2</i>	230
Figure 7.5	Collocates of ' <i>the</i> ' and positions in which they occur	235
Figure 9.1	Task sequence: PUI-2 semantic grouping	286
Figure 9.2	Task sequence (<i>I'd like</i>) a <i>X1 of X2, please</i> .	292
Figure 9.3	Task sequence: PUI-4 repeated sequences	293
Figure 9.4	Example of an activity using keywords	297
Figure 10.1	Syllabus of Eigo Note 1 (MEXT, 2009 pp.2-3)	312
Figure 10.2	Suggested prototype syllabus accompanying the current material <i>Eigo Note</i>	314

LIST OF CONCORDANCES

Concordance 6.1	OLD MAN - leaving and coming home	183
Concordance 6.2	OLD WOMAN – domestic	183
Concordance 6.3	OLD WOMAN in EOCL	184
Concordance 6.4	<i>MONKEY</i> + saw; watched	186
Concordance 6.5	adjectives + <i>MONKEY</i>	187
Concordance 6.6	Animal similes	188
Concordance 6.7	Uncooked Rice	190
Concordance 6.8	Cooked rice (steamed rice, rice cakes, rice ball(s))	191
Concordance 6.9	RICE in EOCL	193
Concordance 6.10	Instances of HONEY in JTCL	193
Concordance 6.11	Instances of HONEY in EOCL	193
Concordance 6.12	Instances of TEMPLE in JTCL	195
Concordance 6.13	SMALL + Quantities	198
Concordance 7.1	Concordance for ‘s’	210
Concordance 7.2	Concordance for ‘d’	210
Concordance 7.3	Concordance for ‘t’	211
Concordance 7.4	<i>I X1 a X2 and I X3 my X4</i>	226
Concordance 7.5	<i>I X1 a X2 and X3</i>	227
Concordance 7.6	<i>person in the world</i>	236
Concordance 7.7	<i>the + *est SUPERLATIVE ADJ + NOUN + in the world</i>	237
Concordance 7.8	<i>the most + ADJ + NOUN + in the world</i>	237
Concordance 7.9	<i>the SUPERLATIVE ADJ + NOUN + in the * except ‘world’</i>	238
Concordance 7.10	<i>the most ADJ + NOUN + in the *</i>	238
Concordance 7.11	Co-selected <i>head</i> and <i>on the X1 of X2</i>	240
Concordance 7.12	Co-selected <i>right</i> and <i>on the X1 of X2</i>	241
Concordance 8.1	‘ <i>I like *</i> ’ in EOCL	253
Concordance 8.2	‘ <i>I don’t like *</i> ’ in EOCL	253
Concordance 8.3	‘ <i>I like *</i> ’ in JTCL	253
Concordance 9.1	Some examples of <i>I’m + ADJ</i>	283
Concordance 9.2	<i>I’m + very/so/too + ADJ</i>	285
Concordance 9.3	<i>a pair of X2</i>	290
Concordance 9.4	<i>a loaf of X2</i>	290
Concordance 9.5	Colligational framework - <i>of X1 and X2</i>	293
Concordance 9.6	<i>do you like + X</i>	303

LIST OF TABLES

Table 2.1	The action plan proposed by MEXT (2003): primary English activities	13
Table 3.1	Five main groups of vocabulary learning strategies	63
Table 3.2	Arguments against a solely incidental learning approach (Sökmen 1997, pp. 237-238)	69
Table 3.3	Criteria for the vocabulary selection and grading for a syllabus	71
Table 3.4	Historical perspectives on the treatment of vocabulary in ELT	87
Table 3.5	Historical review of criteria and methods of identifying 'pedagogically useful' lexical items	96
Table 4.1	Summary of research questions for each survey	110
Table 4.2	Outline of Survey I	111
Table 4.3	Four types of question used in the Survey I	113
Table 4.4	Outline of Survey II	116
Table 4.5	Outline of Survey III	117
Table 4.6	Coding example of answers to open question 1	121
Table 4.7	Coding example of answers to open question 2	122
Table 4.8	Basic data for CCL and each sub-corpus (EOCL and JTCL)	132
Table 5.1	Characteristics of samples: Survey I	142
Table 5.2	Comparison of teachers with training and without training	144
Table 5.3	Teachers' frustrations or problems in conducting English lessons	155
Table 5.4	Comparison of the number of study hours (Course of study 2003 and 2008)	156
Table 5.5	What teachers regarded as important for prospective teachers to learn	159
Table 5.6	Difficult aspects of teaching perceived by prospective teachers	160
Table 5.7	Useful aspects of the training course	161
Table 5.8	Specific aspects of knowledge prospective teachers felt were lacking	162
Table 5.9	Prospective teachers' view towards what should be taught in training courses	164
Table 6.1	Comparison of top 20 positive key words in JTCL and KWD	177
Table 6.2	Keywords in JTCL in comparison with EOCL: People inc. proper nouns	179
Table 6.3	Keywords in JTCL in comparison with EOCL: Animals and creatures	185
Table 6.4	Keywords in JTCL in comparison with EOCL: Food	189
Table 6.5	Keywords in JTCL in comparison with EOCL: Places	194
Table 6.6	Negative keywords in JTCL in comparison with EOCL	197
Table 7.1	The frequency of the top 100 words in CCL (EOCL and JTCL)	209
Table 7.2	Methods of extracting larger strings from corpora	213
Table 7.3	Top 20 2-6-grams from CCL	219
Table 7.4	Top 30 3-6 n-grams in CCL	221
Table 7.5	The occurrence of <i>and X1 and X2 and X3</i> in CCL and BoE	228
Table 8.1	Frequency of '-s', '-t', '-ll', and '-re' in EOCL and JTCL	248
Table 8.2	Different forms of '-s' found in EOCL and JTCL	248
Table 8.3	Contractions of different forms	249
Table 8.4	Frequency of ' <i>I like</i> '; ' <i>I don't like</i> ' in EOCL and JTCL	253
Table 8.5	Patterns following ' <i>I like</i> '/' <i>I don't like</i> ' in EOCL and JTCL	253
Table 9.1	Frequency of Verbs (GO/COME/RUN/WALK) + <i>home/back</i> in CCL	287

CHAPTER ONE

INTRODUCTION

1.1 General aims of this study

The aim of this study is to investigate the possibility of using a corpus of children's literature as a basis for producing a corpus-influenced literacy syllabus for young learners in Japan. In this thesis, 'literacy' is defined as a set of skills including:

pre-reading skills, such as concepts of print and alphabetic knowledge; word-level skills, including decoding, word reading, pseudoword reading, and spelling; and text-level skills, including fluency, reading comprehension, and writing skills. (August & Shanahan, 2008, p.1).

The main enquiry involves a methodological investigation aimed at finding pedagogically useful items from the children's literature corpus. In considering a pedagogical application of any kind, it is important to fully understand the context within which the application is to be framed. The general aims of this study are 1) to identify challenges regarding the introduction of English language teaching (ELT) at primary level for Japanese young learners of English and their teachers 2) to investigate the kinds of contributions a corpus of children's literature can make for this particular context (i.e. primary English language teaching in Japan).

1.2 Research context: Primary English education in Japan

1.2.1 The background of the study

In 2002 and 2003, MEXT (the Japanese Ministry of Education, Culture, Sports, Science and Technology) announced the introduction of English activities into the Japanese

primary school curriculum within the broader focus of ‘international understanding’. There was no specific mention in the guidelines of teaching English literacy, nor of the teaching of English grammar at primary level. Rather, attention was concentrated on so-called ‘communicative’ teaching and ways to make English classes ‘enjoyable’ through activities such as games and songs, as these were thought to be appropriate for primary learners. The course of study was then reviewed in March 2008. It was changed to introduce separate study hours called ‘Foreign Language Activities’ rather than conducting such activities within the ‘Period of Integrated Study’, instigating a primary focus in dealing with English (MEXT, 2008: 2010) and intercultural understanding.

Although MEXT (2008, 2010) suggested that homeroom teachers (HRTs) were to take charge of the class as a main teacher, most schools have been making use of ALTs (Assistant Language Teachers who are native speakers of English) so as to give the pupils a chance to talk to foreign people and be introduced to their culture. Despite the decision-makers’ assumptions, however, a survey conducted by MEXT with children at pilot schools reported that some pupils feel that English activities are beyond them. There have been debates over the introduction of literacy and some researchers (Allen-Tamai, 2010; Butler 2005; Paul, 2003; Torikai, 2006a; Yoshida, 2003) have noted the importance of possessing literacy skills in order to be truly communicatively competent. However, MEXT have not yet established a balanced handling of the various language skills and their integration, especially at primary level.

Moreover, the policy is also often criticised for its lack of consideration for the pupils’ transition to the secondary level of learning. Some researchers (e.g. Chujo and

Nishigaki, 2004) point out that the recent guidelines demonstrate little attention to the later primary years and lack any method of transition to junior high school English education. Others, such as Allen-Tamai (2010, p. 140-141) stress the importance of gradual literacy teaching for the smooth transition to secondary teaching of English. A survey conducted by Benesse reports that nearly three quarters of the primary school teachers surveyed (73.5%) feel that they are not confident in teaching English (Benesse, 2007, pp. 8-9). The situation seems to have changed little according to a recent survey conducted by Benessee in July and August 2010 (cited in Japan Times, 2011).¹ Possible reasons for this situation might have been that there was as yet no official syllabus nor enough teacher training available. Although MEXT issued 'Eigo Note' (English notebook) in March 2009 as supplemental materials, the entire issue is not yet resolved. Moreover, current policy includes a restriction against teaching grammar or rules of vocabulary (MEXT, 2003, 2008 2010), which has led to a somewhat simplified and sometimes incorrect English to be taught. It can be argued that this policy, despite its revision in 2008, leads to a challenge for TEYL in Japan, as a result of lack of understanding of its intentions, and its failure to address the transition to secondary level. It is important to consider how to make pupils aware of structural information while avoiding the explicit teaching of grammar, and consider what kind of syllabus and approach can be suggested for literacy instruction for primary ELT.

1.2.2 The benefits of corpus linguistics to syllabus design

One of the contributions of corpus linguistics to ELT is that it has made a great impact in terms of the way that we look at language. For example, the traditional distinction between lexis and grammar has been challenged and the pedagogical focus has shifted

¹ <http://search.japantimes.co.jp/cgi-bin/nn20110226f1.html>

towards perception of the patterns of lexis and collocation (Zimmerman, 1997, p. 17). On the basis of corpus research findings, Willis (1990, p. 5) argues that there is a contradiction and gap between the grammatical, structural syllabus and a communicative methodology. In using the communicative approach, Willis (1990, p. 14) suggests looking at grammar ‘very much from a lexical viewpoint’ and treating grammar as lexical patterning. This view originates from the suggestions made by Sinclair and Renouf (1988), in which they propose the notion of a ‘lexical syllabus’.

The notion of a lexical syllabus seems useful in informing the potential design of a syllabus for primary English education: avoiding the explicit teaching of grammar in a traditional manner, but providing benefits to learners by presenting lexis together with its dominant patterns, which would effectively provide opportunities at the same time to expose learners to syntactic information. Hunston and Francis (1999) also suggest that ‘the role of pattern is as input and support to the concept of grammatical consciousness-raising’. Furthermore, Hunston (2007, p. 205) suggests that ‘awareness of patterns could facilitate reading’.

In terms of the role of corpora in the design of a syllabus, Sealey and Thompson (2006, p. 27) suggest that ‘corpus evidence is a resource available to those who make pedagogically informed decisions about the language and literacy syllabus.’ As far as can be discovered from the available literature, there has been little research into the implementation of a corpus-based syllabus in primary EFL and especially in associating the design of a syllabus with the teaching of reading, despite the benefits of a corpus-based approach described in Hunston (2007, p. 264). It can be said that there still

remains a need for research into designing a corpus-influenced syllabus which identifies local and contextual considerations for different educational settings. It is important that the corpus should be designed with such consideration of its context, in order to inform pedagogy in practice.

1.2.3. The benefits of storybooks in learning to read

As research into literacy development in SLA, especially with young EFL learners, is in its infancy, there is no solid evidence to suggest the most appropriate way to introduce literacy work at primary level (J. Moon, 2008). One of the most popular ways to expose children to written texts, however, is through storytelling, and there has been some research demonstrating that the use of children's literature helps to increase students' English proficiency (Dlugosz, 2000, pp. 288-290; Ghosn, 2002, p. 172). Stories are often beneficial in the context of EFL, for introducing the world to readers and expanding their imaginary world; this applies in particular to children, who have limited experience of it. Reference to literary worlds is a way of motivating learners, while giving them a meaningful context at the same time. As Teubert (2001, p. 137) also recognises, 'picture book texts play a more important role with regard to the acquisition of word meanings than dictionaries.' Indeed, for elementary learners, the presentation of stories with pictures makes it easier for them to guess the context of the narrative which might in turn lead to their greater understanding of the text. . Moreover, in order to understand a text, not only should one know the words individually, but the contexts in which they appear. This has a strong implication for improving the presentation of language to young learners.

Furthermore, the words or contexts which belong to particular cultures can be used to raise learners' awareness of differences in culture. For instance, stories contain some terms which have figurative meanings which are particular to some culture, which could be used in EFL classes as an aid to cross-cultural comparison (e.g. Nesi, 1995). By contrasting such difference and similarities, it is possible to expose learners to different cultures. This is one of the many benefits that stories and their meaningful contexts provide to language learners.

1.2.4. Aims and scope of the present study

It is clear that more research is needed into methodologies and syllabus design which are suitable for developing EFL young learners' literacy skills. On this basis, the proposed study investigates the potential of a corpus of children's literature for designing a syllabus for EFL literacy for Japan. When planning a policy, a curriculum or a syllabus, it is important to include the viewpoint of educational practitioners and the context in practice (Cooper, 1989; Harwood, 2010, p.14; Ricento & Hornberger, 1996). Therefore as part of this research, three surveys were conducted in order to understand the local context and challenges in primary ELT in Japan. Survey I was conducted in particular to identify 1) how ELT practitioners view the introduction of English literacy at primary level, and 2) what texts they consider suitable for the target learners. Surveys II and III have a particular focus on current teacher training in order to understand the local needs of teachers. The information resulting from Survey I is important, as it can shed light upon the texts which are likely to be encountered by the learners. The corpus proposed in this study is a 'specialised corpus' or 'pedagogic corpus' (Willis, 1990, 2003) which, in this case, consists of a collection of texts written for children, and likely to be read or encountered by target learners. Moreover, it is

important to consider whether texts include aspects which contribute to learners' intercultural understanding, considering the objectives set by MEXT (2003: Section 2.5, 2008: Section 2.2).

The aims of the research can be summarised as an attempt to find out, in relation to the state-funded primary educational sector (over 90% of primary schools) in Japan, and, with reference to the teaching of English literacy skills to learners aged around 10-12:

- 1) what government and policy-makers think learners should learn; and what the practitioners' views are on:
 - (a) introducing students to English literacy;
 - (b) the texts used to introduce English at primary level in practice;
 - (c) what kinds of teacher training are needed;
- 2) In what ways culture can be explored in texts; and
- 3) what role a corpus-based approach might have in the teaching of literacy skills in English.

Given the current situation of the teaching of young Japanese EFL learners and their instruction in L2 literacy, the main objectives are to identify ways in which a corpus-influenced syllabus may be constructed on the basis of the concept of the lexical syllabus advocated by Sinclair and Renouf (1988) and Willis (1990), and also on the notion of a 'pedagogic corpus' (Willis, 1990, 2003). In particular, my enquiry aims to investigate the methodologies which assist the identification of pedagogically useful features of texts in the corpus. This procedure should shed light on the way in which an analysis of the corpus could be applied to and inform the teaching of English literacy in primary English education in Japan, by identifying lexical items and patterns which learners may need for the development of their English literacy.

1.3. Outline of the thesis

This thesis is divided into ten chapters including this introduction. Chapter 2 mainly reviews the current situation of primary English education in Japan, including the background of primary ELT (English language teaching) and the English education policy and guidelines published in 2002 and 2003, along with the modification of guidelines made in 2008 English education in Japan (MEXT, 2008). This is followed by a discussion of current issues and conflicts inherent in primary English education in Japan. Based on this discussion, three areas of challenge in primary ELT are identified and discussed, as well as a review of the relationship between cultural education and the use of literature. Chapter 3 discusses the relationship between corpora and ELT, and in particular their contribution to language teaching syllabuses and material design. Firstly, different kinds of syllabuses available in ELT are reviewed, followed by a discussion of issues in teaching and learning of vocabulary, including criteria for vocabulary selection. Two key issues associated with grammar teaching are also discussed. The lexical syllabus is also discussed as an example of a corpus-based syllabus in order to explore the potential application for primary ELT in Japan, together with a review of existing research using literature corpora and their application to teaching language, and the exploration of cultures. Moreover, by means of a review of the historical methods for identifying vocabulary for teaching, I consider what is pedagogically useful for Japanese learners from SLA perspectives. The chapter summarises remaining challenges and the possibilities for future research. Chapter 4 explains the design of this study from two perspectives: cultural and linguistic issues related to the study, and a discussion of the overall design and process of this research. After presenting the specific objectives of this study, methods for tackling the cultural challenges are

addressed. Issues concerned with conducting three surveys are explained. Methods and procedures for dealing with linguistic challenges are described, including a discussion of a specialised corpus design in relation to its pedagogical application in the Japanese primary ELT context. As one of the focuses of this research is methodological, this chapter describes analytical methods which are used mainly as a starting point for subsequent analysis. Chapter 5 focuses on current primary ELT literacy teaching. It reports findings from Survey I which was conducted with primary teachers in Japan, in particular on their use of written materials and their attitudes towards English literacy teaching. This is followed by a discussion of findings from Surveys II and III regarding current teacher training. Based on the three survey results, local considerations needed for syllabus design are discussed. Chapter 6 reports on the keyword analysis conducted in order to explore how culture is represented in texts. Having described methodological information regarding the comparison of sub-corpora, findings of the comparative analysis are reported in relation to the cultural and linguistic differences observed in the texts, along with a consideration of their pedagogical implications. Chapter 7 deals with the main part of the analysis in terms of finding pedagogically useful features of the corpus, compiled for the current research with application in mind. Three methods are applied and explained: Frequency, N-grams and Colligational framework. Based on this, additional methods are suggested as a Stage 2 analysis of integrated methods for identifying Pedagogically Useful Items (PUIs). Chapter 8 presents a comparative analysis of pedagogically useful items (PUIs) found in each sub-corpus, and discusses the potential implications of PUIs to the context of primary ELT in Japan. Based on this discussion, criteria for deciding pedagogically useful items for primary ELT in Japan are summarised. Chapter 9 mainly discusses the implications

of phraseological features found to be prominent in the analyses of PUIs in Chapters 7 and 8, and their relevance to teaching literacy in primary ELT, as well as the possible design of a primary ELT literacy syllabus. Using the PUIs as examples, this chapter also presents the integration of what is necessary from the point of view of practice and methodology for the teaching of English literacy to young learners in Japan. As this suggestion is mainly concerned with teachers, current issues of teacher training are discussed in the light of the implications of the corpus-based approach. Chapter 10 concludes and summarises the current research as a whole. After reviewing the aim of the research set out in Chapter 1, this chapter provides a prototype syllabus suggested for primary ELT, incorporating English literacy instruction. This is followed by answers to the research questions, and discussion of the possible implications and limitations of this study. Finally, it concludes with suggestions for further possibilities in related fields of research, and their application to education in primary ELT.

CHAPTER TWO

CONTEXT AND CULTURE

2.1 Introduction

The context of this study is English language teaching in primary schools in Japan, with a focus on the introduction of literacy work. It is important to understand and contextualise a particular pedagogical situation, before discussing pedagogical implications of any kind. Therefore, this chapter reviews the literature pertaining to each of a number of areas relevant to this study, including the current debate over the introduction of primary English language teaching (ELT) in Japan. The following sections discuss this question with particular reference to recent policy issued by the Ministry of Education, Culture, Sports, Science and Technology, in Japan (MEXT), the current teaching situation and its challenges. Bearing in mind that the objectives of the current policy include intercultural understanding, its relationship to the cultural context will also be discussed, as will the way in which literature can play a role in the introduction of cultural education.

2.2 Primary English language teaching in Japan

English education in primary schools has a longer history in the private sector as opposed to state-funded schools. According to the survey conducted by MEXT in 1995, 89.9 % of private schools were already conducting English lessons as a part of their curriculum. The Japanese government in 2003 issued a five-year ‘Action Plan to Cultivate Japanese with English Abilities’. It was issued in response to criticisms of the

lack of communicative skills in English among Japanese people in general. This was the first official document to allow the introduction of English instruction at the primary level in state-funded sector and the announcement of this Action Plan provoked heated debate.

In the subsequent sections, in order to contextualise my research, I examine the content of this policy, the 2003 ‘Action Plan to Cultivate Japanese with English Abilities’ (hereafter referred to as the ‘Action Plan’) (MEXT, 2003) with particular reference to primary English education. The policy, its goals and objectives, as well as the context for the teaching initiative will also be discussed. This is followed by an outline of the conflicts raised by the ideological orientation of the Action Plan which reports counter-arguments to the introduction of ‘English activities’ at the primary level. The current situation and the problems which have been reported will be reviewed, in order to focus attention on some possible future directions of English education in Japanese primary schools.

2.3 Action Plan 2003 and Course Guidance 2008 in primary education

In 2002, MEXT issued new ‘Course of Study’ guidelines which proposed to introduce ‘foreign language activities’ in primary education. It was proposed that these should be introduced within the broader focus of education in international understanding which has the following objectives:

1. understanding different cultures;
2. developing a sense of self and being Japanese; and
3. developing [in learners] basic foreign language skills and other communicative skills in order to express themselves and respect others (MEXT, 2002b, p. 121).

In 2003, MEXT issued the ‘Action Plan’, which incorporated ‘foreign language activities’ within the framework of the ‘Period of Integrated Study’ at primary level. This was the first official document in which the Japanese Ministry of Education had specifically mandated primary English education. The content of the Action Plan with regard to the introduction of English activities in primary level is summarised as follows (Table 2.1):

Table 2.1 The action plan proposed by MEXT (2003): primary English activities

Target domains of improvement	Goals	Major action plans
- English activities at the elementary school level	- At least 1/3 of the English activities will be conducted by foreign teachers, junior high school teachers and those who are proficient in English	- Supporting improved teaching methods to conduct English conversation activities - Enhancing teacher training using various human resources including foreigners, junior high school English teachers and other individuals - Promoting research in English at the elementary school level

(Reproduced from Butler & Iino, 2005, pp. 34-35)

It is suggested that the goals of such ‘English activities’ are to promote positive attitudes to communication in English by ‘providing children with exposure to foreign language conversation in an enjoyable manner and by familiarising them with foreign cultures and ways of living’ (MEXT, 2003). It is worth mentioning that, according to MEXT, the introduction of ‘English activities’ in the ‘Period of Integrated Study’ at the elementary level is intended to serve as providing ‘experimental learning activities that are suitable for elementary school students’ (MEXT, 2003 Section 2.5.).

MEXT issued the modified course of study in March 2008 (translation in English was published in October 2010)², which states the introduction of a new subject called ‘Foreign Language Activities’. According to this change, ‘English activities’ are to be introduced inside the study hour, instead of within the ‘Integrated Study’. The overall objective of ‘foreign language activities’ is outlined below:

To form the foundation of pupils’ communication abilities through foreign languages while developing the understanding of languages and cultures through various experiences, fostering a positive attitude toward communication, and familiarizing pupils with the sounds and basic expressions of foreign languages (MEXT, 2008; 2010 English translation).

Moreover, the promotion of pupils’ awareness of their mother tongue (i.e. Japanese) is stressed as follows:

Teachers should enable pupils to deepen their understanding not only of the foreign language and culture, but also of the Japanese language and culture through foreign language activities (MEXT, 2008; 2010 English translation).

The revised policy is slightly different, in terms of recommending that ‘English Activities’ should be done in the hours called ‘Foreign Language Activities’ (MEXT, 2008; 2010, p. 1). In addition, MEXT produced supplementary materials called ‘Eigo Note’ (MEXT, 2009a; 2009b). It is often said that these were produced to indicate the minimum levels to be attained at primary level. The selection of materials and syllabus (whether they use ‘Eigo Note’ or other materials) are still left up to local education boards or individual schools. The production of ‘Eigo Note’ is certainly progress compared to the situation formerly, but it is still unclear from the policy statements what kinds of activity are expected to be undertaken. Given the ambiguity, there has been

² http://www.mext.go.jp/a_menu/shotou/gaikokugo/index.htm

much debate about how they should be interpreted. It is obviously important for practitioners to have clear interpretation of the statements, since the implementation of the policies will depend on this.

2.4 Conflict of ideologies in the Action Plan and its implementation

As has been pointed out, the ambiguity of the statements in the Action Plan allows for different interpretations. The vagueness in the policy statements seems to suggest that there are conflicts of ideological orientation reflected in the Action Plan itself (Butler & Iino, 2005, p. 25). Butler and Iino (2005, p. 25) argue that the Action Plan contains three major conflicts of ideological orientation:

- 1) whether Japan should pursue a policy of multilingualism or one favouring the spread of English;
- 2) whether Japan should emphasise international understanding or simply English education; and
- 3) promoting egalitarianism versus allowing for individualised education.

The two points (2 and 3) above are most relevant for this study, since it is not concerned with the teaching of other languages. In addition, it is important to consider the classroom context with regard to the implementation of English activities at the elementary level. This is because introducing the communicative approach suggested by MEXT requires a classroom context which is different from that of the traditional Japanese classroom (Torikai, 2006a, pp. 209-210). In the following sections, I shall discuss these points in turn, bearing in mind the current teaching situation and context.

2.4.1 International understanding or simply English education?

The introduction of English activities in primary schools in Japan is often criticised, on the basis of SLA (Second Language Acquisition) research, for subscribing to the myth

of the ‘earlier the better’ without any theoretical support (Ohtsu, 2006a; Torikai, 2006b). Nevertheless, some research suggests that it is advantageous for learners to make an earlier start, in particular on the phonological aspects of language (Brown, 2000, p. 58). As pointed out in Section 2.2.1 above, English activities at the primary level, according to MEXT, are not meant to initiate English education at an earlier stage, but to contribute to meeting the goal of promoting ‘international understanding’ (MEXT, 2003: Section 2.5). However, this position is rather weak, in the sense that MEXT suggests elsewhere that the objectives for international understanding are ‘to develop basic foreign language skills and other communicative skills in order for pupils to express themselves and respect others’ (MEXT, 2002b, p. 121). As Butler and Iino (2005, p. 40) point out, these statements seem to contradict each other and raise the question of ‘how can one develop basic foreign language skills without learning the language?’ The confusion among teachers caused by such conflicting statements is reported in Matsukawa (2007); teachers are unsure about whether they are supposed to teach English or not.

In addition, with regard to placing ‘English activities’ within education for ‘international understanding’, Ohtsu (2006b) points out a fundamental problem: a claim of promoting ‘international understanding’ which ironically suggests the promotion of only one language (i.e. English) and the cultures surrounding it. This is an important point to consider in planning a curriculum. Some researchers argue that education for international understanding should be conducted separately from language teaching (Ohtsu, 2006a). Others suggest that foreign language learning is indispensable for achieving international understanding (Butler & Iino, 2005).

Nevertheless, since ‘English activities’ are currently placed in the broader focus of education for ‘international understanding’, it is important to bear in mind what MEXT means by ‘international understanding’. However, as mentioned in Section 2.2.1 above, international understanding seems to be broadly defined. With this breadth, one must accept limitations on what can be implemented in the classroom. It seems, then, that the point to consider is what is taught and how it is taught in order to promote ‘international understanding’, despite the reality that what is taught can never cover every variety of language and culture.

One of the suggestions by MEXT in this regard is promoting communication skills in English. The Action Plan of 2003 was announced in response to the criticism that Japanese English education concentrated on English for entrance exams and the claims that Japanese people are largely incapable of communicating in English as a consequence of this (Butler and Iino, 2005; Ohtsu 2006a). Some believe that Japanese people’s inability to communicate successfully in English in a global context has been part of the cause of the country’s recent economic stagnation (Butler & Iino, 2005; Inoue, 2001; Yanagise, 2006).

2.4.2 Promoting egalitarianism versus individualised education

The second ideological conflict concerns whether education should be egalitarian, or individualised (Butler & Iino, 2005). While the Action Plan is addressed to all Japanese nationals in terms of its goals, individual schools and local governments are encouraged to design their own English education individually in order to meet these. In addition, the Action Plan encourages ‘small-group teaching and streaming of students according

to proficiency in the English classes of junior and senior high schools' (MEXT, 2003: Section 2.1).

Butler and Iino (2005, p. 25) see a positive side to this phenomenon and the flexibility of the Action Plan, which encourages the initiatives of local governments and teachers which they summarise as follows:

Despite the challenges that these conflicting goals present, the Action Plan gives greater autonomy to teachers and local governments and thus may improve English education by enabling them to become active participants in the development of language education policies rather than simply being passive consumers of such policies.

The notion of individualised English education promoted by MEXT (Yoshida, 2003, p. 68) is relevant to what Hiraizumi, a member of the House of Representatives, proposed in 1974: removing English education from the list of compulsory subjects and training only a small number of people as English specialists (Butler & Iino, 2005, p. 31). Nevertheless, the egalitarian ideology of Japanese educationalists at the time prevented the proposal from being passed (Takahashi, 2000). The reasons behind the plan to introduce English as an academic subject in elementary schools, which were given in 2004 (Asahi Newspaper, 2004), were similar: the plan was devised in response to the concern of teachers and parents over the imbalance of provision in primary English education settings: it varied from one school and local government to another, according to the amount of time and effort each chose to spend on 'English activities'.

As reviewed earlier, MEXT decided to set 'Foreign Language Activities' as a compulsory subject (The Course of Study 2008, English translation made available on the 7th October 2010), but as the produced supplementary materials, 'Eigo Note'

(MEXT, 2009a; 2009b) have the ‘supplemental’ status, there remains the issue of designing a curriculum and a syllabus which are applicable to all. Some researchers, such as Chujiyo and Nishigaki (2004), propose that there is a need to teach vocabulary for everyday purposes. Considering the increasing numbers of primary students who are starting to learn English, more careful attention is needed to what they learn and also to what materials they use, bearing in mind the impact which learning English at primary level will have on learning English education at secondary level. The need for specific guidelines is obvious and a solid syllabus may help not only Japanese learners but also their teachers.

2.4.3 Classroom orientation

In addition to the conflicts mentioned by Butler and Iino (2005, p.25), a third ideological conflict exists in the difference between traditions in Western and Japanese pedagogic orientation. Traditionally speaking, the dominant influence in Japanese classrooms is Confucianism, which expects or prefers certain behaviour patterns among students and teachers. In classroom discourse in particular, it is the quiet and attentive pupils who often tend to be referred to as ‘good students’ (Torikai, 2006a). However, the main Action Plan set by MEXT (2003) implies the need to encourage conversations inside the classroom. Thus the proposed ‘communicative teaching approach’ in the guidelines requires a classroom discourse culture wholly different from what has been traditional, and therefore seems to create a self-contradictory position. This is one of the examples provided by Hofstede and Hofstede (2005, pp. 53-55; 178-179), which alert us to the fact that the western style of teaching cannot be transported directly to a different cultural setting. Yoshida (2003) argues that although MEXT (2003, Section 2.5) suggests the promotion of ‘experiential learning activities suitable for elementary

students', the decision about what is suitable will often be made from the standpoint of the educator or policy maker but not from that of the student. It is important to consider the students' point of view and a balanced approach which closely relates to 'the culture and experience of the learners' (Prodromou, 1988, p. 80) is needed. For instance, few researchers or teachers would interpret the use of 'games and songs' exclusively throughout the period of English activities as 'suitable for elementary students'. This is demonstrated in the results of a survey reported by MEXT on the degree of satisfaction among pupils regarding 'English activities'. While 73.9 percent of pupils answered 'I like English activities' including 'I like English if I have to choose between like and dislike', almost 30 percent of students answered 'I dislike English' (Torikai, 2006a, pp. 94-95). The percentages and reasoning among the pupils for disliking English activities are as follows:

'Because I cannot read English well' (50.4%);
'Because I prefer other activities to English activities' (41.9 %);
'Because I cannot communicate and interact with peers in English well' (39.7 %);
'Because I cannot talk well with foreign teachers' (37.5 %);
'Because I cannot sing nor play games well in English' (35.1%)
(Torikai, 2006a, pp. 94-95: My translation)

It is important to note the phrases used in describing the reasons for disliking English activities, since '*cannot* do something (*dekinai*)' is often used, which suggests that students are conscious of their inability. 'Games, songs and dance' are supposed to be 'enjoyable' for children at this stage, but, as the survey demonstrates, this is not always the case. It is important to bear in mind that, although children in general are believed to be less conscious about losing face, not every child likes to dance, sing and play games in English, especially when they are close to puberty and feel particularly self-conscious.

In primary schools in Japan, each class has a homeroom teacher (HRT) who teaches children all the subjects. There is a report on pupils' reaction to English activities which includes a comment about their HRT acting differently in English lessons (e.g. "Only teachers were too energetic" (literal translation)) (Sankei newspaper, 2006). This comment can be interpreted that teachers are somehow trying to impersonate the identity of an English speaker. Historically, it is regarded that "[b]ecoming good speakers of English could mean losing their identity [as Japanese] as implied in the Course of Study issued in 1947 and 1951 by MEXT, referring the objectives of English learning or teaching in Japan" (Morizumi, 2009 p. 84). Morizumi (2009, p.84) further argues that the following quotations from The Course of Study imply that 'learning a foreign language has something to do with forming our mental behaviour'.

To learn English is not to memorise many English words but to make our mind and heart work in the same way as English speaking people do.
(The Course of Study, 1947)

The objectives of English teaching are ... to develop an understanding of, appreciation for, and a desirable attitude toward the English speaking peoples especially as regards their modes of life, manners, and customs.
(The Course of Study, 1951)

It is important to comment that Japan was occupied under the Allied forces from 1947 to 1952 (28th April, 1952). These two quotes reflect the reality of the occupation. By contrast, in the Course of Study 2003, the importance of preserving the Japanese identity was stressed.

Candlin and Mercer (2001 p. 3) assert that, what happens in a classroom is a social reality to learners. In relation to this aspect, it is worth questioning 'to what extent the social conditions and priorities of the social world outside the classroom and the learners' places in that world, affect what learners do in classrooms and how effectively

they can learn' (Candlin & Mercer, 2001 p. 3). An approach needs to be considered which suits Japanese young learners, which will be discussed in detail in the following section.

2.5 Current teaching context and additional issues in primary ELT

2.5.1 In relation to teaching syllabus and materials

Specification of syllabus and materials is an important factor in teaching. Despite the guidelines which were made in 2002 (MEXT, 2002b), there had been no official syllabus nor materials established with regard to 'English activities', and little guidance was offered as to what should be taught and how. Most primary schools have created their own materials and syllabuses where possible. However the imbalance between the demands for the materials prescribed by MEXT and syllabus increased.

In 2009, taking such needs into consideration, MEXT issued a set of supplementary material called 'Eigo Note' in 2009, which accompanies a teacher's book, audio material with CD and a DVD with digital version of the notebook for presentation in the classroom. It is certainly a positive development, compared to the situation where primary teachers had to make materials from scratch.

This is an improvement compared to the situation before then, but still, the teacher's book and the content of 'Eigo Note' has already received some criticism (i.e., the teacher's materials are written for 'English teachers' not for 'inexperienced teachers'

etc.; Takahashi, 2010 p. 11). It should be borne in mind that teachers at primary level in Japan were not expected to teach English until recently.

In the students' workbook, there are few written/spelled words, but mostly pictures of items which are introduced in each lesson. This aspect demonstrates the main belief held by MEXT towards English literacy. Moreover, a study of the content of 'Eigo Note' by Kamiya, Hasebe, Nishina & Hirata (2010) suggests that 'Eigo Note' contains uncommunicative phrases (e.g. "*Where am I from?*" used as an example of introducing oneself).

One of the problems in practice reported by Torikai (2006a, pp. 120-122) is that expressions which are introduced to primary school children are somewhat de-contextualised, sometimes even meaningless or wrong (e.g. repeating expressions such as '**I am apple*'). Some inaccuracies in the information given about English grammar were reported by Koizumi (2011, p.33), who pointed out that 'wrongly simplified' English forms are presented to the learner (e.g. '**This is a elephant., *I want a apple., *This is dog.*'). Torikai (2006a) and Koizumi (2011) both suggest that this provides a typical example of the lack of awareness of what could be problematic for learners later: here the article is missed out or used wrongly, though this is one of the most difficult areas of English for Japanese people, even for advanced learners. It is not necessary to explicitly teach pupils at this stage, but it is necessary to retain the correct models of language.

It is important to include communicatively useful language as well as providing activities which raise children's awareness of some features of the target language. In designing a teaching syllabus and teaching methodologies, there is a need for teachers to know the basic norms of language, especially those which do not occur in the native tongue of learners.

In relation to the issue raised above, little has been mentioned about grammar teaching in terms of English activities at primary level. Moreover, MEXT has set a restriction which explicitly disapproves of traditional grammar teaching at primary level (i.e. the use of the grammar translation method) (MEXT, 2002b). It is reasonable to claim that this method is not appropriate for primary learners as it requires them to have knowledge of grammar and metalanguage. There is no doubt that, at this stage, primary learners do not need to be able to explain the grammar of English; however, an awareness of grammar is still important. As Nattinger and DeCarrico (1992) assert, the use of lexical phrases enables people to speak with fluency. Therefore, without an awareness of such features of the English language, students are unlikely to become communicatively competent. Swan (2005) also points out the importance of realising the purpose of learning English as well as the importance of grammar. If the purpose of ELT in primary schools is to help learners to be communicatively competent, both lexical and syntactic knowledge are necessary, since both are important in understanding and conveying information (Bourke, 2006; Brewster et al. 2002; Paul, 2003; Torikai, 2006a).

Moreover, Yoshida (2003, p. 110) argues that the term ‘communicative’ is predominantly taken to mean ‘concentrating on speaking and listening’ without teaching of grammar, nor the skill of literacy. Morrow and Schocker (1987) suggest the use of texts in communicative teaching for more ‘truly communicative’ activities. In today’s primary schools, as Yoshida (ibid) suspects, little attention in the context of ‘English activities’ has been paid to literacy. One might argue that, since the linguistic content of English and the children’s mother tongue is widely different, it is hard for children to learn different skills in a foreign language. However, focusing on listening and speaking alone, in teaching young learners, could be misleading as it may give learners the wrong message that English has no written form (Paul, 2003). It can be argued, then, that it is necessary to have literacy skills as well as other skills in order to become communicatively competent in any language (Yoshida, 2003). In addition, Torikai (2006a) suggests the importance of promoting ‘intercultural literacy (*Ibunka literacy*)’ which is defined as the ability to understand different cultures, using the ability to read and write and going beyond the comprehension of sentences. If English education is to remain part of ‘international understanding’, it is important to take the notion of ‘intercultural literacy’ into consideration.

However, an important question remains: how to devise a teaching approach which avoids explicit teaching of the grammatical components of the syllabus, while at the same time effectively teaching or making learners aware of the way in which the language is constructed? The current suggestion made by MEXT disregards this aspect, and no current officially-sanctioned materials available at primary school contain this element. This indicates a gap which is worth investigation.

2.5.2 Cultural contexts and Communicative Language Teaching (CLT)

MEXT announced ‘communicative competence’ as one of the most important goals for ELT in Japan. This change is one of the major reasons for communicative language teaching (CLT) to receive great attention in Japan, especially at junior and senior high school levels. This is also in the case for primary level ELT.

In particular, task-based language teaching (TBLT) has recently become a focus of interest in primary ELT. TBLT is one form of communicative language teaching, which uses tasks in order to give learners chances to use the target language in a meaningful way (Slattery and Willis, 2002, Carless, 2003; 2004). Having a strong theoretical basis, TBLT has been supported widely in the Asian teaching context (e.g. Adams & Newton, 2009; Carless, 2009; Deng and Carless, 2009), especially in countries where CLT is promoted in their policies.

Despite its popularity, some challenges are also reported with regard to the adoption of TBLT in Asia (e.g. Adams & Newton, 2009, Deng and Carless, 2009). For instance, Adams and Newton (2009, p. 5-11) outline mainly three areas of challenge: institutional factors, classroom factors and teacher education. According to Adams and Newton (2009, p. 5-11), institutional factors influencing the disfavour of TBLT in practice are closely related to the ‘non-communicative nature’ of entrance exams in Asian countries. In relation to classroom factors, Qi (2007) reports the difficulties in involving students to participate in communicative activities (Qi, 2007), unless they are highly motivated. In terms of teacher education, the lack of awareness and understanding of CLT is highlighted. Carless (2009, p. 50) reports a case in Hong Kong, where TBLT is

supported in curriculum guidelines submitted by Curriculum Development Council, that 'do not necessarily impact on classroom practice'. Teachers often prefer the traditional P-P-P (Presentation-Practice-Production) approach to TBLT (Tang, 2005, cited in Carless, 2009 p. 50).

In TBLT, it is suggested that 'a teacher controlled focus on grammar should come at the end of a task cycle' (Willis & Willis, 2007 xiii). In relation to this task cycle, Deng and Carless (2009, p.130-131) provide a report of peer-observations of TBLT in a teacher training course in China. In their data, some teachers commented that allowing students to use new language without model sentences is too precarious. Some regard TBLT as suitable for 'adults, generally at intermediate proficiency levels, and mostly with English as the target language' (Skehan, 2003 p. 3). Swan (2005, p. 255) argues that TBLT is intended for learners who 'have already been taught more language than they can use'. In other words, TBLT is thought to be more suitable for learners who have already had more input of language by means of a more traditional style of teaching.

Most of the challenges outlined by the above researchers are also relevant to Japanese contexts. In relation to the sequence of instruction, in particular, it was reported that teachers often feel the need to explicitly teach the language first, avoiding ambiguity of the language that their pupils need to produce. On the contrary, in some cases, teachers feel the need to focus on 'communication' in their classroom activity, and as a consequence, rush into activities without giving pupils much input or contextual information about the interaction taking place (Koizumi, 2011, p. 33). Taking into account the perceived difficulties in conducting TBL with beginners (Swan, 2005) as

identified above, and the particular cultural context, there may be a slight modification needed for the task-cycle in order to conduct TBL in an elementary classroom context in Japan.

Despite some problems and challenges related to TBLT, if we are to use CLT, it is necessary for teachers to realise the aim of CLT, to be able to choose and modify tasks which suit their own pupils, and prepare a task cycle which suits their own classroom orientation. Needless to say, a classroom is a social reality to learners, and this is the context where CLT takes place. As has been suggested earlier, it is important to consider the effects of local and social conditions on the learners' behaviour in the classroom and learning outcomes (Candlin & Mercer, 2001 p.3). As Carless (2009) rightly points out, 'what is required is probably what good teachers have already known and done; namely, a balance of a variety of activities and different approaches adapted to the needs of a particular group of students in a specific setting' (p. 64). Whether CLT takes place in practice depends on teachers, thus teacher training is an inevitable issue for its successful application. The issue of teacher education is discussed next.

2.5.3 Teacher Education

Cameron (2003) rightly states that 'the continuing expansion of teaching English to young learners (TEYL) brings challenges to the wider ELT field.' She further reminds us that starting early may not lead to an automatic improvement in L2 language proficiency unless the remaining challenges of both teacher education and secondary ELT are dealt with along with the expansion of TEYL (Cameron, 2003, p.105). This view is also valid for TEYL in Japan, and teacher education seems to be the biggest challenge in the current situation.

As a result of the implementation of 'English activities' (MEXT, 2002a, 2003), more and more primary students are starting to learn English. Primary schools which have started the practice mainly rely on ALTs (Assistant Language Teachers who are native speakers of English) to carry out 'English activities' (Shiozawa, 2005, p. 66). This situation has arisen partly because there is a lack of Japanese teachers who are experienced in teaching English at primary level, for primary teachers had never had responsibility for teaching English until the Action Plan was issued in 2003. A survey reports that 75 percent of the teachers surveyed feel that they are not confident in teaching English (Benesse, 2007, pp. 8-9).

There are good reasons for relying on ALTs, in the sense that the objectives set by MEXT emphasise communicative competence and intercultural understanding. The Ministry of Education (MEXT) also suggests making use of regionally available professionals who are English teachers or are successful in international exchange. However, as the recent budget allocation issued for primary English language teaching in 2010 has decreased, there will certainly not be enough funding to hire ALTs at every school. In the current situation, it is obvious that there are insufficient funds available to hire qualified and trained English teachers for young learners. There is a clear need to rise to the challenges of the new situation.

Teacher training is an issue which should have been addressed more carefully before the implementation of 'English activities'. The criteria for appointing English teachers at primary level may also be in need of modification. In the context of a training system for TEYL teachers in Japan, there is no official course or qualification authorised by

MEXT. Several companies or organizations such as NGOs, local education boards, and universities have been providing workshops and teacher training courses for people who wish to teach at primary level. At national universities, some students qualifying as primary school teachers also take English as a secondary subject, and qualify with an English teaching qualification. This course of study involves taking a proportion of the classes followed by Junior High School English majors, and a period of teaching practice in Junior High School English classes. There are some teacher training programmes available for teaching for young learners in the private sector. These are mostly authorised by the J-Shine NPO (<http://www.j-shine.org/>) and 44 organisations or private schools (e.g. AEON, ALC Ltd., for distance learning) were registered in 2010. Although this organisation claims their qualifications are relevant for teaching English to YLs, their certificate is authorised neither by MEXT nor the local educational boards, which means there is no guarantee for successful participants to be able to work at state-funded primary schools.

Nevertheless, the content of their courses is similar, especially in the fact that the main focus is on oral language. As noted in Chapter 2, MEXT has been very reluctant to introduce English literacy to primary teaching and this might have affected the priorities for those organizations. This aspect is also evident in ‘Eigo Note’ (MEXT, 2009a; 2009b) (see Appendix 2.1).

It is necessary for the Japanese Education Ministry (MEXT) to establish an appropriate evaluation measurement for selecting teachers capable of delivering ‘Foreign Language Activities’ at primary schools, in order to assess their ability to teach YLs. Not many

local governments have set guidelines for the selection of English teachers. According to Butler (2005, p.262), only 5 local autonomous bodies (Chiba, Saitama, Gifu, Yokohama-city and Kawasaki-city) require prospective teachers to conduct demo lessons in their selection procedures. Some other local autonomous bodies set the English ability requirement using existing English tests such as TOEFL, TOEIC and STEP. However these tests are created for a particular reason which does not necessarily reflect the ability to teach in English (Butler, 2005 p.263). One solution may be to create tests specifically to gauge prospective teachers' English ability as well as their teaching ability.

In Taiwan, an assessment system for the English ability (four skills) of primary English teachers was introduced in 2004. In my view, Japanese English education should take into account other countries' evaluation systems and devise a specific assessment system suitable for primary school in Japan.

Having reviewed the current situation, it is clear that there is not much ELT training for ELT given to the YL teachers. Needless to say there is little or none for English literacy instruction. As Moon (2008, p. 400) points out, given the general lack of guidance in ELT materials available both globally and locally, YL teachers are usually not in a position to 'help children to acquire L2 literacy skills' because they have not received enough training in how to teach L2 literacy.

There is certainly a need to formalize the teacher education and the evaluation system to suit to the local context of Education in Japan. This calls for an understanding of the

situation on the ground; and also for input from current teachers, as it is important to focus on the local context and reflect it in suggesting any educational improvement.

2.6 Cultural education and literature

This section focuses on the relationship between literature and culture and discusses the possibility of using children's literature for the introduction of culture in Japanese primary ELT settings. This is important since one of the aims of the 'Period of Integrated Study' (MEXT, 2003) and 'Foreign language activities' (MEXT, 2008) is international understanding, including the understanding of different cultures. Brewster, Ellis and Girard (2002, p. 147) identify a trend in national policies in countries such as Taiwan, or in the European Union (ibid, p. 146) towards putting an emphasis on education in foreign cultures as a result of globalization.

It is a generally accepted view that teaching a language involves teaching its culture and therefore they are inseparable (Dogancay-Aktuna, 2005; Holliday, Hyde, & Kullman, 2004; Nazari, 2007; Scollon & Scollon, 1995). It is important to consider possible ways of making young learners of English aware of cultural elements. With regard to the cultural education of young learners, Jones (1995 cited in Brewster, et. al. 2002, p.146) suggests that cross-cultural understanding is likely to help young learners become aware of their own culture. This seems to have a practical implication, since comparing one's own culture with others' is likely to promote better understanding of one's own. The most popular approach at present, as has been the case for the current context of teaching, is to introduce foreign cultures by means of activities or communication

practices with ALTs (Shiozawa, 2005, p. 66). However it is clear that the degree of variety which can be presented to learners is limited.

According to Kubanek-German (2000 cited in Brewster et. al. 2002, p.147) there is no clear research evidence as yet to show ‘how children form ideas and judgements about other countries and cultures’, ‘the impact of foreign language lessons on young learners’ intercultural awareness’ nor the process by which judgements are formed. This is especially true the further apart the culture of the learner is from that of the target language (Hofstede, 2001). In order to ‘overcome cultural distance’ Swaffar (1992 cited in Kramsch, 1993, p. 225) suggests exposing students to texts from popular literature in the foreign language. Swaffar gives two reasons for this:

[E]asy stories and familiar plots give the students an incentive to read; moreover they can easily recognize the social stereotypes on which these texts are based and thus develop a critical apparatus to perform ... ‘a cultural reading’ of any other, more sophisticated text in L2 (Swaffar, 1992 p.225).

That is to say, instead of presenting individual cultural information, students are likely to learn and form by themselves a notion of the culture by matching ‘checklists of cognitive typologies or universal networks of meaning with specific texts’ (Kramsch, 1993 p.225). A similar approach is suggested by Brewster et al. (2002, p. 148): teachers should take an ‘ethnographic’ view and encourage students to discover a culture by themselves. Incorporating literature in cultural education seems well accepted in this context (Brewster, et al. 2002, pp. 147, 150), especially at the introduction stage. Kramsch (1993, p.175) also suggests a close connection between cultural education and literary texts as follows:

By constructing with the literary text a reality different from that of texts of information, students are given access to a world of attitudes and values,

collective imaginings and historical frames of reference that constitute the memory of a people or speech community. Thus literature and culture are inseparable (Kramsch, 1993, p. 175).

In addition, Wade (1990, p. 15) recognises the importance of stories as a part of natural experience for children. He suggests that storytelling is ‘strongly associated with pleasure and gives satisfaction through its capacity to shape and order’ children’s learning.

Kramsch (1993, pp. 205-206) suggests four different notions with regard to the teaching of language and culture: ‘1) establishing a “sphere of interculturality”; 2) teaching culture as an interpersonal process; 3) teaching culture as difference; and 4) crossing disciplinary boundaries.’ The implementation of these four points is likely to promote what Jones (1995 cited in Brewster et al. 2002, p.146) calls ‘awareness of others’, which helps to develop an understanding of such cultural issues as:

- 1) Similarities and differences between groups of different language communities
- 2) Social conventions
- 3) Things which are unfamiliar within a target language community
- 4) Stereotypes, as perceived by one linguistic group about another
- 5) Language as culture

For foreign language learners in particular, Kramsch (1993, p. 131) suggests the following:

[F]oreign language learners have to be exposed to different types of texts, from the most conventional to the most particular, but if they are eventually to find their own voice in the foreign language and culture, literary texts can offer [learners] models of particularity and opportunities for the dialogic negotiation of meaning.

However, as Kramsch (1993, p. 176) also warns us, it would be ‘reductionist to view culture as constituted only of the particular narrative voices of literary texts’ since literary texts are not to be read as something which represents only a given social and cultural context. She suggests the use of ‘culturally authentic texts’ as well as literary texts in the teaching of the cultural context of language, in the larger sense. As discussed in Section 2.2.3, the importance of a balanced input and various texts in promoting ‘intercultural literacy’ is obvious.

Children are often exposed to fictional texts in their first and foreign language when they are introduced to reading in primary school settings (Linse, 2007). Storybooks serve not only as a source for linguistic input, but also as a source for cultural information, providing a meaningful context at the same time. As Ghosn (2002, p. 173) remarks, good stories for children almost always have a moral which could also cultivate their ‘intercultural and interpersonal attitudes’. Moreover, exposure to fictional texts such as children’s literature tends to transform children’s culture (Trousdale, 2007) and to reveal the way in which children are represented and socialised in the culture. Therefore it can be claimed that it is efficient to use literary texts as a starting point for the introduction of literacy as well as for the cultural information which literary texts might reveal. Knowles and Malmkjær (1996, p.263) state that ‘a fictional world may distort “reality”, [but] the distortion is significant because it imposes this ‘reality’ upon its reader.’ This is another possible use of fictional texts: to create an imaginary world and make children aware that ‘what you read can be a distorted version of reality’ by contrasting it with the ‘reality’ experienced by the young learners themselves.

Obviously, the appropriateness of the content of the culture should be taken into account for different participants. For instance, it would not be appropriate to introduce foul language or swearing in this context, though both might be claimed to be part of an existing culture involving 'real' language and context. In this regard, texts written specifically for children are not likely to include expressions which are inappropriate for them. Moreover children's literature seems to be relevant to choosing what to teach, bearing in mind children's perspectives and their importance, as stressed by Brewster et al. (2002, p. 148).

As Kramsch (1993, p.130) indicates, 'the major difficulty encountered by language learners when they deal with written texts stems from the passage from a more orate to a more literate mode of speech.' This is particularly true in the case of children learning a foreign language. It is important to understand the characteristics of the text when proposing that texts written for children should be incorporated in cultural education. With regard to the use of 'real-life materials', Kramsch (1993, p. 188) raises another important point: 'culture is a reality that is social, political and ideological' and therefore, 'viewing the world from another perspective' can be more difficult than 'grasping another lexical or grammatical code.' However, the ways in which 'culture' is represented in the text should be considered and so should identification of these ways. Therefore the present study will consider those two points when considering the design of a syllabus for young learners. They will be addressed in more detail in Chapter 3.

2.7 Summary

This chapter has reviewed current primary English education in Japan and official policy with regard to the introduction of English language teaching in primary schools. The current teaching context in primary schools was discussed with reference to recent policy, together with the possibility of using children's literature as a part of cultural education as well as in English literacy education. The discussion stressed the importance of paying attention to local and contextual educational settings, since this is directly involved with the successful application of the research findings.

Three areas of challenges with regard to the successful implementation of English education at primary schools are identified. One is to establish a teaching syllabus which comprises not only the focus of spoken English but also the written element of English in order to enhance wider learning chances for YLs. Second is to consider a variety of teaching methodologies suitable for YLs in Japan. The third challenge is related to teacher education. As has been mentioned, unlike other subjects, there had been no teacher education regarding English education at primary level until recently, since teachers did not have responsibility to teach English until 2009. Considering the current teacher training system, it is important to establish an effective training system aimed at all current teachers and prospective teachers. In order to do so, it is necessary to take into account what the local teachers perceive to be problems in practice, in order to identify real challenges and suggest possible solutions.

In particular, considering the lack of literacy education in the existing guidelines provided by MEXT (2002; 2003; 2008), and the constraints inherent in the policy,

corpus linguistics could have an impact on ELT: since it challenges the view of a grammar and lexis boundary (i.e. a phraseological view of language) it might inform effective teaching. Therefore, in the following chapter (Chapter 3), the relationship between corpora and the ELT classroom is discussed.

CHAPTER THREE

CORPORA IN THE CLASSROOM

3.1 Introduction

This chapter reviews the relationship between corpora and English language teaching. In order to explore the possibility of a corpus-influenced syllabus for teaching literacy to young Japanese learners of English, existing literature on syllabus design for ELT will be reviewed first. This is followed by a discussion of the vocabulary learning and teaching in ELT, and a review of how the selection of vocabulary is conducted in designing a syllabus. In order to assess the potential for the design of a corpus-influenced syllabus in primary English education in Japan, issues associated with corpora and syllabus design will be reviewed, including the emergence of a notion of a corpus-based lexical syllabus and its subsequent development in English pedagogy. Some brief discussion of issues related to grammar teaching and learning are also provided. To illustrate the impact of corpus linguistics on the language description and ELT, (in particular, the shift of focus from syntax to lexis), this section includes a brief account of the history of vocabulary teaching in ELT since the 1920s. This was a time when vocabulary was considered to be one of the most important aspects for the successful learning of a foreign language (Richards & Rodgers, 2001, p. 37). After this account, there is a review of the integration of corpus, literature and language teaching by referring to previous studies which used literature corpora in relation to English language pedagogy in both native and English as a Foreign/Second Language (EFL/ESL) settings. Some relevant studies will also be reviewed involving linguistic

investigations of children's literature carried out using corpus-based methods. The end of this section will discuss ways in identifying the pedagogically useful items for language teaching, and suggests the direction of this study.

3.2 Syllabus design

3.2.1 Definition of syllabus - syllabus as a part of materials design

Since the current study aims to investigate the possibility of constructing a syllabus for English literacy teaching for Japanese young learners, it is worth defining what we mean by a syllabus.

White (1988, p.4) provides the following distinction between the terms 'syllabus' and 'curriculum':

In a distinction that is commonly drawn in Britain, 'syllabus' refers to the content or subject matter of an individual subject, whereas 'curriculum' refers to the totality of content to be taught and aims to be realized within one school or educational system. In the USA, 'curriculum' tends to be synonymous with 'syllabus' in the British sense.

Nation and Macalister (2010, p. 5-10; 2011, p.1-2) also view 'syllabus' as a part of curriculum design. They suggest a syllabus consists of the following: 1) goal, 2) content and sequencing, 3) format and presentation; and 4) monitoring and assessing. Similarly, Brown (1995, p. 14) defines syllabuses as 'ways of organizing courses and materials'. Following White (1988), Nation and Macalister (2010; 2011) and Brown (1995) above, the definition of 'syllabus' which places 'syllabus in a subordinate position to curriculum' (White, 1988 p. 4) is employed in this thesis.

White (1988, p. 44) suggests that there are largely two kinds of syllabus for language teaching: 'Type A' and 'Type B' syllabuses. 'Type A' syllabus mainly concerns itself with 'what is to be learnt?' (i.e. product-based) and Type B' mostly relates to 'how is it to be learnt?' (i.e. process-based). There have been many types of syllabus designed and proposed for ELT. He identifies the following types within his two main divisions:

Type A syllabus: grammatical syllabus, functional-notional syllabus, lexical syllabus situational syllabus, topic-based syllabus, skills-based syllabus

Type B syllabus: process syllabus (negotiated syllabus), procedural syllabuses (task-based syllabus).

The following sections, 3.2.2 and 3.2.3, review the major syllabuses of Type A and Type B in relation to their design and development.

3.2.2 Type A syllabuses

3.2.2.1 Structural/grammatical syllabus

One of the classical traditions in Type A syllabus design in ELT was the structural (grammatical) syllabus. This is a syllabus whose organising principle is structural and grammatical information. It consists of a list of structural items, which are arranged in order of the intended instruction. The structural syllabus is 'traditional' (Yalden, 1983) in the sense that the methods which dominated foreign language teaching from the 1840s to 1940s - grammar translation and audio-lingual methods - are based on this syllabus (Richards & Rogers, 2001 p.6).

Although the structural syllabus was long established as the traditional form of syllabus, it received a number of criticisms. It is recognised as not suitable for teaching students how to use language, as the examples in the structural syllabus are often contrived for

the purpose of grammar presentations and do not provide realistic discourse with real communicative functions (White, 1988). Gaies (1977, p. 211) notes that examples contained in the grammatical syllabus are ‘too often semantically divorced in every way from the classroom surroundings or the student’s experience, past and present’. Likewise, Krashen (1980, p.7) claims that a ‘grammatical focus will usually prevent real communication using the second language’. As SLA research progressed, the structural syllabus was criticised for not corresponding to how learners acquire language (Ellis, 1993 p. 91).

O’Dell (1997, p. 259) describes the shift from the grammatical syllabus in the 1970s as follows:

Attention began to be paid to the fact that EFL students might master a grammatical or structural syllabus – and get good marks in exams based on such syllabuses – while still being incapable of carrying out the simplest everyday task in the foreign language; typically they could translate quite complex literary texts and do slot-and-filler exercises but they could not cope with the realities of everyday situations.

Given the inadequacies of the grammatical syllabus, and the move towards a communicative approach in the 1970s and 1980s, the search for alternatives led to various kinds of syllabus being proposed. Nevertheless, as Wilkins (1976, p. 7) notes, ‘[t]he use of a grammatical syllabus can be regarded as the conventional approach to language teaching since the majority of syllabi and published courses have as their core an ordered list of grammatical structures’. It can be said that grammatical and structural information continues to play a part in syllabus design, despite the recent unpopularity of grammatical syllabuses.

3.2.2.2 Topic-based/ Theme-based syllabus

The organising principle of topic-based syllabuses is ‘topic’ or ‘theme’, and such syllabuses integrate language through topics or themes (Brown, 1995 p. 9). Unlike the structural syllabus, ‘a grammatical point to be taught is linked to an interesting theme or practised in a ‘real world’ setting rather than learning mechanically and outside any context’ (McDonough & Shaw, 2003 p.12).

White (1988, p.66) mentions one of the benefits of the topic-based syllabus is related to raising learners’ motivation. McDonough and Shaw (2003, p. 51) also note that topics ‘are the most obvious ways in which learners’ needs and interests can be taken into account.’ In selecting topics, the main criteria include ‘interest, need, utility and relevance’ as well as ‘the interests of the syllabus designer’ (Fein and Baldwin, 1986, p. 2). In relation to the benefits, Bourke (1996) champions the topic-based syllabus for teaching English to young learners. He contends that ‘topics provide a natural context for the integration of language input and skills development’ (Bourke 1996, p. 282). His suggestion is that the topic determines language items such as structures, functions and vocabulary, as well as keeping tasks on-topic. Despite stressing the value of the topic-based syllabus, Bourke (1996, p. 284) also recognises the need for integration of language input and skills development, as well as the importance of grammar teaching.

While there are advantages associated with topic-based syllabuses, they are not without their drawbacks. For instance, since the language is selected around the topic, it can lead to a lack of fine-tuning of language exposure (White, 1998 p.68). Since the term ‘topic’ is a difficult concept to define without ambiguity, unlike grading structures with

linguistic criteria, topics have a less clear theoretical foundation which can be used to grade them. Fein and Baldwin (1986, p.2), in their design of a topic-based syllabus for a pre-sessional course in the USA, applied three criteria for selecting content: 1) pedagogic merit, 2) affective consideration, and 3) practicality.

As McDonough and Shaw (2003, p.51) rightly point out, it is important for syllabus designers to realise the difficulties of ‘meaningful classification’ of the topics and that it is necessary for a syllabus designer to consider the relevance of the topics to the learners and whether the learning context influences the choice of topics. Nevertheless, in spite of some drawbacks associated with the topic-based syllabus, White (1988, p. 65) points out that, notwithstanding the rarity of purely topic-based syllabuses, most Type A syllabus designers have recourse to topic as a category.

3.2.2.3 Situational syllabus

According to McKay (1978, p. 11), ‘situational syllabuses are based on the idea that language is found in different contexts, or situations’. The organising principle of such syllabuses is ‘a series of situations’ and in creating such a syllabus, the following factors are considered in relation to the aspects of situation: 1) ‘the setting’, 2) ‘the participants’, and 3) ‘relevant objects within the setting’. After the situations are specified, appropriate language is selected according to the situations concerned, and it is often sequenced in terms of chronological order or the similarity of the situations (Brown, 1995 pp.8-9; White, 1988 pp. 63-64).

It is clear that the emphasis of the situational syllabuses is on using the language rather than learning grammar. However, this is not to say that the language system should be

dismissed entirely, given its use in ‘generating further utterances of the type represented in the model.’ (White, 1988 p. 64). Similarly, the use of structural information in grading the situations may be used as an alternative, due to the absence of well-defined criteria and the difficulties associated with grading situations.

Brumfit and Johnson (1979, pp. 83-84) point out the difficulties involved in defining ‘situation’ itself, and alert us to the danger of overtly situation-specific language which may alienate some students. Moreover, Wilkins (1976, p.19) criticises the situational syllabus for limiting learners’ exposure to a range of functions by pre-selecting certain situations.

These factors may be the reasons for the rarity of materials which are purely based on situations. Nevertheless, for most syllabus designers, ‘situation’ is one of the important elements in syllabus design, as long as authenticity and practicality are kept in mind (White, 1988 p. 64-65).

3.2.2.4 Skills-based syllabus

According to Brown (1995, p. 11), materials in a skills-based syllabus are organised ‘around the language or academic skills’ that are thought to be necessary for students by syllabus designers. It is an ELT convention for ‘language skills’ to mean the ‘four skills’ which are usually divided into receptive (i.e. listening and reading) and productive (i.e. speaking and writing) skills. These four skills are usually treated almost equally in the syllabuses for English for general purposes. The appropriateness for all learners of this ‘equal weighting’ of four skills has been questioned after the emergence of ESP and the concept of needs analysis (White, 1988, p.68). For instance, White (1988, p. 70)

dismisses the idea that linguistic skills are all one needs in order to be able to use the target language. As a result of research in applied linguistics, psychology and education in the late 1970s and 1980s, new insights from these areas shifted the ‘very nature of reading and writing’. Moreover, there was also a shift in the teaching of reading and writing from product-focused to process-focused. Based on Munby’s (1978) taxonomy of language skills, White (1988, p. 73) offers the second categories of skills (other than the division of ‘four skills’) as follows: ‘language skills, cognitive skills, composition skills, and study skills.’ As Brown (1995, p. 11) points out, the skills are selected based on what syllabus designers consider useful for the learners, and the sequencing is decided ‘based on some sense of chronology, frequency, or relative usefulness of the skills’. In relation to the selections and sequences of the skills, the difficulty of such skills-syllabuses lies in deciding how and what teachers evaluate. White argues that even a syllabus based upon sub-skills grouped together would provide only one part of a language syllabus, as it is necessary to bear in mind other features, such as form and meaning (White, 1988 p. 73).

3.2.2.5 Functional-notional syllabus

Having recognised the problem of students’ inability to deal with everyday situations as a result of grammar-translation teaching, language planners started to focus on the communicative aspect of language. In the 1970s, with the rise of the idea of a notional-functional approach, two important elements were introduced in syllabus design: one was a notional aspect concerned with ‘abstract conceptual categories’ such as ‘*distance, duration, quantity, quality, location, size* and so on’; while the other was a functional aspect concerned with semantic uses, or purposes for using language (e.g. ‘*greeting people, introducing someone, changing a topic*’) (Brown, 1995 p. 10).

The concept of using notions and functions as organising principles was based on the work of the Council of Europe; *Threshold* (Van Ek 1975) and *Waystage* (Van Ek & Alexander 1977). This was a time when syllabus organization would no longer be decided only by grammatical categories, but would also focus on communicative competence.

Wilkins (1972; 1976) proposed a meaning-based syllabus which organized the teaching and materials by categories of basic notions and functions, in order to help learners to gain communicative competence. He identified three categories of notions: 1) semantico-grammatical (e.g. past, future, location), 2) modality (e.g. possibility, necessity, obligation), and 3) communicative function (e.g. asking questions, expressing agreement and disagreement, inviting, accepting, declining). According to O'Dell (1997, p. 259), the lexical content of this syllabus was chosen in order to illustrate the functions and notions so that the contexts were similar to those which might be encountered by the learner in an English-speaking environment.

White (1988, p. 18) points out that Wilkins' work on *Notional Syllabuses* (1976) is merely a list and hardly a syllabus in the sense of 'guide to teaching' without providing important elements such as indications of grading and sequencing. He argues that the selection and grading of functions and notions tends to rely on intuition considering the learners' needs both in classroom and the real world (White, 1988 p. 82). Widdowson (1979, p. 254) is also critical of the notional-functional syllabus, and argues that notional-functional categories provide:

only a very partial and imprecise description of certain semantic and pragmatic rules which are used for reference when people interact. They

tell us nothing about the procedures people employ in the application of these rules when they are actually engaged in communicative activity. If we are to adopt a communicative approach to teaching which takes as its primary purpose the development of the ability to do things with language, then it is discourse which must be at the center of our attention.

Di Pietro (1981, p. 27) points out some problems with notional-functional syllabuses. He points out the difficulties of distinguishing between 'language uses which are ritualized and those which are psychologically and culturally diverse'. He further argues that '[i]n order to adapt the functional-notional syllabus to a structured classroom experience, essential information must be provided on the cultural and psychological relevance of individual speech functions' (Di Pietro, 1981, p. 27).

Moreover, it was suggested that notional and functional syllabuses merely replaced the traditional grammar syllabus (Brumfit, 1980). It is more difficult to specify, describe, systematize and generalize from notions and functions than from structural information. Furthermore, as it is certain that various functions can be expressed by structures in different frequency and complexity, there is an inevitable conflict between what is clear and learnable in terms of structures, and what is useful and relevant in terms of functions (Brumfit, 1980; White 1988).

Peters (1983) suggests that SLA findings on how learners process and store functional unanalysed chunks and sequences in memory separately from grammatical analysis, thus using two different 'strands', may have implications for the language teaching syllabus. White (1988, p. 83) also suggests that providing learners with a hybrid syllabus involving both functional and grammatical work could be a valuable

compromise in dealing with the drawbacks of notional-functional syllabus design for the following reasons:

[T]he learner would be presented with contextualized chunks of unanalysed language which could be used for communicative purposes in interactive practice. Subsequently, the student would be introduced to exercises in which the chunks already known and used would be analysed grammatically, and opportunities for producing new variations on the elements concerned would be provided.

As seen in the review of other syllabuses, it seems that any approach to ‘pure’ syllabus design in an absolute vacuum would appear impractical. Around the same time as the development of the functional-notional syllabus, new perspectives of looking at language highlighted by corpus analysis influenced a new kind of syllabus. The next section reviews the lexical syllabus.

3.2.2.6 Lexical syllabus

The lexical syllabus, as stated, is a kind of corpus-based syllabus; its basic concept was suggested by Sinclair and Renouf (1988) and was further developed by Willis (1990) (Hunston, 2002, p. 189). Sinclair and Renouf (1988, p. 159) distinguish the lexical syllabus from other syllabuses as having the characteristic of independence from other methodologies in controlling explicitly ‘what is to be taught in a language course’ (ibid, p.159). They state that the following three points should be focused on in the syllabus:

- (a) the commonest word forms in the language;
- (b) the central patterns of usage;
- (c) the combinations which they usually form (ibid 1988, p. 148).

It has been suggested that what determines the choice of certain word forms in the content of the lexical syllabus is mainly their frequency (Willis, 1990). However, Sinclair and Renouf (1988, p. 160) also suggest that words which are related to ‘everyday domestic reality’ (e.g., colours, days of the week, etc.), specifically, words

which learners are likely to encounter in everyday life, should also be added to the list, regardless of their level of frequency (Sinclair & Renouf, 1988, p. 151). Willis (1990, p. vi) defines the lexical syllabus as follows:

The lexical syllabus does not identify simply the commonest words of the language. Inevitably it focuses on the commonest patterns too. Most important of all it focuses on these patterns in their most natural environment. Because of this, the lexical syllabus not only subsumes a structural syllabus, it also indicates how the 'structures' which make up that syllabus should be exemplified. It does this by emphasising the importance of natural language.

The lexical syllabus, therefore, deals with the most frequent words in a corpus as well as their associated meanings and the patterns of the lexical items (Sinclair & Renouf, 1988, p. 148; Willis, 1990, p. vi). Hunston (2002, pp. 190-191) also provides a useful overview of the description of a lexical syllabus, as follows:

A description of the syllabus would, in effect, be a description of the corpus. If the syllabus was expressed as a list of items, it would be as a list of the most frequent word-forms in the corpus, along with their most typical phraseologies.

This view is related to the notion of 'pattern grammar' (Hunston & Francis, 1998, 1999; Hunston, Francis & Manning, 1997), which also counters the traditional distinction between lexis and grammar. Patterns are closely related to the way in which words appear in structures, since 'grammar patterns' can be identified in respect of all verbs and most nouns. Moreover, words belonging to the same pattern can normally be classified into groups of words which are similar in meaning (Hunston & Francis, 1999, p. 270). On these principles, Hunston and Francis (1999, p. 262) suggest that 'the role of pattern is as input and support to the concept of grammatical consciousness-raising'.

Therefore, Willis's suggestion of treating grammar as lexical patterning seems reasonable, since 'patterns attach to all lexical items in the language [and hence] learning the lexis means learning the patterns and therefore the grammar' (Hunston, 2002, p. 190). Willis (2003, p. 48) introduces the idea of a 'grammar of class' as a 'grouping of words which are linked by the way they behave and the patterns in which they occur'. This view is also supported by Aston (2000, p. 9), as follows:

Language teaching needs to draw attention to the specific colligational patterns associated with particular semantic categories and with particular lexical items, rather than attempting to provide 'slot and filler' combinatory rules which are based on grammatical categories alone.

If a syllabus was designed following the suggestions made by Sinclair and Renouf (1988) and Willis (1990), it should subsume a structural syllabus, as well as describing how the 'structures' which make up the overall syllabus are used in practice. Moreover, as has been suggested, the lexical syllabus should be an 'independent syllabus, unrelated by any principles, to any methodology' (Sinclair & Renouf, 1988, p. 155). This means that whatever kind of methodology is employed, the syllabus could be adapted to it and this would also encourage flexibility in the design of the teaching materials based on the syllabus in question.

Hunston (2002, p. 190) comments that corpus linguistics would effectively demonstrate what should be taught to learners if the corpus consisted of 'authentic language that contains instances of the most frequent patterns of the most frequent words'. It is important, then, to be sure what to include in a corpus (Willis, 1990, p. vii). In creating a corpus-based syllabus, there are a number of points which should be carefully considered by the syllabus designer, all of which will be discussed later in this thesis (Chapter 4).

In addition, a lexical syllabus embodies an attempt to consolidate the learner's lexical knowledge through recycling and reviewing previously taught prefabricated units, lexis and phrases (Zimmerman, 1997, p. 16). This is one of the most important principles of the lexical syllabus, stressing the need for reviewing and recycling lexical items taught throughout a language course (Harwood, 2002, p. 139; Willis 1990, p. vi).

3.2.3 Type B syllabus

As mentioned in Section 3.2.1, Type B syllabuses can be sub-divided into two types: the process-based syllabus and the procedural syllabus (White 1988, p. 94). White notes that these are 'methodology-based syllabuses' relating to one another not in terms of content but rather in methodology.

3.2.3.1. Negotiated syllabus

One form of the Type B syllabuses is the negotiated syllabus, which is also known as the process-based syllabus. White (1988, p. 96) points to the educational rather than linguistic priorities which form the basis of such a syllabus. The process syllabus (i.e. negotiated syllabus) is often regarded as 'learner-led', as it places importance on the students' learning preferences.

Candlin (1984, p. 34) argues for the importance of creating a syllabus 'which is social and problem-solving in orientation' as opposed to one built around 'preselected and often pre-digested knowledge'. According to Candlin (1987, p. 6), teachers and learners plan the direction of a course jointly which leads to the design of a syllabus containing the following: 1) learning, 2) content, and 3) action.

Breen (1987, p. 166) explains the two syllabus components as the following:

- (i) a plan relating to the major decisions which teacher and learners need to make during classroom language learning, and
- (ii) a bank of classroom activities which are themselves made up of sets of tasks

Clarke (1991, p. 15) states that '[a] negotiated model is thus different in kind to *all* other syllabus types in that its content is entirely unknown prior to its creation.' Depending on the negotiation with learners, the content might turn out to be something similar to the other types of syllabuses, or a combination of them.

Although this learner-centred and process-based approach to syllabus design seems ideal in supporting learners depending on their needs, it is not without its drawbacks. In an extreme model of a negotiation syllabus, learners are expected to get involved fully, from planning to evaluation. White (1988, pp. 101-102) identifies five problems with the negotiated syllabus. The first problem is the difficulty of evaluating such a syllabus. Secondly, he points out that teachers need to be very professional and confident in implementing a process syllabus. The third criticism relates to a potential clash of the negotiated syllabus with cultural norms of teacher and student dynamics. This is particularly the case in a so-called high power distance society (Hofstede and Hofstede, 2005, pp. 53-55) such as Japan. The fourth problem is that not only the teachers but also the students need to change their roles. In addition, some teachers also may be unwilling to abdicate large areas of authority. Fifthly, the negotiated syllabus could render textbooks unusable. Moreover, objectives are often decided by authorities other than teachers, as is the case in Japan, and this places both teachers and learners in a position whereby they cannot participate in the process of decision making, thus negating the practicality of the process-syllabus in most situations.

In the light of the constraints above, Clarke (1991, p. 26) stresses the importance of the concept of negotiation to the design of a syllabus by allowing students a certain degree of choice and self-expression. Therefore, he proposes that negotiation may be incorporated in each syllabus component, rather than being rejected entirely.

3.2.3.2. Task-based syllabus (procedural syllabus)

The development of the procedural syllabus is greatly influenced by the work of Prabhu (1987) on the 'learning centred' approach to language teaching. This procedural syllabus is based on the idea that form-related learning should take place in the context of close attention to meaning. The main difference between the process-based syllabus and the procedural syllabus is that the procedural syllabus is influenced by SLA research stressing the concept of 'learning-centred' as opposed to 'learner-centred', and uses tasks to involve students in a process focusing on an achievement (i.e. task completion) as opposed to explicit language learning. According to Prabhu (1987, p.86), the procedural syllabus can be defined as the following:

The syllabus is a form of support for the teaching activity that is planned in the classroom and a form of guidance in the construction of appropriate teaching materials. It is concerned, from this point of view, with what is to be done in the classroom, not necessarily with what is perceived to be taught or learnt thereby; its role is essentially to make it possible for one teacher to draw on the experience of another.

In the task-based syllabus, learners are expected to focus on 'meaning-content' of the task and the focus on the language is planned to occur in completing the task.

According to Prabhu (1987, p. 69-70), the task based teaching

operates with the concept that, while the conscious mind is working out some of the meaning-content, some subconscious part of the mind perceives, abstracts or acquires (or recreates, as a cognitive structure) some of the linguistic structuring embodied in those entities, as a step in the development of an internal system of rules. The intensive exposure

caused by an effort to work out meaning-content is thus a condition which is favourable to the subconscious abstraction – or cognitive formation – of language structure.

Moreover, language errors produced by students are corrected through ‘incidental correction’ (e.g. repairs) (Prabhu, 1987 p. 61), instead of providing learners with a rule of grammar.

Task-based language teaching has been supported by a number of writers (e.g. Nunan, 1989, 2004; J. Willis 1996; Skehan 1998; Lee 2000; Bygate et al. 2001; Ellis 2003). Nunan (2004, p. 11) suggests that task-based syllabuses ‘do not rely on prior analysis of the language into its discrete points’. Nevertheless, he goes on to emphasise the importance of the teachers’ own judgement in selecting tasks, so as to maximise the variety of structures, especially in the case of elementary level teaching.

According to Nunan (2004, p. 12), task-based teaching places ‘experiential learning’ as an important conceptual basis, and uses ‘the learners’ immediate personal experience as the point of departure for the learning experience.’

Long and Crookes (1992, p. 47) point out three problems inherent in the task-based syllabus. The first lies in the difficulty of differentiating tasks (and the subtasks which make them up), raising questions about their ‘finiteness’ and ‘generative capacity’. Another is concerned with the issue of the difficulty of tasks, that is in determining the criteria for grading and sequencing. Long and Crookes (1992, p. 47) point out that these are the unresolved issues for synthetic syllabuses, despite many past discussions over criteria such as frequency and difficulty. The third problem is related to the lack of

thorough field evaluation of task-based teaching, which they feel is being hindered by lack of funding.

In the light of the problems described above, the task-based syllabus is still under development and more research needs to be done on foreign/second language acquisition and language teaching in order to overcome the challenges which still persist.

3.2.4 Summary: Integrated syllabus and teacher training

While the types of syllabus reviewed in the above section are all different, each syllabus type has benefits and drawbacks. As is often the case with any kind of teaching syllabus, it is hardly likely to produce courses or materials which universally suit different needs and requirements. One common feature of all of the approaches to syllabus design outlined above is the recognition of some form of integration, including attention to grammar learning (e.g. Bourke 1996; White 1988; Wilkins 1976). As O'Dell (1997, p. 272) rightly notes:

Topics, notions, functions and grammar, therefore, all have a useful part to play in the vocabulary syllabus and it seems misguided to focus too much on one to the possible exclusion of the others. (O'Dell 1997, p. 272)

Brown (1995, pp. 12-14) also recognises that 'divergences' from the pure forms of syllabuses occur largely in two ways: 'mixed syllabus' and 'layered syllabus'. He describes mixed syllabuses as those produced by authors who intentionally mix two or more kinds of syllabus. A layered syllabus, on the other hand, consists of a dominant syllabus with other forms of syllabus operating beneath it. White (1988, p. 111) also notes:

Language teaching is concerned with more than a choice of the ‘best’ syllabus, and an excessive amount of time and effort can be devoted to the selection and ordering of content, while giving insufficient attention to questions of methodology; to the numerous factors which should influence choices of design and procedure; to the practical issues of implementation.

In relation to the practical issues of implementation of a syllabus, as reviewed in Section 2.5.2, it is important to remember that the direct implementations of what was successful with adults may not be the same when working with young learners. It is necessary to pay attention to not only the specification of the content, but also the practical and methodological aspects which suit the target learners and their learning environment as well as their cultures. It is important for syllabus designers to consider the practicality of the syllabus, as well as meeting the targeted learners’ needs.

Although most teachers may not need to (or, indeed, have the authority to) establish a syllabus from scratch, knowing about the origin of different syllabuses helps to give them insight into what they are doing, and make adaptations to the material they are given to work with as suggested by Brown (1995 p. 14). Research suggests (e.g. Tsui, 2003) that novice teachers often find it difficult to adapt materials to suit learners’ individual needs. Block (1991) suggests that material design is part of the responsibility of each teacher who needs to be accountable for what happens in their class, also enabling them to contextualize the language in a more accessible way for learners. It is necessary for teachers to be equipped with the ability to rearrange the content or materials in order to adapt them to the needs of learners (Samuda, 2005 p. 235) as, clearly, they are in the best position to assess the individual requirements of their own students. In considering how the content of a syllabus can be put into practice by teachers, incorporating syllabus and material design in teacher training programmes, as

suggested by many researchers (e.g. Block 1991; Harwood 2010; Hutchinson & Torres, 1994; Tomlinson 2003), is another aspect of teacher education which has been overlooked so far in Japan, but should not be ignored.

3.3 Learning and teaching of vocabulary

As basic learning at primary level is concerned mostly with learning vocabulary, issues associated with teaching and the learning of English vocabulary are reviewed here. Firstly, I will review how the learning of vocabulary is viewed in ELT by exploring incidental and explicit learning of vocabulary, how this relates to gaining knowledge of meaning and grammatical knowledge, as well as the strategies involved in the learning of vocabulary. I will then review the different approaches suggested for the teaching of vocabulary. In order to explore the potential of a corpus-influenced syllabus for the teaching of young EFL learners in Japan, I will also review the emergence of corpora influenced language description, which is challenging the traditional distinction of lexis and grammar, while offering a historical perspective on the treatment of lexis in ELT. After this account, the selection of vocabulary for a syllabus will be discussed by revisiting the syllabuses reviewed in 3.2 and seeing how each syllabus dealt with lexis. In particular, corpus-based syllabuses (e.g. lexical syllabuses) will be reviewed, with particular reference to the way in which the work of corpus linguistics has been reflected in the design of the ELT syllabus.

3.3.1 Learning of vocabulary

There are largely two main processes involved in vocabulary acquisition. One is ‘explicit learning’, in which learners learn lexical items ‘through the focused study of words’. The other is ‘incidental learning’ in which learners learn words ‘through

exposure when [their] attention is focused on the use of language, rather than the learning itself' (Schmitt, 2000 p. 116). A variety of research has shown that learners, whether they are learning their L1 (e.g. Jenkins et al., 1984; Nagy et al., 1987; Nagy et al., 1985; Shu et al., 1995) or L2 (e.g. Day et al., 1991; Dupuy & Krashen, 1993; Hulstijn 1992; Pitts et al, 1989; Waring & Takaki, 2003), may obtain knowledge of meaning by reading texts. According to Nation (2001), the majority of vocabulary learning in the L1 is through incidental learning. In the case of vocabulary in L2, some researchers (e.g. Laufer, 1991, 2001; Laufer & Paribakht, 1998; Schmitt, 2000; Webb, 2008) stress the need for explicit learning. Nevertheless, researchers agree that encouraging and incorporating both incidental and explicit learning of vocabulary is necessary in L2 learning (e.g. Nation, 2001; Schmitt, 2000; Waring & Takaki, 2003).

In considering how words are learned, it is necessary to consider the acquisition of different lexical aspects. The most studied aspects are word meaning and grammatical/morphological knowledge. The three stages in which L1 children process meaning acquisition in their L1 are summarised in Aitchison (2003, pp.188-199) as: 1) labelling, which involves linking a word to a concept; 2) categorization, in which children group words into categories; and 3) network building, in which they connect related words.

After the acquisition of core meaning, children learn additional word knowledge, such as extension of word meaning and semantic boundaries, by further exposure to the target word in context (Schmitt, 2000 p. 123). Schmitt argues that this process is a continuous one, as each exposure helps learners define the word boundaries. Schmitt

further explains how children learning their L1 categorise words by ‘likeliness’ and the categorisation criteria are altered according to their development from ‘perceptual similarities’ (i.e. shapes) to ‘coordinate relationships’ (i.e. similarities in concepts). Through enough exposure to words and such a process, L1 learners clarify word boundaries, but the process of ‘learning new applications for [already known] words’ continues as a constant process even for native adults who are already literate. According to Schmitt (2000 p. 124), this is partly to do with the fact that ‘[m]any or most words in English have two or more meaning senses (polysemy)’ and people constantly encounter new meaning senses.

In terms of vocabulary learning in L2, the core meaning sense of the word seems to be learnt before figurative meaning senses as ‘progression in conceptual development impacts on vocabulary learning’ (Cameron, 2001 p. 94). L2 learners can make use of the knowledge of concepts acquired in their L1 in defining L2 word boundaries. However, L2 learners may have difficulties in learning related words as their L1 and L2 may not always have a clear one-to-one correspondence, but rather exhibit differing meanings as a result of cultural differences (Cameron, 2001 p. 74). In this case, learners are required to ‘disambiguate similar words that may have almost the same core meaning, but subtle differences’ (Schmitt 2000, p. 125). In the light of such complications involved in vocabulary development, Cameron (2001, p. 94) also stresses the importance of giving learners ample opportunities to use and consolidate the words for more effective early vocabulary learning.

As for the acquisition of grammatical knowledge, different types of morphology are acquired at different times. According to a study by Berko (1958, p.176-177), it seems that inflections and compounding are acquired prior to the acquisition of derivational suffixes. This may be explained by the fact that inflections are rule-based and thus easier to learn compared to derivational suffixes (Schmitt, 2000 p. 126). Moreover, Schmitt (2000, p. 126) adds that ‘children simply have more exposure to inflections than derivational suffixes’ as the latter appear more in the written language. Thus he further argues that children learn derivational suffixes incrementally as they become familiar with more written language. Although there is some evidence of the acquisition of derivations and inflections being facilitated by the pre-existing knowledge of a base-form (e.g. Freyd & Baron, 1982; Sandra 1993), Schmitt (2000, p. 126) argues that acquiring derivations should not be assumed to be easy or automatic. He points out that even native-speaker adults have difficulty in mastering morphology fully, and the likelihood of difficulties associated with morphological acquisition for L2 learners is obvious. Schmitt and Meara (1997) also found that their L2 learners (i.e. 95 Japanese learners of English in secondary and post-secondary schools) coped well with producing inflections in comparison to their performance on derivational suffixes. In addition, they also found that these learners generally lacked knowledge of suffixes, even for verbs that they recognised.

Nevertheless, in relation to the acquisition of lexical aspects, Schmitt (2000, p 127) points out that although there is some understanding of how word meanings and grammatical knowledge are acquired, he describes the acquisition of other lexical knowledge (e.g. acquiring register and collocation) as ‘unexplored mysteries’. Thus, it

can be said that there remains much to be done in the area of understanding how learners come to acquire full 'knowledge' of a word.

A number of vocabulary learning strategies (VLS) have been suggested over the years, along with the shift from teacher-oriented language teaching to the responsibility of learners for their own acquisition. Schmitt (1997 pp. 208-217; 2000, pp. 133-136) provides a detailed summary of VLS. Strategies are largely divided into two types: 1) discovery strategies, and 2) consolidation strategies. Discovery strategies are used for finding the meaning of the word, and consolidation strategies are used for remembering the word for future use. He classifies the strategies into five groups: 1) Determination strategies (DET); 2) Social strategies (SOC); 3) Memory strategies (MEM); 4) Cognitive strategies (COG) and 5) Metacognitive strategies (MET). The following table (Table 3.1) summarises these five groups of strategies described in Schmitt (1997 pp. 208-217; 2000, pp. 133-136).

Table 3.1 Five main groups of vocabulary learning strategies

Groups	Description	Examples
Determination strategies (DET)	Discovering the meaning of new words without depending on other people's expertise	<ul style="list-style-type: none"> • Guessing from context • Guessing from L1 equivalent • Guessing from structural knowledge • Use of reference materials
Social strategies (SOC)	Interacting with others in order to gain knowledge about language	<ul style="list-style-type: none"> • Asking teachers or classmates (translations, synonyms, etc)
Memory strategies (MEM)	Integrating previously gained knowledge with the word that needs to be remembered	<ul style="list-style-type: none"> • Grouping • Using images • Incorporating physical movement (i.e. Total Physical Response method by Asher 1977).
Cognitive strategies (COG)	Not only manipulating and transforming the target language, but also including repetition and mechanical ways to study vocabulary	<ul style="list-style-type: none"> • Keeping record in vocabulary notebooks. • Flashcards, wordlists which are to be used for initial input
Metacognitive strategies (MET)	Involve a conscious overview of the learning process and making decisions about planning, monitoring, or evaluating ways to optimise study. Reflecting on the learning process and searching for the best ways to study by planning, monitoring, or evaluating the process.	<ul style="list-style-type: none"> • Increasing exposure to input • Discovering efficient ways to study and review • Testing for improvement • Deciding the order of priority of words to be learnt or remembered

In relation to the DET (Determination Strategies) above, Webb (2008) researched whether contexts play a role in his learners' gaining knowledge of meaning through incidental learning of vocabulary, by contrasting the number of encounters with target words with those of non-target words. His findings suggest that while the quality of context has a greater effect on the learners' acquisition of meaning, the number of encounters contributed to the knowledge of form (Webb, 2008 p. 238-239). Elley (1989) also conducted a study investigating the vocabulary learning from stories used with 7-8 year old children in their first language. He reports that there are different factors which influence the learning (e.g. number of occurrences of new words; whether

the new words are accompanied with pictures), and suggested that what children learn from certain stories is affected by their level of involvement in the stories.

Nevertheless, one should bear in mind that learning words is a continuous process of increasing one's familiarity with 'previously partially known words' (Cameron, 2001 p. 92), and that it is necessary for learners to meet words a number of times before fully acquiring them, whether it be in the form of explicit or incidental learning (Schmitt, 2000 p. 138).

It is true that learning can take place outside the teacher-oriented classroom and is more effective if learners take control of their learning by incorporating different vocabulary learning strategies by themselves. Nevertheless, Schmitt (2000, pp. 136-138) points out that it is necessary to be sensitive to learners' needs in incorporating the VLS into a curriculum by taking account of different factors (e.g. age, motivation, proficiency, learners' desire) so that learners are equipped to choose appropriate strategies for their learning of vocabulary. In relation to the TEYL context, Cameron (2001, p. 93-94) also stresses the importance of modelling and providing explicit training in how to use strategies with young learners in classroom tasks and of teaching sub-skills which are necessary in order for learners to use certain strategies to help themselves learn and remember vocabulary. This aspect is linked to vocabulary teaching, which is discussed in the next section.

3.3.2 Teaching of vocabulary

There are largely two kinds of approaches in vocabulary teaching: an explicit approach and an incidental learning approach, and both these approaches are regarded as necessary in all well-designed vocabulary programmes (Schmitt, 2000 p. 145).

An explicit approach is necessary for the most basic and frequent words in any second language, 'since they are prerequisites for language use' (Schmitt and McCarthy, 1997 p. 3). Sökmen (1997, p. 239-256) discusses ways of implementing the explicit teaching of vocabulary and suggests the following themes:

- build a large sight vocabulary
- integrate new words with old ones
- provide a large number of encounters with words
- promote a deep level of processing
- facilitate imaging and concreteness
- use a variety of techniques
- encourage independent learner strategies

Similarly, Cameron (2000, p. 87-89) summarises how a teacher can help young learners memorise words in the longer term and suggests organising vocabulary: 1) thematically (e.g. shopping); 2) 'through relations of wholes to parts (e.g. body – arms/ leg – fingers/ toes)'; 3) 'in general to specific hierarchies (e.g. food – vegetables – cabbage)'; 4) 'through words and antonyms (e.g. hot – warm – cold)'; or 5) 'in 'ad-hoc' categories (e.g. things to take on a picnic)''.

Although there are a number of ways to carry out explicit vocabulary instruction (e.g. Nation, 1990 p. 51), learners might cross-associate words depending on how they are introduced (Nation, 1990 p. 47). It is often the case that antonyms, synonyms and semantically related words are taught together, which causes cross-association. In order

to prevent this, Nation (1990) suggests that it is important to introduce the most frequent and useful words first and their counterparts should be introduced after the initially introduced ones have been established.

Furthermore, Schmitt (2000 p. 147-148) suggests that it is important to teach ‘the underlying meaning concept of a word’ in order to enable learners to understand words in different contexts. In addition, teaching word families rather than individual word forms is considered to be important, as this helps learners to acquire the habit of considering derivations of a word.

In L2 learning, ease or difficulty with vocabulary learning can be affected by different factors. According to McKay (1965), there are five factors in deciding the learnability of a word. The first factor involves whether the word in the L2 is similar to the equivalent in the learners’ L1. Secondly, whether the meaning is easily demonstrated can also affect the teachability. The third factor is brevity: it is assumed that shorter words are easier to learn than longer ones. The fourth factor is the regularity of the form. Finally, whether the words contain elements which are already familiar to the learners from their previous learning can also affect learnability. These factors are largely divided into two: 1) intralexical factors (i.e. ‘inherent in the word itself’) (Laufer, 1997 p. 154), and 2) crosslinguistic factors (i.e. ‘how well the learner’s L1 matches the L2’) (Schmitt, 2000 p. 148). As for intralexical factors, Laufer (1997, p. 154) lists the following factors which affect the learnability of words: ‘pronouncability (phonemes, combinations of phonemes, stress); orthography; length; morphology; inflexional and derivational complexity; deceptive morphological transparency; synformy; part of

speech; abstractness and specificity/ register restrictions; idiomaticity and multiplicity of meaning.’ In dealing with the difficulties caused by crosslinguistic factors, Swan (1997, p. 178-179) considers that making learners aware of the similarities and differences between their L1 and L2 can help them to realise the ‘nature and the limits of crosslinguistic correspondences, and to become more attentive to important categories in the second language which have no mother-tongue counterpart’.

While an explicit approach focuses on the linguistic forms that learners should be aware of, in an incidental learning approach, it is necessary to make sure that learners are exposed to the language as much as possible (Cameron, 2001 p. 94; Schmitt 2000, p. 149). Although it is assumed that putting learners in an L2 situation is the most efficient way to learn the language (e.g. Milton and Meara, 1995), not all learners have opportunities to experience language in such a situation. This is the case for most Japanese learners. Therefore, it is necessary to find ways to provide learners with enough exposure to the target language. One suggestion is to make learners read more. Reading has been regarded as useful in providing a context for incidental learning. Schmitt (2000, p. 150) explains the reason why written discourse provides more opportunities for vocabulary learning than spoken discourse:

Of course, many words can be learnt incidentally through verbal exposure, but considering that spoken discourse is associated with more frequent words and lower type-token ratios than written discourse, ... it would be optimistic to expect to learn a wide vocabulary from only spoken discourse. Written discourse, on the other hand, tends to use a wider variety of vocabulary, making it a better resource for acquiring a broader range of words.

Although it is recommended to use written texts in reading, a number of separate exposures are needed in order for learners to acquire the vocabulary. It is suggested that what is necessary is extensive reading, such as courses making use of graded readers, which provide learners with continuous contact with texts designed so that the vocabulary is introduced systematically and recycled (Schmitt, 2000 p. 151). Extensive reading is thought to be appropriate for beginners, giving the best amount of input, providing them with input consisting largely of already-known lexis together with repeated exposure to partially-known lexis (Schmitt, 2000 p. 151). Furthermore, for intermediate students who are almost ready to start reading authentic texts, ‘narrow reading’ is suggested. In this ‘narrow reading’, readers read different authentic articles on a same topic, so that topic-specific vocabulary can be encountered repeatedly (Hwang and Nation 1989). Nevertheless, as reviewed in the previous section, it is important to encourage active involvement in the stories in the case of young learners, when incorporating an incidental learning approach (Elley, 1989).

While there are a number of benefits to the incidental learning approach as described above, Sökmen (1997, pp. 237-238) explains why relying solely on an incidental learning approach is potentially problematic as summarised in the table below.

Table 3.2 Arguments against a solely incidental learning approach (Sökmen 1997, pp. 237-238)

	Reasons	Supporting research
1	Acquiring vocabulary mainly through guessing words in context is likely to be a very slow process.	Sternberg (1987) Carter and McCarthy (1988) Scherfer (1993)
2	Inferring word meaning is an error-prone process	Pressley, Levin and McDaniel (1987) Kelly (1990)
3	Even when students are trained to use flexible reading strategies to guess words in context their comprehension may be low due to insufficient vocabulary knowledge.	Haynes and Baker (1993)
4	Putting too much emphasis on inference skills when teaching vocabulary belies the fact that individual learners have different, yet successful, styles of acquiring unfamiliar vocabulary.	Hulstijn (1993)
5	Guessing from context does not necessarily result in long-term retention.	Parry (1993) Mondria and Wit de-Boer (1991)

Sökmen (1997, p. 238-239) contends that it is important to balance explicit and incidental learning approaches. As Stahl and Fairbanks (1986) also suggest, combining both explicit (i.e. definitional information), and implicit (i.e. contextual information) approaches in teaching proves more effective in having students achieve a deeper processing of the target vocabulary.

3.3.3 Selection of vocabulary items for a syllabus

As discussed in Section 3.2, there are different types of syllabuses, mainly divided into two: Type A and Type B. In the 1950s and 1960s when the grammatical syllabus was a mainstream of language teaching, vocabulary was regarded as of secondary importance to structural information (O'Dell, 1997, p.259). Lexical items were selected based on the structures which were to be introduced at points in the teaching programme. The 1970s was also a time when vocabulary was considered as 'subservient to the other more important elements of learning and was usually introduced in ways that suited the

presentation of grammar or functions, or through texts used for various structural or communicative purposes' (O'Dell, 1997 p.259). In the 1980s, attention to lexis increased. As Dublin and Olshtain (1986, pp.111-112) put it,

Lexis has failed to receive enough attention either in older grammatical syllabuses, or in more recent communicative approaches. In fact, lay people believe that 'knowing a language' consists of knowing words, while modern linguistic theories have placed little emphasis on vocabulary, focusing more on structures, functions, notions and communication strategies. However, it may be the case that possessing a good vocabulary stock is what enables many learners to use their knowledge of the language effectively and in ways which fit their specific needs.

With the development of computer corpora, attention to lexis increased, as can be seen in the suggestions for a lexical syllabus by Sinclair and Renouf (1989), Willis (1990) and Lewis (1993).

In a Type A syllabus (i.e. a product-based syllabus, Nunan 1988), the grammatical structures or vocabulary to be taught will be decided in advance. White (1988, pp. 48-50) lists seven criteria which need to be considered in planning the lexical content of the syllabus. These criteria are: frequency, coverage, range, availability, learnability, opportunism and centres of interest. By contrast, the Type B (i.e. process-based) syllabuses (Nunan 1988, pp. 40-50) 'tend to focus attention more on learning skills and strategies than on content' (O'Dell 1997, p. 275), and the selection of vocabulary is not often mentioned. O'Dell (1997) argues that the key criteria necessary for the selection of vocabulary in a Type A syllabus are also appropriate for that of a Type B syllabus, 'given that the teacher plays at least some part in the negotiation of the syllabus content' (O'Dell, 1997, p. 261). The criteria suggested by White (1988, pp. 48-50) are summarised in the following table.

Table 3.3 Criteria for the vocabulary selection and grading for a syllabus

	Criteria	Suggestion
1	Frequency	Words which appear frequently should be selected and introduced. However, there are difficulties associated with using frequency as a basis for selection and grading. For instance, highly frequent words have multiple meanings which calls for a specification of which of the meanings is the most frequent.
2	Coverage	Words which have a broader coverage should appear first. The idea of coverage is to prioritise words which can cover wider uses (e.g. 'boat' covers 'ship', 'yacht', 'canoe' etc).
3	Range	Words used in different types of texts should be introduced before those which only appear in a certain text. Words which have a high range are found in a large number of texts within a corpus. This helps to distinguish the highly frequent words which are representative across different texts and those which are only frequent in certain texts in a corpus.
4	Availability	Words which are commonly used by native speakers should be presented regardless of their frequency. (e.g. 'salt and pepper').
5	Learnability	Words which are easier to learn should be introduced before those which are harder to learn. (See section 3.3.2 for review of McKay 1965, Laufer, 1997 & Swan 1997)
6	Opportunism	Words which are relevant to the learners' immediate situations (e.g. classroom vocabulary such as <i>blackboard</i>) despite their low frequency, coverage, range etc.
7	Centres of interest	Words which are likely to interest students should be included. (e.g. transport, food, clothing, work, travelling, entertainment).

O'Dell (1997, p. 270) suggests that all of the criteria suggested by White (1988) are useful in selecting lexical items for a syllabus for beginners. She further mentions that the first six criteria highlighted by White (i.e. frequency, coverage, range, availability, learnability, opportunism) gradually become less significant for more advanced learners, and centres of interest become more important as the words required by the advanced learners are more likely to be for technical or professional purposes.

3.3.4 Lexical syllabus

The lexical syllabus is one of the Type A syllabuses, as reviewed in Section 3.2. It is worthwhile to discuss the benefits which it offers (Section 3.3.4.1), as well as its constraints and the difficulties involved in its practical implementation (Section 3.3.4.2) in order to determine the possibility of applying some of the key principles to the TEYL context in Japan. It is also worth recalling how the lexical approach has been viewed in practice, by referring to the evaluation of the Collins COBUILD Course (Willis and Willis, 1987-8), the only commercial coursebook designed to exploit the notion of lexical syllabus, as suggested by Sinclair and Renouf (1988) (Harwood, 2002, p. 148).

3.3.4.1 Benefits of a lexical syllabus

With the emergence of a corpus-based approach to linguistics and its research findings, this distinction of grammar and lexis was challenged and a shift in pedagogical focus from grammar to lexis was suggested (Lewis, 1993; Sinclair, 1991; Sinclair & Renouf, 1988; Willis, 1990). This view of language has been supported by studies of phraseology (e.g. Aijmer, 1998; Cowie, 1998; Glaser, 1998; Howarth, 1996; 1998) and these findings have informed ELT materials, stressing the importance of teaching memorable readymade chunks of language, known as lexical items or as prefabricated language i.e. strings of words which tend to be found together (Harwood, 2002, p. 140). Some research of this kind has criticised the textbooks used in ELT for their limited representation of authentic English (e.g. having no efficient coverage of lexical variations and predominant patterns), since most coursebook writers tend to rely on their intuition (Holmes, 1988; Hyland, 1994; Lewis, 1997, p. 10; Mindt, 1996, p. 232; Williams, 1988; Willis, 1990, p. vii). It has been suggested that the use of a lexical approach can avoid giving unrepresentative input to learners, as an approach which

makes use of corpora (or a corpus) consists of ‘authentic English’ as a database (Sinclair & Renouf, 1988, p. 159). Therefore in the lexical syllabus, corpora serve as a ‘database for lexically based inquiry and instruction’ (Richards & Rodgers, 2001, p. 131), which effectively informs the pedagogical materials and acts as the basis of ‘new syllabuses’ (Sinclair, 1991, p.78). In other words, once a corpus is designed to be representative of the authentic language, the lexical syllabus derived from the corpus should provide ‘authentic models’ for learners.

The lexical approach derives from a ‘view of language in which lexis plays a central role’ (Harwood, 2002, p. 39; Richards & Rodgers, 2001, p. 133), such as would allow learners to be aware of not only ‘what words occur in natural use, but [also] in what forms and in what patterns they typically occur’ (Carter & McCarthy, 1988, p. 150). In this way, the lexical syllabus provides access for learners to a meaningful context, which assists them in understanding meaning and recognising associated patterns of lexical items. As discussed in Section 3.2.2.1, this is related to the notion of pattern grammar (Hunston and Francis, 1999), which suggests that words belonging to the same pattern are normally similar in meaning and therefore teaching patterns can effectively contribute to the learner’s understanding of the meaning of lexical items (Hunston & Francis, 1999, p. 270). Cook (1998, p. 60) warns against the possibility of a ‘bewildering refusal to teach grammar’ if the course content is too preoccupied with lexis. However, as has been reviewed, the lexical syllabus treats grammar as lexical patterning in the belief that ‘patterns attach to all lexical items in the language, [and] learning the lexis means learning the patterns and therefore the grammar’ (Hunston, 2002, p. 190). In other words, it is possible to effectively teach grammar by having the

focus on lexis (Wray, 2000). This view is also supported by Bolander (1989, p. 85) who proposes the following:

[W]hen the number of prefabs stored in memory is large enough, syntactic rules derived as help for the memory force a reorganization of the language data into sub-units suitable for fast and flexible manipulation.

This echoes the notion of a lexical approach, claiming that ‘language production is not a syntactic rule-governed process but is instead the retrieval of larger phrasal units from memory’ (Moudraia, 2001). Nattinger (1980, p. 341) has a similar view:

Perhaps we should base our teaching on the assumption that, for a great deal of the time anyway, language production consists of piecing together the ready-made units appropriate for a particular situation and that comprehension relies on knowing which of these patterns to predict in these situations. Our teaching, therefore, would center on these patterns and the ways they can be pieced together, along with the ways they vary and the situations in which they occur.

In addition, research on formulaic language (e.g. Aijmer, 1996; DeCock, 1998; DeCock, Granger, Leech, & McEnery, 1998; R. Moon, 1997, 1998; Wray, 2000; Wray & Perkins, 2000) and on meta-discourse (e.g. House, 1996; Hyland, 1998, 1999, 2004, 2009) supports the advantages of the lexical approach, emphasising the importance of teaching prefabs and lexical items with information derived from phraseology rather than the traditional approach of teaching ‘words’ as individual units. Sinclair and Renouf (1988, pp. 155-156) also discuss the efficiency of teaching lexical variation, as one of the benefits of a lexical syllabus.

3.3.4.2 Constraints and difficulties in the implementation of a lexical syllabus

As we have seen, theoretically speaking, the lexical approach has its benefits. However, it appears that the terms ‘lexical approach’ and ‘lexical syllabus’ either have not been fully understood, or at least have not been successfully implemented, in practical ELT

contexts. In this section, I shall start with a discussion of the constraints and difficulties involved in the implementation of a corpus-derived lexical syllabus by examining the example of the *COBUILD English Course*.

The *Collins COBUILD English Course* (Willis & Willis, 1987-8) was created by making use of the Bank of English corpus and is the only ‘general’ ELT course to have been fully derived from the notion of the lexical approach (Harwood, 2002, p. 148). The central claim of the *COBUILD English Course* is that ‘the lexical database provides a rich input of real language’ giving authenticity and context to the tasks and exercises’ (McDonough & Shaw, 2003, p. 48). In their design of a lexical syllabus, word frequency information is used to itemise language and identify what is to be learnt in the target language (Willis, 1990, p. viii; 15). Having the basic concept of treating grammar from ‘a lexical viewpoint’ (Willis, 1990, p.15), the syllabus is specified by prioritising lexical items, rather than structures (ibid, p. 27), as units of the syllabus. A word identified in the corpus is analysed through concordances which make the students picture ‘not only ... its meaning, but also ... the syntactic patterns with which it is associated and its use in discourse’ (ibid, p.31). The information gained in the previous stage was edited so as to look like a series of dictionary entries and also edited as a ‘data-sheet’ which contained a categorised number of occurrences, examples of occurrences under the categories and information on phrases and collocates.³ The information on the ‘data-sheet’ was used to design teaching materials. The course has three levels covering in total about 2,500 words, all of which ‘occurred at least 120 times in the corpus’ (Willis, 1990, p. 28) and designed to recycle lexical items

³ For examples, see Willis (1990, p. 32-33)

throughout the course (Willis, 1990, p. vi). As Carter and McCarthy (1988 p. 59) state, the construction of a lexically-graded syllabus made it possible to design 'lexically-authentic' materials.⁴ Nevertheless, with the course and concept behind it being so innovative at the time of publication, it did not sell well. Willis and Willis (cited in Schmitt & McCarthy, 1997, p. 323) recognise one of the reasons for the COBUILD course's lack of commercial success stemmed from its presentation of grammar (e.g., its presentation as lexical grammar) which differed greatly from the presentation used in the traditional treatment of grammar more familiar at the time to teachers and students. This raises the question of how the syntactic information should be presented in a lexical syllabus. A more recent application of a lexical syllabus seen in a general English course is in the textbook *Innovations* (Deller & Hocking, 2000 cited in Harwood 2002, p 148) which makes partial use of the lexical approach combined with traditional 'grammar' components. However, Harwood (2002, p. 148) argues that this is not dissimilar to a standard coursebook, so it therefore hardly realises the full implications of the lexical syllabus.

Apart from what is described above, the application of the lexical approach seems to be found especially in the field of teaching English for Specific Purposes (ESP) or English for Academic Purposes (EAP) (e.g. Jabbour, 1998) and in the application of 'specialised' corpora in teaching (e.g. Flowerdew 2001; Johns, 1989, 1991; Stevens, 1991; Tribble 2001). In such contexts, teachers or course designers are aware of the clear needs and wants of their learners, which help to inform the content of the syllabus.

⁴ For examples, see Willis (1990, p. 34)

It is important to note that there is a crucial difference between general and ESP applications of a corpus-based syllabus, which is reflected in the content of the corpus. One of the factors hindering the implementation of a lexical syllabus in general English courses may result from this requirement to ensure that the corpus be representative of ‘natural’ English. While the lexical syllabus for a general ELT course needs to ensure that the corpus represents ‘a wide cross-section of registers and genres’ (Flowerdew 2001, p. 72), the corpus used in teaching English for special purposes needs to represent the language related mainly to the discipline in focus, taking learners’ needs into consideration.

Reviewing the practical implementations of the lexical syllabus available to the commercial market makes it clear why the application of a lexical syllabus has not so far been widely accepted. As discussed in section 3.2.2.1, in applying the notions of a lexical syllabus suggested by Sinclair and Renouf (1988) and Willis (1990), the use of a corpus cannot be avoided. Applying the corpus data in ELT requires ‘adjustments’ (e.g. removing untypical or inappropriate lexical items) which inevitably presents some difficulties for teachers or syllabus designers (Harwood, 2002, p. 142). It has been suggested in section 3.2.2.2, that the lexical approach can theoretically provide efficient lexical variations and predominant patterns, unlike most textbooks used in ELT (Harwood 2000, cited in Harwood 2002 p.146). However, there are also criticisms of the use of corpora in selecting lexical items. For instance, Cook (1998, p. 61) warns against the danger of giving inappropriate prominence to particular words or expressions by relying only on frequency as a criterion, rather than that of effectiveness or usefulness. As Harwood (2002 p.142) states, ‘implementing a lexical approach

requires a delicate balancing act' between what appears to be frequent and the usefulness of the item: ensuring the inclusion of the 'most useful' lexical items and patterns for learners in the syllabus. Other researchers also suggest that it is important to ensure the 'saliency' (Hunston, 2002, p. 191), 'learnability' (Lewis, 1993, p. 109), 'teachability' (Harwood, 2002, p. 142), as well as the 'usefulness' of the lexical items (Kaprowski, 2005). Therefore it can be said that one of the difficulties in the implementation of a lexical syllabus is that of deciding which lexical variations and patterns to teach within the constraints of time for preparation, as well as course length. Even with these suggestions available when considering what should be included in the syllabus, however, there has been little enquiry into what features can be labelled 'salient' and 'pedagogically useful', in what ways they serve and how they can be identified, other than by frequency information alone.

Another related important aspect in implementing a lexical approach is the question, when selecting which corpora to use, of knowing the contents of the corpus (or corpora) which is to serve as the database for a lexical syllabus. As discussed, course designers must be aware of learners' needs and wants and reflect them in the choice of corpora, or in deciding the content of the corpus, if they make their own. In addition, in the case of using publicly available corpora, a significant problem is that of the cost involved, mainly due to the commercial interests of the publishers who hold the copyright (Harwood, 2002, p. 144; Kaprowski, 2005, p. 331). In the case of ELT practitioners, making their own corpus for their course is time-consuming and difficult, since they have to go through all the procedures necessary for obtaining the texts which they want to include. Moreover, setting criteria for the content of a corpus is inevitably

problematic, since there are more varieties of English recognised by some scholars (e.g. Jenkins, 2003; Kachru, 1992; Kirkpatrick, 2007) than was previously the case, in the light of the recent development in English as a Lingua Franca (ELF), an area which also questions the application of a corpus-derived lexical syllabus, especially one derived only from native-speaker texts. For instance, it becomes a crucial issue if a learner aims to learn so-called World English rather than English based on standardized native speaker norms.

Another difficulty of the lexical approach relates to issues of pedagogy. The notion of a lexical syllabus suggested by Sinclair and Renouf (1988) and Willis (1990) stresses the importance of recycling the lexical items. However, research by Harwood (2000, cited in Harwood 2002) suggests that the recycling of prefabricated language or metadiscourse is not given any importance in most current ELT textbooks. He points out the failure of ELT textbooks to reflect the research findings in applied linguistics and claims that this hinders the systematic recycling of prefabricated language in practice (Harwood, 2000 cited in Harwood 2002 p.146). The benefits of recycling lexis tend to be underestimated by teachers as well as by ELT textbook writers (Harwood, 2002 p.146) as Lewis (1993: 51) states that recycling in a pedagogical context is viewed merely as: “‘doing the same thing twice” [which] is still widely considered [to be] time-wasting and potentially boring’. This comment should remind course book writers (whatever theoretical rationale their syllabus is based on) to include recycling activities which are carefully designed in order to avoid the charge of being boring and to engage learners and maximise their learning. This also relates to the question of how the

research findings in such areas as phraseology can be incorporated into the syllabus, while ensuring its systematic representation.

3.3.4.3 Dealing with criticism and difficulties: potential applications

According to the above, the importance of the lexical syllabus is supported by findings about the language, such as phraseology gained from corpus-related research (Hunston et al., 1997). Some researchers suggested that the lexical syllabus could replace current English language teaching syllabuses in general (Sinclair and Renouf, 1988; Willis 1990). However, as noted earlier, there are both advantages and disadvantages in constructing and implementing a lexical syllabus for ELT. While the notion of a corpus derived syllabus has been accepted and implemented more in the context of ESP and EAP (Coxhead, 2000; Jabbour, 1998), it has rarely been carried out in a general holistic syllabus. Presumably, one of the main factors hindering the spread of the lexical approach is the limited guidance available in the literature on the design of a lexical syllabus (Harwood, 2002). As O'Dell (1997, p. 278) notes,

While some have suggested that a lexical syllabus may replace all other EFL syllabuses, it seems more appropriate that the lexical syllabus should simply occupy a more significant position than has previously been the case. Within the lexical syllabus, it is important for students to be given a variety of approaches to lexis with the hope of achieving a breadth of coverage. Some approaches have a greater significance at elementary levels and some at more advanced levels. The treatment of collocation and learner training are of particular significance in the construction of all levels of the syllabus.

Considering the difficulties and constraints involved in implementing a lexical syllabus, it seems more reasonable to suggest that the process should start by implementing some aspects of the lexical syllabus in more sections in a course than has so far generally been the case. As Moudraia (2001, p.2) puts it:

[I]mplementing a lexical approach in the classroom does not lead to radical methodological changes. Rather, it involves a change in the teacher's mindset. Most important, the language activities consistent with a lexical approach must be directed toward naturally occurring language and toward raising learners' awareness of the lexical nature of language.

Therefore, in considering the possible design of a corpus-influenced syllabus for the teaching of EFL young learners in Japan, the difficulties identified in Section 3.2.2.3 should first be overcome. The main difficulties concern: 1) corpus selection, or corpus design and compilation; 2) lexis selection; and 3) pedagogic considerations including teacher training. As we have seen in the prior discussion, the lexical syllabus has so far been more successfully deployed in the case of ESP/EAP. Therefore, in a later section of this thesis, I shall discuss the potential of employing a corpus influenced syllabus by comparing the proposed research context with the design of ESP/EAP syllabuses, with particular reference to the use of specialised corpora and a pedagogic corpus suggested by Willis (1990, 2003) and in relation to the SLA research findings of instruction in using prefabricated language (e.g. Nattinger & DeCarrico, 1992; Perera, 2001). This is relevant to the benefits of raising awareness of patterns, which is likely to contribute to literacy, as described in Hunston (2007, p. 264).

3.4 Teaching and learning grammar

The previous section highlighted the change of focus towards vocabulary teaching since computer corpora were made available for more accurate language description. This does not mean that the knowledge of grammar has been disregarded in teaching and learning of languages, but that it can act as a 'stepping stone to learning and using

grammar' (Cameron, 2001 p. 72). In this section, two key issues associated with grammar instruction and learning about grammar are discussed.

3.4.1 Noticing

Traditional grammar teaching concentrated on mostly providing learners with explicit instructions on the use of specific grammatical structures and giving them opportunities to produce them (Ellis, 1995 p. 87). However, this traditional approach has been criticised over the years, for its ineffectiveness in learners' acquisition, as a result of its disregard for how learners learn grammatical structures. For instance, Ellis (1993, p. 92) criticises the traditional approach to grammar often found in the structural syllabus, which often involves 'practice and production' suggesting that it is not well-suited to how learners acquire grammar.

In the light of the fact that learning of grammar takes place as a slow and gradual process, Ellis (1993, pp.108-109) explains the importance of raising learners' awareness regarding grammar. He suggests that learners who are explicitly aware of grammatical structures can deal with future encounters by noticing them. Ellis (1995) further proposes a way to develop learners' awareness of grammatical structures without teachers having to explicitly tell them the rules. He suggests that teachers can do so by means of using interpretation tasks which aim at providing learners a target structure through input, 'and that enable them to identify and comprehend the meaning(s) of this structure' (Ellis, 1995 p. 88). In other words, this process helps learners to notice grammatical features in the language presented to them as input. Ellis (2003, p. 346) explains 'noticing' as 'a cognitive process that involves attending to linguistic form in the input learners receive and the output they produce'.

Similarly, Schmidt (1990, p.149) argues that ‘noticing’ is an inevitable condition for any learning to take place, suggesting that ‘what learners consciously notice’ is the intake that learners gain from the input. Schmidt (1993, p. 26) contrasts ‘noticing’ with the term ‘understanding’ as follows:

I use *noticing* to mean registering the simple occurrence of some event, whereas *understanding* implies recognition of a general principle, rule or pattern. For example, a second language learner might simply notice that a native speaker used a particular form of address on a particular occasion, or at a deeper level the learner might understand the significance of such a form, realizing that the form used was appropriate because of status differences between speaker or hearer. Noticing is crucially related to the question of what linguistic material is stored in memory... understanding relates to questions concerning how that material is organized into a linguistic system. (Schmidt, 1993 p. 26)

Schmidt (1990, p. 149) highlights the importance of the task design in achieving learners’ implicit language learning through a task. He suggests that task should be designed so that it ‘demands focus attention on relevant features of the input’ and facilitates the learners’ noticing through input.

Long and Robinson (1998, p. 40) point out that effects on students’ learning through noticing might not be immediate, ‘given the generally slow, non-linear, and partial nature of much L2 learning’. For this reason, the effectiveness of noticing cannot be measured in short-term studies by means of production measures which demand learners’ immediate performance on target forms (Long & Robinson, 1998 p. 41). Long and Robinson (1998 p. 24) mention that the term ‘noticing’ means ‘an intended outcome of focus on form’, and learners process linguistic material by recalling noticed features. The notion of focus on form is reviewed in the following section.

3.4.2 Focus on form

Long (1988; 1991) was the first writer to distinguish ‘focus on forms’ from ‘focus on form’. Focus on forms ‘characterises earlier, synthetic approaches to language teaching that have as their primary organizing principle for course design the accumulation of individual language elements’ (Doughty & Williams, 1998 p. 3). On the contrary, focus on form requires engagement in meaning before giving any attention to specific linguistic features. There are two definitions of focus on form, proposed by Long (1991) and Long and Robinson (1998).

Focus on form ... overtly draws student’s attention to linguistic elements as they arise incidentally in lessons whose overriding focus is on meaning or communication. (Long, 1991, p. 45-46)

Focus on form often consists of an occasional shift of attention to linguistic code features – by the teacher and/or one or more students – triggered by perceived problems with comprehension or production. (Long & Robinson, 1998 p. 23).

The first definition is theoretical, and the second one is operational, in a sense that it provides practical implications (Doughty & Williams, 1988 p. 3). The advantage of the focus on form as opposed to traditional grammar teaching, is that ‘learner’s attention is drawn precisely to a linguistic feature as necessitated by a communicative demand’ (Doughty & Williams, 1988 p. 3). In other words, focus on form is a process by which learners pay attention to form when engaging in the comprehension or production of communicative contents of messages.

Long and Robinson (1998, pp. 24-26) explain ways in which material designers or teachers may carry out focus on form, providing three examples of how focus on form can be achieved. The first example is to engage learners in a problem-solving task. The

task is designed to require learners to read texts on the topic. By means of frequent encounters with the lexical items used in the texts (sometimes these are italicised or underlined in order to make them more salient), it is more likely for the learners to notice these lexical items. The activity can be followed by an additional task which encourages learners to use those lexical items (Long & Robinson 1998, p. 24-25).

Secondly, Long and Robinson (1998, p. 25) suggest that the teacher may interrupt the task if they find some learners are making the same errors repeatedly. It is then that the teacher gives 'explicit negative feedback' by 'draw[ing] attention to the problem, using pedagogical devices appropriate for students of the age, literacy level and metalinguistic sophistication concerned'.

Long and Robinson's (1998, pp. 25-26) third suggestion is to conduct focus on form by giving learners implicit negative feedback. It has been suggested that learners are more likely to notice the grammatical information in recasts (i.e. 'corrective reformulations of children's utterances that preserve the child's intended meaning') provided by teachers (Long & Robinson 1998, p.25). Studies conducted with L1 learners (Baker & Nelson, 1984; Farrar, 1992) and L2 learners (Oliver, 1995; Ortega & Long, 1997) all suggest that the use of recasts is more effective for the acquisition of grammatical information than the models presented to the learners. Long and Robinson (1998, p. 26) suggest that recasting can be a way for teachers or material designers to provide focus on form without disturbing the lesson's main focus on meaning.

As described above, it is possible to make learners notice the important grammatical forms or rules without making the grammar itself a focus of the lessons. This way of

drawing learners' attention to form has also implications for TEYL in Japan, as it is not officially encouraged to teach the grammar explicitly in primary school English class.

3.5 Corpora and English language teaching

In order to explore the potential of a corpus-influenced syllabus for the teaching of literacy to young EFL learners in Japan, in this section I will first review how the emergence of corpora influenced the language description which challenged the traditional distinction of lexis and grammar by offering a historical perspective on the treatment of lexis in ELT. This is followed by a review of how available research on literature corpora and their use in language teaching, together with suggestions of how finding from such research could be used to reveal cultural aspects of children's literature.

3.5.1 Historical perspectives on vocabulary teaching

The traditional view of language has a tendency to divide grammar from vocabulary. The importance of vocabulary in language learning and teaching has been recognized by many researchers over the years (Faucett, West, Palmer, & Thorndike, 1936; Richards, 1976; West, 1930, 1953). The following table briefly summarises the historical perspectives on vocabulary teaching since the 1920s.

Table 3.4 Historical perspectives on the treatment of vocabulary in ELT

Period / Year	Author	Emphasis/ Focus of research
1920s – : Vocabulary is important in especially in reading skills		
1929	Coleman	Reports deficient foreign language skills of American students; emphasises the importance of vocabulary in reading skills (Zimmerman, 1997, p. 9)
1930	West	Examines the use of English in India in the 1920s and emphasises the importance of vocabulary in language learning and teaching. Recommends the use of word-frequency lists as the basis for selection and order of vocabulary in materials for students. ⁵
1936	Faucett et al.	The Interim Report on Vocabulary Selection
1944	Thorndike and Lorge	<i>Teacher's Word Book of 30,000 Words</i> (Thorndike & Lorge, 1944)
1953	West	<i>A General Service List of English Words</i> (West, 1953)
1945-1970: Vocabulary is as not as important as grammar or phonology (Richards, 1976, p. 77)		
1945	Fries	'the problem of learning a new language was not learning its vocabulary, but mastering its sound system and its grammatical structure' (Fries, 1945, p. 39 cited in Carter and McCarthy 1988, p. 40)
1961	Gleason	'In learning a second language, you will find that vocabulary is comparatively easy, in spite of the fact that it is vocabulary that students fear most. The harder part is mastering new structures in both content and expression' (Gleason, 1961, p. 7 cited in Carter 1998a p.184)
Late 1960s – early 1970s: Vocabulary gains more weight in language teaching		
1970	Sinclair et. al.	Vocabulary gained more weight in language teaching, due to the lexical research findings reported in the late 1960s and early 1970s (e.g. Sinclair, Daley, & Jones, 1970 cited in Carter 1998a, p.185)
1972	Wilkins	Wilkins (1972, p. 111) incorporates the insights of lexical semantics into language teaching and learning, in the belief that, while, 'without grammar very little can be conveyed, without vocabulary <i>nothing</i> can be conveyed'. He introduced two categories of meaning which are involved in communication: notional and functional. His emphasis on meaning in discourse contributed to 'a major feeding ground for vocabulary practitioners in the 1970s and 1980s' (Carter & McCarthy, 1988, p. 42).
1980s –: 'Grammar and lexis distinction challenged' – increasing attention to phraseology		
1980s	COBUILD ⁶ project	Uses a corpus of twenty million words to demonstrate language use and allow for more accurate language description.
1985	Sinclair	'Major reorientation in language description' (Sinclair, 1985) arising from corpus research

In the 1980s, the 'major reorientation in language description' (Sinclair, 1985) arising from corpus research brought linguists and ELT practitioners to realise more about the

⁵ The attempts to use a scientific rationale for the selection of vocabulary in course content were among the first efforts to establish principles of syllabus design in language teaching (Richards & Rodgers, 2001, p. 37)

⁶ Collins-Birmingham University International Language Database

‘nature of language’ and the role of vocabulary in language: the notion of ‘chunks’ of language larger than lexical items, such as lexical phrases and prefabricated units was introduced. For instance, Nattinger and DeCarrico (1992, p. 1) call the ‘chunks’ ‘lexical phrases’ and define them as follows:

Lexical phrases are ‘chunks’ of language of varying length that have an idiomatically determined meaning (e.g., *as it were*, *on the other hand*, *as X would have us believe*, etc.). The phrases have slots for various fillers (e.g., *a month ago*, *a year ago*) and are each associated with a particular discourse function.

Some other corpus-based studies (Erman & Warren, 2000; Kjellmer, 1987) also suggest that the aggregate of spoken and written English produced by a native speaker contains significant numbers of ‘prefabricated items’ (Harwood, 2002 p. 140). Lewis (1993) claims that ‘language consists of grammaticalised lexis, not lexicalised grammar’ (Lewis, 1993, p. 89); therefore it is suggested that lexical items should be the focus of language teaching (Lewis 1993; 1997; 2000; Sinclair & Renouf, 1988; Willis 1990; 1993).⁷

In addition to the challenged distinction between lexis and ‘traditional grammar’ (Singleton, 1997, p. 213) as a result of corpus findings, the traditional view of word boundaries (i.e., vocabulary as a series of ‘single’ words) was also challenged, with an emphasis on the fact that learners need to perceive and use patterns of lexis and collocation (Zimmerman, 1997, p. 17). As Widdowson (2003, p. 77) describes it:

[N]ow that linguists have discovered this basic descriptive fact ... it can and should now be prescribed as the content for English courses. Whereas previously we had to depend on the evidence of our own intuitions about English, or the unreliable responses of informants, we now have a record.

⁷ Lewis defines a ‘lexical item’ as ‘a word or word phrase that operates as a “socially sanctioned independent unit”’ (1993, p. 90).

It is a fair supposition that learners would benefit by paying attention to the corpus research findings of real language usage and that the findings of phraseology should be reflected in ELT materials (Howarth 1998; Nattinger & DeCarrico, 1992). However, some questions still arise as to how far corpus-based descriptions can represent the real language and also as to how and to what extent these ‘real descriptions’ can be of use to learners (Cook, 1998; Owen, 1993, 1996; Widdowson, 2000). Those questions indicate that applying corpus evidence and the phraseology to ELT is far from straightforward.

Nevertheless, corpus-based language description and its attention to lexis and phraseology holds considerable implications for future language pedagogy, in particular in designing syllabuses and teaching methodology (Bernardini, 2000; Carter, 1998b; Clear, 2000; Groom, 2005; Hoey, 2000; Hunston & Francis, 1998; 1999; Hunston, Francis & Manning, 1997; Nattinger & DeCarrico, 1992; Sinclair & Renouf, 1988; Willis, 1990; 1993; 1994; 2003). I now turn to a discussion of corpus-based syllabuses and their principles, and how the research findings of corpus linguistics have been reflected in the design of syllabuses for English language teaching.

3.5.2 Corpora, literature and language teaching

Findings from corpus linguistics studies have been influential in understanding language and have provided new insights into not only linguistic evidence but also into the teaching of language in EFL, ESL as well as the first language. The following section will review the research in the use of literature corpora and their application to ELT.

3.5.2.1 Literature corpora in language teaching

One of the early applications of corpus-based materials to ELT is the data-driven learning (DDL) suggested by Johns (1991). This approach is rather inductive, having the characteristic of relying mostly on ‘learners’ intelligence to find answers’ (Johns, 1991) by consulting evidence from corpora (i.e., concordance lines in the Key-Word-In-Context format). On these grounds, it was assumed to be appropriate for adult learners (Johns, 1991, p. 12). Some research reports the effectiveness of DDL in EFL settings in facilitating learners’ consciousness of pattern and lexical items (Lee, 2006; Sripicharn, 2002, p. 403). Sripicharn (2002, p.408) suggests future research to test DDL with students at a lower level of proficiency, with a suggestion that the DDL materials should be rewritten to suit the level of the students. Moreover, Lee (2006) suggests that the use of literature corpora along with DDL or other methods such as CALL (Computer Assisted Language Learning) can promote students’ understanding of vocabulary and extended texts, and thus improve their reading skills.

For first language pedagogy, Sealey and Thompson (2004, p. 90; 2006, p. 27; 2007) identify the possibility of using a literature corpus (CLLIP) as a part of the literacy curriculum for British primary school children. They use concordances produced from their CLLIP corpus, which was used to promote metalinguistic awareness among children in the L1 context, in order to lead children to discover and identify patterns in language. The concordances were initially colour-coded according to word classes by means of the concordancing software integrated with the BNC, SARA (Dodd, 2001). It was suggested that colour seemed to help children to notice the grammatical words and

describe the difference between lexical words in concordance lines (Sealey and Thompson 2004, p. 88).

Louw (1991) suggests the incorporation of literature and language teaching, using the notion of a lexical syllabus suggested by Sinclair and Renouf (1988). In order for such integration to succeed, Louw (1991) notes the importance of including factors which are common in both disciplines. In the case of my research, the benefit of using children's literature or stories is recognised by both approaches (that is, in relation to the literary standpoint and from the nature of language). Moreover, Sealey and Thompson (2006, p. 27) recognise that '[c]orpus evidence is a resource available to those who make pedagogically informed decisions about the language and literacy syllabus.' However, as far as can be discovered from the available literature, there has been little research on the implementation of a corpus based syllabus in TEYL in Japan.

3.5.2.2 Implications for primary ELT in Japan

The studies mentioned above are not based in classroom contexts using a lexical syllabus, but they do indicate some prospects for introducing a corpus-based approach to the teaching of young EFL learners. Although this study is not concerned with applying DDL directly to the classroom situation, for instance, the study by Sealey and Thompson (2004) clearly indicates how much additional support there was from the layout (i.e., the colour-coded concordance lines), which made the concordances 'accessible' to children. As Thompson and Sealey (2007) point out, the analysis of a corpus should help to highlight the kinds of patterns in language which learners may need to familiarise themselves with as they learn. Tsui (2004, p. 40) suggests that teachers use corpus findings concentrating on high frequency words, i.e., which are

‘usual’ rather than ‘exceptional’ usages, especially for elementary and intermediate levels.

With a different environment from first language education and the characteristics of teaching young learners, it is important to take into consideration the aspects of ‘learnability’, ‘usefulness’ and the ‘accessibility’, as well as the choice of text when considering a pedagogical application of corpus findings. Considerations such as those mentioned above are particularly important when designing a corpus-influenced syllabus for young EFL learners. There still remains a need for more research into designing a corpus-influenced syllabus with these local and contextual considerations.

3.5.2.3 Corpus analysis and culture

Previous research on fictional texts written for children has suggested that such texts play an important part in the socialisation of children (Hunt 1992; Stephens 1992). As Halliday (1978) emphasises, language acts as a very powerful socialising agent, through which people learn about customs, hierarchies and attitudes. The language presented to children through literary texts is therefore regarded as ‘a particularly effective agent’ in promoting and reinforcing the adoption of values and accepted attitudes (Knowles & Malmkjaer, 1996 p. 44). Stephens (1992) analyses the ideological potential of children’s literature, in terms of the use of narrative techniques and intertextuality. Stephens (1992, p. 8-9) maintains the view that every book has implicit ideological aspects such as beliefs and values which are often taken for granted. As ideology is the collectively shared assumptions, beliefs, and values in social groups (Simpson, 1993 p. 5), the fact that it is often taken for granted in society makes it difficult to uncover underlying assumptions; often the analyst him/herself may have internalised similar values. Studies

in critical linguistics reveal the ways in which ideologies are embedded in literary texts, suggesting that children's literature is inherently ideological (e.g. Hunt 1992; Sealey, 2000; Stephens 1992).

The studies mentioned above have little to do with the corpus approach; however, there is some research on children's fiction which makes use of the corpus approach. For instance, Stubbs (1996, pp. 68-70) discusses the analysis done by Baker and Freedry (1989) who analysed a corpus of elementary reading books, in search of a constructed view of childhood.

Using a corpus of children's literature, Knowles and Malmkjaer (1996) demonstrate how the information of the choices of language and textual structure could be a key to understanding writers' intentions in terms of the effect their work will have on their readers' society. They stress the importance of looking at linguistically substantial data, that is using a corpus method; 'since no linguistic item or structure is 'ideological' in and for itself, it cannot be assumed that the presence of one form or another in a given text is an indicator that ideology is in operation' (Knowles and Malmkjaer 1996, p. 46).

This position is similar to that of Louw (1997, pp. 249-250) who also suggests that the corpus approach to literary criticism often 'provide[s] constant evidence of readers' blindness not only to the presence of phraseological meaning within the language but also to an author's intention to involve it in a specific instance.'

Thompson and Sealey (2007) also investigate texts written for a child audience, but they focus on ‘the ways in which the world is represented to the child reader’ (Thompson and Sealey, 2007 p. 4). They conducted the analysis by comparing corpora of fictional texts written for children and those for adults. While their analysis revealed similarities in linguistic characteristics, the close analysis of the words or phrases in the concordances revealed differences in the respective representations of ‘the world and the human relation to the world’ in the corpus of fictional texts written for children.

Having reviewed the available corpus-based investigation of children’s literature and the culture represented in the texts, it can be said that texts written for children are, as Hunt (1990, p. 2) suggests:

Culturally formative, and of massive importance educationally, intellectually, and socially. Perhaps more than any other texts, they [novels for children] reflect society as it wishes to be, as it wishes to be seen, and as it unconsciously reveals itself to be.

One might argue that literary texts are not the only ‘socialising agent’ as pointed out by Knowles and Malmkjaer (1996, p. 61), and point out that most of the classic and popular children’s literature are often televised, or made into films and watched by children. Knowles and Malmkjaer (1996, p. 61) suggest that the medium of television ‘can encourage subsequent reading’ by the children viewing the story. Though it is difficult to prove how the literary texts influence children’s behaviour, Steedman (1982 p. 11) reports that children used structures and narrative techniques which reflect those appearing in the literary text to which they were previously exposed.

Therefore it seems worthwhile to investigate the possibility of incorporating a literature corpus with language teaching which illuminates cultural differences to children in EFL education. In relation to this, the selection of texts is especially important and the cultural impact that they carry is also worthy of attention, considering the policies set by MEXT.

3.6 Identification of pedagogically useful items

3.6.1 A historical overview of methods of extracting vocabulary items for teaching

In Sections 3.3.2 and 3.3.3, the criteria often used for vocabulary selection for designing a syllabus (e.g. frequency, coverage, range, availability, learnability, opportunism and centres of interest) are introduced and discussed in detail. Moreover, under the influence of corpus linguistics which has challenged the boundaries of grammar and lexis, as reviewed in section 3.5.1, subsequently the criteria for extracting vocabulary items for pedagogical purposes have shifted. The table below (3.5) summarizes what researchers have said about lexical items in terms of their criteria for selection for teaching and their methods of defining them. I have also identified each researcher's purpose in identifying lexical items, whether the research was intended for any pedagogical application and whether researchers employ an explicitly phraseological outlook in their research. If not, their approach is considered to be word-based.

Table 3.5 Historical review of criteria and methods of identifying ‘pedagogically useful’ lexical items

Research	Description	Criteria	Data/ Methods	Pedagogical Application	Phraseological view
Thorndike & Lorge (1944)	‘The Teacher’s Word Book of 30,000 Words’ contains 30,000 lemmas, or about 13,000 word families (Goulden, Nation & Read, 1990), based on a count of an 18,000,000 word written corpus.		Manual counting	✓ Valuable in size – but old	– (wordlist)
West (1953)	‘General Service List’ contains 2,000 headwords. The frequency figures for most items are based on a 5,000,000 word written corpus	‘frequency’ ‘range’ ‘difficulty of learning’ ‘necessity’ ‘cover’ ‘stylistic level’ ‘emotional words’	Manual counting	✓	– (wordlist)
Carrol, Davis & Richman (1971)	American Heritage Word Frequency Book	Frequency (according to grade/levels)	Corpus of US texts used in schools (levels, grades) 5,000,000 running words		-
Francis & Kucera (1982)		‘frequency’	Brown and LOB (Lancaster-Oslo/Bergen) corpus		
COBUILD	Reveals how lexical items act differently according to their form. They also take differing collocates which express ‘different shades of meaning’ (Sinclair, 1991).		Corpus driven	✓	✓
Nagy & Anderson (1984)		Likelihood of learners encountering the word			

Irujo (1986)	Multi-word items	‘frequency’ ‘need’ ‘transparency’ ‘semantic simplicity’ (p.300)			✓
White (1988)	Vocabulary selection (White, 1988, p. 49).	‘frequency’, ‘range’ ‘coverage’, ‘availability’, ‘learnability’, ‘opportuni sm’ and ‘centres of interest’			
Sinclair and Renouf (1988)	‘lexical syllabus’ based on frequency	‘frequency’	Bank of English corpus	✓	✓
Willis (1990)	‘lexical syllabus’ The most frequent 700 words of English constitute 70% of English text. The most frequent 1500 words constitute 76 %. The most frequent 2,000 words constitute 80%.	‘frequency’ ‘lexical coverage’	Bank of English corpus	✓ COBUILD English Course (1988)	✓
Hirsh & Nation (1992)	The authors claim that a 2,000-3,000 word vocabulary can form a sound basis for language use.	2,000 words – 3,000 words for language use	Novels for teenagers and young learners		
Bauer & Nation (1993)	Word families	‘frequency’ ‘regularity’ ‘productivity’			
Lewis (1997)	Concentrating on nouns, adjectives and verbs with a high meaning content and taking into account their ‘learnability’ from the viewpoint of students.	‘learnability’(p.110) ‘frequency’ communicative power			
Moon (1997)	Multi-word items.	‘frequency’ should be complemented (p.62)	Bank of English corpus		✓

Nation & Waring (1997)	Vocabulary size, text coverage and wordlists	frequency, range representativeness word families idioms and set expressions Other criteria			✓
Jabbour (1998)	Designing a syllabus for EAP	‘frequency’ ‘collocation’	A specialised corpus of medical research articles		✓
Coxhead (2000)	Academic word list	‘frequency’ ‘range’	3.5 million word corpus		
Hunston (2002)		‘frequency’ ‘saliency’ ‘how highly valued an item is’ (p.193)	Checked against existing research findings (e.g. frequency information supplied by earlier studies on ‘frequently occurring sequences’ such as ‘lexical bundles, formulae and phrases.’)		✓
Harwood (2002)		‘teachability’ (p.142)			✓
Chujyo & Nishigaki (2004)	Created a wordlist for young learners - 600 words identified as ‘most relevant to students’ daily lives’	‘range’	‘RANGE’ (Nation and Heatley, 2003) is used for the analysis of a corpus of 20 picture dictionaries.	✓	
Kaprowski (2005)	Investigates the usefulness of multi-word items (lexical phrase in the course book)	frequency and range	Frequency and range Reference to the Bank of English corpus		✓

3.6.2 Identification of PUI from SLA (=English taught) research perspectives

The previous section reviewed different methods and criteria for selecting vocabulary for language teaching. As Granger (2005, p. 3) states, it is important to consider whether items ‘identified by quantitative methods are pedagogically valuable’ for particular learners. SLA has a long tradition in terms of identifying problems of learners’ productive usage (either spoken or written). The difference in L1 is an important indicator as to what kind of areas are likely to be difficult for learners. As Izumi and Isahara (2004 p. 63) suggest:

In foreign language education, it is important for teachers to know their students’ acquisition order of major linguistic items in the target language.

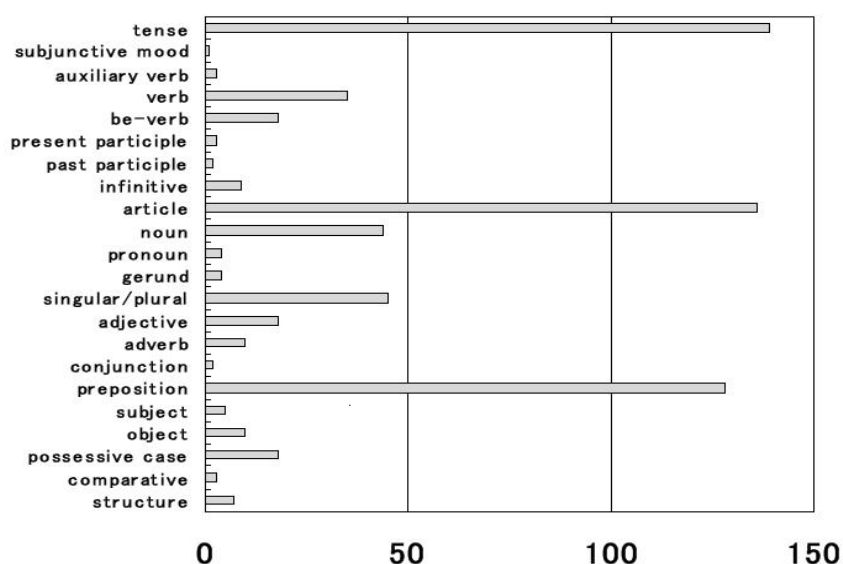
This enables them to teach these items more effectively in language classrooms.

In foreign language education, it is necessary to consider learners’ first language and the findings from SLA research when deciding what to include, and in what order, in a syllabus. Therefore, in considering what is ‘useful’ in terms of teaching English to Japanese young learners, it is necessary to know what kind of challenges English presents to Japanese learners in general, as they may also present challenges to young learners in their later learning. In this section, I discuss existing literature on English usage by Japanese learners.

There is ample research available in the area of error analysis dealing with Japanese learners. For instance, Ano (2005) studied spoken production of Japanese high school

learners and found 644 errors, which he broke down into 22 categories. The following table shows the frequency of errors in each category in his study.

Figure 3.1 Frequency of grammatical errors (Ano, 2005, p.15)



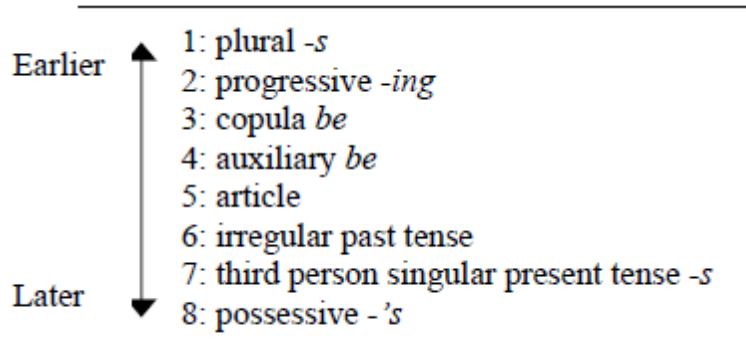
As can be seen in the table above, it is obvious that the most frequent grammatical errors made by his subjects were tense, articles, and prepositions. In terms of tense, he emphasises ‘the confusion between past and future tense with present tense’ as the most frequent error type. Ano (2005, p.15) speculates that ‘the learners could not distinguish the changes in tense of verbs or auxiliary verbs’, despite the fact that they were aware of the time difference.

In relation to prepositions, Ano (2005) gives an example of the use of unnecessary prepositions by learners because of set phrases or chunks they have previously learnt (e.g. “go to shopping”* as opposed to “go shopping” as a result of having learnt ‘go to’ as a chunk). Kaneko (2006) also claims that articles and prepositions are the most challenging features of English for Japanese learners. She researched exclusively the

use of vertical axis prepositions (i.e. *above*, *over*, *under* and *below*) by Japanese learners, using the NICT JLE (Japanese Learner English) corpus, and provided more detailed analysis of each item and its use, along with the kind of errors made by the learners. Kaneko (2006) emphasises the importance of introducing lexico-grammatical features of prepositions.

In addition, there exists some research on the order of acquisition of specific items which also presents some interesting findings of relevance in deciding what is ‘useful’ for learners. Dulay and Burt (1973), examining Spanish-speaking children (aged 8+), presented the following acquisition order in terms of grammatical morphemes.

Figure 3.2 Acquisition order of children whose L1 was Spanish (Dulay and Burt, 1973)



In further research (Dulay and Burt, 1973, 1974), they formulated a hypothesis that ‘a common acquisition order exists across learners of different backgrounds’ (cited in Izumi and Isahara, 2004). This hypothesis was supported by other researchers (Krashen, 1978, Krashen, Seliger, & Hartnett 1974).

In contrast to this, studies which dealt only with Japanese subjects, such as Terauchi (1994) and Tono (2002) showed contradictory findings. For instance, Tono (2002) examined acquisition order of grammatical morphemes in the Japanese EFL Learner Corpus (JEFFL Corpus). His findings suggest that ‘articles and plural -s were mastered later, and possessive -’s was acquired earlier’ in comparison with the order presented by Dulay and Burt (1973) (cited in Izumi and Isahara, 2004, p.64).

The main difference between the research evidence carried out by Dulay and Burt (1973); and by Tono (2002) was that the subjects who participated in the latter were all Japanese learners. Izumi and Isahara (2004) revisited the hypothesis that the order of acquisition is influenced by the learners’ background, and examined the acquisition order using the NICT JLE Corpus, in which the learners’ errors in grammar and lexis were grouped into 47 types and tagged manually. They found no correlations between the existing research on the ‘universal’ second language acquisition sequence presented by Dulay and Burt (1973) and other researchers, but found that their results showed a high level of correlation with research on acquisition order hypothesis for Japanese learners and the findings presented by Tono (2002). Izumi and Isahara (2004) presented their findings from the NICT corpus as below.

Figure 3.3 Ranking of morpheme scores in the NICT JLE corpus (Izumi and Isahara, 2004, p.69)

1: possessive -’s
2: progressive -ing
3: copula be
4: third person singular present tense -s
5: plural -s
6: auxiliary be
7: irregular past tense
8: article

Based on the findings of Tono (2002), and with further confirmation from their own data, Izumi and Isahara (2004) conclude that ‘Japanese learners acquire articles and plural -s in (sic) a later stage’. They explain this error as L1 transfer, as the Japanese language does not have an equivalent system to English articles and plural –s. The results from both studies suggest that the learning sequence is not necessarily universal.

With the existing research in mind, we can summarise two criteria which might contribute to deciding what kind of features of English are likely to cause a challenge to Japanese learners as follows:

- a) the most frequent errors made by Japanese learners
- b) the acquisition order hypothesis relating to grammatical morphemes

Based on the criteria and reported research results above, it seems that the following areas present particular challenges for Japanese learners.

- tense
- articles
- prepositions
- plural –s.
- other grammatical morphemes (e.g. possessive ‘s)
- word order difference (SVO vs. SOV)

One thing to note about these SLA studies is that their focus is on ‘productive’ language, rather than on receptive language. Considering literacy skills, both aspects are included (i.e. writing as productive, and reading as receptive). Nation and Newton (2009, cited in Nation 2009, p.1) outline the four strands of a language course and stress the importance of balancing the following: ‘meaning-focused input, meaning focused output, and language-focused learning and fluency development’. In learning a

language course for beginners who are learning to read in another language, ‘Meaning-focused output’ should involve different skills such as speaking and writing (Nation 2009, p.7). Taking this into account, it can be said that the SLA findings are also relevant for the ‘receptive’ aspect of skills in language learning. Nevertheless, how to devise such a language course is another matter, which will greatly influence the teaching methods and tasks or activities used in class. This aspect should also feature when considering the pedagogical implications of a corpus-influenced syllabus.

3.7 Summary and direction of this research

This chapter has reviewed issues mainly concerned with syllabus design, corpora and ELT. In Chapter 2, attention was drawn to the lack of literacy education in the current guidelines provided by MEXT (2002; 2003; 2008) and the constraints inherent in the policy. Bearing this in mind, it was suggested that corpus linguistics could have an impact on primary ELT and syllabus design in Japan, especially in the light of its lack of attention to literacy teaching.

As an example of a corpus-influenced syllabus design, the chapter discussed the idea of a lexical syllabus, with its principles, benefits, and constraints. Because successful applications of the lexical approach are most often found in the area of ESP, it is suggested that comparing the proposed research contexts with that of ESP might, in the use made by both types of specialised corpora, shed light on the potential for the design of a reading syllabus for young EFL learners in Japan. I also reviewed some literature corpora and their application to teaching, with a discussion of some relevance to current

teaching. It is suggested that a close analysis of children's literature by means of a corpus approach is likely to contribute to the understanding of not only the linguistic characteristics of the texts, but also the ideology and culture represented in texts written for children.

A historical overview of what some researchers regarded as 'pedagogically useful' was also presented together with an explanation of their criteria and methods of selecting items. With the influence of corpus linguistics which has challenged the boundary of grammar and lexis, the phraseological view of language seems to exert an influence on the choice of pedagogical items. This shift seems to imply some potential for effective teaching. Nevertheless, criteria such as saliency, learners' needs, teachability and learnability, in particular, calls for intuitive judgments on the part of researchers, and it was pointed out that the process of defining what should be identified as 'pedagogically useful' is not yet systematised. Section 3.6.2 attempted to define 'pedagogically useful items' for Japanese learners from SLA perspectives. It was pointed out that the information from SLA evidence is useful as it can point towards what is problematic for learners.

In the light of the review and discussion in Chapters 2 and 3, there does seem to be some potential for applying a corpus-influenced syllabus, derived from a corpus of children's literature, to the teaching of EFL to young learners in Japan. However, in order to be successful and practical in the pedagogical implementation, I want to stress that not only should the current educational situation be understood, but also a clear rationale for the choice of texts should be established for the purpose of teaching

literacy to children. It is also necessary to investigate which lexical items and patterns appear to be pedagogically important in the corpus, alongside the methodological issues involved in finding them, and finally to consider a syllabus which reflects the findings derived from the analysis, as well as the teaching methods for pedagogical application for this particular group of learners.

Therefore, for the successful implementation of a corpus-influenced syllabus, there are largely two areas needed to be investigated and understood concerning primary ELT: one is related to cultural issues; and the other to linguistic issues. The next chapter explores the design of this study in reference to these two areas.

CHAPTER FOUR

STUDY DESIGN

4.1 Introduction

Chapters 2 and 3 reviewed the existing research relevant to this study and discussed the possibility of a corpus-influenced syllabus, based on a corpus of children's literature, for the teaching of EFL to young learners in Japan. It was pointed out, however, that it is necessary to deal with challenges in two contexts, for the successful implementation of such a syllabus.

This chapter firstly describes the two challenges faced in considering a corpus-influenced syllabus in primary ELT, which influenced the overall design of this study (Section 4.2). Specific objectives are summarised in Section 4.3. The subsequent sections describe the methodologies used to formulate research questions of this study (Sections 4.4 and 4.5).

4.2 Overall design of this study

Based on the review of existing literature relevant to this study (Chapters 2 and 3), it was suggested that two contexts, one related to culture, and the other to more linguistic aspects, need to be investigated and understood, in order to achieve the successful implementation of a corpus-influenced syllabus. These two contexts may be summarised as follows:

1) Cultural Context:

1.1) Current literacy teaching in primary ELT in Japan

1.2) Teachers' and prospective teachers' point of view on:

- the current educational situation in primary ELT
- their attitudes towards literacy education (criteria for the choice of texts for the purpose of teaching literacy)
- current teacher education

1.3) Suggestions for improvement of teacher education

1.4) Methodologies most suited to Japanese pupils

2) Linguistic Context:

2.1) Investigation into which lexical items and patterns appear to be pedagogically important in the corpus and how these are relevant for EFL young learners in Japan.

2.2) The methodological issues involved in identifying the above items and patterns.

2.3) How to design a corpus-influenced syllabus which reflects the information obtained from the corpus.

2.4) Methodologies which are most suited to teaching children based on the findings.

Although the two contexts stated above were described separately for the purpose of presentation, it is important to note that the results need to be closely interrelated in discussing any pedagogical implications. As well as the case for the design of the proposed corpus for this study, it is necessary to consider both cultural and linguistic aspects in deciding the teaching methodology.

4.3 Specific objectives of this study: research questions

The following key questions form the foundation for the research.

Research question 1: The cultural context of Japanese primary ELT

- RQ 1.1** What are the views of primary teachers on introducing the teaching of English literacy in elementary schools?
- RQ 1.2** What texts are being used in starting to introduce English in Japan and what reasons do the teachers have for choosing them?
- RQ 1.3** What challenges do current teachers believe that Primary ELT in Japan presents for teacher education?
- RQ 1.4** What challenges do prospective teachers believe that Primary ELT in Japan presents for teacher education?
- RQ 1.5** What are the local, contextual considerations which need to be taken into account in planning a corpus-influenced syllabus for young learners in Japan?

Research question 2: Pedagogic corpus initiatives

- RQ 2.1** In the light of Qs 1.1-1.5, what texts should be included in a pedagogic corpus?
- RQ 2.2** What, if any, differences and similarities can be identified between a sub-corpus of texts written originally in English and those translated into English from Japanese?
- RQ 2.3** To what extent does each sub-corpus of texts written for children (English & Japanese) reveal aspects of culture?

Research question 3: Pedagogic corpus analysis methods and implications

- RQ 3.1** How can relevant pedagogically useful features of the texts be identified?
- RQ 3.2** How, if at all, might the data from a corpus of children's literature be used in designing a literacy syllabus in an EFL educational setting?

As mentioned in the previous section (4.2), the cultural and linguistic contexts are closely related. The objectives of RQs 1.2 and 1.5 are relevant to the discussion of RQ 2.1 (see Chapter 5), as those results feed into the design of the corpus (i.e. the

‘ethnographic approach’ of corpus creation (Flowerdew, 2002, p.111)). The details are explained further in Section 4.5.1.

Having identified the areas which need further investigation, the following sections (4.4 and 4.5) describe the methods which were adopted to pursue these two lines of enquiry.

4.4 Methods - 1: cultural challenges

In Chapter 2, it was pointed out that it is important to understand the local context (i.e. the specific context of Japanese TEYL) rather than trying to address the global context (i.e TEYL in general). Three surveys were designed in order to understand the local context of 1) current teachers’ experience and attitudes towards English literacy teaching, including the use of texts (Survey I); and 2) current challenges and needs in teacher education (Surveys II and III). The relevance of each survey and the research questions are summarised as follows:

Table 4.1 Summary of research questions for each survey

Survey	Relevant research questions (RQs)	
Survey I	RQ 1.1	What are the views of primary teachers on introducing the teaching of English literacy in elementary schools?
	RQ 1.2	What texts are being used in starting to introduce English in Japan and what reasons do the teachers have for choosing them?
	RQ 1.5	What are the local, contextual considerations which need to be taken into account in planning a corpus-influenced syllabus for young learners?
Survey II	RQ 1.3	What challenges do current teachers believe that Primary ELT in Japan presents for teacher education?
	RQ 1.5	What are the local, contextual considerations which need to be taken into account for young learners?
Survey III	RQ 1.4	What challenges do prospective teachers believe that Primary ELT in Japan presents for teacher education?
	RQ 1.5	What are the local, contextual considerations which need to be taken into account for young learners?

As can be seen, all three surveys contribute to the answer to RQ 1.5. The following section provides further details about each survey, including: 1) the objectives, 2) the design, 3) types of questions and 4) sampling and challenges.

4.4.1 Survey I: Survey on Japanese primary teachers' views

4.4.1.1 Main objectives and outline of Survey I

Survey I was designed in order to answer the following research questions: RQ 1.1) RQ 1.2) and RQ 1.5). The following table provides an overview of the survey.

Table 4.2 Outline of Survey I

Type of survey	Questionnaire	Duration	Dec. 2006 – Dec. 2007
Respondents	Primary school teachers in Japan who teach English	Sampling method	Non-probability sampling. Judgement/snowball sampling (Bryman, 2001)
Sample size	108	Distribution method	Technology-based delivery by email (Gorard, 2001)

4.4.1.2 Design of the questionnaire: procedure and methods

The questionnaire needed to be kept brief and clear, and the questions unambiguous (Brace, 2004, pp. 118-119) in order to allow respondents to complete it without difficulty. Most of the questions were closed, answers being multiple choice or yes/no. To identify the texts which are used by or are familiar to Japanese teachers, a list of texts was provided for their convenience, based on recommendations by researchers who promote Extensive Reading research (Furukawa & Itoh, 2004). Additional space was provided for the respondents to list books or texts which they used but which were not in the provided list. Two open questions were included, to which I sought to avoid

biased responses by providing choices. For details of question types and actual questions, see Section 4.4.1.3 below.

Another important point to consider when using a survey of this kind relates to ethical issues (Brace, 2004, p. 172). These should be carefully considered as a responsibility towards respondents. To each respondent I addressed a cover letter including information about my subject matter to give general information, as well as promising confidentiality of their personal data, assuring them that their names, email addresses and places of work would not be revealed, in order to ensure anonymity (Appendix 4.1). The explanation was kept brief in order not to bias responses by revealing the precise nature of the survey.

Piloting the questionnaire before the actual survey is an important part of the process (Brace, 2004, p. 163; Brown & Dowling, 1998, p. 67; Gillham, 2000, p. 42). My questionnaire was piloted informally with three Japanese teachers of English, in order to measure the average length of time taken to complete it and also to test the clarity and order of the questions. The pilot study indicated that the questionnaire took approximately 10 minutes to complete.

After the content and time needed were checked, a further small pilot survey was conducted with five primary teachers in one prefecture (Fukuoka, Japan) in December 2006, to test the quality of the questionnaire, as well as to indicate the kinds of responses most likely to be obtained from this survey. An important finding which emerged from this was that these teachers make use of Japanese storybooks translated

into English. Therefore, the original questionnaire was revised and further questions on the use of the translated Japanese books were added to the final version of the questionnaire (see Appendix 4.1, for a translated version of the questionnaire).

4.4.1.3 Survey I: Types of questions

The questions used in the survey are categorised into four types, as shown below:

Table 4.3 Four types of question used in the Survey I

Type of questions	Examples
1. Closed question Yes/No questions	Q5: Do you use storybooks in your classroom?
2. Yes/No questions with contingency questions	Q2: Have you ever attended any training programmes before starting to teach English at primary level? Q3: If you answered “YES”, to Q2, please name the institution or organization which gave you training. Q7: Do you use Japanese stories translated into English in your classroom? Q8: If you answered “YES”, please list them. Q9: If you answered “NO” to Q7, would you use translated Japanese storybooks or texts, if they were available?
3. Multiple-choice questions which allow multiple answers	Q1: In what kind of institution(s)/school(s) do you teach? Q4: Which age group(s) of students do you teach? Please select all that apply. (a. 0-3, b. 4-5, c. 6-7, d. 8-10, e.11-13) Q6 (a) Please look at the list of titles below, and indicate which books you are familiar with by putting “X” in the first left hand column. (b) Please indicate which books you have used by putting “X” in the second left hand column.
4. Open questions	Q10: Please would you clarify the reasons why you chose these specific reading materials, and what other factors influence your choices generally. Q11: Please give your opinion regarding the introduction of English literacy work to children at primary school level.

The method of analysis differs according to the type of question. This will be further discussed in Section 4.4.4.

4.4.1.4 Sampling and challenges encountered

Judgement sampling (Bryman, 2001) was employed first in selecting which prefectures to survey, based on information about the pilot schools to which MEXT assigned experimental English activities.⁸ The reason was that they were the only schools that were officially carrying out English classes at the time of investigation. The Educational Boards in the prefectures concerned were contacted directly by phone. Some did not agree to cooperate, but others were willing to help in conducting the survey. Those in charge of the primary education sector distributed the questionnaire to each of the teachers in the state funded primary schools. Some respondents offered to distribute the questionnaire to their acquaintances or colleagues after they completed their own questionnaire (i.e. snowball sampling, Bryman, 2001). Bryman (2001, p. 99) points out that ‘the problem with snowball sampling is that it is very unlikely that the sample will be representative of the population’. However, in the case of the situation above, the respondents knew that the survey was directed to primary teachers who undertook English teaching activities, so this method of distribution would have been unlikely to affect the representativeness of the survey or the results. It did, however, mean that an accurate response rate could not be calculated, as the number of questionnaires eventually distributed is unknown.

In addition, since English activities at primary school result from a relatively recent change in policy, those prefectures which have pilot schools were deliberately chosen in the first stage. This may have affected the degree of generalizability of the results, which may not be truly representative of the views of primary school teachers as a

⁸ http://www.mext.go.jp/b_menu/shingi/chukyo/chukyo3/015/siryo/04070501/002.htm

whole. However, considering the aim of the survey, the results can be regarded as sufficient.

4.4.2 Survey II & Survey III: Surveys on teacher education

4.4.2.1 Main objectives and overviews of Survey II and Survey III

The results of Survey I, conducted in 2007, indicated that there seemed to be a lack in teacher education. This result informed the decision to look at the current status of ELT teacher education for primary teachers. Like other Asian countries which have introduced similar policies which focus on communicative language teaching, there may be a gap between what is suggested in curriculum guidelines and what happens in actual classrooms (e.g. Adams & Newton, 2009, Carless, 2009, Deng & Carless, 2009). It is important to know what kind of training teachers have received in general, as well as their views on what should be done based on their experiences in the classroom.

In July 2009 and January 2010, two surveys (Survey II and Survey III) were conducted locally in Fukuoka prefecture, Japan. Fukuoka prefecture was chosen for its convenience, being the prefecture in which I work as a teacher trainer. Survey II was conducted with current primary teachers in order to discover:

- current teachers' attitudes towards their English ability and training they have received;
- what kind of frustrations and problems primary teachers face in conducting English teaching; and
- what current teachers expect prospective teachers (those enrolled in teacher training courses) to learn at university level before coming to teaching practice.

These relate to answering RQ 1.3, RQ 1.4. and RQ 1.5. The outline of Survey II is summarised below.

Table 4.4 Outline of Survey II

Type of survey	Questionnaire	Duration	July 2009 – Jan. 2010
Respondents	Primary school teachers in Fukuoka	Sampling method	Non-probability sampling Criterion sampling (Dörnyei 2007 p. 128)
Sample size	61 (28 in July 2009, and 33 in January 2010)	Distribution method	Delivery by post (Gorard, 2001)

The questionnaires were posted to ten primary schools (five in July, 2009 and five in January, 2010) in Fukuoka prefecture. These ten primary schools were chosen because they have been accepting university students (prospective teachers) as a part of their teacher training course, and they are therefore likely to have more knowledge and experience of the requirements of prospective teachers (in other words, this decision made use of criterion sampling; Dörnyei, 2007 p. 128). For each distribution, a cover letter, questionnaire sheets and a stamped self-addressed envelope were included. Eight questionnaire sheets were included in each distribution, which means that 80 questionnaire sheets were distributed in total. The cover letter included was addressed to curriculum coordinators explaining the objectives of the survey, and asking them to hand out questionnaires to teachers who have experience in conducting English activities. After the questionnaires were completed, each curriculum coordinator collected them and sent them back to me. The overall response rate was 76.3 percent.

Survey III was conducted with students who enrolled in a teacher training course for Teaching English to Young Learners (TEYL), and who have completed their teaching

practice at primary schools. This survey was conducted as a part of their ‘reflections’ on their teaching experience, and also in order to find out:

- what kind of aspects of teaching prospective teachers found difficult as well as what kind of knowledge they thought they lacked in teaching children at primary school;
- what aspects of their training course were useful in teaching pupils; and
- what kind of training they think should be given in conducting teaching practice at primary schools.

The Outline of Survey III is summarised in the Table 4.5 below.

Table 4.5 Outline of Survey III

Type of survey	Questionnaire	Duration	July 2009 – January 2010
Respondents	Prospective teachers (students who took a teacher training course at Fukuoka Jo Gakuin University, Japan.)	Sampling method	Non-probability sampling Criterion sampling (Dörnyei, 2007 p. 128)
Sample size	79 (24 in July 2009 and 35 in January 2010)	Distribution method	Paper-based delivery (Gorard, 2001)

The questionnaires for Survey III were distributed in class, so the return rate was 100 percent.

Some parts of the results obtained in these two surveys are compared with each other, in order to discern similarities and differences between the perspectives of current teachers and those of prospective teachers.

4.4.2.2 Design of the questionnaire: procedure and methods

Surveys II and III were designed to find out what kind of training in primary ELT is being provided for teachers in Japan, as well as what kind of training the teachers and prospective teachers feel that they need.

In relation to questionnaire design, Dörnyei (2007, p. 112) suggests that previously obtained materials such as ‘qualitative exploratory data gathered in interviews or student essays focusing on the content of the questionnaire’ can be used to draw inspiration for what could be included in the questionnaire. In the case of Survey II, the responses gained from eight primary teachers on the feedback form on the teacher training courses used in May 2009 were used for these purposes. The respondents were all course coordinators and people who were in charge of English activities at the time. Based on these responses to the kinds of training which they consider useful, the form of some questions was modified in order to ask what current teachers think prospective teachers need to learn, based on their own experiences. The university students’ responses were used as the basis for the multiple choices offered on the questionnaire. Then the revised survey was piloted with two cooperative teachers, and their feedback on the contents of each question and multiple choice items was taken into account in the final design. For example, one mentioned that ‘children’s psychology’ is an important aspect of what students should learn, so it was added to the multiple choices. The section in Survey II asking what kinds of training teachers had received was changed into a multiple-choice question so as to clarify for respondents what kinds of answer were required. Moreover, a section for ‘Others’ was added in case respondents needed

to add more. As primary teachers are usually very busy, the questionnaire was kept brief, taking only 7-10 minutes to complete.

Survey III was intended to elicit prospective teachers' views towards the training course and the usefulness of its contents. As a part of the course, students were constantly submitting teaching reflection forms, which were collected after each lesson (Appendix 4.4). These served as a basis to identify 'an item pool' (Dörnyei 2007, p. 112) for Survey III (Appendix 4.3).

4.4.2.3 Types of questions: Survey II and III

In order to identify the local needs of teacher education, it was necessary to find out what kind of training current teachers had received, and what prospective teachers felt they needed when they carried out their teaching practice. For these reasons, similar questions were asked in both of the surveys.

Survey II included three types of questions: closed questions, multiple questions which allow multiple answers, and open questions (Appendix 4.2). Survey III, which was conducted with prospective teachers, includes two types of questions: multiple choice questions which allow multiple answers and open questions (Appendix 4.3).

4.4.2.4 Sampling and challenges encountered

Non-probability sampling i.e. 'criterion sampling' (Dörnyei, 2007 p. 128) was employed for both Survey II and III. As for Survey II, first of all, schools were chosen based on the criterion that they had experience in accepting university students for their teacher training. After the schools were identified, I contacted head teachers to obtain

their permission for conducting the survey. Once permission was obtained, the questionnaires were sent to curriculum coordinators at each school, and they were asked to distribute the questionnaires to teachers who had conducted English lessons.

A slight problem occurred in the responses of open answers, as some teachers misunderstood the questionnaire as a form of feedback towards the actual students' achievements in the teaching practice. The comments were too general to include as a response to the open question. Therefore such comments were not included in my analysis for this investigation.

For Survey III, students who had completed their teacher training course were asked to answer the questionnaire. They were asked immediately after they completed their teaching practice, so that their memory was fresh. They were given enough time to complete the questionnaire in the class, so that they could answer open questions without rushing.

4.4.3. Methods of analysis of surveys

4.4.3.1. Recording data from closed and multiple choice questions

The following procedure was followed for the analysis of the data of the surveys, once collected. Each participant was coded with a survey number followed by T(number) (e.g. I-T1, I-T2...), so that no individual would be identified. Each item of data was categorised according to the type of question (closed, multiple choice questions or open questions) as illustrated in Sections 4.4.1.3 and 4.4.2.3. For closed and multiple choice questions, a coding method was created for data administration. The data was entered

into a Microsoft Excel spreadsheet for simple counting, calculating percentages, and cross tabulations. The data was thoroughly re-checked after entering. Answers to the closed questions were counted and the percentages calculated. The individual questions in the surveys were not designed to explore issues such as the influence of variables on outcomes, or the effects of characteristics of the respondents on their opinions. Furthermore, the size of the sample for each survey is small, so it was neither necessary nor appropriate to carry out any tests of statistical significance on the data. Analysing the open answers required different steps, which will be discussed in the next section.

4.4.3.2. Clustering the open answers - coding

In analysing the answers to the open-ended questions, content analysis is necessary in order to group the answers into types (Gillham, 2000, p. 63). Once the answers were categorised and coded, the percentages were calculated quantitatively for each category as well as qualitative analysis being carried out for each answer (Gillham, 2000, p. 69). Each answer to open questions was translated, and these translated answers were checked by a native English-speaking researcher who understood Japanese so as to ensure the accuracy of the translations. Examples of coding (taken from Survey I) are presented as follows:

Table 4.6 Coding example of answers to open question 1

Q10: Please would you clarify the reasons why you chose these specific reading materials, and what other factors influence your choices generally.		
Examples (my translation)	Attribute	Code
familiar content, plot, storyline	Familiarity	1
vocabulary, linguistically, simple structure	Level	2
repetition, repeated expressions, phrases	Repetitive expressions	3
illustration, colourful picture, design etc	Attractive design	4
N/A	Other/ No response	5

Table 4.7 Coding example of answers to open question 2

Q11: Please give your opinion regarding the introduction of English literacy work to children at primary school level.		
Examples (my translation)	Attribute	Code
“My pupils were eager to learn how to write what they say, or at least see the letters after spending three months of listening and speaking. It should not be restricted.” “Some pupils are motivated when they see what they hear is written in my storytelling activities.” “Pupils can do more than we anticipate, therefore we should not limit the possibility of their further learning.” “Linking the letter and sound (phonics) is important, so teaching literacy should be encouraged.” “I think it is worth teaching literacy after they have received some instruction.” “I think it is important” “Some children are frustrated for not being provided a chance to see how it is written” “Children seem confused with the Roman letter and English spelling. I myself do not know the way, but it is useful if we can teach them the distinction”	In favour	1
“It is up to the decision of the Ministry of Education.” “I have no specific opinion about it.”	Neutral	2
“I do not see the need.” “It is not too late to start this at secondary school.” “We (teachers) ourselves are not competent, therefore we should not teach English at all.” “It could interfere with the development of their Japanese.” “We must concentrate on Japanese teaching.” “English literacy is not needed.”	Disagree	3
N/A	No response	4

After each answer was coded and counted numerically, percentages were calculated. For detailed qualitative analysis, exact wordings used in the survey responses were reported.

4.5 Methods 2: linguistic challenges

As explained above, the surveys were conducted to establish the ‘cultural’ context which any corpus-influenced teaching of English to young learners in Japan would need

to take into account. In particular, the texts to be included in the corpus were identified using the results of Survey I. This section discusses decisions made for the design of the corpus proposed for this study. Bearing in mind the purpose of using the corpus, specific issues associated with the actual design of a corpus are discussed, including representativeness, balance and size, and copyright issues (Section 4.5.2). Descriptions are given in relation to the necessary procedure taken for the compilation of the corpus used and challenges encountered during the collection (e.g. obtaining permission from the copyright holders) and digitisation of texts (e.g. scanning, and removing irrelevant information). Section 4.5.3 gives a summary of the basic statistical data for each sub-corpus, and Section 4.5.4 describes the main procedure of the corpus analyses. The methods described in this section are relatively established methods that are relevant to answering RQ 2.2, RQ 2.3 and RQ 3.1. The comparative analysis of sub-corpora is conducted in order to answer RQ 2.2 and 2.3. In addition, one of the main research tasks for this study is to explore methods for identifying pedagogically useful features of the texts (RQ 3.1). As the main answer to this question is more methodological than substantive, it will be further discussed along with the results in Chapter 7 and 8. Therefore, this chapter describes only the initial methods used for the analysis of the corpus.

4.5.1 The corpus design: a ‘pedagogically informed corpus’

4.5.1.1 Principles of a specialised corpus

While ‘any body of text’ can be referred to as a ‘corpus’ (McEnery & Wilson, 2001, p. 197), in corpus linguistics the term usually means a ‘principled collection of texts’ (Biber, Conrad & Reppen, 1998) with the intention of distinguishing it from other

collections of texts, such as archives (McEnery, Xiao, & Tono, 2006, p. 13). It is important to know the principles and criteria for corpus design, for ‘the design criteria of a corpus allow us to assess its representativeness’ (O’Keeffe et al. 2007, p. 1). Many researchers have discussed the criteria used in corpus design (Biber, 1993; Biber et al., 1994; 1998; McEnery et al., 2006) and the importance of having a set of criteria in creating one’s own corpus. Most writers agree that the corpus designer should pay attention to sampling, representativeness, balance and size.

Decisions about these issues are largely influenced by the purpose for which the corpus is to be used (Kennedy 1998; Wynne 2005; Adolphs 2006). One of the main purposes of the corpus suggested for the present study is to assess its pedagogical implications, especially for syllabus design for teaching English literacy to young learners. Many researchers have outlined the implications for syllabus design of findings from large general corpora (Bernardini, 2000; Clear, 2000; Hoey, 2000; Hunston & Francis, 1998, 1999; Hunston et al., 1997; D. Willis, 1990, 1993, 1994). However, as discussed in Chapter 3, it is difficult to reflect a general English course as a whole and to design a prototypical lexical syllabus for it. Moreover, the findings from a general corpus are not necessarily always relevant to all learners and their needs.

Specialised corpora, in contrast, tend to be easier to incorporate into pedagogy in a way which will satisfy ‘the specific needs of particular groups of learners’ (Aston, 2000). Specialised corpora allow us to define more precisely the lexical and grammatical patterns found in particular types of text. These findings often directly inform the application to language teaching (e.g. designing syllabuses and materials) as well as

contributing to the ‘understanding [of] producing particular text types’ (Biber et al, 1994; J. Flowerdew, 1993; J. Flowerdew, 1996; O’Keeffe et al., 2007). In the area of ESP (as well as EAP and LSP), there has been ample research which makes use of specialised corpora, especially with regard to linguistic analysis (Bernardini, 2004; Biber & Finegan, 2001; Collins, 2000; Conrad, 2001; Coxhead, 2001; Gledhill, 2000; Harwood, 2005a 2005b 2005c; Thompson, 2000; Tribble, 2000; Williams, 1998) and the implications of such findings for teaching have also been suggested (Harwood & Hadley 2004, Hunston, 2009). Learner corpora containing samples of learners’ writing have also been suggested for use in ELT (Granger & Tribble, 1998; Nesselhauf, 2004; Seidlhofer, 2000; Tanko, 2004).

4.5.1.2 Approach to representativeness, balance and issues of size

The representativeness of the data is an important issue in corpus design. McEnery et al. (2006, p.19) suggest that sampling is closely related to the representativeness and balance of a corpus. Biber (1993, p. 243) defines representativeness as ‘the extent to which a sample includes the full range of variability in a population.’ It is true, however, as Kennedy (1998, p.52) suggests, that, as it is difficult to be aware of all types of text or their distribution in a population, ‘a “representative” sample is at best a rough approximation to representativeness, given the vast universe of discourse.’ This notion of ‘approximation’ toward representativeness applies to any case of corpus design, whether it is for representing general or specialised discourse.

The corpus used in this study can be categorised as a ‘specialised corpus’. It conforms to the notion of ‘a pedagogic corpus’ suggested by Willis (1990), one which contains

texts or materials which learners are likely to encounter in a particular pedagogical situation. It is important at the same time to ensure that the full variety of language of a ‘specialised’ discourse is represented (McEnery et al., 2006, p.15). In order to ensure the representativeness of the corpus, ‘an ethnographic approach’ (L. Flowerdew, 2002, p.111) was followed, as an additional study, identifying the texts used by, or familiar to, the teachers involved in teaching this curriculum, by means of the survey (Survey I, to be discussed in Chapter 5). Also taken into account were the recommendations of texts suggested by research in Extended Reading for Japanese children (Furukawa and Itoh, 2004). In relation to pedagogical considerations with specialised corpora, the texts which are identified as suitable for the target learners were collected with careful attention to their authenticity.

There are many discussions available on the optimum size of a corpus for a specific purpose. The size of the corpus suggested for this study is not ‘big’ by any standard, as can be seen in Section 4.5.3. However, as far as corpus-based language learning is concerned, ‘bigger is not necessarily better’ (Tribble, 2001). Aston (2000, p. 10) states that analysing specialised corpora can expose ‘unsuspected regularities’ in certain types of text. He goes on to highlight the benefits of working with smaller but specialised rather than general corpora, pointing out that ‘it is easier to relate lexicogrammatical features to particular pragmatic functions, given the recurrence of these functions, often in similar positions within texts of the same kind’ (ibid, p.10). O’Keeffe et al. (2007, p. 198) suggest that it is easier to define and delimit ‘the pedagogical goals in terms of how they are used and applied’ with specialised corpora.

Perez-Paredes (2003) suggests that it is not necessary for corpora for specific purposes to be as large as a general corpus as long as the texts that go into the corpus are well considered. Bowker and Pearson (2002, p. 48) also favour this position, suggesting 'even corpora of between a few thousand and a few hundred thousand words have proved useful for language for special purposes (LSP) studies.' In other words, a small corpus which is well designed can be representative enough for the investigation of the specific area in question. Hunston (2002, p. 26) also states that a corpus cannot be assessed as 'good' or 'bad', but should be assessed in terms of its suitability for the purpose of research. L. Flowerdew (2002, p. 111) suggests that, as researchers such as Gledhill (1996), Hyland (1998, 2000), and L. Flowerdew (2000) promote utilising information obtained from 'specialist informants' in a specific discourse community in compiling and analysing corpora, it is more important to create a corpus which is representative and reliable, and which suits the purpose of the research. That is to say, a specialised corpus is more reliable than a general corpus in terms of representativeness, if the sampling standards are clearly designed. Similar points are made by Hunston (2002, p. 202) and O'Keeffe et al. (2007, p. 198).

Not only self-compiled corpora such as those used in many ESP or LSP studies (Beaugrande 2001; Coniam 2004; Chaeng and Warren, 1999; Harwood, 2005a 2005b 2005c, Simpson 2000), but also sub-corpora or portions of texts included in large corpora are often used in order to suit the purpose of specific research. For instance, Sealey and Thompson (2004, 2006) created the CLLIP (Corpus-based Learning about Language In the Primary-school) corpus, which contains 40 texts written for children extracted from the British National Corpus (BNC). Moreover, Thompson and Sealey

(2007) conducted comparative analysis of three corpora (CLLIP corpus, COMP corpus, which comprises fictional texts written for adults, and Newspaper corpus), all of the texts of which were extracted from the BNC. In this second study, the number of texts in the CLLIP corpus was reduced to 30 texts, removing all the examples of non-imaginative fiction, in order to achieve one of the researchers' purposes: to compare imaginative fiction between texts aimed at children and those for adults.

As Engwall (1994, p. 51) mentions, '...no scientific criteria exist for determining the size of any corpus. It has to be decided simply with reference to a balance of depth and breadth, but the lack of resources sometimes restricts the desired design', which is especially true in the current study due to the fact that the materials desired were all in copyright (details are discussed in the following section). Moreover, in relation to the size of the corpus, Biber et al. (1998) stress the importance of paying attention to the number and range of texts for the compilation of a representative corpus. Based on the study by Biber (1988), Biber (1993, p.245) estimates that at least 59.8 texts are needed in order to retrieve reliable linguistic features pertaining to high-frequency items. On the other hand, he reports that much larger sample texts (1,190 texts) are required in order to study infrequent grammatical constructions, with each text sample of 2,000 words in length.

Considering the discussion of representativeness, size and balance the corpus used in this study would appear to be suitable for the investigation of linguistic challenges for the literacy teaching of primary ELT in Japan. Specifically, the suggested specialised corpus in this study is suitable for the investigation of patterns that may have

implications for the teaching of reading, by providing answers to the question of ‘what to teach’. The following section discusses the actual procedure of corpus creation in detail.

4.5.2 Corpus building procedure: decisions and challenges

4.5.2.1 Specifying the content of the lists

Survey I generated a preliminary list of texts – i.e. those which were identified by teachers. As Survey I was conducted with teachers who were relatively active in the TEYL field in Japan (on the recommendation of various local educational boards, and some private schools), and also considering the fact that literacy teaching was not being conducted in many schools at the time of data collection, it was decided to list all the texts mentioned as being ‘used’ by the teachers. As for familiar texts, five teachers have to recognise them in order to be included as a candidate for the corpus of this study.

4.5.2.2 Collection of texts: dealing with copyrighted materials

Since the texts in the lists are mostly contemporary, all of them are in copyright. To reproduce them in electronic format, it was necessary to obtain permission from copyright holders. Due to the commercial interests of the publishers, as also pointed out in Kaprowski (2005), it is often difficult to obtain permission to use the whole text or even obtain a reply. At the initial stage of the data collection, it was considered whether it is enough to use 5 per cent of each text, since this would allow the use of texts without breaching copyright. However, this percentage was considered to be too small, considering that children’s storybooks tend to be short in the first place. Therefore it was decided to use whole texts as far as possible, and letters and emails were sent to

request the relevant permission (see Appendix 4.5). It was stressed in this correspondence that the texts would be used only for non-commercial research purposes. Permission was obtained for 200 texts (130 English original texts and 70 Japanese translations) out of approximately 330 texts requested. Once such permission was obtained from the copyright holders, the texts were converted into text files in electronic format.

4.5.2.3 Formatting, naming and mark-up of files

Since the texts were not readily available in electronic format, each text was scanned and converted into text files in Microsoft Windows text format (.txt), using the OmniPage⁹ OCR (Optical Character Recognition) programme. Each text was proofread after the scanning and converting was completed, since, although OCR software was used, sometimes the programme could not recognize some letters and misread (e.g. - was recognised as @; ‘ as ^ ; ‘H’ as ‘II’ etc.). These mis-readings were post-edited during the proofreading procedure and checked individually against the original text.

Each file was given a name which encoded the following: 1) original language of text: whether originally in English (E) or translated from Japanese (J); 2) source text (group, series); 3) serial numbering in the group; and 4) brief title. In addition, the texts which had been identified by Japanese teachers in the survey were marked: U meaning ‘used by teachers’, or F ‘familiar to teachers’. For example, J_CFS_1_U stands for a Japanese original (J), belonging to the Children’s Favourite stories series (CFS), number 1 text in the group, and used by teachers.

⁹ Copyright NewSoft Technology Co., Ltd.

Non-textual elements (i.e. pictures and graphics) can carry meaning, as they do in the case of picture books. However, the non-textual features are not under focus in this study and, furthermore they are much less susceptible to analysis using corpus methods. Therefore the descriptions of these aspects are excluded.

There are several mark-up schemes available (e.g. COCOA, TEI, CES) and mark-up languages, such as Standard Generalized Mark-up Language (SGML), Extensive Mark-up Language (XML) (McEnery et al., 2006, pp. 23-27; Sperberg-McQueen & Burnard, 2002). However, in this corpus, the texts need only to include headers with basic information about the text itself. This is marked simply by entering information between angle-brackets as <head> and </head>¹⁰, such as the title, chapter title (where appropriate), the names of authors and translators and the intended audiences. It might have been possible to mark linguistic and/or textual features (i.e. sentences, paragraphs, sections). Available software such as Xaira or MLCT (McEnery et al., 2006, p. 75) can be used for automatic mark-up of sentences and paragraphs. However, as the corpus used in this study is relatively small and my questions do not require them, such automatic mark-up schemes were regarded as superfluous to needs.

The Unicode (UTF-8) format was adopted when preserving these data since some concordance applications cannot handle other formats; Unicode is the most versatile text format for any corpus concordance program.

¹⁰ See Oxford WordSmith Tools (Scott, 2004) Help: Markup and Tags

The corpus was divided into two sub-corpora: English original texts and Japanese original translated texts , and I shall hereafter use the term CCL (Corpus of Children's Literature) to refer to the entire corpus, EOCL to refer to the English Original Children's Literature sub-corpus and JTCL to the Japanese Translated Children's Literature sub-corpus.

4.5.3 The corpus: basic statistical information

The corpus was compiled according to the principles discussed above (Section 4.5.1; 4.5.2). The basic answer to RQ 2.1 (i.e. In the light of Qs 1.1-1.5, what texts should be included in a pedagogic corpus?), is the texts included in the corpus I compiled, and these are listed in Appendix 4.6. Table 4.8 illustrates the basic statistical data for each sub-corpus used in terms of tokens, types, S-TTR (Standardised-Type Token Ratio), Mean Word Length and Word length standard deviation.

Table 4.8 Basic data for CCL (entire corpus) and each sub-corpus (EOCL and JTCL)

	EOCL	JTCL	CCL
No of texts	132	70	202
Token	331,407	44,803	376,210
Types (distinct words)	12,763	3,778	3,690
S-TTR	69.45	66.54	69.11
Mean word length (MWL)	4	4	4
Word length std. dev.	2.03	1.82	2.01

Token indicates the number of words used for word lists created by Oxford WordSmith Tools ver. 4.0. (Scott, 2004). Oxford WordSmith Tools ver. 4.0 gives two figures for tokens, one being running words used in text, and the other words used for word lists. The above table shows the latter. Types show the distinct words used in the corpus. The S-TTR indicates the degree of variety of vocabulary in a corpus. In calculating this score, the type/token ratio is calculated for each 1000 words (this is the standard value)

in the entire corpus, and a running average is computed. If this value gives a low number, it means that many of the same words are used repeatedly. If the value gives a high number, the texts include a wide variety of words, and fewer words are used repeatedly (see Help in Oxford WordSmith Tools ver.4.0).

These basic scores show the quantitative tendency in each corpus. For instance, Type/Token Ratios (TTR) show ‘the mean frequency of each different word form’ (Barnbrook, 1996, p. 54). Barnbrook (ibid, p. 54) suggests the usefulness of TTR as follows:

Distribution of tokens between the types in a text can provide a useful measure of the degree of lexical variety within it and may even provide a starting point for examining lexical differences between different types of texts.

However, TTR is only useful when the texts or corpora compared are the same length. The S-TTR (Standardized Type/Token Ratio) score allows texts or corpora which are different lengths to be compared, and helps to measure the variety of vocabulary in a corpus. ‘The standardised type/token ratio (STTR) is computed every n words as Wordlist goes through each text file’ (for details, see Oxford WordSmith Tools ver 4.0, Help). By default, the number is set to 1,000, which means that if the number of words in each text is less than 1,000, the score of S-TTR does not indicate anything useful. In employing the measure in the analysis for this study, the default setting of the number needed to be changed. Given that the minimum word count among the texts is 339 (which is the word count in one of the Japanese translated texts), the number has been set to 100 (n=100) and computed using the *Wordlist* function in WordSmith. Both sub-corpora of the S-TTRs were calculated with a Standardised-TTR basis of 100. The

overall score for each sub-corpus should serve as a useful indicator when comparing one with the other. The result shows that there are some differences in the score of EOCL and JTCL (Table 4.8). The average S-TTR of EOCL is 69.45 and of JTCL it is 66.54, indicating that the texts included in JTCL tend to repeat words slightly more than those in EOCL do; and texts in EOCL use a slightly more varied vocabulary than those in the JTCL.

4.5.4 Corpus analysis: basic concepts and procedure

4.5.4.1 Wordlists, concordances and collocates

‘Wordlists offer an ideal starting point’ for the lexical investigation of texts (Scott & Tribble, 2006, p. 31). Using the corpus, a basic wordlist according to frequency was compiled, using the *Wordlist* function in Oxford WordSmith Tools version 4.0 (Scott, 2004). It is possible to compile a lemmatised wordlist (e.g. the base form ‘go’ would include all the inflected forms *goes*, *going*, *went*). However, the basic frequency lists are compiled according to types, i.e. morphological forms (Hoey, 1991, p. 240), since the lemmatized word can reveal only the partial usage of the word (Sinclair, 1991). For example, Sinclair (1991) reports on the difference and the meanings in the usage of *yield*, *yielding*, and *yielded*. As indicated by this example, it is important to remember that ‘not only different words, but different forms of a single lemma, have different grammatical distributions’ (Stubbs, 1996, p. 38). A simple example found in the corpus of this study also suggested that when ‘honey’ started with a lower case, it is usually referring to the food, whereas when starting in upper case, it is referring to a name or used as term of endearment. Such specific usage of a particular word might not have

been revealed if it was lemmatized. Therefore, it was decided not to lemmatize the corpus for this study.

Once the frequency lists were created, selected lexical items were analysed in context, using concordance lines, the *Collocates* function, and the *Pattern* function in Oxford WordSmith Tools. The *Pattern* function is useful, since it represents ‘the items which are most frequently found to left and right of a search word’ (Scott & Tribble, 2006, p. 40). This method is useful for finding candidates of ‘collocations’ in a window. It is common to look at ‘-4 +4’ or ‘-5 +5’ (4 or 5 words to the left or to the right). The window does not have to be symmetrical, and the span can be arranged in order to suit the purpose of analysis. It is important to remember, however, that as Sinclair et al. (2004, xxvii) point out, ‘the wider the span, the lower is the significance in general’. For my analysis, both ‘-4 +4’ and ‘-5 +5’, spans were tried, but mostly the span of ‘-5 + 5’ is employed.

4.5.4.2 Multi-word units: N-grams

Many studies have suggested that ‘multi-word phenomena are a fundamental feature of language use’ (O’Keeffe et al., 2007). For instance, Erman and Warren (2000) state that more than half of the lexical items (55 %) in any text are part of ‘prefabricated multi-word units’. One way of identifying multi-word units or sequences is to look at repeated n-grams (n to be specified as a number, e.g. 2-grams, 3-grams, depending on the number of words in the sequence). As one of the methods in identifying multi-word units and their frequency information from the corpus of the present study, *Collocate* (Barlow, 2004) and the *Cluster* function in Oxford WordSmith Tools (ver.4.0) (Scott,

2004) were used to calculate frequent n-grams in each sub-corpus. Detailed discussion of multi-word units will be provided in Chapter 7 (Section 7.3.2.1) prior to the analysis in Sections 7.3.2.2 and 7.3.2.3.

4.5.4.3 Multi-word units: Colligational frameworks

One of the advantages of using an automated program to find recurring sequences is that, in the first place, it can identify language usages (linguistic patterns) which are subliminal and not immediately accessible to human intuition (O’Keeffe et al. 2007, p. 64). However, the methods used for clusters and lexical bundles are reported as problematic, as they do not allow identification of MWUs which are of low frequency (Adolphs, 2006: 43).

Danielsson (2001; 2003; 2007) suggests an objective method which identifies multi-word units. Her approach is based on the belief that ‘words are distributed non-randomly in text, and that this non-random distribution carries information about meaning’ (Firth, 1957; de Saussure 1918/1959, cited in Danielsson, 2007). While there are ample examples of research in the area of identification of sequences, Danielsson (2007) further argues that it is ‘counter-productive’ to set the length of a sequence in advance, such as in many algorithms in identifying n-grams. She further explains that, as the analyses of the output of sequences with pre-set length soon tells us ‘there is no length specification to units of meaning’ (Danielsson, 2007). On this basis, she suggests a method which outputs meaningful units based on frequency information in a purely automatic manner, followed by subsequent analyses in concordance lines using the

information of downward/upward collocations (Danielsson, 2001; 2003). The details are presented as follows:

The downward/upward distinction of collocates was used with the cut-off point set at the frequency of the node word. ... The original argument for using the frequency of the node word as the cut-off point lies in the assumption that less frequent words carry more meaning. Hence any collocate appearing less frequently than the node word can be expected to carry more meaning. Such an approach assumes that the entire vocabulary of a corpus will be run through the software; certainly a lengthy process. (Danielsson, 2007).

Danielsson (2007) suggests a simplified method, which uses ‘an arbitrary cut-off point which may symbolise the distinction between function words and content words’. Overall, one of the advantages of the method suggested by Danielsson (2001, 2003, 2007) is that it is able to capture low frequency but nevertheless meaningful multi-word units, unlike the usual n-gram extraction methods. The details of the steps taken for the analyses are presented in Chapter 7 (Section 7.3.3.2 and 7.3.3.3)

4.5.4.4 Comparison of sub-corpora: Keywords and Database of Key Key-words

Cultural phenomena can be revealed by a corpus-based approach to texts, as Stubbs (1996) demonstrates through examples of ‘cultural keywords’, that is, ‘words which capture important social and political facts about a community’ (Hunston, 2002, p. 117). Baker and Freebody (1989) conducted a study of children’s books and found that certain collocations reveal ‘ideological messages about the social world’ (cited in Stubbs, 2000 p.169). Moreover, it has been reported that the close analysis of concordance lines enables world views to be discovered in texts written for children, as well as revealing the characteristics of these texts per se (Knowles and Malmkjær, 1996; Thompson and Sealey, 2007). This suggests that the corpus of this current study can be

further used to investigate the texts themselves in order to identify their characteristics and the views represented in them, as well as the prominent patterns used in texts written for children. Considering the benefits of children's literature, the use of a corpus of children's literature might help to build bridges between cultures and the design of teaching materials or tasks which contribute to intercultural competence in learners.

To answer RQ 2.2 (i.e. What, if any, differences and similarities can be identified between a sub-corpus of texts written originally in English and those translated into English from Japanese?); and RQ 2.3 (i.e. To what extent does each sub-corpus of texts written for children reveal aspects of culture?), a comparison was made between the English original (EOCL) and Japanese original (JTCL) sub-corpora. In comparing the frequency counts and occurrences between EOCL and JTCL, normalised figures (per million words) are used as the sizes of these two corpora are unequal.

Using the Wordlist comparison tool in WordSmith, I extracted the significantly frequent keywords by comparing the wordlists from the sub-corpus of children's literature translated into English from Japanese (JTCL) with a reference corpus – in this case the English sub-corpus (storybooks originally written in English, EOCL). This procedure helps to identify words which are positively key and those which are negatively key. The concept of 'positive and negative keyness' is defined as follows (i.e. see WordSmith Help file, Definition of keyness):

A word which is *positively* key occurs *more* often than would be expected by chance in comparison with the reference corpus. A word which is *negatively* key occurs *less* often than would be expected by chance in comparison with the reference corpus.

Once keywords are identified, they are further investigated in concordances to see why they are identified as key.

In addition, a list of ‘key keywords – words that are key in multiple texts in the corpus’ (Baker, Hardie, & McEnery, 2006, p. 97) was compiled for each sub-corpus, to ensure that a word selected as key is a representative key of its whole corpus (i.e. ‘Database of Key Key-Words’¹¹). This is important, since the comparison is made not between one text and another, but among all the texts in each sub-corpus. The following procedure was followed in order to create the database: 1) a wordlist was produced for each text in each corpus; 2) the keyword-list was then created for each file using the keyword function; 3) all the created key-word lists were then saved in one folder; 4) a database was created in an Excel spreadsheet, using the *Key Key-Word Database* function in WordSmith Tools.

In regard to RQ 2.3, the *Keyword*, *Collocate* and *Concord* functions in WordSmith were used to identify the particular words and phrases which illustrate the keyness of each sub-corpus in the search for words which have particular ‘cultural significance’ (Stubbs, 2002, p. 145). Stubbs (1996, p. 157) suggests that the ‘cultural connotation of the words’ can be found by looking at fixed and semi-fixed phrases and expressions. For this reason, the patterns around the keyword were also examined in the concordance.

¹¹ For details of the Database of Key Key-words, please refer to WordSmith tools Help (KeyWords index).

4.6 Summary

This chapter presented the objectives and the overall design of this study. The methods related to two contexts, the cultural and the linguistic, and these were described. Methodological issues and decisions made for the design of this study were also presented. This chapter has also explained the procedure used in creating the corpus and the methods of analysis, taking account of advice from previous research. In order to extract meaningful data from a self-compiled corpus, various standards should be adhered to. While I am by no means claiming that the corpus created in the current research is representative of all children's literature, I explained how I created it to represent as far as possible the texts – and therefore the linguistic patterns - which Japanese primary learners and teachers are likely to encounter.

Sections 4.5.4 to 4.5.6 presented the corpus analytical methods which were used as the starting points of the analysis. It has to be mentioned that one of the main objectives of this study is methodological, i.e. an attempt to establish methods for using a corpus to identify items with the potential to be pedagogically useful ('Potentially Pedagogically Useful Items', or PPUIs) particularly in the development of literacy skills. A detailed explanation of methods will be presented prior to explaining what was revealed by the application of different methods to the corpus, in Chapters 7 and 8.

In the following chapters 5 and 6, cultural contexts are addressed. Chapter 5 reports the findings of the three surveys and considers the local and contextual needs for primary ELT. Chapter 6 discusses the comparison of cultural aspect of sub-corpora used in this study.

CHAPTER FIVE

ENGLISH LITERACY TEACHING AND TEACHER EDUCATION IN JAPAN

5.1 Introduction

In Chapter 2, it was pointed out that little attention has been paid to English literacy teaching at primary level in Japan, as there is often a misinterpretation of ‘communicative’ to mean ‘predominant focus on speaking and listening’ among practitioners. Moreover, a number of issues were identified in the existing literature, especially the lack of teacher education and training for TEYL was highlighted.

This chapter reports on the results of three surveys conducted in order to explore the cultural contexts discussed in Chapter 4. Section 5.2 especially describes the issue of introducing English literacy teaching to Japanese EFL young learners, based on the results obtained from Survey I. Section 5.3 mainly addresses the issue of teacher training for primary ELT based on the results from Surveys II and III. In the light of the results and discussion in Sections 5.2 and 5.3, I attempt to examine the local and contextual considerations needed for a potential corpus-influenced literacy syllabus in Japan in Section 5.4.

Finally, for the qualitative accounts of the analysis of the surveys, direct quotes are used with the following convention throughout the chapter: [Survey number – T(teacher) or PT(pro prospective teacher) number] (e.g. [I-T8] means teacher 8 in Survey I; [III-PT8] means prospective teacher 8 in Survey III).

5.2 English literacy teaching in primary ELT: based on Survey I

The main purpose of Survey I is to discover 1) teachers' attitudes to and opinions on introducing English literacy work at primary level; 2) the current use of children's literature and the reasons for the primary teachers' choice of texts; and 3) local and contextual considerations needed for Japanese primary ELT. In this section, the acquired data is analysed and discussed in terms of the three points raised above, after the characteristics of respondents have been reported.

5.2.1 Characteristics of respondents and their attitudes towards literacy work

Below is the summary of the respondents, indicating whether they have received some kind of training or attended workshops, listed according to the kind of institutions they work at (e.g. state-funded, private, other).

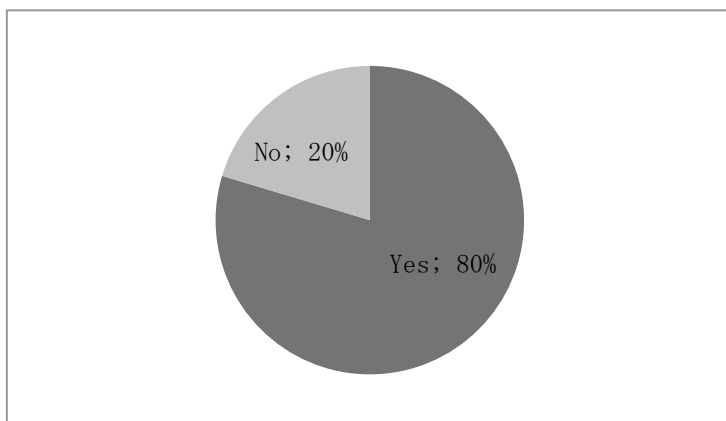
Table 5.1 Characteristics of samples: Survey I

	With training and/or workshop	Without	Numbers of respondents
State-funded	85	17	102
Private	5	0	5
Other	1	0	1
Total	91	17	108

5.2.2 The use of storybooks in any form of teaching

In response to the question whether they use storybooks or not, those who use some kind of written materials or books accounted for 80 % (86 people) and those who do not use them accounted for 20 % of respondents (22 people).

Figure 5.1: Use of written materials and/or storybooks



The above result (Figure 5.1) indicates that the majority of teachers use some kind of written materials and/or storybooks in their classroom. For those who answered ‘yes’, (i.e. using storybooks), the kind of activities they carry out are storytelling, drama, picture-story-show etc. 20 % answered that they do not use storybooks. Of these 20 %, 15 people provided their reasons for not using any written materials. Those reasons are categorised into two categories. One is related to policy, and the other is related to their concern about their skills. 11 people gave reasons for not using the storybooks or written material similar to the examples given below:

‘The MEXT guideline would not encourage us to teach using written materials’ [I-T3]

‘Communication should be stressed, therefore we do not encourage students to read’. [I-T18]

While the majority of those who provided their reasons emphasised the fact that they are legitimately following the rules set by MEXT as above, a few teachers expressed their concern about their method of teaching or teaching English in general as follows:

‘I am not generally confident to teach English’ [I-T6]

‘I fear that I teach them to read in an incorrect way.’ [I-T32]

‘I do not know the way to teach reading. I am not trained as an English instructor therefore I should not be teaching pupils English’. [I-T81]

It has to be mentioned that, generally speaking, primary teachers were not responsible for teaching English until the Action Plan was issued in 2003. Therefore, even if their workplace might have been chosen as pilot schools, this does not mean that they are trained for or confident or enthusiastic about teaching English.

In addition, out of the 22 people who do not use storybooks in the classroom, 14 of them were those who had never received any prior or in-service training. In order to investigate whether there is any difference between the groups and the choice of using storybooks in their classroom (those who had training, and those who never received training), a chi-square test was performed using Microsoft Excel 2007.

Table 5.2 Comparison of teachers with training and without training

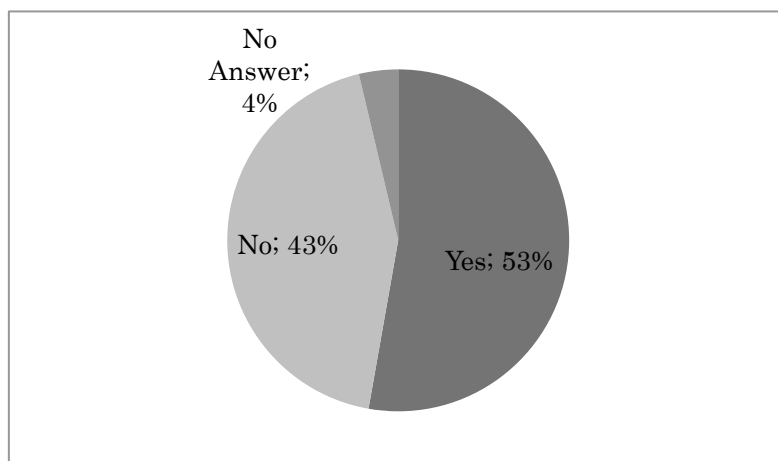
	With training and/or workshop	Without any training	Total
YES	83 (91%)	3 (18%)	86
NO	8 (9%)	14 (82%)	22
Total	91	17	108
$\chi^2=9.734$ $\chi^2_{Yeats}=46.051$ $df=1$ $p=47.5477E-12$ ($p<.001$)			

From the result above, it can be said that there is a difference between those who had training and their likelihood of using/introducing storybooks in their lessons.

It is worth mentioning that institutional variations may have influenced some of the data presented here. However it is not within the scope of this study to compare the differences in training the respondents have received and its influences. Nevertheless, as the result above (Table 5.2) shows, teachers who received some form of pre-service or in-service training are more in favour of using storybooks.

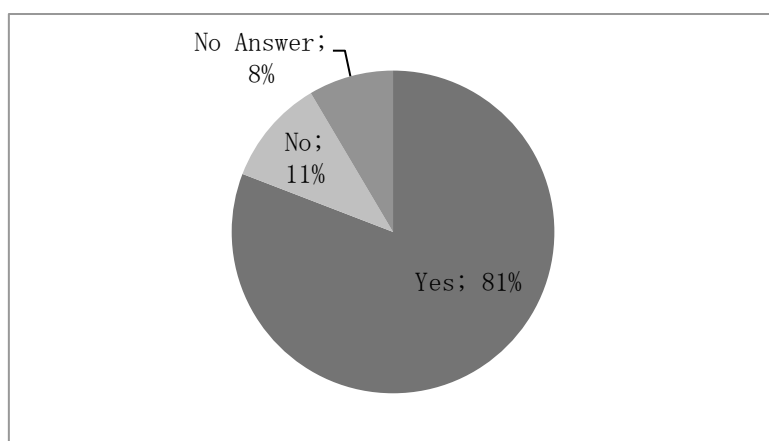
In relation to the use of Japanese translated storybooks, as the chart below shows, 53 % (57 respondents) of respondents answered that they use some kind of storybooks which were originally written in Japanese, either in the form of translation or retelling.

Figure 5.2 The use of translated Japanese storybooks



In the case of those who answered 'no' to the use of Japanese translated storybooks (47 respondents), a further question was asked, namely whether they would like to use them if they were available to them for classroom use. 80 % (38 people) answered 'yes', and 11 % (5 people) answered 'no', while 9 % (4 people) left the question unanswered (See Figure 5.3. below).

Figure 5.3 Use Japanese translated texts, if available?



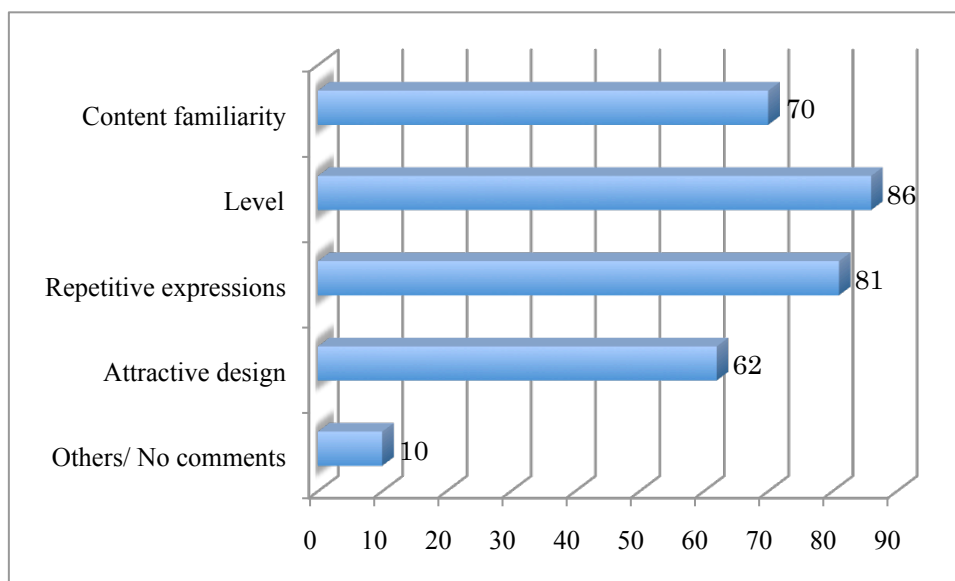
From these results, it could be speculated that the main reason for not using Japanese translated storybooks is just the matter of availability. These results indicate that most teachers would be keen to use such materials if they were available. The 11 percent who disagree with the use of Japanese storybooks all answered ‘no’ to the more general question about the use of storybooks. Furthermore, those who answered ‘no’ all seem to be against the whole idea of the introduction of English at primary level. Some provided the reasons for this by means of a strong opinion such as *‘introducing English at an early stage would impede the development of their Japanese language ability’* [I-T74]. This opinion seems to be mainstream among those opposed to the introduction of English at the primary stage, as it has also been reported by other studies, such as those carried out by Benesse (2007).

5.2.3 Reasons for the choice of books/storybooks for the classroom: summary and key-points

Including those who answered ‘no’ to the use of storybooks, teachers were asked the reason for their choice of texts, and what they considered to be important when selecting books for teaching or presenting to learners. Some of those who answered ‘No’ to the use of textbooks, provided answers such as those reported in 5.2.2.

The keywords which came up in their answers are coded and counted (see Chapter 4, Section 4.4.4.2), the result is presented in the following figure (Figure 5.4). The number in the figure indicates the number of times those keywords were mentioned. It should be noted that some participants listed more than one reason.

Figure 5.4 Reasons for choosing books for teaching young learners



Below are the details of keywords.

1) Familiarity

The 70 responses which were coded ‘familiarity’ were of two types. Most frequently mentioned was the content of the stories, plot, storyline. Others reported the importance of ensuring the familiarity and appropriateness of the cultural information treated in the texts. This might explain why it appears to be popular to choose translated Japanese storybooks for learners as shown in Figure 5.2.

Matsuka (2001) mentions that children feel comfortable when encountering information with which they are familiar in stories. In addition, the use of picture books is also effective for the introduction of cultural information (Matsuka, 2001, p.171). For this reason, she especially recommends picture books which are published both in Japanese and English.

2) Level

86 responses were coded as 'level'. Responses included measurements of the level, such as degree of sophistication of the story itself, as well as linguistic features such as manageable vocabulary and syntax.

Cameron (2001, p. 92) points out 'the need to match stories to children's interests and language levels' as it may be difficult for learners to make sense of a story with a lot of new words. Level is the most frequently mentioned aspect of what teachers consider when they select their materials. However, how the teachers measure the level is another question, and subjective opinions may contribute to their measurement of their written material selection.

3) Repetitive expressions

81 responses were coded as 'repetitive expressions' which includes 'use of repetitions, repeated expressions or phrases' in texts. This result confirms that repetition is perceived to be important among teachers, as has also been suggested by existing research on language teaching and learning. Repetition and the personal involvement and participation in the story itself creates a stimulating and encouraging learning environment for pupils. It has a positive effect on vocabulary acquisition for children by participating in the discourse in context (Cameron, 2001 p. 95). Many researchers and practitioners in Japan (e.g. Matsuka 2001, p.170; Kuno, 1999, p 53; Iizuka, 1997 pp.140-141) stress the importance of repetition and creating opportunities to interact with written material in teaching children to read. Matsuka (2001, p.170) also suggests that integrating different activities (i.e. not only reading but also watching, listening to

the stories, or touching some objects) in relation to the storyline is important. These data confirm that this idea is shared by a lot of teachers.

4) Design

The design of the book code occurred 62 times. Qualities such as illustration, colourful pictures, and design which attract pupils' attention, were mentioned. This is a particular attribute of writing for children. For beginning readers, pictures are often useful in giving clues to children when they are interpreting the story (Hawkins, 1991, p.181). The process of interpreting pictures, discussions about them and so on, is important in learning to read and write (Hawkins, 1991 p.181). Indeed, children begin to write by means of drawing at an initial stage, as suggested by Vygotsky (1978, p.114-115).

5) Others/ No comments

One respondent mentioned the appropriateness of morals treated in the story. There were 9 respondents who made no mention of their reasons. This could mean that they had no opinion about the choice of texts, or they do not use storybooks at all hence they had no opinion. On the other hand, they may have been disinclined to write an answer, due to the open nature of this part of the questionnaire.

5.2.4 Books in use or familiar to teachers

Another aim of Survey I is to find out what kind of books or storybooks are used and known by Japanese primary teachers, the result of which is presented in the following table. Teachers were asked to identify books which 1) they use or used in the classroom, and 2) they are familiar with. The books which are recognised as familiar (at least by 5

teachers) are listed in Appendix 5.1 (texts used and recognised as familiar by primary teachers) with the number of teachers who ticked each column. Many teachers are familiar with the books listed initially but few of them are actually used. This may be to do with availability and also of the influence of the guidance provided by MEXT.

Appendix 5.2 shows the list of books which are not mentioned in the initial list in the Survey I (Appendix 4.1) but which respondents provided additionally. Although the number of teachers who responded in this way may be somewhat limited, this information is valuable in the sense that teachers positively recommend these. These lists are valuable results as teachers may be willing to extend the use of these books, if they are using them already in their classroom in some way or another. It is also important to notice that many of them are books originally in Japanese. As the results of the teachers' comments on the choice of texts represents, this result also confirms that teachers value the quality of familiarity in choosing books for teaching.

5.2.5 Conclusions and limitations of the findings of Survey I

Although Survey I is a worthwhile component of the present study, it is important to make clear that this is intentionally limited in scope, and that the research does not seek to explore in great detail the contemporary situation of TEYL in Japan. Moreover, while the number of respondents is too few to generalise the results, it should be noted that some Japanese teachers still feel uncomfortable with the idea of teaching English to their pupils, even though most of them have had or have begun to have some kind of teaching practice and training. Therefore, this survey posed the further question of whether teachers are receiving enough training, which is necessary or appropriate for

them in order to conduct English lessons. This aspect is further discussed in the following section (Section 5.3).

In relation to the introduction of literacy to young learners in an EFL environment, as Cameron (2001) points out, ‘there is much that remains unknown’ and therefore she suggests that it is necessary to rely on ‘clear thinking and carefully monitored practice as guides in the classroom’ (Cameron, *ibid* p.123). It can be said that this survey which sought teachers’ opinions and facts about practice, despite the small sample, was useful in revealing the relevant issues or defects underlying the current situation of primary English education in Japan. Further discussion is presented in the following section together with the results from Surveys II and III (Section 5.4).

5.3. Teacher education for TEYL in Japan: Survey II and III

As reviewed in Chapter 2 (Section 2.6.3), it was pointed out that the system for teacher training for primary ELT is somewhat limited. The main focus of this survey is to investigate the kinds of training teachers received and what needs to be improved in the current teacher training courses provided at university level. First, the results and discussion of Survey II, which was conducted with current teachers, are presented. This is then followed by a description of results and discussion of Survey III, conducted with prospective teachers.

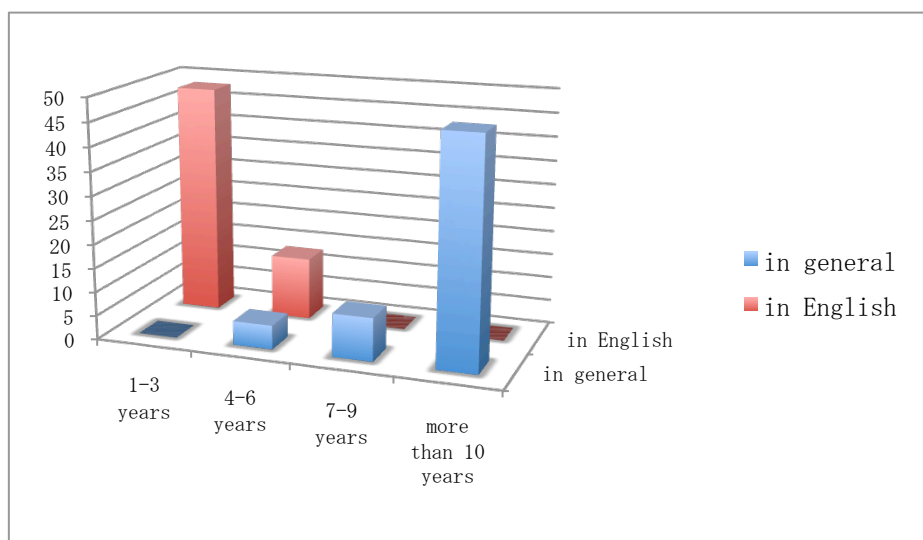
5.3.1 Views from current teachers: based on Survey II

Survey II was conducted with 61 teachers who are carrying out English lessons at primary schools. This survey was conducted in order to identify the needs from the current situation.

5.3.1.1 Teachers' teaching experience in general and in teaching English

Wide variations are found in teaching experience in general, as opposed to that in teaching English. As the results in the figure below (Figure 5.5) show, most teachers have less than 4 years of experience in teaching English, while a number report more than 10 years of general teaching experience.

Figure 5.5 Teachers' teaching experience in general and in teaching English

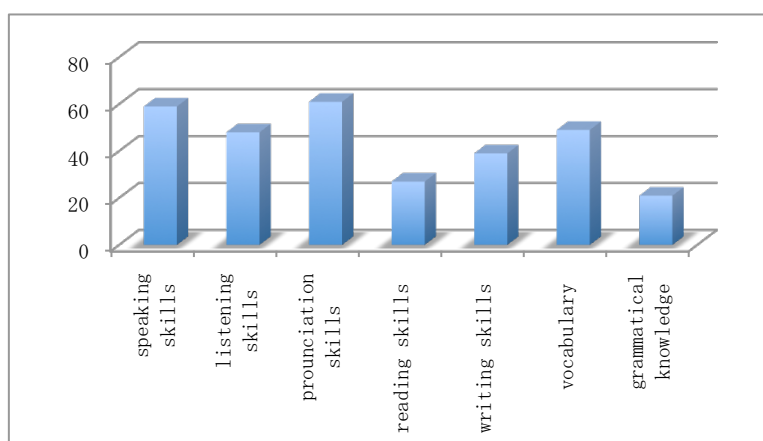


5.3.1.2 Teachers' perceptions of their English competence

In response to Q2 (To what extent are you confident in using English?), nobody selected 3= cannot say, 4 = fairly confident nor 5= very confident, rather mostly selected either 1= not confident at all, or 2= not very confident.

20 % (12 respondents) chose ‘not confident at all’ and 80 % (48 respondents) chose ‘not very confident’. Although this result might partially be due to humility on their part, it can be said that no one claimed to be confident. It is interesting to note that, by contrast, the majority of respondents answered ‘speaking skills’ (96.7 %) and ‘pronunciation skills’ (100 %) in response to Q3, which asked the areas of skills they feel they need to improve (i.e. Q3. In terms of your English ability, which of the following do you feel you need to improve? – a. speaking skills, b. listening skills c. pronunciation skills, d. vocabulary, e. grammatical knowledge, f. pronunciation skills, g. writing skills, h. reading skills, i. other?). This could be due to the fact that the objectives of English activities set out by MEXT heavily emphasise the spoken aspect of communication, and teachers are expected to use so-called classroom English, which seems challenging for primary teachers.

Figure 5.6 Teachers’ perception of the skills they need to improve



5.3.1.3 Kinds of training teachers received

As for the kinds of training, (Q5: So far, what kind of training have you received? multiple choices permitted), the areas of training teachers have received are;

- Material design for teaching young learners 75.4 %
- Teaching methodology & techniques 85.2%
- Theories about second language acquisition 63.9%
- Educational and language policy 6.5%
- Lesson planning 83.6%
- Observation at schools 9.8%

At first glance, the above results seem to suggest that most of the teachers have received some kind of training. However among those who answered, only two of them had received more than three sessions of training and the rest of them only received one session or a one-day workshop (Q6 How many times did you attend training sessions? - a. once, b. twice, c. three times or more).

5.3.1.4 Teachers' frustrations and specific problems

Most teachers gave the following responses to the open question of Q7 (What kind of frustrations or problems do you have in conducting English lessons? Please provide the details below.) The responses mostly fall into the three categories described below.

Table 5.3 Teachers' frustrations or problems in conducting English lessons

Keyword code/ theme	Examples	Percentage
Lack of time	<i>'Not enough time for making materials, meetings between teachers, planning and discussing lessons'</i> <i>'No time for evaluating our lessons'</i>	95%
Lack of training	<i>'Not much instruction nor support'</i> <i>'Not quite sure whether we are doing the right thing in our own ways of teaching'</i> <i>'Lack of instruction on how to use Eigo Note'</i>	86.8%
Lack of English skills	<i>'Communications between Japanese teachers and ALTs'</i> <i>'It is difficult to conducting English lessons as I am not confident about my pronunciation or speaking in English'</i>	96 %
Choice of materials	<i>'We developed materials over the years but now it seems that we need to use Eigo Note.'</i>	47.5 %

Most respondents provided short answers such as in the table above (Table 5.5), some provided detailed answers, which highlight the current problems and issues in primary ELT. Below is a discussion of each problem.

In terms of the lack of time, some teachers expressed very frank views such as the following:

English lessons are partly 'accessories' to us and we do not have time to put too much effort into them, simply because we need to have responsibilities and duties in teaching all subjects. As in English lessons children are not evaluated, we need to focus on other subjects which are subject to evaluation. [II-T1]

A major problem raised by the majority of teachers was the lack of time. In Japan, the educational reform of 'Yutori kyoiku' (low-pressure education) has received severe criticism, the main one being that this reform has caused a decrease in the level of intellectual standards of young Japanese people (Chujo et al, 2008, p.57-58). In the light

of this criticism, recent government policy has tended to revert school level teaching (primary, junior high and high school) to something close to what it used to be in the past. This new educational policy of increasing the number of study hours has in turn made it difficult for teachers to cope with all the extra teaching within the limited time. This effect may be at the root of comments such as the one above.

The following table (Table 5.4) shows the difference in the total of course per subject between the 2003 course of study and that of 2008, illustrating the overall increase in hours and the addition of ELT (MEXT, 2003; 2008). ELT at primary school is part of the ‘Foreign language activity’, with the main focus on English.

Table 5.4 Comparison of the number of study hours (Course of study 2003 and 2008)

	Subjects									Moral Education	Special Activity	Foreign Language	General Education
	Japanese	Social Science	Mathematics	Science	Life Science	Music	Arts and Crafts	Home Economics	Physical Education				
New course of study 2008	1461	365	1911	405	207	358	358	115	597	209	209	70	280
Previous course of study 2003	1377	345	869	350	207	358	358	115	540	209	209	0	430
Difference	+84	+20	+142	+55	0	0	0	0	+57	0	0	+70	-150

For this reason, ELT at primary school often tends to be regarded as something additional and receives little priority. The result of teachers’ responses regarding the lack of time seems to be evidence of this current situation.

While many teachers provided short comments regarding their lack of English ability (such as the one presented in Table 5.3), the following comment from a teacher [T2] concerns both the lack of English ability and lack of time.

ALT teachers only come by the terms of their contracts and arrive ten minutes before the start of the class. So there is not enough time to discuss with them about the content. Therefore we tend to rely on them as we do not have the ability to conduct lessons in English on the spot without preparation. [II-T2]

It is possible to infer the sense of burden this teacher [T2] feels from this comment, which also indirectly illustrates the lack of training in the current situation. Similarly, there were specific concerns related to the introduction of literacy expressed by two teachers.

Pupils learn the Roman alphabet in Japanese classes in their fourth grade, and they try to write what they hear in English lessons using their knowledge of Roman alphabetical writing (kunrei-shiki). Of course, there are often mistakes in them, and I feel I need to correct them but I do not know how to. I wonder if it is OK for me to neglect them just because the course guidance does not specify. [II-T57]

Now that students have 'Eigo Note', sometimes they want to write in their notebooks what they hear. They use katakana in writing. When we were students, we were criticised for pronouncing English words in the way we do in katakana. I feel that we need to know how to teach this. Otherwise the children might have difficulties when they enter junior high schools. [II-T48]

These comments made by teachers demonstrate the conflicts between what they faced in class and the course guidance. The expression ‘...but I do not know how to.’ found in comments made by teacher 57 [T57] suggests the lack in training, especially when it comes to literacy. Similar experiences are also described in Survey III in the next section (Section 5.3.2.3).

86.8 % of teachers mentioned the lack of in-service teacher training and 96 % mentioned their English ability as major problems. There were mentions of materials by 47.5 % of the respondents. While some teachers acknowledge the material published by MEXT, 'Eigo Note' (MEXT, 2009a; 2009b) as a step forward in primary ELT, others considered it 'extra work' for them to get used to. An example of this is seen in the following response:

We have conducted English activities within the 'Integrated Period of Study' for many years and established our own courses through trial and error. Now, we feel the need to switch to 'Eigo Note' completely because we must also consider the even level of pupils when they go into the junior high schools. [II-T5]

Here, this teacher has a dilemma over whether to use what they are used to, or to use 'Eigo Note' as the majority of the schools do. As mentioned in Chapter 2, 'Eigo Note' (MEXT, 2009a; 2009b) was introduced as a set of supplementary materials for English activities. This means that it is up to the schools or local educational boards to decide whether to use 'Eigo Note' or other materials. As for 'Eigo Note', even though it comes with a teachers' book which is written in Japanese, it is often criticised for the lack of training or explanations provided in English for ALTs. Of course, experienced teachers can devise appropriate activities which suit the pupils' learning even with limited resources. However, most Japanese homeroom teachers (HRTs) feel uncomfortable speaking in English, as can be seen in the results of their confidence in English (See Section 5.3.1.2).

In terms of training for current teachers, training sessions and workshops are usually provided by the local education boards. However, as the results show, even among

those who have attended training sessions, the majority have only attended once or twice. It is very rare to find teachers who have been to more than three sessions. Indeed, two thirds of teachers pointed out the ‘lack of training’ as one of the current problems.

5.3.1.5 Current teachers’ expectation of prospective teachers

In response to Q8 (What are the most important things that prospective teachers need to learn at university before starting to teach English?), teachers selected the following from the multiple choices.

Table 5.5 What teachers regarded as important for prospective teachers to learn

Choices	Percentage
a. Development of children’s psychology	3.2%
b. Knowledge about policy	13.1%
c. Lesson planning & teaching demonstrations	63.9%
d. Teaching methods (effective use of games/ materials/activities)	98.3%
e. General English skills	100%
f. Syllabus design and material design	75.4%
g. Others	8.1%

For the ‘others’ category, five respondents provided the following, which could also be categorised in the choice of ‘teaching methods’

- *Use of songs and phonics* (1)
- *Use of picture books, Use of story-telling* (2)
- *Effective teaching of writing (spelling), How to teach spelling* (2)

As demonstrated by the perception of their English competence, current teachers feel that it is important for prospective teachers to have basic English skills (100% of respondents), in addition to the knowledge of teaching methods (as identified by 98.3% of the respondents). The third highest response is ‘syllabus design and material design’ (as identified by 75.4% of the participants). The reason behind its importance can be inferred from the fact that teachers are required to plan a teaching syllabus for a year,

whether they use ‘Eigo Note’ or other materials, thus they consider it as one of the most important things for prospective teachers to learn.

5.3.2 Views from prospective teachers: based on Survey III

Survey III was conducted with 79 students in total who undertook a training course for TEYL. This survey was conducted in order to find out the needs from prospective teachers and their opinions about training received based on their experience in teaching at primary schools.

5.3.2.1 Difficult aspects of teaching pupils perceived by prospective teachers

In response to Q1 (What were the most difficult aspects of teaching pupils? (multiple choice permitted)), most prospective teachers found it difficult to manage time (81 %), and they had difficulties in dealing with unexpected circumstances (88.6 %).

Table 5.6 Difficult aspects of teaching perceived by prospective teachers

Choices	Percentage
a. communication with children	67%
b. time management	81%
c. the use of activities which suit pupils	21.5%
d. keeping children’s attention	70.8%
e. dealing with unexpected circumstances	88.6%
f. using English in classroom	82.2%
g. others	N/A

As can be seen, the aspects pointed out by prospective teachers were often concerned with their experience of teaching in general, which is understandable in the sense that they are still learning to teach.

5.3.2.2 Useful aspects of the training course

The following table (Table 5.7) presents the aspects of the training course found useful by the prospective teachers in answering Q2 (i.e. What were the most useful aspects of the training course when you taught pupils?).

Table 5.7 Useful aspects of the training course

Choices	Percentage
a. examples of activities used in classes at primary schools	67%
b. children's development	86%
c. foreign cultures	56.9%
d. use of classroom English	74.6%
e. current situations and issues in TEYL in Japan and other countries	89.8%
f. teaching practice at primary schools	94.9%
g. teaching methods suitable for pupils	84.8%
h. teaching methods suitable for different skills	84.8%
i. observations at primary schools	87.3%
j. lesson planning and teaching demonstrations	91.1%
k. theories about second language acquisition	44.3%
l. material design and syllabus design	84.8%
m. educational and language policy	11.3%
n. phonology	15.1%
o. others	N/A

The practical aspects of the training (such as teaching practice, observations at primary schools, lesson planning and teaching demonstrations) received the highest percentages of acknowledgement from students. Information on the current situation in other countries (mostly in Asia) received a more positive response from the prospective teachers compared to the in-service teachers. It is often said that Japanese primary ELT is somewhat behind other Asian countries, which might have provoked their interest in what is happening in other countries.

5.3.2.3 Specific aspects of knowledge prospective teachers felt were lacking

Question 3 asked prospective teachers to write comments on some specific aspects of knowledge that they felt were lacking when they did their teaching practice. The responses were categorised into seven keywords as below (Table 5.10).

Table 5.8 Specific aspects of knowledge prospective teachers felt were lacking

Keywords/themes	Examples	Percentage
Classroom management	Communication between children, Getting the attention of children	78.4%
	Making children listen to our instruction	
Literacy	I did not know how to teach pupils without using katakana	21.5 %
Lesson planning	Planning a lesson suitable for pupils	60.7%
	Planning a lesson which has links between activities	
General English skills	Speaking skills	70.8%
	Pronunciation skills	
	Differentiating Japanese sounds and English sounds	
Flexibility	I could not adapt to unexpected circumstances	82.2%
	I was in a panic when pupils did not do what I wanted them to do.	
	When we finished the lesson early I was not able to think of another activity at hand.	
Time management	Finishing every activity in time	68.3%
	Sometimes we finish lessons too early.	
Children's psychology	It was pointed out by teachers that I was not considerate towards children's psychology even when we play games	88.6%
	Activities we assumed suitable for children were in fact boring as it was too easy for them.	
	Choosing the right activities which suit pupils	

Similar to the current teachers in Survey II, prospective teachers also felt their English skills were not enough.

Survey III with prospective teachers revealed some difference in needs from the perspectives of current teachers. Especially, classroom management seemed to have been a problem for them. For instance, the following comments refer to the discipline of the pupils.

Sometimes the choice of activity was not suitable, because pupils got too excited and we could not manage the class well. [III-PT9]

Even if we tried to ask them to keep quiet, they did not listen to us and kept doing the activity or talking. [III-PT58]

We could not control the students. If we want to be a teacher we should be able to keep them quiet and make them pay attention to what we say. [III-PT65]

C. Kennedy (1986, p.169) suggests that ‘attitudes towards teaching and learning are related to cultural preconceptions of what constitutes good learning and good teaching’. One possible reason that prospective teachers responded as above could be that they have preconceptions of how pupils should react towards teachers’ instructions. In other words, they have the belief that ‘not listening to the teachers and their instruction’ is a form of bad behaviour.

5.3.2.4 What prospective teachers feel university students need to learn

As reported in previous sections, the majority of prospective teachers found it difficult to maintain children’s attention and interacting with pupils (5.3.2.1) or to be flexible when something unexpected happened in class (Section 5.3.2.3). Naturally, the majority pointed out that learning about the development of children’s psychology is one of the most important aspects that prospective teachers need to learn at university level (Table 5.9).

Table 5.9 Prospective teachers' view towards what should be taught in training courses

Choices	Percentage
a. development of children's psychology	89.8%
b. knowledge about policy	15.1%
c. lesson planning & teaching demonstrations	64.5%
d. teaching methods & techniques (effective use of games/ materials/activities)	65.8%
e. general English skills	79.7%
f. syllabus design and material design	58.2%
g. others (observation at schools)	7.5%

While 64.5 % of prospective teachers consider lesson planning and teaching demonstrations to be a necessary part of their training courses, some commented on the difference between teaching 'real children' as they are required to do in their teaching practicum, and the teaching demonstrations with 'classmates who are pretending to be pupils'. This may be partly due to the fact that while adults can sense what teachers want them to do, we cannot always anticipate how children will react to teachers' instructions.

5.3.3 In relation to literacy teaching

Although surveys I and II were not aimed to find out about teachers' attitudes towards literacy teaching, there were some comments relating to literacy teaching in the responses of the two surveys (Survey II and III). Some similarities are found in the responses given by current teachers (Survey II) and prospective teachers (Survey III). In terms of ELT at primary level in Japan, the focus has been heavily on oral English. In terms of the medium of instruction, the following quotes exemplify the limitation of teaching pupils by 'only oral English'. Two teachers reported as follows (my translation):

Although we have been conducting English activities only orally, we felt the need to present some kind of written English as children got older to go with whatever expressions we are using in class. [II-T31]

In some cases when pupils showed interest in how English is written, I tended not to prohibit this. It seems that, for some pupils, written forms help them remember what we teach. [II-T56]

Moreover, with reference to the avoidance of written materials, one teacher expressed the following concern:

It seems my students remember words or expressions by sounds. I believe it is not a bad thing in itself, but sometimes it is difficult to teach distinctions between singular/plural. (e.g. Do you like ~? lions, cucumber). [II-T15]

Such concerns, relating to common structures (for example, the fact that the Japanese language does not have singular/plural forms), have some implications for the introduction of corpus findings into the classroom, which is discussed in Chapter 9.

In addition, prospective teachers pointed out the same problems with regard to pupils making connections between the spoken form and written form. One student reported an incident she experienced when she had a teaching practicum, as follows:

The pupils asked me how to write the pronunciation of some words or phrases in katakana [(i.e. a syllabary used for foreign or borrowed words)] at the end of the lesson. I did not know whether it is appropriate to teach them in katakana. [III-P18]

It can be said that pupils try to transfer their literacy knowledge (in this case, *katakana*) in their L1 into their learning of English. Cameron (2002, p. 137) mentions that '[t]he way the child is being, or has been taught to read in the first language will create

expectations about how foreign language will be taught'. The following reports from two prospective teachers, illustrates the point raised by Cameron (ibid).

Children seemed to be interested in writing down what we taught. I think it is good to give some time to write down, even if it takes time to write it in their notebook. [III-P6]

When we showed the flashcards of our names when we taught 6th grade pupils, they understood how to read them out. We thought it might be difficult as we had learnt that MEXT cautions against the introduction of written English, but pupils might want to have words written. [III-P69]

As both teachers and prospective teachers (undergraduate students) above pointed out, it seems clear that the details of the language cannot be taught only orally, especially in an EFL situation where the learners' mother tongue and the target language are greatly different. For instance, prospective teachers have reported that pupils often failed to pronounce plural forms or inflected forms. This may have resulted from oral instructions being given without showing any letters as they remember only by what they could hear from the teachers or other pupils. One might say that there is no need to pay attention to minute details such as this. However, in Japanese there are no plural forms of nouns. This is one of the areas where there is a need for compromise in instruction, even though the focus should be on oral English.

The importance of remembering expressions at a discourse context level is highlighted as can be seen in the following report from a teacher:

When I asked my students 'How are you?', I expected them to answer with phrases such as 'I'm fine,' 'I'm so-so,' or 'I'm thirsty' etc. One of the students answered 'I'm hot chicken.' It might have been the influence of how to introduce themselves (e.g. I'm Hanako.)*
[II-T25]

This example somewhat resembles the one mentioned by Butler (2005) (i.e. 'I'm apple' as discussed in Chapter 2). It shows the need to present sequences not only at word-by-word or phrase level, but also to present expressions at a discourse level, showing the pairing of what goes with which expressions (Cameron, 2001). Otherwise some pupils may get confused even with a question like the above.

Considering there seems to be a need among some teachers for training in dealing with written English, more research in these areas is needed. Most practitioners who have been active in the field of TEYL in Japan strongly support the idea of using written materials (even just for presentational purposes) at the early stage of learning (e.g. see Iizuka, 1997, p.140-141; Kuno, 1999, p.53; Nakata, 1993, p. 128). In particular, Nakata (1993, p.128) points out the fact that delaying the exposure to written materials could result in confusing learners, as Japanese and English have completely different systems of writing.

5.3.4 Conclusions and limitation of the findings of Survey II and III

I am by no means attempting to claim that this small sample may have led to significant findings, but rather that views held by participants in these surveys may reflect current problems in the teacher training system.

As the result of Survey II has indicated, there is a lack in training which is one of the reasons for some teachers feeling that 'Foreign Language Activities' is a burden, even though it is going to be officially a compulsory subject in April 2011. For instance, while all teachers have attended some kind of training workshop, most teachers had it only once or twice. Some teachers expressed their worries over whether their way of teaching is appropriate for their pupils. As the results showed, teachers are experienced as primary teachers, but there was a big contrast between the number of years they are experienced in general teaching and their experience in English teaching. It was pointed out that further needs in enriching the training system for conducting English language teaching for current teachers is necessary as well as improving the teacher training system at university level for future teachers.

A more relevant finding was that, while neither the Survey II nor III was aimed at finding out about their attitudes towards literacy teaching, some teachers commented on the limitation of oral-only instruction, and expressed the need for some writing and reading when pupils are ready. There was also some disagreement expressed towards the policy provided by MEXT, as observed both by teachers and prospective teachers, suggesting that presenting written samples of language may not be a burden for pupils.

In summary, it seems that both current teachers and students who experienced English teaching at primary school feel that there is a limit to what teachers can deal with by only oral means. As other researchers such as A. Koda (2000, p.61) warn, purely spoken-focused instruction might cause pupils to dislike English when they enter junior high school, in which they start to be exposed to the explicit teaching of grammar. It is

obvious that there is a need for more research into EFL literacy teaching with some real empirical evidence from practice.

The results of Survey III conducted with prospective teachers are useful in identifying newly appointed teachers' needs. As is often the case, it is difficult for prospective teachers to turn their knowledge into practice. Moreover, while current teachers tend to focus on skills (in teaching, or in their English skills), prospective teachers seem to have a tendency to feel they are lacking in understanding children's psychology and they feel the need to gain knowledge of it.

Overall, the results highlighted that it is necessary to establish a practical teacher training course which is adapted to practical needs. Moreover, as for the content of the training, more than 90 % of respondents selected 'prior observations' (in Survey III), 'teaching practice' (in both the Survey II and III) and 'syllabus and material design' as useful or necessary components of teacher training. Since these elements rely on collaboration between in-service teachers and those involved in teacher training courses at university level, these results seem to indicate that close collaboration between the two sides is necessary for the success of primary ELT in Japan. As far as the research into TEYL in the Japanese classroom context, much remains unknown and there is certainly scope for further development.

5.4 Local and contextual considerations for literacy syllabus design

In the light of results suggesting that there are opposing ideas about English literacy teaching among teachers, it is worth discussing English literacy teaching and Japanese

teaching in relation to educational policy in Japan and SLA findings. This may help to clarify the local and contextual considerations for teaching English literacy in primary ELT in Japan.

Benessee (2007) conducted a much larger survey in relation to the introduction of primary English education in 2006. The aim of the survey was to clarify viewpoints of primary teachers and parents on the change of policy and its impact on the current education system. It is a contrastive analysis of perspectives held by teachers and those held by parents. The research found that there is a huge gap between the expectations from parents and the perception of the change (i.e. introduction of English education at primary level) from the side of practitioners. For instance, according to the result, while 76.4 percent of parents believe that primary English education should be introduced at an early stage as a compulsory subject, the figure for teachers is 36.8 %. In addition, it was reported that some teachers believe that it is better to wait until the pupils have fully established Japanese language skills and identity. This result seems to support the comments in the previous section, where some teachers expressed less enthusiastic opinions about introducing English at primary level.

In addition, there is strong opposition among some writers to the introduction of English literacy education in Japan (e.g. Ohtsu 2006; and Torikai 2006) as was also found in the survey conducted as a part of the current study. For instance, Kageura (1997) argues that introducing phonics with letters simultaneously puts pressure on language learners and can be an obstacle for learners, using an example of junior high school students disliking English. Indeed, it is true that the MEXT does not encourage the teaching of

literacy at primary school level, or even the introduction of phonics until junior high school (i.e. until the age of 13). However, many practitioners and researchers oppose this idea (see Butler, 2005: Noro, 2007 p.102) implying that the suggestion by MEXT seems to be based on ungrounded assumptions or speculation. This is also indicated by the way some respondents to Survey I expressed their opinions on the introduction of English:

'It seems rather late, and will never happen if one says you must wait until the children acquire Japanese completely' [I-T26]

'We learn our language on a day-to-day basis, and it is not certain as to the relationship between Japanese and English. It might be a rather pessimistic view if we restrict the learning opportunities' [I-T38]

'If children show interest, it is important to let them read – not to restrict the learning' [I-T87]

The local and contextual considerations for teaching English literacy needs to be carefully established, and not to be decided merely because MEXT opposes its introduction.

Judging from the answers about the reasons for teachers' choice of texts as identified in Survey I, it could be suggested that making use of authentic texts from both languages with the content familiar to pupils is necessary, along with some detailed attention towards level and the use of repetition. In terms of the treatment of cultural aspects, some teachers commented that it is important to maintain Japanese identity. In respect of this point, as discussed briefly in Chapter 2, it is often said that one tends to realise more about one's native culture or identity by contrasting it with those of others. Therefore, it is necessary to conduct a contrastive analysis of the sub-corpora in this

study in order to find some contrastive cultural aspects which might benefit pupils' learning of English. This will be discussed in Chapter 7.

Moreover, it is also necessary to think of the methodological aspect of literacy teaching. As expressed by some teachers in the surveys, it is necessary to consider what teaching method in literacy is appropriate for Japanese learners in particular, as their literacy knowledge in their L1 is not directly transferrable. As the empirical evidence from the Japanese classroom itself is currently very limited in the case of TEYL, it is important to value each response from teachers in considering the local and contextual considerations necessary for the design of the syllabus. This sort of discussion can in turn inform teacher training for TEYL.

5.5 Conclusion

As a result of the surveys, a number of challenges have been raised which, it is suggested, should be the focus of this study. The results from Survey I indicated some implications for the teaching of English literacy and the design of a corpus-influenced syllabus for primary ELT. It was also revealed that teachers who received some training tend to think that the teaching of literacy is necessary with the right timing. The choice of texts made by teachers and the reasons for their choice as identified in Survey I must feed into the design of a corpus (Chapter 4) and the choice of pedagogically useful items for particular learners. Nevertheless, as the results from Surveys II and III indicated, some teachers expressed their concern about the method of teaching literacy to their pupils.

Taking this into account, it was pointed out that it is necessary for material or syllabus designers to consider the teaching methodologies as well as identifying pedagogically useful items for learners, when making recommendations about teaching (Chapter 9). In addition, it was suggested that investigating the cultural differences might reveal some items which are worth presenting to learners, making use of children's literature (Chapter 6).

In summary, the discussion of local and contextual challenges highlighted the need to investigate the methods used in identifying items necessary for teaching English literacy for young learners in Japan, but methodological challenges also remain. In addition, considering the fact that most teachers feel uncomfortable with their level of English use, it seems to be worth investigating the implication of the findings from the corpus in teacher training (Chapter 9) (i.e. how the corpus findings can be integrated into a teaching approach).

CHAPTER SIX

COMPARISON OF SUB-CORPORA: CULTURE

6.1 Introduction

In Chapter 2, it was reported that ‘Foreign Language Activities’ is placed within the objectives of international understanding, according to the policy issued by MEXT (2003, 2008). Specifically, the course of study 2008 states that ‘Foreign Language Activities’ are aimed at helping pupils to have the ability to express their own cultural information to people in other cultures (MEXT, 2008, p. 8). Taking this into account, the potential of a corpus of children’s literature was suggested as a means of investigating the culture from which it originates (See Section 2.4.1 for details). Since cultural phenomena can be revealed by a corpus-based approach to texts (Stubbs, 1996, 2002), exploring the difference between EOCL and JTCL by means of the *Keyword* function available in Oxford WordSmith Tools (ver. 4.0) (Chapter 4, Section 4.5.4.4) was proposed. This section attempts to address the two following research questions:

- RQ 2.2** What, if any, differences and similarities can be identified between a sub-corpus of texts written originally in English and those translated into English from Japanese?
- RQ 2.3** To what extent does each sub-corpus of texts written for children (English & Japanese) reveal aspects of culture?

The following section reports on the methods and procedure of analysis (6.2.1), and the results of the comparative analysis (6.2.3). Finally, Section 6.3 summarises the results and discusses the relevance and implications for TEYL in Japan.

6.2 Cultural and linguistic differences between EOCL and JTCL

6.2.1 Methodology and procedure

The corpora used for this comparison are EOCL and JTCL (for the details of statistical data for these corpora, see Chapter 4 (Section 4.5.3)).

When investigating texts and the representation of relevant social aspects in those texts, it is often difficult to decide which lexical items to start investigating from (Mahlberg, 2007a, p. 196). One of the approaches is to make use of the *KeyWord* function in the software Oxford WordSmith Tools (Scott, 2004) as can be seen in existing research (e.g. Stubbs 2007, Mahlberg 2007b). This function uses two wordlists (one list usually smaller than a reference corpus) in order to identify significantly frequent keywords (those which are positively key and those which are negatively key) (see WordSmith Tools Help section: Definition of Key-ness). In comparing two wordlists, a statistical calculation is applied in order to extract the ‘keyness’. WordSmith Tools provides two statistical measures: chi-square and log-likelihood. Log-likelihood is often used in corpus linguistics, as the value is not influenced by the size of corpora (Dunning, 1993). Therefore, log-likelihood is chosen as a statistical measurement of comparison for this study.

In conducting this keyword analysis, it is necessary to choose a reference corpus. A reference corpus is a set of data which is necessary for the analysis of a text or set of texts using frequency-based techniques, which is compared against the corpus to be studied (i.e. target corpus). In selecting the reference corpus, it is important to bear

one's research questions in mind. Although there is still some debate on the best size for a reference corpus (Berber Sardinha, 2004; cited in Scott & Tribble, 2006, pp. 64-65), it is important to make sure that the comparison corresponds with the initial research questions. The main purpose of this comparison is to identify whether there is any difference in linguistic and cultural features between the English originals (EOCL), and the English in the translated versions of Japanese stories (JTCL). Given the purpose of the comparison, it is justifiable to use EOCL as a reference corpus in order to extract statistically significant items from JTCL. The following summarises the procedure for the comparison of two corpora in this section.

- 1) Wordlists of each of the EOCL and JTCL corpus were created.
- 2) JTCL was set as a target corpus, and EOCL as a reference corpus
- 3) Based on the created wordlists in 1), a list of keywords was compiled using the *KeyWords* function in WordSmith Tools (Scott, 2004).
- 4) The list of Key-Key Words, which lists 'representative keywords' was also compiled as described in Chapter 4 (Section 4.5.4.4).
- 5) Items were selected using the information obtained above, and analysed in concordances.

Knowles and Malmkjær (1996, p. 97) used the information of collocates to determine the features of certain representations of characters, or how some notion (e.g. friendship, family) are represented in stories from children's literature.

6.2.2 Comparison of keywords and key-keywords

First, the keywords were extracted from the comparison of JTCL and EOCL as a reference corpus (see Appendix 6.1 for the list of keywords). As mentioned in the previous section, in theory, it is necessary to compile a Key-Keyword Database (KWD) in order to look for so-called 'representative keywords'. The following table (Table 6.1) shows the comparison of the top 20 positive keywords in JTCL and the Key-Keyword

database (KWD). The reference corpus of these comparisons is EOCL in both cases. The keyword database (KWD) column shows the following information: 1) key word; 2) the number of texts which have the identified word as a keyword; and 3) the percentages of the keyword frequency in the corpus.

Table 6.1 Comparison of top 20 positive key words in JTCL and KWD

JTCL (positive key words)				Key-Keyword Database		
N	Key word	Freq.	%	Key word	texts	%
1	GURI	87	0.19	MONKEY	5	26
2	GURA	79	0.17	MAN	4	21
3	SAYS	108	0.24	OLD	4	21
4	RICE	72	0.16	RICE	4	21
5	KEN	57	0.12	VERY	4	21
6	MONKEY	90	0.20	WOMAN	4	21
7	CHUG	52	0.11	GURA	3	15
8	MAN	156	0.34	GURI	3	15
9	KINTARO	44	0.09	HE	3	15
10	THE	3071	6.86	ISLAND	3	15
11	CLOTH	50	0.11	SEA	3	15
12	FINALLY	62	0.13	ASKS	2	10
13	OLD	216	0.48	CAKES	2	10
14	AMY	38	0.08	CLOTH	2	10
15	BARNEY	35	0.07	HAPPY	2	10
16	SEA	60	0.13	LET'S	2	10
17	SUHO	32	0.07	MORTAR	2	10
18	SO	376	0.84	NEW	2	10
19	DOG	49	0.10	PRINCESS	2	10
20	SABURO	30	0.06	SAYS	2	10

The right column of Table 6.1 shows the rank order of keywords by keyness, extracted by comparing the wordlist of JTCL and EOCL. As the above table shows, it is normal for the positive keywords to be the words which are particular to specific texts in a corpus, such as words for the characters' names (e.g. *guri*, *gura*). The left column shows the Key Key-Words taken from the KKW database, which reveals the keywords which are representative of multiple texts and only those of positive keywords (see Help in Wordsmith ver 4.0). As can be seen above, one obvious difference between KW and KKW lists above is the rank order of identified keywords. Scott and Tribble (2006) comment that it is important to consider how many times those words occur across how

many texts, and therefore key-keywords are useful in identifying the representative keyword of a whole corpus. Taking this information into account, the items identified from the comparison of JTCL as a target corpus and EOCL as a reference corpus (Appendix 6.1) are checked against the key-keywords in the database in order to avoid making conclusions only by particular instances from a particular text.

6.2.3. Results and Discussion: comparison of keywords

The keywords in the original list (Appendix 6.1) are divided largely into five categories: 1) characters' proper nouns and nouns referring to people; 2) animals and creatures; 3) food; 4) places; 5) others (linguistic features). It is natural for keywords to be somehow particular to the texts such as the proper nouns, animals and creatures which appear as characters in the story. The focus here is on 1) finding out how these characters are represented; and 2) investigating whether the keywords which appear both in EOCL and JTCL, project any cultural difference. In this section, keywords which appeared in both EOCL and JTCL are compared, to see if there is any cultural difference or difference in their representation. Keywords occurring in both JTCL and EOCL are highlighted in tables for the purpose of clarity in the presentation of data.

6.2.3.1 Characters including proper nouns and nouns referring to people

The following table (Table 6.2) shows the keywords which are nouns referring to people including proper nouns. The gray-shaded rows show the keywords which appeared both in EOCL and JTCL. As can be seen here, there are lot of proper nouns which only occur in JTCL. Indeed, some of the names or characters obviously reflect the origin of the texts. Of course, choice of names can greatly be influenced by the original culture (i.e. Japanese stories use more typical names in Japan: e.g. '*Taro*',

'Misuzu'). It is interesting to note, however, that there are some considerations given in translation (e.g. 'Ellie' which should be 'Eri' according to the Japanese Roman spelling rule; some names were modified: 'Aki' becomes 'Amy', 'Kon' becomes 'Ken'). One can assume that consideration may have been given for the intended readers. Translated Japanese stories are usually translated for those who have not much knowledge of the Japanese language and culture.

Table 6.2 Keywords in JTCL in comparison with EOCL: People inc. proper nouns

N	JTCL as TC			EOCL as RC		LL	P
	Key word	Freq.	%	Freq.	RC. %		
1	GURI	87	0.19	0		370.5	2.1E-17
2	GURA	79	0.17	0		336.4	2.8E-17
5	KEN	57	0.12	0		242.7	8.8E-17
9	KINTARO	44	0.09	0		187.3	2.22E-16
14	AMY	38	0.08	0		161.8	3.83E-16
15	BARNEY	35	0.07	0		149.0	5.25E-16
17	SUHO	32	0.07	0		136.2	7.48E-16
20	SABURO	30	0.06	0		127.7	9.71E-16
24	BABA	28	0.06	0		119.2	1.29E-15
26	GROOMPA	27	0.06	0		114.9	1.51E-15
30	MOMOTARO	25	0.05	0		106.4	2.1E-15
31	MIKI	25	0.05	0		106.4	2.1E-15
36	TARO	23	0.05	0		97.9	3.06E-15
37	NAOMI	23	0.05	0		97.9	3.06E-15
43	MOM	21	0.04	0		89.4	4.7E-15
44	JEEPER	21	0.04	0		89.4	4.7E-15
45	PRINCESS	52	0.11	54	0.01	88.2	5.02E-15
48	BROTHER	37	0.08	24		81.8	7.3E-15
53	JUNKMAN	17	0.03	0		72.3	1.41E-14
55	SID	16	0.03	0		68.1	2.01E-14
56	YOSAKU	16	0.03	0		68.1	2.01E-14
57	SHIRO	16	0.03	0		68.1	2.01E-14
58	WOMAN	59	0.13	100	0.03	66.8	2.26E-14
67	INCH	24	0.05	10		63.5	3.11E-14
69	MEW	14	0.03	0		59.6	4.76E-14
70	ELLIE	14	0.03	0		59.6	4.76E-14
76	JAPAN	15	0.03	1		56.6	6.89E-14
82	PRIEST	13	0.02	0		55.3	8.22E-14
86	SHOBEI	12	0.02	0		51.0	1.6E-13
87	MOMMY	12	0.02	0		51.0	1.6E-13
88	LOBSTER	12	0.02	0		51.0	1.6E-13
89	BEANIE	12	0.02	0		51.0	1.6E-13
91	BUMBUKU	12	0.02	0		51.0	1.6E-13
92	CASEY	12	0.02	0		51.0	1.6E-13
93	FISHY	12	0.02	0		51.0	1.6E-13
94	BLACKIE	12	0.02	0		51.0	1.6E-13
99	BOO	17	0.03	6		47.5	3.25E-13
101	SUDSY	11	0.02	0		46.8	3.78E-13
102	KITTEN	11	0.02	0		46.8	3.78E-13
103	ZOOF	11	0.02	0		46.8	3.78E-13
104	GOVERNOR	11	0.02	0		46.8	3.78E-13

105	MISUZU	11	0.02	0		46.8	3.78E-13
106	FRIENDS	42	0.09	74	0.02	45.7	4.97E-13
110	BABY	28	0.06	33		43.4	9.54E-13
115	WARRIORS	10	0.02	0		42.5	1.27E-12
120	FISHERMAN	11	0.02	1		40.2	3.35E-12
124	JIZO	9	0.02	0		38.3	9.93E-12
126	KATIE	9	0.02	0		38.3	9.93E-12
127	WEN	9	0.02	0		38.3	9.93E-12
128	ELEPHEE	9	0.02	0		38.3	9.93E-12
139	WOODCUTTER	15	0.03	9		34.3	1.58E-09
141	STOREKEEPER	8	0.01	0		34.0	2.41E-09
142	ZIPPY	8	0.01	0		34.0	2.41E-09
149	DOUSER	8	0.01	0		34.0	2.41E-09
162	RAINMAKERS	7	0.01	0		29.8	4.49E-08
164	HERDBOY	7	0.01	0		29.8	4.49E-08
173	PRINCE	5	0.01	223	0.06	-29.7	4.73E-08
179	GRANDMOTHER	5	0.01	255	0.07	-36.4	5.15E-11

This aspect in the spelling of proper nouns has some implications for teaching YLs, and will be discussed in Chapter 9.

Stubbs (2002, p. 145-169) demonstrates how significant words identified from corpus methods in certain texts can provide ‘systematic evidence’ of certain cultural significance. The Oxford WordSmith Tools (Scott, 2004) has a function of displaying words which appear frequently around the node word i.e. collocates, and displays information of collocates which are most frequent at different positions (L1, L2... or R1, R2...) within the specified span (for my research the span is set at -5 + 5). Those collocates (displayed in red) are selected and the number of times the collocates appear in certain positions are also indicated in brackets below.

Pearce (2008) examines how men and women are represented using the British National Corpus (BNC) by focusing on the three grammatical relations: ‘MAN/WOMAN as subject, MAN/WOMAN as object, and attributive adjectives associated with MAN and WOMAN’ (Pearce, 2008 pp. 6-7). By examining the collocational patterns, he discovers differences in representations of men and women in five domains: power and deviance,

social categorisation, personality and mental capacity, appearance and sexuality. He reports that the collocates of MAN and WOMAN often represent stereotypical images across these five domains. In addition to such stereotypical images, he also reports that a ‘sociological discourse’ is revealed by the collocational evidence of descriptions of women: ‘women are represented as the objects of sociological enquiry within a discourse which acknowledges their subordinate status and attempts to redress it’ (ibid, p. 21). Pearce (2008, p. 20) mentioned that the use of a large corpus enables one to uncover ‘culturally-prominent patterns of representation’ since ‘texts are a product of discourse’.

In JTCL, the keyword WOMAN (59 occurrences in JTCL; 100 in EOCL) was identified as a positive keyword with a log-likelihood score of 66.8. This means that *woman* appears significantly more frequently in JTCL in comparison to EOCL. As can be seen in Table 6.3, *man* appears frequently as one of the collocates at L2 of the node WOMAN; *man* was also searched in comparison.

In the case of MAN as a node, some paradigmatic oppositions were also observed in the L1 position in JTCL: *old-young*, *kind-mean*. It is common to include an educational point of view (e.g. moral, values) in children’s literature. In JTCL, the contrast of ‘*kind old man*’, and ‘*mean old man*’ was made in a story. The comparison of collocates (Table 6.3) shows ‘*old man*’ and ‘*old woman*’ occur in both JTCL and EOCL.

Figure 6.1 Positional frequency of collocates for *woman* and *man*, span 5:5

WOMAN	L3	L2	L1	N	R1	R2
JTCL 59	so (2) then (2) saw (3)	the (36) man (10) then (2)	old (39) and (10)		was (5) were (7) said (3)	to (8) very (6)
EOCL 100		the (17) this (6) my (3) an (2)	a (14) old (15) that (4) good (8) on (2)		and (6) was (6) in (5) said (5) with (3) had (2)	
MAN	L3	L2	L1	N	R1	R2
JTCL 156	that(3)	the(88), a(5), kind(11) , an(10), mean(7) asked (2)	old (122) young (8)		and (23) was (8) of (4) said (8) very (8) had (6) in (3) who (6) took (5)came (5) went (4) rice (1)	his (10) woman (10) in (3) up (2) him (3) as (1) it (3) home (4) this (2) himself (3)
EOCL 216	was (11) by (4) is (5) for (3) at (4)	the (47) an (10) as (2) what (2) one (3) my (4) replied (3)	a (36) old (36) that (8) little (16) this (5) conceited (9) hairy (3) cloud (6)		of (7) who (19) in (9) with (8) I (3) on (3) had (3)	had (3) be (2) him (3) up (2)

In JTCL, concordances of OLD MAN (122 instances) include instances of ‘*old man and woman*’ (11 instances), ‘*old man and his wife*’ (2 instances), and ‘*old man and his old wife*’ (3 instances), which shows that old couples were mentioned together 16 times. On the contrary, in EOCL, the instances of *old man* and *old woman* occur separately.

In terms of how these ‘*old man*’ and ‘*old woman*’ are described in JTCL, ‘*old man*’ is often the person who leaves the house (for work) and comes home as the examples from Concordance 6.1 show (e.g. ‘*came home*’, ‘*went home*’, ‘*went into the mountains*’ ‘*goes home*’). On the contrary, as the concordances 6.2 exemplify *old woman* tends to stay at home (e.g. *old woman was waiting for him by the fire*) or near the house for some

domestic reason (e.g. ‘*went to the river to wash some clothes*’, ‘*making lunch*’, ‘*collecting firewood for cooking or baking*’).

Concordances 6.1 OLD MAN - leaving and coming home

112	em in the old man's lunch box. Then the old man left the house. He went far into
113	ed him out of the forest. So the second old man went sadly home with two wens on
114	cause they had no children. One day the old man went into the mountains to cut f
115	ers to weave cloth! That night when the old man came home, the little girl came
116	nto the woods. A little while later the old man came home and found the sparrow
117	r that night. Late in the afternoon the old man came home, and the old woman sai
118	d it flew away into the sky. After the old man got home and was telling his wif
119	very fine mochi for New Year's. And the old man never had to go out in the snow

Concordances 6.2 OLD WOMAN – domestic

1	t his tail. Presently the monkey saw an old woman who was gathering wood. Some o
2	tail. Presently the monkey saw another old woman who was baking cookies. Now, t
3	way home. When he reached his house the old woman was waiting for him by the fir
4	the mountains to cut firewood, and the old woman went to the river to wash some
5	l the rice they could possibly eat. The old woman made rice cakes for herself an
6	e some rice cakes for my lunch." So the old woman made rice cakes and put them i
7	man gave him a sword and armor, and the old woman fixed him a good lunch of mill
8	e it all and eat it right away." So the old woman brought a big knife from thc k

In the search of WOMAN in EOCL, there are some evaluative adjectives found preceding WOMAN. These adjectives can be categorised into two: 1) physical quality
2) personal quality

1) *enormous, slim, frail, huge, skinny, little;*

2) *haughty, silly, nasty, cheeky, extraordinary*

At word-level, ‘old’ may not have a negative connotation itself, but, in the case of OLD WOMAN, it is used in an abusive way as the following examples show:

“*You nasty cheeky old woman!*” “*You’re a silly old woman!*”

Overall, rather as Pearce (2008) found western stereotypical images through investigating the patterning around MAN and WOMAN, the results from JTCL seem to indicate the stereotypical gender roles in traditional Japanese culture.

Concordance 6.3 OLD WOMAN in EOCL

1	'You gave me too much!' snapped the old woman. 'We had to, Grandma.' 'And
2	nasty trick. One night, when the old woman was asleep, he crept out of bed an
3	ight drink more easily. Now this old woman was a fairy, who had taken the for
4	oe. 'Plymouth . . . ' croaked the old woman. 'That rings a bell, too . . . Yes
5	d trick he was going to play on the old woman. Suddenly, as Mr Twit tipped th
6	ndmother said. "You nasty cheeky old woman!" shouted Mr Jerkins. He started f
7	enness that made everyone jump, the old woman sat bolt upright in bed and shoute
8	or another story. "Tonight," the old woman said, "I am going to tell you how
9	her and began weeping with joy. The old woman pushed her aside and said, 'What,
10	'Of course I don't,' gibbered the old woman. 'Nor would you if you were as old
11	ound to stare at this extraordinary old woman in the sky. Mr Twit G
12	" Mrs Jenkins said. "You're a silly old woman!" "I am trying to tell you as g
13	e day when she was at the spring an old woman came up and begged for a drink.
14	taggering about with an eighty-year-old woman and you'd drop her like a ton of b
15	mall baby suddenly exploded into an old woman, and Charlie all at once found him

In the case of collocates which appear in L1 position to the node MAN in EOCL, these include some instances which are clearly found in only one particular text. For example, the phrase '*little man*' recurs as the description of one of the characters in a particular story ('Charlie and the Chocolate Factory' written by Roald Dahl). Most instances of OLD MAN were '*the/a old man*' without any particular evaluation attached, except the following instances with somewhat negative evaluative adjectives used with *old man*: '*an extremely horrid old man*' or '*a foul and smelly old man*'.

In the JTCL *old man* and *old woman* refer to simply aged people and described their roles, which are more traditional and gender-specific as used to be the case in Japan. EOCL had some examples which seem to have negative connotations, especially in the use of *old woman* and *old man*.

The word *friends* appears as a keyword in JTCL (42 occurrences in JTCL and 74 occurrences in EOCL). Knowles and Malmkjær (1996, p. 97) find the feature of 'friendship' by looking at the significant collocates which appear with *friend*, *companion* and *comrade*. It was found that possessive pronouns such as *his*, *my*, *our* (in

JTCL) and *my*, *his*, *her* (in EOCL) were observed. Evaluative adjectives were also observed in both sub-corpora, but different items *good* (in JTCL) and *dear* (in EOCL). There were no particular differences observed in terms of how *friends* are presented in both EOCL and JTCL. However, some characters or name of occupation can be related to a particular gender. For instance, the keyword *woodcutter* was mentioned referring to a male in both EOCL and JTCL.

6.2.3.2 Animals and creatures

The following table (Table 6.3) shows the keywords which are related to animals and creatures. Some unusual characters such as *elf*, *eel*, *tengu* etc. appeared in JTCL. However, there are more overlaps in the occurrence of items between EOCL and JTCL, as highlighted in gray in Table 6.3.

Table 6.3 Keywords in JTCL in comparison with EOCL: Animals and creatures

N	JTCL as TC			EOCL as RC		LL	P
	Key word	Freq.	%	Freq.	RC. %		
6	MONKEY	90	0.20	42	0.01	228.8	1.08E-16
19	DOG	49	0.10	19		132.9	8.27E-16
22	ELF	29	0.06	0		123.4	1.12E-15
25	TURTLE	40	0.08	12		117.1	1.39E-15
27	SPARROW	30	0.06	2		113.2	1.6E-15
34	BEAR	44	0.09	28		98.2	3.02E-15
35	KITTY	25	0.05	1		98.2	3.02E-15
42	EEL	21	0.04	0		89.4	4.7E-15
51	BOAR	18	0.04	0		76.6	1.03E-14
71	HORSE	44	0.09	63	0.01	58.3	5.52E-14
79	RABBIT	24	0.05	14		55.7	7.81E-14
88	LOBSTER	12	0.02	0		51.0	1.6E-13
96	CRAB	13	0.02	1		48.3	2.68E-13
97	JELLYFISH	13	0.02	1		48.3	2.68E-13
102	KITTEN	11	0.02	0		46.8	3.78E-13
121	OGRES	12	0.02	2		40.1	3.5E-12
122	PHEASANT	12	0.02	2		40.1	3.5E-12
130	DRAGON	23	0.05	25		37.8	1.45E-11
134	DEMON	11	0.02	2		36.1	7.73E-11
143	CRANE	8	0.01	0		34.0	2.41E-09
144	OCTOPUS	8	0.01	0		34.0	2.41E-09
148	DEVIL'S	8	0.01	0		34.0	2.41E-09
151	TENGU	8	0.01	0		34.0	2.41E-09
156	OGRE	24	0.05	36	0.01	30.5	2.96E-08
157	WOLF	23	0.05	33		30.4	3.13E-08
159	ANIMALS	18	0.04	19		30.1	3.63E-08
160	CAT	30	0.06	57	0.01	30.0	3.83E-08

The word MONKEY appeared as the 2nd key-keyword in the KW database, occurring in 26 % of JTCL. Stubbs (1996, p.176) shows that ‘collocations can reveal associations and connotations, and thus assumptions embodied by cultural keywords.’ With this in mind, the first strong collocate of the word MONKEY was searched for in order to investigate the depiction of ‘*monkey*’ in the texts. On the first right (R1) of the node word ‘*monkey*’, ‘saw’ was observed as one of collocates. There were 7 instances in which ‘saw’ was used in the sense of ‘noticed’. Moreover, ‘watched’ which is similar in meaning, was picked up from the concordances. All the instances below describe the monkey either being attentive and noticing something, or being very curious.

Concordances 6.4 MONKEY + saw; watched

1	tting the tip of it right off! When the monkey saw this, he became very angry. "
2	full of notches and scratches. When the monkey saw this, he became very angry. "
3	ewood was all burned to ashes. When the monkey saw this, he became very angry. "
4	carrying his razor with him. When the monkey saw him, he said: "Please, Mr. Ba
5	ven think about his tail. Presently the monkey saw another old woman who was bak
6	ven think about his tail. Presently the monkey saw an old man who was carrying a
7	ven think about his tail. Presently the monkey saw an old woman who was gatherin
8	make it short enough to carry home. The monkey watched her a little while. He wa

There was another instance with MONKEY and ‘saw’ which was not from the same text and was used to mean ‘realised’

(6.1) *The monkey finally saw there was no escape.* (from *Crab and Monkey*)

As for collocates for MONKEY appearing in the first left (L1) position, the adjective ‘*foolish*’ was identified. In order to see how ‘*monkey*’ is described, words which come up in L1 position in the concordances were searched in order to see what other adjectives precede MONKEY. Overall, five adjectives which directly preceded MONKEY were found, as follows: ‘*clever*’ (1 from *Crab and Monkey*), ‘*brave*’ (2 from *Bobtail Monkey*) and ‘*foolish*’ (7 from *Bobtail Monkey*), ‘*little*’ (1 from *Little mouse’s red vest*), ‘*poor*’ (2 from *Kintaro’s Adventures*). It is important to note that instances of

‘*brave*’ were always negated (i.e. *wasn’t brave*). As can be seen in the above, since all these adjectives were used in limited numbers of texts only, it is difficult to determine whether they occurred because of the nature of the particular texts in the corpus, or whether they exemplified some specific cultural meanings attached to monkeys.

Concordances 6.5 adjectives + MONKEY

1	see this foolish monkey wasn't a brave monkey at all. Just then a barber came w
2	ber, this foolish monkey wasn't a brave monkey at all. So when he saw the razor
3	dea one little bit. Being a very clever monkey, he said: "Wait a minute! Wait a
	all stacked in a bowl. Then the foolish monkey went walking away through the for
	e monkey her firewood. Then the foolish monkey went walking away through the for
	the monkey his razor. Then the foolish monkey went walking away through the for
	e thorn out. But remember, this foolish monkey wasn't a brave monkey at all. So
	g very loudly, for you see this foolish monkey wasn't a brave monkey at all. Jus
	'll take the cookies." Then the foolish monkey took the gong and climbed to the
	but he was so proud of how the foolish monkey sang, looking very silly as he wa
4	at's a nice vest you're wearing, Little Monkey! May I try it on?" "Sure!" "It's
5	ming as fast as he could after the poor monkey. It was a terrible struggle, but
	eds of buzzing hornets flew at the poor monkey, and he was so surprised that he

Thus, no definite conclusion can be drawn from these observations, but it is interesting to note that in these three stories, the monkey is always presented as a curious, and even a rather cunning and short-tempered character, who has a tendency to trick or have fights with others. Two of the four stories end with the monkey being punished for his behaviour in the end.

It is a common trope in children’s literature for animals to be linked with certain qualities. As Nesi (1995) suggests, animals are often used in metaphors, and some animal metaphors denote different meanings in different cultures.

In EOCL, the node DOG, following collocates are found in left to the node DOG:

Left: a, the, of, as, like, white, large, smelling

Right: droppings, and, or, cats

It is difficult to arrive at any firm conclusions from the collocation results, and some of them were the properties of one single text (e.g. *White dog, dog's droppings*). However, in the case of the collocates, *like*, some use of similes are found (e.g. *like a large dog, lapping up the chocolate like a dog*).

Other similes or metaphors in EOCL and JTCL were searched. The total instances of *like* in both EOCL and JTCL of course include other usages of 'like' (e.g. the verbal usage of *like*). Therefore, the concordances were checked manually one by one. Out of 25 explicit similes found in the PREP category, 10 were referring to the following animal terms: an owl, an otter, a monkey, a sparrow, and a turtle.

Concordances 6.6 Animal similes

1	like the otter, and even see in the dark like an owl. When Kintaro was about eigh
2	rivers, run as fast as the rabbit, swim like the otter, and even see in the dark
3	woman began to chatter and jump around like a monkey, while the old man held hi
4	nkey, and the old an dancing and acting like a sparrow. "Here, here!" he said. "
5	old man held his hands out and chirped like a sparrow. First the old man sang:
6	old man held his hands out and chirped like a sparrow. First the old man sang:
7	a man's dance should be bold and manly, like a monkey's dance. Not the other way
8	dance should be graceful and lady-like, like a sparrow's dance, and a man's danc
9	he saw the old woman dancing and acting like a monkey, and the old an dancing an
10	head. She curled up into a tight ball, like a turtle in its shell, hugging Chub

Similarly, '*as * as*' was also searched in both JTCL and EOCL. Some similes referring to animals were found such as the following:

JTCL: *run as fast as a rabbit, climb any steep cliff as easily as the mountain goat*

In the case of EOCL, there were 371 instances of '*as * as*'. The following examples were identified relating to animals.

EOCL: *as big as robins, dig as quick as a fox, feel as frisky as a froghopper,
as thick as a whale, as fast as a mouse, as angry as a bee,
as loud as a bee's sneeze, as quiet as a mouse, as nimble as a deer*

In describing something ‘fast’, *rabbit* was referred to in JTCL (e.g. *as fast as a rabbit*), and *mouse* in EOCL (e.g. *as fast as a mouse*).

Although these differences were not significant, the fact that some understanding of animal qualities can differ among different cultures (Wang and Dowker, 2007) and that metaphorical expressions can sometimes present difficulties to learners (Nesi, 1995), makes this aspect worth further investigation in terms of its implications for teaching (see Chapter 9).

In terms of the genders of animals, the use of CAT in EOCL and JTCL showed some difference. In EOCL, there are instances of *Mr Cat*, whereas instances such as *Mrs Cat*, *lady cat*, *mother cat* appear in JTCL indicating that CAT is often associated with female characters in JTCL. This again may simply be the influence of particular texts, however, it is interesting that these differences exist.

6.2.3.3 Food

The following table (Table 6.4) shows the keywords which are related to food. The word RICE appeared in the 4th rank as a positive keyword (appearing in 14 texts out of 70). On the other hand, the keyword which is negatively key in JTCL is HONEY, indicating that HONEY occurs unusually frequently in EOCL in comparison to JTCL.

Table 6.4 Keywords in JTCL in comparison with EOCL: Food

N	Key word	JTCL as TC		EOCL as RC			
		Freq.	%	Freq.	RC. %	LL	P
4	RICE	72	0.16	4		276.3	5.62E-17
85	WINE	16	0.03	3		52.3	1.3E-13
137	COOKIES	18	0.04	15		34.9	4.29E-10
147	YAMS	8	0.01	0		34.0	2.41E-09
185	HONEY	3		380	0.11	-74.0	1.25E-14

In Japanese, there are at least three lexemes of the word ‘rice’: *ine* (which suggests ‘rice growing in the field’), *kome* (which suggests ‘uncooked rice’), *gohan* or *meshi* (which suggest ‘cooked rice’). Both are modifications of lexemes, whereas in English it seems that there is only one lexeme for ‘rice’ and usually there is a preceding modifier in order to differentiate the meanings (i.e. ‘uncooked rice’, ‘boiled rice’).

In the light of the differentiation of uncooked and cooked rice noted above, the instances are divided into three categories: 1) uncooked rice (original form); 2) cooked rice or rice made into something else; and 3) other. Apart from a couple of instances, most of the instances of *rice* and of its use refer to, or are related to, eating or drinking. For each category, analyses are presented below.

Concordance 6.7 Uncooked Rice (bag of rice, bales of rice)

1	en went home, carrying the small bag of rice with him. When his old wife heard h
2	d home with the mortar. "Please give me rice. We need rice." So saying, he groun
3	gold." But when he began pounding, the rice turned to terrible smelling garbage
4	the old man began pounding than all the rice turned into gold! Now the kind old
5	ing through the window and had seen the rice turn to gold. He still wanted some
6	of steamed rice and began pounding the rice to make rice-cakes. But no sooner
7	of rice. There was so much that he gave rice to everyone in the village. "This i
8	hem. After that they always had all the rice they could possibly eat. The old wo
9	And all at once out came rice, bales of rice. There was so much that he gave ric
10	e mortar. "Please give me rice. We need rice." So saying, he ground the stick in
11	is old wife heard his story and saw the rice, she said: "Humpf! That won't make
12	akes." But when she started pouring the rice out, they were surprised to discove
13	o his wife. "When I begin pounding this rice, it'll turn to gold." But when he b
14	a kind of wine that they had made from rice. It smelled so good that the old ma
15	Why, yes, I've "mochi" cakes of finest rice. It made the old man and woman fee
16	in the mortar. And all at once out came rice, bales of rice. There was so much t
17	h they poured out of it. It was a magic rice bag, a wonderful present that the m
18	y after tomorrow. And we don't have any rice at all. So we won't be able to make
19	em. Then with the money we can buy some rice and make some mochi." So the old ma
20	w our thanks we'll give you this bag of rice." And the mouse gave the old man a
21	happy. They were able to buy plenty of rice and make very fine mochi for New Ye
22	uld take it off the fire, put it on his rice, and eat it up, smacking his lips a
23	e mouse gave the old man a small bag of rice about the size of a fat coin purse.

The close analysis of each concordance line suggests that rice is used to make some ‘*mochi*’ (line 15, 18, 19, 21) which is ‘rice-cake(s)’ in English translation.

Concordances 6.8 cooked rice (steamed rice, rice cakes, rice ball(s))

1	rtar home and filled it full of steamed rice. "Now watch," he said to his wife.
2	nk. Then they filled it full of steamed rice and began pounding the rice to make
3	's friends waiting to drink tea and eat rice-dumplings with them. There was the
4	That won't make more than two or three rice cakes." But when she started pourin
5	nd thing like that than to have all the rice-cakes in the world. We'll get along
6	t the kettle on a wooden stand, and the rice-cakes on another stand. Then, with
7	nly cried: "Oh, my!" because one of the rice cakes had fallen out of the box, an
8	ully to the treasure house, and put the rice-cakes beside him. It is said that B
9	tle cakes would be enough. Without some rice-cakes we can't even celebrate New Y
10	ome firewood today. Please make me some rice cakes for my lunch." So the old wom
11	after tomorrow. How I wish we had some rice-cakes to eat on New Year's Day! Eve
12	sold these hats, why don't you buy some rice-cakes and bring them back with you?
13	r. Then with the mortar we'll make some rice-cakes in memory of Shiro." So the
14	the song the tiny voices were singing: Rice cakes, rice cakes, Nice, fat rice c
15	" the old man said, and he kept rolling rice cakes down the hole until they were
16	ice and began pounding the rice to make rice-cakes. But no sooner did the old ma
17	could possibly eat. The old woman made rice cakes for herself and the old man e
18	es for my lunch." So the old woman made rice cakes and put them in the old man's
19	s of field mice. They had eaten all his rice cakes and now they were singing aga
20	re quietly forever, always feeding him rice-cakes like these I've brought and n
21	biggest and most beautiful and freshest rice-cakes the old people had ever seen.
22	of money and some of Bumbuku's favorite rice-cakes to the temple. When they got
23	ging: Rice cakes, rice cakes, Nice, fat rice cakes, Rolling, rolling, rolling -
24	"Remember how Shiro used to love to eat rice-cakes? Let's cut down that big pine
25	nkman would give Bumbuku some delicious rice-cakes to eat. The junkman sold so m
26	w, for some of the old lady's delicious rice cakes." Then he suddenly cried: "O
27	ry happy. They still give him delicious rice-cakes to cat every day and never, n
28	n, and he might never give me delicious rice-cakes." "Just leave everything to m
29	"Thank you very much for the delicious rice cakes, old man," the leader of the
30	e tiny voices were singing: Rice cakes, rice cakes, Nice, fat rice cakes, Rollin
31	the world. We'll get along without any rice-cakes for New Year's." By this time
32	y that he hadn't been able to bring any rice-cakes. "My! that was a very kind th
33	ll, because now they had this wonderful rice-cake to eat
34	e stone Jizo who brought this wonderful rice-cake to us." The old woman said: "Y
35	ve them your hats, so they brought this rice-cake to show their gratitude. The o
36	d see." After he had dropped the second rice cake into the hole, he put his ear
37	" he asked himself. "I'll drop one more rice cake down and see." After he had dr
38	onkey loved persimmons even better than rice-balls, and once he was up the tree
39	finally persuaded the crab to trade the rice-ball for the persimmon seed. The mo
40	simmon seed. The monkey quickly ate the rice-ball. The crab couldn't eat the per
41	y rustle the new leaved trees. With our rice-ball lunches in our knapsacks, Our
42	d a pesimmon seed, and the crab found a rice-ball. The monkey wanted the crab's
43	they were singing again as they pounded rice. "Thank you very much for the delic
44	de of stone, the kind used for grinding rice or wheat into flour. "This is a mag

The above examples are all of RICE made into something else such as ‘*rice cakes*’ (*mochi*), ‘*rice-ball*’, and ‘*steamed rice*’. Two out of 72 instances of RICE occurred with the word ‘*steamed*’. It is interesting to note that the verb ‘*to steam*’ clearly indicates the way the rice is cooked in Japan (i.e. there were no instances of ‘*boiled*’ rice). In addition, the modifiers used with RICE seem to show that rice in general has positive connotations, as it goes with such modifiers as: *delicious*, *wonderful*, *freshest*, and *favourite*. This notion is also represented in the example: ‘*rice turning to gold*’.

The highlighted lines in the above concordance lines (lines 9, 11, 13, 22) indicate another aspect on how ‘*mochi; rice-cake(s)*’ is perceived in Japanese culture. ‘Mochi’ is often used for celebration (e.g. New Year’s Day as in line 9 and 11), or to show one’s respect towards the dead on a memorial day, as can be seen in Line 13 and 22. Lines 43 and 44 suggest ‘rice’ made into something else or made into small pieces (e.g. pounded rice, grinding rice). The rice can be processed this way in order to make different kinds of rice-cake.

‘Rice’ growing in a field is referred to in the expressions ‘*rice paddies*’ or ‘*rice field*’, as the examples below show:

(6.2) *we can see rice paddies!*

(6.3) *man was walking along the road beside a rice field*

Moreover, there is an instance of ‘*rice-bowl*’. Rice-bowl is usually an item of crockery used for serving individual portions of rice. In this story, it is used as a boat, as the main character is only an inch in size.

(6.4) *he went floating down the river in his rice-bowl boat, paddling with
his chopsticks*

Another interesting instance of *rice* is found below - ‘*rice man*’:

(6.5) *they sang another song: Nice man, rice man, Nice, fat mice man ...*

This is an instance of play on sounds. As can be seen, it does not make sense as a morphological form, but phonologically rhymes. Such rhymes can be combined with the teaching of spelling.

In contrast, there are only 4 instances of RICE in EOCL, all of which are shown below.

Concordances 6.9 RICE in EOCL

1	arwigs cooked in slime, And mice with rice - they're really nice When roasted
2	each one about the size of a grain of rice. They were extraordinarily beautif
3	in the bread, or the cornflakes, or the rice-pudding or whatever food you see
4	rice, 'Today it's scrambled eggs and rice.' Now every day, week in week ou

The comparison between the two corpora shows particular types of rice recipe in each country.

JTCL: rice-cake (*mochi*), rice-balls (*onigiri*), steamed rice

EOCL: rice-pudding

'Honey' was the least frequent keyword in the food category in JTCL and the following concordance lines show the 3 instances found in JTCL.

Concordance 6.10 Instances of HONEY in JTCL

1	p yourself.'" The fox licked up all the honey. "But if I leave the basket empty,
2	come along." So the bear left a jar of honey in the basket. The donkey was stil
3	presents of berries, fruits, nuts, and honey. Each time, Father Bear would brin

In contrast, 380 instances of HONEY are observed in EOCL. Out of all the instances (380) of HONEY, 373 instances are part of 'Miss Honey' and 1 was of 'Miss Jennifer Honey'. The rest of the instances are:

Concordance 6.11 Instances of HONEY in EOCL

1	he sand is the color of honey. And that honey color was making me happy, too. Wh
2	ly. At sunrise the sand is the color of honey. And that honey color was making m .
3	tchen cupboard and feast on raisins and honey!" "B-b-but B-B-Bruno!" stammere
4	was in the parlour Eating bread and honey, The footmen and the servants s
5	more money. 'I have had my taste of honey. 'I'm wishing for a decent man
6	ld dance the polka like you, my little honey-bun!' she said to Thunder Karlson

In line 5, HONEY is used in a metaphorical sense of referring to a sweet memory. In line 6, HONEY is used in the expression ‘*my little honey-bun*’, showing endearment towards a person.

In summary, HONEY was certainly significantly frequent in the EOCL corpus, but this is because of 3 texts which featured a character called ‘*Miss (Jennifer) Honey*’. Overall, as for HONEY as a food category EOCL had 2 instances (lines 3 and 4), and there were 3 instances in JTCL, so the difference in frequency is not significant. However, this cultural meaning of ‘honey’ in terms of endearment seems a particular usage found only in EOCL, which deserves attention when teaching children.

Overall, as for the food category, it can be said that RICE and HONEY could reveal a difference in each culture, even though the numbers of instances are limited due to the size of the corpus.

6.2.3.4 Places

The following table (Table 6.5) shows the keywords which are related to places. As can be seen, almost all the keywords about places appear in both EOCL and JTCL, except TEMPLE.

Table 6.5 Keywords in JTCL in comparison with EOCL: Places

N	Key word	JTCL as TC		EOCL as RC			
		Freq.	%	Freq.	RC. %	LL	P
16	SEA	60	0.13	35	0.01	139.3	6.84E-16
21	MOUNTAIN	46	0.10	18		124.3	1.08E-15
32	SKY	64	0.14	71	0.02	103.7	2.36E-15
40	FOREST	38	0.08	20		92.1	4.07E-15
52	HOME	113	0.25	298	0.08	73.2	1.32E-14
62	ISLAND	25	0.05	11		64.9	2.71E-14
80	TEMPLE	13	0.10	0		55.3	8.22E-14
98	MOUNTAINS	22	0.14	15		47.5	3.24E-13
131	VILLAGE	27	0.08	37	0.01	37.1	2.48E-11
172	HOSPITAL	9	0.25	2	0.08	28.3	9.6E-08

Although I have been looking at the keywords which appeared in both EOCL and JTCL, with regard to the names of places, TEMPLE occurred only in the JTCL corpus.

Concordance 6.12 Instances of TEMPLE in JTCL

1	till there in the treasure house of the temple today, where he is very happy. Th
2	ave all the money to the priest for the temple. Then he said: "So will you pleas
3	kes to the temple. When they got to the temple the junkman explained to the prie
4	of Bumbuku's favorite rice-cakes to the temple. When they got to the temple the
5	d and I would like to live quietly in a temple. But that old priest might put me
6	ed. So why don't I take you back to the temple, where you can live very quietly?
7	of it." Just then a junkman came by the temple. So the priest took the kettle ou
8	on't want anything like that around the temple. We must get rid of it." Just the
9	is three young pupils, who lived in the temple. "Just look what a fine kettle I
10	So he bought it and took it back to his temple. He polished the kettle until all
11	Presently he came to a temple. At the temple he saw a beautiful girl dressed i
12	toward the city. Presently he came to a temple. At the temple he saw a beautiful
13	and the Little One-Inch went to visit a temple near the lord's house. Suddenly a

These instances seem to indicate the aspect of Buddhist culture rooted in Japan. The instances are too few to generalise the result, but temples are regarded as somewhere safe and peaceful as can be seen in the concordance lines 1, 5, 6, and 12. In addition, the culture of donating something (money or food) at the temple is also expressed in line 2, and 4.

6.2.3.5 Some linguistic features identified from KW Analysis

In addition to the content words mentioned above, some differences in the use of verbs were discovered, for example in the usage of SAYS, CRIED and SHOUTED.

SAYS

In JTCL, SAYS is used significantly more frequently (108 times) in comparison to its use in EOCL (37 times). In the JTCL, all the instances occurred in the narration. The present tense is often used in order to make the narrative seem more immediate (i.e. it helps to make the story more vivid for the reader). In contrast, in EOCL, 28 out of 37 instances of *says* were used within speech marks, which indicates the presence of

characters. In the case of instances in JTCL, the frequent use of *says* may have resulted from the translators' choice of words. The verb to utter something in Japanese '*iu* (*itta*: past tense)' can be translated as 'call' or 'say' (calls, says, called, said, calling, saying. For instance 'so-called' can be translated as '*to iu*'). Detailed discussion would require insights from translation studies, which is beyond the scope of this study.

Related to cultural aspects, some instances indicated British or American culture in the EOCL, as can be observed in the examples below.

(6.6) 'Let's play Simon Says...' she cried.

(6.7) Simon Says ... put your feet in the air!

'*Simon Says*...' is a type of game which children play. This is a game which was originally a rhyme 'Simple Simon' and has its origin in England.¹² There is no equivalent game available in Japan, although 'Simon Says' is often used in primary English classrooms in Japan as a part of Total Physical Response (TPR) exercises.

The following table (Table 6.6) shows the keywords which are identified as 'negatively key': keywords which are significantly infrequent in JTCL but significantly frequent in EOCL as a result of comparison. KNOW, SAID, CRIED and SHOUTED appeared as negative keywords in JTCL, which indicated that CRIED and SHOUTED are significantly infrequent in JTCL in comparison to EOCL.

¹² See for detail http://www.rhymes.org.uk/simple_simon.htm

Table 6.6 Negative keywords in JTCL in comparison with EOCL

N	Key word	JTCL as TC		EOCL as RC		LL	P
		Freq.	%	Freq.	RC. %		
175	KNOW	30	0.06	572	0.17	-34.2	1.88E-09
177	SHOUTED	16	0.03	407	0.12	-35.0	6.24E-10
183	CRIED	35	0.07	782	0.23	-58.2	5.62E-14
186	SAID	339	0.75	4027	1.21	-80.7	7.83E-15

CRIED

An especially interesting finding is observed in the instances of CRIED (EOCL 782 times in total). Of all the instances, 779 instances are used in the sense ‘cry out’ (i.e. ‘shout’) rather than the prototypical sense of ‘cry’ (i.e. ‘shed tears’), which was used only 3 times out of 782 instances. According to the Collins COBUILD Advanced Learners Dictionary, which lists the word sense according to the frequency in sense, the prototypical sense of ‘cry’ comes at the top as the first sense and the second is the sense of ‘shout’. From this finding, it can be said that what is prototypical in a ‘general corpus’ may not be prototypical in a specific corpus such as a collection of Children’s Literature in the current study. From the results above (i.e. *cried* is used more to mean ‘cry out = shout’) and the fact that SHOUTED was also listed as significantly frequent in EOCL, it can be said that there seems to be lot of ‘shouting’ in EOCL.

LITTLE

In the course of analyses presented in section 6.2.3.1, the word LITTLE appeared frequently as a collocate of identified keywords in both EOCL and JTCL. In relation to the word LITTLE, Stubbs (2002, p.162-163) reports the results of comparative analysis of LITTLE and SMALL which appear in a 200-million-word corpus (Cobuild 1995, cited in Stubbs, 2002, p.162). He found that ‘the most frequent noun to co-occur with *little* is *girl*, and the most frequent adjective to co-occur with *girl* is *little*’. Moreover,

‘the phrase *little girl(s)* is 30 times as frequent as *small girl(s)*, whereas *little boy(s)* is less than 5 times as frequent as *small boy(s)*.’ He explains this results using the collocates that immediately precede and follow these two adjectives. In his analysis, ‘small most often collocates with formal words concerning quantities. These rarely occur with little’. He states that the combination of common words can reveal aspects of culture like the following:

In terms of communicative competence, all words, even the commonest in the language, contract collocational relations. In terms of cultural competence, culture is encoded not only in words which are obviously ideologically loaded, but also in combinations of very common words. (Stubbs 2002, p. 164)

In EOCL, out of 403 instances of SMALL, 295 were complementing objects such as door. There are 8 instances of small complementing quantities as shown below (Concordances 6.13). The rest of them (99) were referring to ‘*Mr Small*’ who appears in the Mr Men Series.

Concordances 6.13 SMALL + Quantities

1	e a bull. But it does have in it a very small amount of the rarest and most magi
2	ried out. "How shall we dispose of this small pile of filth?" They're talking
3	ggled with the invitation, and when the small hours of morning had come, and th
4	-Vite doesn't work at all when given in small doses. You've got to throw everyth
5	ren never seem to be able to take it in small doses. They want to sit there all
6	ion myself. I suppose it's all right in small doses, but children never seem to
7	nd Vegetable Store in Shanghai,' said a small distant voice. 'Mr Wing speaking.'
8	sterday I am personally prrree-paring a small quantity of the magic formula in o

Moreover, in EOCL corpus, the following number of the combinations ‘*little girl(s)*, *small girl(s)*, *little boy(s)*, *small boy(s)*’ were found.

little girl(s) 32	little boy(s) 52
small girl(s) 10	small boy(s) 22

The proportions of the *little* and *small* collocating either with *girls* or *boys* and those presented by Stubbs (2002, p. 162) are different. In EOCL, *little* co-occurs more with

boy(s), than with *girl(s)*. However, the most frequent adjective to co-occur with *girl(s)* and *boy(s)* is the adjective ‘little’. As Stubbs (2002, p. 163) suggests:

[T]he words which immediately precede *little* often convey the speakers’ evaluation: either approval that someone or something is “cute”, or disapproval that someone or something is “strange” and/or repulsive in some way. These do not occur before *small*.

The examination of concordances of ‘little boy(s)/girl(s)’ and ‘small boy(s)/girl(s)’ revealed that while ‘little boy(s)/girl(s)’ is preceded by other evaluative adjectives (e.g. *smelly*, *loneliest*, *disagreeable*, *brave* etc.), ‘small boy(s)/girl(s)’ only occurred either with articles, numbers or demonstrative pronouns (e.g. *that*). The claim made by Stubbs (2002, p. 163) was the following: ‘paradigmatic oppositions...are usually thought of as being permanently available in the structure of the vocabulary, but words are co-selected, and this limits choices in syntagmatic strings.’ The results reported above (i.e. the occurrences of ‘little boy(s)/girl(s)’ and ‘small boy(s)/girl(s)’ found in CLC) mirrors what Stubbs (2002) reported using a 200-million-word corpus. This suggests that even a small corpus such as the one used this study can provide results which are similar to those of a larger (a 200-million-word) corpus.

6.3 Summary and implication of cultural keywords

Although the results of this chapter are somewhat limited, due to the size of the corpus, this comparative analysis suggested some features of the Corpus of Children’s Literature, which are categorised into two: 1) cultural, and 2) linguistic features used in the Corpus of Children’s Literature (EOCL and JTCL) as a whole.

(1) Cultural features

Not many significant cultural findings were revealed from the analysis. However some specific details relating to culture were identified.

- Proper nouns and characters (e.g. differences in spelling of names)
- Representation of gender and their roles (e.g. *old man* and *old woman*, *woodcutter*)
- Representation of some animals (e.g. MONKEY, CAT)
- Cooking procedure and recipes which use RICE (e.g. *rice pudding* and *rice cakes*, *rice-balls*)
- Places and customs which are particularly associated with Buddhist culture (e.g. TEMPLE, donating food or money)
- Use of metaphors (e.g. as fast as a rabbit, as fast as a mouse, as white as snow, as white as papers)

(2) Linguistic features

As for linguistic features, verbs were analysed to identify differences and similarities between EOCL and JTCL. Differences observed are summarised as follows:

- SAYS

In EOCL, it is used mostly in speech marks rather than in the narrative.

In JTCL, all of the instances were observed in narration.

- CRIED

In both EOCL and JTCL, the instances showed the opposite frequency of senses from these listed in the dictionary.

It is important to stress that any differences or similarities identified in this analysis could be limited to the sets of texts included in the corpus, and the results cannot be generalised easily. Some aspects of culture and stereotypical views were also identified through the comparative analysis of collocates and concordances. However, this also of course largely depends on the texts chosen in the corpus for both types, as Stubbs (2002, p. 165) states ‘... certain representations circulate widely in a society at a given time. It is individuals who produce individual texts, but these texts are not produced with complete freedom’. It would be possible to extract keywords in a more quantitative way and these could be explored further in order to obtain more reliable results, if there were similar types of, but much larger corpora available.

In research design, one of the most important factors is to have a clear rationale and criteria for choosing specific texts, especially when investigating something relevant to a particular society, or ideology (Adolphs, 2006, p.81). As Mahlberg (2007, p.196) puts it:

The way in which an analysis of corpus data can be related to social situations depends on the information that is available on the origins and contexts of the texts. If the texts in a corpus are selected according to transparent criteria and information on their contexts is stored together with the texts, corpora can provide useful insights into meanings that are relevant to a society and indicative of the ways in which society creates itself.

It is necessary to mention, however, that these corpora were not created for the purpose of representing the particular ‘culture’ as such. The corpora which I used for investigating the difference and similarities of English and Japanese texts were built from books written for children which have been identified and recommended for the

target age group (children aged 10-12) (For a detailed discussion of the representativeness of each corpus, please refer to Chapter 4). It can be said, however, that the *Keyword* function in WordSmith ver 4.0 made it easier to narrow the search to a starting point, and helped to indicate some keywords related to each culture, which are identified through the subsequent analysis.

Although the methods of choosing the items from the list, and the methods of analysis were not all systematic, as Mahlberg (2007a, p.197) notes, ‘for words that are intuitively selected a corpus analysis can still make an important contribution to the characterization of their meanings’.

It can be argued that even these limited results have some implications for reading instruction, and comparative analysis can be a suitable starting point for providing some insights into the pedagogical considerations necessary for teachers when designing a reading syllabus, materials and activities suitable for TEYL. Pedagogical suggestions using these results will be discussed in Chapter 9 (See Section 9.3.2.5 and 9.3.2.6).

CHAPTER SEVEN

USING A CORPUS OF CHILDREN'S LITERATURE TO IDENTIFY PEDAGOGICALLY USEFUL ITEMS

7.1. Introduction

One of the central research questions for this study relates to the methodological issue of identifying pedagogically useful features within the corpus, in order to use them in teaching. As reviewed in Chapter 3, there are many different criteria for selecting which lexical items to teach and in what order. Research using corpora has certainly influenced the selection of lexical items and it can be argued that the suggestions vary according to each researcher's theoretical perspective (see Chapter 3, Section 3.3.4). Nevertheless, criteria such as saliency, learners' needs, teachability and learnability, in particular, calls for intuitive judgments on the part of researchers. In Chapter 3 (Section 3.3.5), I attempted to define what is meant by 'pedagogically useful items' for Japanese learners from SLA perspectives. It can be seen that the process of defining what should be identified as 'pedagogically useful' is not yet systematised.

It should be made clear that the main attempt is not to identify 'what items' to teach, but to identify 'how' the corpus approach assists us in finding pedagogically relevant items (as discussed in Chapter 3 – considering the need for teaching them according to the concept of a lexical syllabus). In other words, the main aim is to identify a procedure for informing a syllabus based on corpora used in this study for the specific purpose of

applying it in primary English education (syllabus design will be further discussed in Chapter 9), in order to enable other researchers or teachers to replicate the procedure.

There are largely two stages of methods applied to the corpus: Stage 1 is described in Section 7.2; and Stage 2 in Section 7.3. Section 7.2 comprises a review of methods for extracting lexical items by means of frequency, n-grams and colligational framework, with a review of their advantages and drawbacks, together with the methods applied to the corpora used in this study (7.2.2 – 7.2.3). These methods identify some potentially pedagogically useful items (labelled as PPUI -1, -2, -3...) and these are further discussed in Stage 2.

In the subsequent sections (7.2.1-7.2.3), I illustrate how a corpus of writing for children can be analysed in order to explore the possibility of using it in teaching. In order to answer the main research question, I illustrate in some detail procedures of analysis for identifying useful lexical items. The overview of methods and criteria suggested by other researchers (c.f. Chapter 3), showed that corpus methods changed perceptions of lexico-grammatical relations, and that phraseology seems to be at the centre of the description of language from a corpus perspective. However, I should stress that I employ the term ‘lexical items’ here to show that I have no preconceptions about what is and is not useful.

Section 7.3 then provides details of methods used in Stage 2, which incorporates co-selection of collocates along with the methods described in Stage 1. Section 7.4 gives a summary and discussion of the integrated methods proposed in this chapter.

This is followed by an account of limitations identified through the analysis (7.5) and a conclusion (7.6).

7.2 Methods of extracting ‘pedagogically useful’ lexical items

7.2.1 Frequency list

7.2.1.1 Treatment of frequency lists in the existing research

One of the significant contributions of analysis using digitised texts, which distinguishes it from the analysis of paper-based texts, is that it facilitates accurate calculations of word frequencies. Wordlists, usually produced either in alphabetical order or by raw frequency, can ‘[provide] a general picture of a text or collection of texts and are a good starting point for subsequent searches of individual items at the concordance level’ (Baker et al., 2006 p.76). Apart from the raw frequency wordlist, Adolphs (2006) mentions that lemmatised lists are also useful in ELT ‘where it can be beneficial to teach all forms of one lemma together and to give priority to the most frequently used form’.

The benefits of frequency information in the context of teaching have been recognised. For instance, frequency-based wordlists are useful in ELT in informing material design, since they provide information about raw frequency and distribution (Adolphs, 2006, p. 40). As mentioned in Chapter 2, the first corpus-driven material design and syllabus design were undertaken by Willis and Willis (1988), for which they utilized the resources of the Bank of English (COBUILD, at that time 20 million words) (D. Willis,

1991, p. vii). The steps which they went through are described in detail in D. Willis (1990). It is often wrongly assumed that Willis's proposal for a *Lexical Syllabus* relied only on a frequency wordlist (Hunston, 2002). Though he compiled a list of the 700 most frequent words to form the basis of an ELT syllabus for Level 1, his suggestion was not to present words alone (i.e. a word-based approach), but to present also the patterns surrounding these frequent words (D. Willis, 1990, p. vii).

More recently among wordlists derived from corpus research, Coxhead (2000) produced her *Academic Wordlist* based 'to some degree' on the frequency wordlist in specialised corpora. In addition, Anthony (2007a, p.116-117) suggests that 'learners should study high frequency items' and 'learners should learn low frequency items through incidental learning (e.g. extensive reading)'. This suggests that it is important to focus on 'high frequency words'. The same view is shared by Tsui (2004, p.40):

In EFL and ESL situations, learners do not have the same amount of exposure to the target language as they do in L1 situations. Therefore, they are unlikely to acquire the language efficiently without systematic guidance on linguistic forms. By focusing on words which have a high frequency of occurrence and by concentrating on the usual rather than the exceptional, teachers can help learners acquire the language more efficiently, especially at elementary and intermediate levels. The findings of corpus analysis can be used as a basis for selecting and sequencing linguistic content, as well as determining relative emphasis. (Tsui, 2004 p. 40).

Hancioğlu and Eldridge (2007) also discuss the usefulness of a frequency list in the teaching context and explain 'how [it] could assist [them] maximally as practising teachers' (ibid, p.331).

As discussed, there are a number of advantages in using frequency lists, but it has been noted that one should be cautious about making frequency the only principle for choosing useful items. Granger (2005) points out that:

...frequency based information should be complemented with insights drawn from other disciplines as not all units identified by quantitative methods are pedagogically valuable.

This is especially applicable in the case of specialised corpora; frequency provides only a partial solution to specifying particular features of those corpora. As other researchers, such as Baker et al. (2006, p.76) suggest:

The [f]requency of words do not explain themselves: concordance-based analysis is therefore required in order to explain why certain words are more frequent than others.

Moreover, the need for caution in relying on frequency lists is endorsed as follows:

[A] frequency list is generated by counting the number of instances of a particular form and is not able to distinguish between different meanings of the same form. This can be a problem when it comes to the analysis of homographs (i.e. words that have the same spelling but different meaning) (Adolphs, 2006, p. 41).

As Hancioğlu and Eldridge (2007, p.340) also emphasise, '[s]ingle-item frequency ... cannot be the only determinant of when and where an item is introduced into the teaching and learning process'.

To sum up the benefits and drawbacks of frequency lists, we can say that they are useful but cannot be relied upon alone. Rather, opinions on the drawbacks of frequency lists indicate that information from such lists must be interpreted with caution and the use of concordances is vital so as to allow the repetitions in such lists to be open to further qualitative examination; this is demonstrated in Sinclair's pioneering lexicographic work, with its notion of phraseology.

7.2.1.2 Implementation of Method (1) frequency

Having reviewed the treatment of frequency above, this section investigates the value of a frequency list. In considering the corpus approach to syllabus design, it is important to consider the usefulness of the frequency list obtained from the corpus used in this study.

Below are the details of the methods applied to my corpora (EOCL and JOCL) in terms of extracting frequency information.

- Step 1: e-texts were prepared in text files (.txt)
- Step 2: these were uploaded to WordSmith Tools ver. 4.0, using the *Wordlist* function;
- Step 3: a wordlist was generated and sorted by frequency.
- Step 4: The 100 most frequent words in the list were then analysed in the concordances for their use in context.

The next section examines lexical items from the quantitative and the qualitative viewpoints illustrated with examples.

7.2.1.3 Analysis and Discussion (1) Frequency

Below are the 100 most frequent items from the joint corpus, following the procedures described above (Table 7.1).

A quick observation of the wordlist reveals that ‘*the*’ is most frequent, as is almost always the case in any other frequency list. I will label this point ‘Potentially Pedagogically Useful Item 1’ (referred to hereafter as ‘PPUI #1’). Most of the words appearing at the top are grammatical words, as is also often the case. There are some instances of idiosyncratic items such as *Mr*, *Miss* and *Wonka*. This is because of specific

longer texts (*Charlie and the Chocolate Factory*) and series (*Mr Men* and *Miss Little*) which are contained in the corpus.

Table 7.1 The frequency of the top 100 words in CCL (EOCL and JTCL)

Order	Word	Freq.	Order	Word	Freq.	Order	Word	Freq.
1	the	18157	21	her	2355	41	so	1274
2	and	10630	22	all	2339	42	were	1263
3	to	8418	23	is	2047	43	not	1257
4	a	8027	24	had	1023	44	there	1246
5	of	6312	25	at	1004	45	But	1236
6	I	6028	26	for	2003	46	them	1215
7	was	4853	27	with	1999	47	little	1209
8	you	4615	28	And	1973	48	this	1204
9	in	4419	29	up	1908	49	are	1186
10	said	4361	30	as	1875	50	Miss	1169
11	it	4169	31	they	1714	51	him	1155
12	he	3900	32	out	1701	52	do	1090
13	s	3504	33	It	1622	53	You	1080
14	that	3467	34	be	1582	54	She	1020
15	Mr	3356	35	He	1536	55	into	1015
16	on	2807	36	my	1459	56	like	1015
17	she	2710	37	one	1381	57	can	1013
18	t	2677	38	have	1379	58	from	986
19	his	2451	39	me	1369	59	down	983
20	The	2380	40	very	1366	60	what	974

Order	Word	Freq.	Order	Word	Freq.
61	went	916	81	ll	700
62	They	906	82	time	687
63	we	894	83	Then	683
64	about	871	84	going	667
65	but	856	85	Wonka	667
66	see	840	86	What	665
67	cried	814	87	no	656
68	then	793	88	will	649
69	could	782	89	day	642
70	over	778	90	right	632
71	now	753	91	old	627
72	who	748	92	re	617
73	go	739	93	We	614
74	when	739	94	if	608
75	just	729	95	an	603
76	your	720	96	know	602
77	would	709	97	came	601
78	their	707	98	There	600
79	by	704	99	Oh	593
80	back	701	100	don	590

In addition, it is clear from these lists, as pointed out by Baker et al. (2006), that the usage of each word cannot be revealed only by a frequency list. For instance, in the case of the occurrence of *s* and *t* in Table 7.1 above, we cannot comprehend their use until we see them in concordance lines. The following shows the sample of concordances, randomly selected from 380 instances of ‘s’.

Concordance 7.1 Concordance for 's'

1	Endless pleasure is one small children's dining-chair, Elizabethan, very rare. It
2	and gobbled up the lot? But wait! That's not the worst of it! Now comes the
3	Take a walk along the street 'Until it's cool enough to eat.' He adds, 'An early
4	'Oh cripes! How hot this porridge is! Let's take a walk along the street 'Until it's
5	Now look above the roof. There's light. It's the moon! Good evening, dear Moon!
6	me, but I look good in it, right?" "That's a nice vest you're wearing,
7	LOW! RED YELLOW GREEN! Wait a minute, that's strange. The lights are meant to
8	one goes about with egg in this hair. There's not a bald head to be seen. Once
9	y, and if I still don't listen, then there's a thrashing to be had, believe
10	must try to forgive me and remember that it's only because I've been a little
11	who at last had found her tongue. 'Yes, it's very wicked,' said Pippi, even
12	repeated the pig looking worried. "There's nothing that's fat and pink and goes

A search of these in the concordances reveals that, in the case of *s*, instances were either the genitive form (e.g. *Anne's*; *New Year's*) or contractions (shortened form of *is* or *has*) (e.g. *There's a lot of work to do*; *That's where I live*; *He's been to the hospital.*..).

Being so frequent in texts, it can be said that this is one of the aspects that learners need to be taught, as learners will certainly encounter them. In addition, it is notable that there are more contractions of present tense forms than of past forms (e.g. *had*, *would*), considering that the narratives tend to be in the past tense. There is less scope for contractions in past forms, because 'was' does not contract, whereas 'is' does. However, in the case of past tense contractions, the likeliest contraction is apostrophe-'d which represents both 'had' and 'would'. There are 41 instances of 'd' used as contraction, and the following concordance shows a random sample.

Concordance 7.2 Concordance for 'd'

1	Right there in front of you, you saw A pig who'd built his house of STRAW? The Wolf,
2	goodness that she got away!' Myself, I think I'd rather send Young Goldie to a sticky
3	even hint it. Nobody would ever print it.) You'd think by now this little skunk Would
4	oldilocks in jail. Now just imagine how you'd feel If you had cooked a lovely
5	To hitch a ride in to the city, And there she'd got a job, unpaid, As general cook
7	ride. The king said with a shifty smile, 'I'd like to give each one a trial.'
8	beast!' said Mr Fox. 'I should have guessed we'd find you down here somewhere.' 'Go
9	to let you go up there and face those guns. I'd sooner you stay down here and die
10	g?' 'Couldn't we make a dash for it, Dad? We'd have a little bit of a chance,
11	nd blast that lousy beast!' cried Boggis. 'I'd like to rip his guts out!' said
12	who stood below looking up at them, that she'd better put the ladder back, or

The ambiguity involved with ‘had’ and ‘would’ may explain why writers avoid past tense contractions and hence why present tense contractions are more frequent in comparison.

The frequency of ‘-t’ is also accounted for by contractions (i.e. ‘not’ contracting to ‘n’t’) as can be seen below. Below shows the randomly selected samples of ‘-t’.

Concordance 7.3 Concordance for ‘t’

1	back to the woodpile. Oh, if only he hadn't slipped and fallen and dropped that
2	The house stayed up as good as new. 'If I can't blow it down,' Wolf said, 'I'll have
3	l need,' Pig said, 'a lot of puff, 'And I don't think you've got enough.' Wolf huffed
4	'Oh Wolf, you've had one meal! 'Why can't we talk and make a deal?' The Wolf
5	put on Grandma's clothes, (Of course he hadn't eaten those.) He dressed himself in
6	Wolfie wailed, 'That's not enough! 'I haven't yet begun to feel 'That I have had
7	the Baby Bear, 'My porridge gone! It isn't fair!' 'Then go upstairs,' the Big
8	Before they clamber into bed. But Goldie didn't give a shred. Her filthy shoes were
9	a bunk. But no. I very much regret She hasn't nearly finished yet. Deciding she
10	eciate antiques. She doesn't care, she doesn't mind, And now she plonks her fat
11	razen little crook. Had I the chance I wouldn't fail To clap young Goldilocks in
12	and this in grumpy tones) 'I wish there weren't so many bones.' 'By Christopher!'

The examples given above (contracted form of ‘not’) with regard to the occurrence of contractions may be obvious as examples of usage, but their great frequency and their usages presented in concordances indicates the importance of including contractions (-’t, -’s, -’d) and the genitive form ‘-s’ for beginners, for the potential usefulness for learners of knowing these phenomena. I will label this point ‘Potentially Pedagogically Useful Item 2’ (referred to hereafter as ‘PPUI #2’).

Certainly their frequency shows how important it is for learners to know these items. However, as reviewed in the previous section, it is also apparent that presenting the word with only the criterion of frequency cannot be of much help, despite the strength of the frequency information. The reason for this is that statistical methods such as frequency are a good indicator of which lexical items to teach, with the proviso that the

lists which they produce cannot be taken to cover the wider contextual usage of such items.

In addition, as a result of corpus-related research, there is a trend in current ELT practices of examining beyond-the-word ‘lexical phrases’ (Nattinger and DeCarrico, 1992), ‘lexical bundles’ (Biber et al. 1999) and ‘patterns’ (Hunston and Francis, 1999). In the following section, therefore, I discuss these phenomena in existing research, together with methods of extracting them and their potential contribution to the present research.

7.2.2 Phraseology

7.2.2.1 Treatment of phraseology in the existing research

As pointed out in Chapter 3, the use of wordlists based on frequency is well established, despite some issues being left as matters requiring the researchers’ intuitive judgement in selecting items for further searches. The identifications used in phraseological description are, in comparison, less well organized. Different researchers use different terms to describe the phenomenon of ‘frequently occurring sequences’. The interpretation of those terms, of course, affects the way in which potential strings are classified. Wray (2000) opposes the claim made by Weinert (1995, p. 182) that ‘while labels vary, it seems that researchers have very much the same phenomenon in mind’. She argues as follows: ‘Every label has its own history and implications and neutral and inclusive reference is impossible without a clear and dedicated one’ (Wray, 2000, pp. 464-465). For instance, some researchers accept as their definition of strings ‘sub-phrasal and clausal strings ... alongside pragmatically meaningful sentence-frames’

(e.g. Althenberg, 1998; DeCock, 2000). This closely relates to the design of an automatic retrieval system of multi-word expressions, which calls for a more restricted definition of multi-word sequences (Stubbs, 2002 p. 228). Several attempts have been made to identify phraseology from corpora. The methods are summarised below:

Table 7.2 Methods of extracting larger strings from corpora

Research	Perspectives/ methods	Examples
Sinclair (1991) Hunston (2002, p.144) Stubbs (2002)	Node+ collocates	e.g. Hunston (2002, p. 144 Table 6.2)
Renouf and Sinclair (1991)	Collocational frameworks	<i>the of; a of</i>
O’Keefe et al. (2007)	n-grams (multi-word units)	Using CANCODE corpus
Scott (1997, p. 41)	Clusters	
Biber et. al (1999, p. 993)	Lexical bundles	Frequency driven approach Common recurrent sequences of words
Gledhill (2000)	Keywords (Scott, 1997) and then looking at grammatical words only.	e.g. <i>the, of, we, and</i>
Stubbs (2002, p. 230)	Chains	
Mason and Hunston (2004)	Grammar patterns	Automatic recognition of grammar patterns
Mason (2006) Stubbs (2002)	Frames + chains	Frames – based on collocational frameworks (Renouf & Sinclair, 1991) Still require manual analysis
Mason (2007)	Frames + n-grams	Produces candidate MWUs
Baker et al. (1998) Ruppenhofer, Baker and Fillmore (2001)	FrameNet database	Multi word expressions Collocational information

As noted above, automatic methods of identifying phraseology are available, as seen in, for instance, Biber’s approach to lexical bundles. This is ‘purely quantitative’, since it takes a frequency-driven approach and pays ‘less attention in the first instance to the relationship between form and meaning’ (O’Keefe et al., 2007, p. 61). The advantage

of this approach is that the result is ‘objective’, since it allows the automatic identification of MWUs and chunks, which intuition barely notices, by setting cut-off points of frequency to define ‘recurrent sequences of words’ as lexical bundles. Clusters, N-grams and Chains also use similar statistics-based automatic recognition.

However, there are problems with using these automatic methods to identify clusters and lexical bundles, including the fact that: 1) ‘some sequences are parts of other sequences’ and need further editing of the frequency-based output; 2) ‘many of these sequences do not appear to have the psycholinguistic quality of phrases in the sense that they are not retrieved from memory as single lexical units’; 3) they do not allow us to capture the low frequency MWUs (which tend to be idioms and formulae) because they are produced entirely on the basis of frequency information (Adolphs, 2006, p.43); and 4) the corpus used for this study is not large enough to retrieve prototypical strings.

In contrast, Renouf and Sinclair (1991) suggest that collocational frameworks should be invoked in studying phraseology. They define collocational frameworks as ‘discontinuous sequences of two grammatical words ‘somewhere between a word and a group’ (ibid: p.129). It is documented that there is a ‘tendency of these frameworks to enclose characteristic groupings of words’ (Stubbs, 2002, p. 128). In Sinclair’s view (1999, p.162), ‘frequent words play a major role in the composition of recurrent phrases ... there are many multi-word units which are themselves frequent and which contain high frequency grammatical words’ (Stubbs, 2002, 227). Summers (1996, p. 262-263) makes a similar point. The high incidence of grammatical words strengthens the case for extracting phraseology. Taking the notion of a collocational framework, Marco

(2000) identifies the set of lexical items restricted to the frameworks of '*the ... of, a ... of* and *be ... to*' as used in medical research articles. He then argues that 'the selection of specific collocates for these frameworks is conditioned by the linguistic conventions of the genre' (Marco, 2000 p. 63). Gledhill (1995, 2000) analyses phraseology through grammatical items and shows the relationship 'between phraseological patterns and the most typical rhetorical functions of the research article sections' (cited in Marco, 2000, p. 64). Groom (2006) also takes the same approach as Gledhill (2000), calling it the SGW (salient grammatical words) approach. He suggests that this 'provides analytical depth [and] complements the breadth of coverage provided by n-gram approaches', but it all depends on whether the aim of the research is to look for phrases or phraseology (Groom, 2006).

Although the approach taken by Sinclair and Renouf relies on frequency and recurring patterns, it can be said that it includes the perceptual salience criterion, which cannot be seen in Biber's approach to lexical bundles. For instance, as one example of the method taken by Sinclair and Renouf, the node word is taken 'with the collocates of the frequent word being considered first, then the collocates of the next frequent word and so on' (Hunston 2002, p. 147). As a consequence, this analysis can produce 'not a "phrase" but evidence of idiomatic phraseology' (ibid. p.147). This relates to the above distinction between phrases and phraseology pointed out by Groom (2006).

Nevertheless, as Mason (2006) describes, the methods used by researchers to identify patterns manually are time-consuming and laborious. For this reason, he suggests that when large amounts of data are involved it would be useful to make the analysis

automatic. As can be seen in Table 7.2, automatic recognition systems are developing. For instance, Stubbs (2002) and Mason (2006; 2007) both make use of the notion of 'Frames', which has its origin in the 'collocational frameworks' suggested by Renouf and Sinclair (1991). Barnbrook and Sinclair (1995) employ a local grammar of English, as suggested by Gross (1993). Devising a local grammar enables a computer program to identify elements, if the local grammar is mapped onto patterns correctly (Hunston, 2002, p.157). As Hunston (2002, p.157) comments, 'local grammars remain one of the ways that an emphasis on phraseology will contribute to different descriptions of English in the future'. Similarly, Mason (2007) also uses the notion of local grammar and Mason and Hunston (2004) have established a system for the automatic recognition of grammar verb patterns, based on the list of verb patterns by Sinclair et al. (1996).

Wray (2000, pp. 466-467) shows some scepticism about whether these multi-word units or formulaic sequences can be identified by a 'definitional criterion', such as that used in corpus-based research; the reason is that there may be some sequences which are widely known to native speakers but which do not appear frequently in the automatically identified data. She further alerts us, in regard to using frequency as a means of identifying formulaicity, that this approach may wrongly give the impression that any repeated sequence is formulaic. O'Keeffe et al. (2007, p. 61) also find that lists of sequences of words automatically generated by computer software are incomplete, because they 'cannot distinguish between strings which recur but which have no psychological status as units of meaning ... and those units which have a semantic unity and syntactic integrity, even though they may be less frequent'.

In summary, the present state of computer software and the sort of automatically generated data which it produces still require a manual and qualitative interpretation on the part of the researcher before deciding whether they can be pedagogically applied. The next section presents the methods and results of ‘n-grams’, as applied to my corpus.

7.2.2.2 Implementation of Method (2) N-grams by automatic extractions

The statistical methods (t-score, log-likelihood, and Mutual Information) are used to retrieve collocations in measuring how closely two words are related. Log-likelihood, which is also sometimes referred to as the G-score, does not presuppose normal distribution. Rather, a score can be described as comparable between two corpora, even if the size of the corpus is small and the frequency of the words is low. The software *Collocate* (Barlow, 2004) allows us to set log-likelihood as the statistical measurement. These different statistical measures can be selected in *Collocates*. The MI score is not very useful if the frequency of the words which are measured is low.

There are different tools available for extracting frequently occurring sequences. Most of the software, such as *kfNgram* (Fletcher, 2007) and *AntConc* (Anthony, 2007b) can generate a list of N-grams of different length according to their frequency (Baker, et al., 2006: 98). The *Cluster* function of Oxford WordSmith Tools is also an equivalent way of extracting n-grams of different length. Similarly, other software, such as *Collocate*, developed by Barlow (2004), uses frequency information and statistical analysis (t-score, log-likelihood, MI) for retrieving lists of:

- collocations with a specified search word and within a phrase of set size (e.g. four words),
- n-grams (lexical bundles) of different lengths and
- collocations extracted from the corpus as a whole

(Barlow and Römer, 2007)

One of the advantages of using statistical measures with an automated program in the first instance is that it can identify the language usages which are subliminal and not immediately accessible to human intuition (O’Keeffe et. al. 2007, p.64). In this case a procedure is followed of using the recurrent strings from the corpus to generate rank-order frequency lists of n-grams (two-, three -, four -, five- etc. word sequences from the entire corpus). A frequency cut-off point is established for the present purposes, which, given the size of the corpus, is set as at least 3 occurrences in the 376,210 word corpus.

Collocate (Barlow, 2004) was used for the extraction of n-grams. As mentioned in Section 7.3.2.1, the shorter n-grams are often parts of longer ones, so selected n-grams were manually checked using concordancers. The analysis is traced in detail in the following section.

7.2.2.3 Analysis and Discussion (2) Automatic extraction of n-grams

The following table (Table 7.3) lists the 30 most frequent 2-6-grams in the corpus.

Table 7.3 Top 20 2-6-grams from CCL

Rank Order	N-grams	Freq.	Rank Order	N-grams	Freq.
1	in the	1659	16	into the	456
2	of the	1537	17	I m	447
3	to the	933	18	Little Miss	446
4	on the	816	19	It was	439
5	Mr Wonka	596	20	he said	433
6	and the	578	21	she said	398
7	at the	576	22	going to	390
8	out of	572	23	it was	381
9	don t	526	24	in a	370
10	said Mr	520	25	Miss Honey	354
11	was a	517	26	of a	347
12	all the	572	27	he was	328
13	it s	473	28	I am	310
14	to be	464	29	for the	278
15	said the	458	30	of them	276

As can be seen from Table 7.3, the top 30 items are all 2-grams, and longer n-grams are less frequent. Stubbs (2002, p. 233) writes of the importance of checking whether the patterns appear to be frequent only because of some idiosyncrasy of the corpus. Taking his point, I ignored as idiosyncratic the n-grams which occur in one text only or which came from a single author. These are mostly proper nouns (e.g. *Mr Wonka*, *Little Miss*, *Miss Honey*).

The difficulty involved with this search is, of course, that it still yields a large amount of data, which can be time-consuming to deal with if every item is looked up in the analysis, using the concordances. However, as we can see in Table 7.4, the combination of prepositions and definite article ‘*the*’ appear frequently in the extracted n-grams. Further down in the list, items such as ‘in a’ (24th), ‘of a’ (26th) are observed. This result could be predicted, given the fact that both items are frequent in the corpus, as shown in Section 7.2.1.3 above. However, it is more interesting to note that these combinations

display two of the most difficult items for Japanese learners to use: prepositions and articles (indefinite or definite), as reported by several authors (Namaoka, 1995; 1996, Hayashi 2001; Tono, 2008). One of the accepted explanations of this difficulty is the lack of an equivalent system in Japanese. To my knowledge, no research has been made available about the use of articles and prepositions by young Japanese learners of English as a foreign language. Nevertheless, since they are reported as difficult for adult Japanese learners, it seems quite likely that young learners may also have problems with them at an earlier stage of their learning. For this reason, it seems reasonable to infer that these two aspects deserve attention in a syllabus for young learners. In Sinclair's view (1999, p.162), 'frequent words play a major role in the composition of recurrent phrases ... there are many multi-word units which are themselves frequent and which contain high frequency grammatical words' (cited in Stubbs, 2002, p. 227). This explains why the high incidence of grammatical words, such as prepositions and articles, can be a strong reason for identifying phraseology containing them. Moreover, it is important to investigate whether the corpus approach can improve understanding and lead to better ways of presenting and teaching those items. Therefore in this section, I label the n-grams which contain a preposition and an article PPUI # 3, as a starting point for detailed analysis.

In Chapter 3 (Section 3.6.2), it was pointed out that Japanese learners of English have problems in the use of prepositions and articles based on the findings from SLA studies. For this reason, in the search for PPUIs, n-grams including prepositions or articles should be analysed manually in the concordance lines. In order to demonstrate this, firstly, the concordances of '*in the*' were searched and 1,822 instances of '*in the*' were

found in the corpus as a whole. Concordances present results containing too many instances, which can be problematic as it is often time-consuming to identify the boundaries of a unit. In order to see longer sequences automatically, 3-6 grams (Table 7.5) are extracted from the corpus. Having compared the results of the top 2-6 grams (Table 7.4) and 3-6 grams (Table 7.5), we can see which pair appears most frequently: ‘*in the*’ is sometimes included in the 3-grams ‘*in the world*’ (rank 10); ‘*in the air*’ (rank 20), ‘*in the middle*’ (rank 26); ‘*in the middle of*’ (rank 30) etc. This illustrates the idea that shorter n-grams are often fragments of longer n-grams. The same applies to the case of ‘*of the*’ as in ‘*end of the*’ ‘*one of the*’. Moreover, the list does not at first glance provide information on the kind of phrase.

Table 7.4 Top 30 3-6 n-grams in CCL

Rank	N-grams	Freq.	Rank	N-grams	Freq.
1	out of the	235	16	the top of	93
2	said Mr Wonka	215	17	the end of	85
3	I don t	197	18	the middle of	85
4	Grand High Witch	156	19	There was a	84
5	the little prince	148	20	in the air	82
6	The Grand High	146	21	don t you	80
7	The Grand High Witch	146	22	I can t	80
8	Miss Honey said	129	23	cried Mr Wonka	79
9	my grandmother said	114	24	the old man	78
10	in the world	109	25	Tommy and Annika	78
11	on to the	106	26	in the middle	76
12	there was a	106	27	in front of	74
13	said to Pippi	97	28	It s a	73
14	one of the	95	29	end of the	67
15	It was a	94	30	in the middle of	67

As pointed out by many researchers (e.g. Adolphs, 2006, p. 136), extracting n-grams by computer, which relies on frequency, does not provide a complete solution to the challenge of identifying meaningful units, even though it appears to have some pedagogical value.

Moreover, it was demonstrated that automatic extraction of n-grams is useful in capturing words which frequently occur next to each other, though n-grams are inefficient with regard to recurrent sequences with variations. For instance, such items as ‘a bag of’ ‘a loaf of’ and ‘a bottle of’, can be identified as belonging to the same group of items, with some variation in the second word. However, in terms of n-grams, these three items are considered as three separate 3-grams. Therefore, for instance, if ‘a loaf of’ occurs below the minimum occurrence setting, then this sequence is missed out.

As Stubbs (2007: 174) mentions:

[N]o tools do everything, and they provide only intermediate data. [Methods using n-grams] can find recurrent multi-word-strings which have a pre-set length and which consist of adjacent words within a fairly short span.

The researcher’s subjective judgment plays a major part in deciding the length of multi-word units, However, Danielsson (2007) questions whether ‘the mental concept of a multi-word unit has a restriction in length, and subsequently if such a length restriction is realized in the actual language, do human beings have an intuitive understanding of what constitutes a unit of meaning, or do we need to be “trained” to see it?’ In the following section, in order to explore the objective method of identifying units, a method suggested by Danielsson (2001; 2003) is discussed in relation to the automatic identification of patterns.

7.2.3 Colligational framework – automatic identification of patterns

7.2.3.1 Treatment of the colligational framework in the existing research

As was also mentioned in earlier sections, the methods used for clusters and lexical bundles are reported as problematic, not least because they do not allow the capture of low frequency MWUs (which tend to be idioms and formulae) since such things as n-grams, clusters and lexical bundles are produced on the basis of frequency information (Adolphs, 2006: 43).

Stubbs (2002) suggests that a different approach can identify different aspects of phraseology, including collocation and colligation. Danielsson (2001, 2003) introduces the methods which extract meaningful units based on frequency information, but which at the same time can also capture low frequency MWUs. This method is designed to point ‘to the prototypical units of meaning (used in representational prosody) as well as to productive patterns of units of structures’ (Danielsson, 2001 p.87). The program created by Danielsson (2008) is designed to show the candidate frameworks which contain a high frequency word (e.g. *and, the, to, of*) followed by a low frequency word (e.g. a content word) which is shown here as ‘X’. This procedure outputs lists for candidate frameworks with three Xs (e.g. *and X and X and X*) and two Xs (e.g. *and X and X*). These candidate frameworks are named ‘colligational frameworks’ (Danielsson, 2008, personal communication), rather than ‘collocational frameworks’ (Sinclair and Renouf, 1991), because it is ‘colligation’ which refers to ‘the pattern of co-occurrence of words with particular grammatical items’ (Adolphs, 2006, p.136), such as this framework is designed to output.

In the previous section on MWUs, the problems involved in automatic identification were discussed as follows: 1) the fact that ‘many of these sequences do not appear to have the psycholinguistic quality of phrases in the sense that they are not retrieved from memory as single lexical units’ (Adolphs, 2006); and 2) the questionable effectiveness ‘of using language users as the arbiters in evaluating the performance of mwu extraction methods’ (Danielsson, 2007). In addition, Danielsson suggests that we should ‘acknowledge the corpus findings as proof of something that is important in language, based on the argument that if it were not important it would not be found repeatedly in a corpus’ (Danielsson 2007).

Although it has great advantages in capturing low frequency items and the outputs contained, as well as more frequent ones, this framework yields much more output to examine. In addition to the number of outputs, Danielsson (2007) discusses the difficulties in evaluating the output, in the following terms:

The question of evaluation remains: how should the results of this method of extracting multi-word units be assessed? Possible evaluators would be the linguist (are these relevant units of language?), or the lexicographer (would you put these in a dictionary?) or the developer of a computational language analysis tool (would you have these units in your lexicon?). The answers to all these questions would probably be ‘no’.

Danielsson (2007) further discusses the tendency for speakers to treat MWUs as things which can be replaced by a single word, whereas the computer also retrieves much longer units. She argues that ‘[t]his poses the question whether the mental concept of a multi-word unit has a restriction in length and subsequently if such a length restriction is realised in the actual language’.

Therefore, in searching for colligational frameworks, similar difficulties to those involved in the MWUs emerge regarding decisions over the length of multi-word sequences and the time needed for these decisions.

7.2.3.2 Implementation of Method (3) Colligational frameworks

Considering the difficulties of identifying multi-word units with no pre-set length (as discussed in the previous section), Danielsson (2001, 2003, 2007) introduces methods which extract meaningful units based on frequency information, but which at the same time can also capture low frequency multi-word units. This was the method which I initially employed. The process of pointing ‘to the prototypical units of meaning (used in representational prosody) as well as [pointing] to productive patterns of units of structures’ (Danielsson, 2001, p.87) can be summarised as follows:

- Step 1: A frequency list was created
- Step 2: The top 95 words, excluding content words, were extracted
- Step 3: The wordlist of the 95 words was uploaded on a program
(designed by Danielsson, 2008)
- Step 4: The list of recurrent candidate sequences was created

The program created by Danielsson (2008) is described in Section 7.2.3.1, above. Each candidate framework was further analysed in concordances, using Oxford WordSmith ver. 4.0, as described in more detail in the following section.

7.2.3.3 Analysis and Discussion (3) Colligational frameworks

In this section, I present potentially useful items, with some discussion, and report on how effective this method is. I also point out instances which need further analysis in Stage 2, which will be discussed in Section 7.3.

PPUI # 4: *I X1 a X2 and I X3 my X4*

The framework *I X a X and I X my X* is the second most frequent candidate structure in the corpus, occurring 10 times, as shown below. While some of the recurring sequences need further categorisation, the following is one of the examples of a sequence successfully identified without any further search.

Concordance 7.4 *I X1 a X2 and I X3 my X4* (10 examples identified from a single text)

1	knees. Can you do it? I can do it! I am a donkey and I kick my legs. Can yo
2	y hips. Can you do it? I can do it! I am a camel and I bend my knees. Can yo
3	y back. Can you do it? I can do it! I am a crocodile and I wriggle my hips.
4	chest. Can you do it? I can do it! I am a cat and I arch my back. Can you d
5	hands. Can you do it? I can do it! I am a gorilla and I trump my chest. Can
6	y arms. Can you do it? I can do it! I am a seal and I clap my hands. Can you
7	ulders. Can you do it? I can do it! I am a monkey and I wave my arms. Can yo
8	y neck. Can you do it? I can do it! I am a buffalo and I raise my shoulders.
9	y head. Can you do it? I can do it! I am a giraffe and I bend my neck. Can y
10	ked up his clothes and went home. I am a penguin and I turn my head. Can y

This result confirms that this method can identify recurrent patterns effectively. However, it should be noted that all these 10 examples occur in the same text. As pointed out earlier, it is necessary to ensure that identified patterns are not due to ‘some idiosyncrasy of [the] corpus’ (Stubbs 2002, p. 233). Nevertheless, this confirms the feature of children’s texts, as identified by many researchers (e.g. Darnton 2002; Slattery and Willis, 2001), of including a high proportion of repetitive sequences. Therefore, the fact that this particular structure occurs in only one text, even though the instances are not representative of a whole corpus, should serve as an example of how this method demonstrates the capture of repetitive sequences which allow variations in Xs. In practical terms, the second personal pronouns (in this case ‘I’) often tend to be omitted after ‘and’ when the pronoun refers to the same person. In order to investigate whether this is a particular feature of CCL or the particular text, the sequence *I X1 a X2 and X3* was searched in CCL (Concordance 7.5 below).

Concordance 7.5 *I X1 a X2 and X3*

1	over to their farm to play with them. I am a girl and I am eight years old. P
2	'Not me!' cried the Centipede happily. 'I am a pest and I'm proud of it! Oh, I
3	ot a slimy beast,' the Earthworm said. 'I am a useful and much loved creature.
4	t that I made no sound at all as I ran. I was a swift and silent mover. And quit
5	o the ends of my toes! I felt as though I was a balloon and somebody was twistin
6	he sun was shining. I felt good because I was a frog. And I felt good because

Two instances of the above (line 1 and 2) repeat the personal pronoun (i.e. ‘I’ in this case). Two instances were observed (line 3 and 4), in which the second pronoun is omitted due to referring to the same person (e.g *I am a useful and much loved creature. I was a swift and silent mover.*). No definite conclusion can be drawn from this, but when teaching pupils to learn to read, this aspect may need attention.

PPUI #5: *and X1 and X2 and X3*

This framework *and X1 and X2 and X3* is the most frequent sequence in EOCL, appearing in 168 instances across 26 texts out of 74. The words identified in the slots (Xs) are mostly nouns and proper nouns (the names of characters in the stories). The analysis of concordance lines revealed the tendency of the words appearing in the *X* position to belong to similar semantic categories. The categories identified are summarised below.

Figure 7.1 Semantic grouping of *and X1 and X2 and X3 (Nouns)*

<i>X1 and X2 and X3</i>	
[Semantic grouping]	Examples
Position:	<i>down and back and forth; up and down and round; round and up and round</i>
Food:	<i>cheese and butter and ham; tea and bread and margarine; peas and potatoes and gravy; sweets and milk and cornflakes</i>
Body parts:	<i>arms and neck and hair; hands and legs and neck</i>
Natural features:	<i>hills and trees and valleys</i>
Places:	<i>fields and farms and towns</i>

Moreover, some instances of gerund (n-ing pattern – ‘noun formed from a verb which refers to an action, process, or state.’ (Collins COBUILD, p.605)) – were also found.

Figure 7.2 Semantic grouping of *and X1 and X2 and X3 (Gerund)*

X1 and X2 and X3	X1, X2, X3 = Gerund
[Semantic grouping]	Examples
Operation on substances	<i>bubbling and mixing and testing</i>
(Limited range of) movements:	<i>nodding and opening and shutting</i>
Expressing emotions; Communication:	<i>yelling and clapping and shouting</i>

From this observation, we can hypothesise that this structure may have a restricted pattern, which tends to contain words belonging to the same semantic category.

The framework *and X1 and X2 and X3* seems to be a usage peculiar to writing for children, or perhaps to fiction, since it is unlikely that academic discourse, for example, would use the conjunction ‘and’ three times in a sentence. In order to confirm that this structure is a particular feature of texts written for children, this result was compared with evidence from an adult corpus, in this case, a subcorpus (British books) in Bank of English (BoE) (Table 7.5 below).

Table 7.5 The occurrence of *and X1 and X2 and X3* in CCL and BoE

	CCL (EOCL & JTCL)	Bank of English (Brbooks)
No. of occurrences	179	704
No. of texts	33	172
Total corpus tokens	376,210	25,361,508
Total no. of texts in the corpus	202	343
Normalised frequency	475.8	27.8

In the BoE British books subcorpus, *and X1 and X2 and X3* occurs 708 times in 172 texts. In order to make comparison easier, the results are normalised because the sizes of the two corpora are different. In order to normalize the figure, the raw frequency (number of occurrences) was divided by the total number of words in each corpus (total corpus tokens), and the results were multiplied by 1,000,000. Taking into account the relative sizes of the corpora, the frequency of this framework is much higher in EOCL and JTCL (475.8) than in BoE (27.8), indicating that this framework is more likely to be found in writing for children. The usefulness of this structure in terms of pedagogy will be further discussed in Chapter 8.

PPUI # 6: sequences which contain prepositions

Through the examination of concordances of the candidate framework, *the X1 of X2*, it was found that the preposition ‘at’ occurs frequently in L1 (first left to the node ‘the X1 of X2’). The following gives examples taken from the concordances.

Figure 7.3 Colligational framework *at the X1 of X2*

	X1		X2
at the	age	of	six, seventy four (2), four, sixty
	arrival	of	such a tiny girl
	back	of	William’s classroom, the room, the classroom

The example below shows the candidate frameworks with the preposition ‘on’.

Figure 7.4 Colligational framework *on the X1 of X2*

	X1		X2
on the	arm	of	Mr Robinson’s favourite <u>chair</u>
	back(s)	of	it, her neck, my neck, their necks, the turtle, the cow, the sleeping boar
	bank	of	chocolate <u>river</u> , Limpopo <u>river</u>
	books	of	the geographer
	bottom	of	each chair leg
	bridge	of	the Queen Mary
	corner	of	Main Street [2], the main street
	day	of	the forty-fourth, sunset, your visit, the wedding, my birth
	door	of	the dreaded private study[2], my house, Mistletoe Cottage
	edge	of	the bed [4], the big bed, the hole, the village [2], the town, the crater, that dark patch, it
	eleventh	of	November
	end(s)	of	his tail, the tow line, Mr Nosey’s nose [3], a long bit of wire, them

Observation showed that if the preposition ‘of’ has the role of combining two nouns or noun phrases, the first noun (X1) identifies the features of the second noun or noun phrase (X2). Pedagogical use of this framework could be related to teaching the use of prepositions which are generally difficult for Japanese learners of English. I labelled the sequences containing prepositional phrases PPUI #6.

The limitation of this colligational framework is, as can be seen from the way in which I have presented the material in the X2 (within brackets), that it sometimes requires more than one word for the framework to produce a meaningful unit. This is rather problematic, as the identified frameworks are not always the unit of meaning if X2 is confined to only one word (e.g. *end of his* as opposed to *end of his tail*). This

phenomenon calls for a systematic method of identifying the boundary of sequences, when dealing with occurrences like those in the examples. Therefore, in the following section, a method of co-selection is proposed in order to deal with such ambiguity in the boundaries.

7.3 Methods of extracting pedagogically useful items: Stage 2

7.3.1 Rationale for having Stage 2 analysis

Stage 1 described useful methods for extracting instances which are ‘potentially’ useful lexical items. As pointed out in Section 7.2, the application of automatic methods threw up items which are not always meaningful and further searches were needed. The following gives a summary of problems encountered with each method applied:

1) Frequency list: single wordlist

Frequency information is crucial in the analysis of corpora. However, the frequency wordlist itself does not provide full information on the usage of a word, and it often is impossible to construe the meaning of the items merely by looking at them out of context. This shows the inefficiency of using frequency lists as a single method in identifying pedagogically useful items, as it fails to provide the usage of the actual instance (e.g. as identified in Chapter 6, instances of ‘cried’ in the CCL did not include the first prototypical sense ‘shedding tears’ as listed in the COBUILD dictionary. This would not have been discovered if we had access to the wordlist only).

2) N-grams

The concept of N-grams is useful in identifying recurring sequences in specified lengths of text. However, low frequency sequences are missed out by this method and variations in sequences are also not brought to light (i.e. *a bag of*; *a loaf of*; and *a bottle of*) unless they are all equally frequent.

3) Colligational framework

In the process of analysing the environment around the candidate framework which included prepositions, evidence was identified which indicates the importance of paying attention to larger units in inferring meaning. However some difficulties were also identified in specifying the boundaries of the sequences.

Of course, each method has advantages and disadvantages (as discussed in 7.2.1.1; 7.2.2.1; and 7.2.3.1); the identified problems are related to the disadvantages of each method. In order to deal with the problems summarised above, the next section gives a description of a proposed method in Stage 2.

7.3.2 Implementation of Method

The notion of collocation which is accepted in contemporary literature on corpus linguistics originates from Firth's (1957, p. 54) description: collocation as 'actual words in habitual company'. The methods used to calculate collocation by means of WordSmith (ver 4.0.) are based on the idea of finding how closely each collocate relates to the search word (i.e. the node word) near which it was found. It is defined as 'a word which occurs within the neighbourhood of another word' (Baker et al. 2006, p. 36-37).

The information of frequency is involved with the selection of collocations, as the technology for retrieving frequency is crucial in identifying collocates.

Danielsson (2007) points out that, although multi-word phenomena in language are acknowledged across disciplines (i.e., modern linguistics and computational linguistics), multi-word units are still seen as longer than one word but serve as a single lexical unit and are therefore treated as ‘instances of the same categories as words’ (e.g. ‘in order to’ as a preposition, and ‘estate agent’ as a noun). However, Danielsson (2007) further argues that ‘if we are to fully embrace and understand the role of words in language, we may need to be aware of the possibility that the construction of multi-word units has much more to do with conventionalization than anything from the grammatical system,’ and suggests a method for identifying multi-word units.

First, all occurrences of the target node word (N) are identified in a large corpus and the most frequent collocate of the node (F1) is calculated within a span of 9 words (i.e. 4 words right and 4 words left to the N) ‘Function’ words are discarded. (They are identified by having an arbitrary cut-off point at a high position in the overall word frequency list for the corpus in question). Then all the lines which contain both N and F1 are selected, irrespective of the position of F1 relative to N. Taking just those lines, the most frequent word in them (apart from N and F1) is identified as F2. The process is repeated until occurrences of the new collocate identified fall to below 5 (Danielsson, 2007).

The procedure suggested in Stage 2 (co-selection) is summarised as follows (adapted from Danielsson 2007).

- 1) A PPUI (as identified in the first stage) is searched in concordances using WordSmith which allows the calculation of collocates around the searched PPUI.
- 2) From the calculated (within the 9-word window) list of collocates, the most frequent lexical item is chosen (disregarding the function words

and grammatical words)

- 3) The concordancer co-selects those two items (i.e. PPUI and the collocate chosen)
- 4) The concordances are examined
- 5) The procedure from 1) - 4) is repeated, until there are no fully lexical collocates
- 6) The sequence is identified.

7.3.3 Analysis and Discussion Stage 2: co-selection

In Stage 1, automatic methods were applied to the corpora, and potentially useful items (PPUI) were selected. The method described above was implemented for each PPUI. Below are examples of the way in which the identified potentially useful items are analysed in the proposed Stage 2. The following items can be proposed as pedagogically useful items, which are identified through the methods described in this section.

1) Potentially pedagogically useful items (PPUI # 1) from frequency

In the discussion of frequency above, it was shown that '*the*' is the most frequent word in the corpus. As pointed out, frequency information has some benefits, but it was also argued that the usefulness of an item should not be determined solely by frequency. For instance, '*the*' appears 28,461 times but this does not reveal how it is used in context. The co-occurrence of several words is one of the ways to observe phraseology in the corpus, and, as stated by Sinclair, grammatical words can support the extraction of phraseology. The following figure (Figure 7.5) shows the collocates of '*the*' and their frequency according to where they appear.

Figure 7.5 Collocates of ‘the’ and positions in which they occur

N	Word	Total	Left	Right	L5	L4	L3	L2	L1	Centre	R1	R2	R3	R4	R5
1	THE	28461	3937	3937	875	1184	1694	181	3	20587	3	181	1694	1184	875
2	AND	6420	3384	3036	601	701	686	666	730	0	0	1354	581	489	612
3	OF	5105	2440	2665	280	290	218	69	1583	0	0	1739	433	229	264
4	TO	3741	2390	1351	291	356	346	428	969	0	0	318	312	365	356
5	IN	3477	2466	1011	162	256	136	94	1818	0	0	304	270	221	216
6	WAS	2165	1046	1119	203	233	250	139	221	0	0	475	282	188	174
7	A	2061	975	1086	267	329	330	49	0	0	0	54	434	320	278
8	ON	1951	1388	563	121	162	61	167	877	0	0	152	132	128	151
9	HE	1501	996	505	212	230	304	250	0	0	2	142	167	102	92
10	ALL	1399	1000	399	109	127	75	127	562	0	0	63	114	125	97
11	SAID	1330	783	547	69	80	60	97	477	0	0	255	144	78	70
12	THAT	1319	710	609	144	134	114	111	207	0	0	186	174	121	128
13	AT	1305	962	343	79	139	74	31	639	0	0	78	93	97	75
14	IT	1145	685	460	134	131	171	224	25	0	0	45	130	164	121
15	UP	1034	705	329	97	65	91	297	155	0	2	48	96	92	91

It is clear from the list that ‘the’ co-occurs with several prepositions in a particular position relative to the node ‘the’. For instance, ‘of’ co-occurs with ‘the’ in second right to the node position, which is ‘the * of’. The same is true for ‘to’ in the first left position, namely, ‘to the’.

2) Potentially pedagogically useful items (PPUI # 3) from N-grams

The concordances of ‘in the’ were searched and it was found that there are 1,822 instances of ‘in the’ in the corpus as a whole. Furthermore, *world* was identified as a collocate in R1 position (126 instances) which suggests the tendency of ‘world’ to occur on the first right of ‘in the’.

Therefore, as a next step, ‘in the+world’ was searched in concordance lines. There were 126 instances of ‘in the world’. In a similar way, the first lexical collocate, *PERSON* (identified as a collocate in L1 position 16 times) is chosen for a further search; ‘person in the world’ was found 16 times in the concordances, as shown below. The

concordances were sorted by L2 (two words to the left of the node) as a first sort and L1 (one word to the left of the node) as the second.

Concordance 7.6 *person in the world*

1	r Miserable, and I'm the most miserable person in the world." "Why are you so
2	d that she should be the most beautiful person in the world; the next, that she
3	ss Whoops really is the most accidental person in the world . . . well not quit
4	very tidy person. Probably the tidiest person in the world. She lived in Twopin
5	person in Bigtown, perhaps the richest person in the world, but now he has lots
6	dred and three?' 18 The Oldest Person in the World 'We return in tri
7	ucket to his wife. 'You'll be the first person in the world to change her father
8	d as he ran. 'I'm going to be the first person in the world to be sent by televi
9	uregards! Our little girl is the first person in the world to have a chewing-g
10	fellow. Actually, he was the clumsiest person in the world! That very same m
11	o," cried Mr Clever. "I'm The Cleverest Person In The World!" "Oh good," said
12	e cleverest person ever. The Cleverest Person In The World! And, he knew it!
13	't feeling anything like The Cleverest Person In The World, Mr Clever decided
14	her worm, "is Mr Clever, The Cleverest Person In The World, on his way home to
15	o," cried Mr Clever. "I'm The Cleverest Person In The World!" "Oh good," said
16	o," cried Mr Clever. "I'm The Cleverest Person In The World!" "Oh good," said

The above sorted concordances for '*person in the world*' as a node reveal that the sequence is always preceded by a superlative adjective. The use of superlative adjectives suggests that the statement tends to be evaluative (Stubbs, 2007, p.171). Hence, in order to see the variability of noun choice (apart from *person*) in the sequence, the two following queries were searched in the concordances: 1) *the* + **est* superlative ADJ + NOUN + *in the world* (Concordance 7.7); and 2) *the most* + ADJ + NOUN + *in the world* (Concordance 7.8).

Concordance 7.7 *the + *est SUPERLATIVE ADJ + NOUN + in the world*

1 which she took from her mother, who was the best creature in the world. No so
2 here before him was an amazing sight. The biggest room in the world! The fl
3 a tiny row-boat underneath the stern of the biggest ship in the world. The Space
4 es. They will make a good place for me, the biggest thing in the world, to sleep
5 was like being pinned down underneath the biggest waterfall in the world and
6 Greedy! "Hello," cried Mr Clever. "I'm The Cleverest Person In The World!" "O
7 by now he wasn't feeling anything like The Cleverest Person In The World, Mr C
8 getful! "Hello," cried Mr Clever. "I'm The Cleverest Person In The World!" "O
9 r was quite the cleverest person ever. The Cleverest Person In The World! And,
10 replied the other worm, "is Mr Clever, The Cleverest Person In The World, on h
11 Happy! "Hello," cried Mr Clever. "I'm The Cleverest Person In The World!" "O
12 a very clumsy fellow. Actually, he was the clumsiest person in the world! Th
13 rdon; but they were dealing with one of the cruellest ogres in the world. Far fro
14 as Mr Tall, tying your shoelaces isn't the easiest thing in the world. As you c
15 dreds of 'em when I was a boy. Mice is the fastest breeders in the world, did y
16 uicy, caught-in-the-web bluebottle was the finest dinner in the world - until I
17 w because I personally have tasted all the finest foods in the world!' Whereupo
18 k I know. Listen. We have just launched the finest hotel in the world. Right?'
19 have a safe tunnel leading to three of the finest stores in the world!
20 eth? Little Miss Giggles could! She was the giggliest girl in the world, and she
21 n little children were near her she had the greatest difficulty in the world to
22 importance! This is, to me, the loveliest and saddest landscape in the world
23 dangerous!' James answered. 'They're the nicest creatures in the world!
24 aren't you? Not in Loudland. They had the noisiest mice in the world. "SQUEAK
25 than a hundred and three?' 18 The Oldest Person in the World 'We re
26 Mr Uppity. Mr Uppity, incidentally, was the richest man in the world! I just
27 ppity was also rich. Very rich! One of the richest people in the world, if not
28 the richest person in Bigtown, perhaps the richest person in the world, but no
29 e story. You see, Mr Uppity was one of the rudest people in the world, if not
30 Man, yes," said Pippi. "But I'm the strongest girl in the world,
31 strong enough to hold it up?' 'It's the strongest glue in the world!' Muggle
32 do it," said Annika. "Why, that's the strongest man in the world!
33 ingers and whisked them away. To climb the tallest tree in the world! He click
34 s Neat was a very tidy person. Probably the tidiest person in the world. She liv

Concordance 7.8 *the most + ADJ + NOUN + in the world*

1 the house. Little Miss Whoops really is the most accidental person in the world
2 he youngest ordained that she should be the most beautiful person in the world;
3 along the riverside with his daughter, the most beautiful princess in the world
4 ! Nothing but thick jungles infested by the most dangerous beasts in the world ?
5 r loom all day long every day. She wove the most delicate stuff in the world. It
6 Well, my name is Mr Miserable, and I'm the most miserable person in the world.
7 ery much to welcome you about it. It's the most miserable place in the world. M
8 mond stomacher, which is far from being the most ordinary one in the world."
9 d. 'It's quite safe. And we're going to the most wonderful place in the world!'
10 ndpa Joe, taking a long deep sniff. All the most wonderful smells in the world s

In order to see whether there are any variants in the position of 'world', *the* SUPERLATIVE ADJ + NOUN + *in the* * was searched in the concordances. In total 45 instances were found: 13 instances with 'world'; 'in the whole wide world' (line 11 below); and the instances with something other than 'world', which are listed below (Concordance 7.9):

Concordance 7.9 *the SUPERLATIVE ADJ + NOUN + in the ** except 'world'

1	importance! This is, to me, the loveliest and saddest landscape in the
2	ne left for me!â€™ Rat was perched upon the highest shelf in the cellar, peering
3	wealthy man, and in fact her father was the richest man in the country. So the ba
4	the gong and climbed to the very top of the highest tree in the forest, way up w
5	was sounded and the word went out that the biggest bomb in the history of the
6	to the bathroom upstairs. It had been the messiest room in the house, but now
7	s that the princess should be placed in the finest apartment in the palace, upon
8	eediest croc in the whole river." "I'm the bravest croc in the whole river," sa
9	reedy," the Notsobig One said. "You're the greediest croc in the whole river."
10	t. "If we brought this off, it would be the greatest triumph in the whole history
11	s the story of Mr Strong. Mr Strong is the strongest person in the whole wide world
12	b up that tree?" he asked, pointing to the biggest tree in the wood. "I can do

The instance in line 11, 'whole wide world' can be classified in the same group as 'world'. Semantically, they are categorised into two types: 1) place (*cellar, country, forest, house, palace, river, wood*) and 2) period of time (*history*).

Concordance 7.10 *the most ADJ + NOUN + in the **

1	put the key in the keyhole. 'This is the most important room in the entire factory
2	e's a murderer," my grandmother said. "She's the most evil woman in the entire world!"
3	VERY UPON ALL EIGHT OF THESE GALLANT FLIERS. THE MOST IMPORTANT PERSONS IN THE LAND
4	Shobei had hapned to pick up. Shobei became the most important man in the village.
5	'Oh, yes,' cried Charlie, 'I think it's the most wonderful place in the whole world!'
6	shall I do?" cried Toad. "These must be the most frightened seeds in the whole world.
7	which she had woven that day. It was the most beautiful cloth in the whole world

The above instances can be explained in Sinclair's model of extended units (1998). The detail of the model is described below.

- Collocation is the relation between the node word and individual word-forms which co-occur frequently with it.
- Colligation is the relation between the node word and grammatical categories which co-occur frequently with it.
- Semantic preference is the relation between the node word and lexical sets of semantically related word-forms or lemmas.
- Semantic prosody is the discourse function of the unit: it describes the speaker's evaluative attitude.

(Sinclair, 1998 cited in Stubbs, 2007, p.178)

Following the model, the example with superlatives can be analysed as follows:

- Collocation (in the + nouns (place or period of time))
- Colligation (the + 1) superlative adjective + noun + in the world)
- Semantic preference (place or period of time; semantically related e.g. tree-forest, croc-river)
- Semantic prosody (evaluation influenced by the superlative adjective)

It should be noted that the instances show an artefact of a particular collection (the ‘Mr Men’ and ‘Little Miss’ series) and could be a limitation of this kind of corpus. However, as demonstrated in the framework of Sinclair’s model above, this sequence as a whole serves to add the evaluation of a particular character by the use of superlative constructions. It could be argued that the nouns and adjectives are sometimes variable, but this method clearly leads to specifying the longer sequences in a different way from the n-grams often analysed.

3) Potentially pedagogically useful items (PPUI # 6) from colligational frameworks

As seen in 7.2.3.3, with the example of the candidate framework ‘*on the X1 of X2*’, the manual procedure yielded a long list of instances. In order to find the prototypical meaning (most common phrases) attached to this framework, the following procedure was adopted, following the methods illustrated by Danielsson (2007).

First, all the concordance lines with the sequence ‘*on the X1 of X2*’ as a node were identified. From these lines the most frequent collocates of the node were calculated and from the collocate lists the function words were disregarded. The most frequent identified collocates within the span specified are: *house, right, bed, head*.

For each collocate (excluding the function words), concordance lines for the node and the collocate were generated, where the collocate occurs within a span of 4 words to the left and 4 words to the right from the node word (9 word window). Below are the concordances for *head* and *on the X1 of X2* co-selected (Concordance 7.11).

Concordance 7.11 Co-selected *head* and *on the X1 of X2*

1	So he took it off his head and tied it on the head of the last Jizo. Then he we
2	he middle of the lawn with a bird-bath on the top of his <u>head</u> .' 'We could try
3	t I think it is ridiculous to have ears on the sides of one's <u>head</u> . It certainl
4	lion dollar budget balanced beautifully on the top of his bald <u>head</u> . Everyone cl
5	amin H, because it makes you grow horns on the top of your <u>head</u> , like a bull. Bu
6	closer look. She was now kneeling right on the edge of the hole with her <u>head</u> do
7	rridors, with his black top hat perched on the top of his head and his plum-colo

The concordances (Line 2, 3, 4, 5, and 7) show that the first noun in X1 is either ‘*top*’ or ‘*sides*’, when head is followed by a possessive determiner (*his*, *one’s*, *your*) and the noun *head*. This exemplifies the earlier assumption that what appears in X1 in ‘*on the X1 of X2*’ is restricted, because what appears in X1 is decided by the following noun or noun phrase in X2. In other words, if the preposition ‘*of*’ has the role of combining two nouns or noun phrases, the first noun (X1) identifies the features of the second noun or noun phrase (X2), as shown in the observation of concordances. Pedagogical use of this framework could be related to teaching the use of prepositions, which are generally difficult for Japanese learners of English to master. Moreover, it indicates that prepositions should not be taught in isolation but in context.

The next context word, *right*, was co-selected with *on the X1 of X2* (collocate calculated within 9 word windows: 4 words to the left and 4 words to the right from the node) and the concordances are as follows (Concordance 7.12):

Concordance 7.12 Co-selected *right* and *on the X1 of X2*

1	and had finally landed with a big thump right on the back of the sleeping boar.
2	in her armchair so that she was sitting right on the edge of it. Both her hands
3	running kick at William. She caught him right on the point of her toe and sent h
4	id, pointing. 'It's that little cottage right on the edge of the town, the tiny
5	get a closer look. She was now kneeling right on the edge of the hole with her h
6	e air and brought it down with a crash right on the top of the wretched Bruce
7	e peeped round. "BANG" went a hammer right on the end of Mr Nosey's nose.
8	the wall. "SNAP" went a clothes peg right on the end of Mr Nosey's nose.
9	"SPLASH" went a very wet paint brush right on the end of Mr Nosey's nose

The collocate around *right on the X1 of X2* within the specified window showed no fully lexicalised collocate. Danielsson (2007) claims that ‘[when] there are no collocates occurring with sufficient frequency cut-off point... we may claim to have achieved the maximal unit’. Moreover, Danielsson (2007) suggests that it is important to check whether ‘there are any paradigmatic variations (i.e. [by testing] each word in the unit to see if the unit allows for alternatives)’. Regarding this structure, the concordance search revealed that the words around the framework are grammatical words and do not have any alternatives. Danielsson (2007) gives the reasons for this as follows: ‘[Frequency words] are often said to provide structure rather than meaning, and as such they seem to object to modification and variation’ (Danielsson, 2007).

This procedure and method shows how to specify the length of the framework with the use of collocates and frequency information. It also demonstrates that this method can to some extent deal with positional variation which the n-gram approach fails to handle.

7.4 Summary of integrated methods: Stage 1 and Stage 2

The analysis was carried out in order to evaluate each method in identifying pedagogically useful items for young Japanese learners. Some automatic identification was tried, but what is also needed is qualitative interpretation of the data, because

automatic identification relies on frequency or statistical methods alone and therefore items which are found by these methods are sometimes not meaningful. Moreover, it should be noted that these automatic extractions of sequence do not allow the extraction of any variation of the similar framework (e.g. *a bit of*, *a lot of* – the “*a * of*” framework) and some examples might be missed because they do not appear as frequent occurrences. In addition, n-grams identified only by statistical methods did not always yield meaningful units, and required further analysis through the use of concordances.

In order to prevent the omission of infrequent structures and variations within the multiword units, a program by Danielsson (2001; 2003) was employed to extract candidate colligational frameworks. As demonstrated in the analysis, this makes more and more candidates available, tending to confirm the view that meanings are often created through the patterning around the words.

Stage 2 analysis, which involves a further search in concordances by co-selecting a collocate together with a node, seems to have solved the problem of determining the length of sequences, and finding the variation within the frameworks. This shows that the methods employed in Stage 1, with the further search of co-selection with collocates (Stage 2), can provide the empirical evidence from a corpus and can therefore contribute to the further exploration of sequences and identification of pedagogically useful items. It should also be noted that, as shown in this section, while the findings from existing research are helpful in pointing out which are the useful items, the evidence is stronger when it can be demonstrably correlated with some empirical identification of the items and sequences which are important for learners.

7.5 Limitations

The methods described above are inter-related or overlap (e.g. frequency is involved in every search) but the difference arises when items are viewed as ‘word-based’, as longer sequences or as frameworks which allow variations in slots.

There is also a limitation with regard to the nature of the corpus used in this study: the size of the corpus, the kinds of text included in the corpus (as seen in instances from a particular series). Therefore, there is a possibility, if different texts are used, that the identified items may not always be the same. However, as mentioned at the beginning of this chapter, the main aim of this section was to demonstrate methods for selecting useful items for reading pedagogy, to allow the presented methods to be implemented even if the texts included in other corpora to be studied are not the same as those in my corpus.

7.6 Conclusion

My main aim in this chapter was to investigate methodologies which can assist in the identification of pedagogically useful features of the texts in the corpus. The two stages of enquiry below were proposed for indentifying pedagogically useful items.

Stage 1: automatic methods (frequency, n-grams, colligational frameworks)

Stage 2: co-selection using information from collocates

In Stage 1, the methods available in corpus linguistics which applied to the corpora of this study were summarised, with a discussion of their strengths and weaknesses, followed by the evaluation of outputs in terms of their potential in selecting pedagogically useful items. Items which are identified as ‘potentially useful’ were labelled PPUI -1, -2, -3 etc... The method using colligational frameworks allowed me to successfully identify some repeated patterns with variations (i.e. 7.3.3.3; *I X1 a X2 and I X3 my X4*), but further qualitative investigation in concordances was mostly necessary after the automatic extraction of the items. In Stage 2, the co-selection method was applied to those potentially useful items in order to determine the usefulness of the items and sequences, as well as the appropriate length and variations of the items.

It can be concluded that this pair of integrated methods (Stage 1 and Stage 2) demonstrate a way to identify pedagogically useful items (PUI). This confirms that both quantitative and qualitative aspects of the investigation are needed in order to determine whether something is potentially useful for teaching. The methods presented here show that they complement each other, each compensating for the weaknesses of the other to some extent (e.g. to identify the variations in slots). Overall, this chapter provided examples of the way in which a corpus analysis fulfils the aim of identifying which lexical items or sequences to teach, how it is used to find them, what sort of problems present themselves and how they can be solved.

Moreover, although I started my analysis with a single-word search and employed ‘lexical items’ so as to show no preconceptions about what is useful (Section 7.2), the detailed analysis provided more evidence of the lexico-grammatical features of language used in the collection of texts written for a child audience (e.g. the prominence

of repeated sequences). This, in turn, suggests the relevance of phraseology to teaching in the context of teaching children to learn to read (since it helps to determine which phraseological items are useful for a syllabus). The next major part of the challenge is to explore how these items could be of use in teaching young learners in the context of TEFL; the findings need to be discussed from the perspective of applying them in EFL primary literacy education. This discussion is the main theme of the following chapter (Chapter 8). After that, most importantly, we should ask how data derived from the corpus by means of these methods can be utilized in syllabus design. This will be discussed in detail in Chapter 9.

CHAPTER EIGHT

PEDAGOGICALLY USEFUL ITEMS FOR JAPANESE YOUNG LEARNERS

8.1 Introduction

Chapter 7 first described how different methods and ways of integrating them were used in analyses of the corpus, and next discussed the findings. The suggestions for integrating methods led successfully to the identification of Potentially Pedagogically Useful Items (PPUIs), where needed.

The initial motivation for my research comes from the belief that primary level English teaching in Japan needs a syllabus for the teaching of literacy. Section 7.2 has explored criteria for establishing what is ‘pedagogically useful’ (i.e. Pedagogically Useful Items – PUIs) to Japanese learners. The present chapter discusses the implications of the PPUIs identified in Chapter 7, and discusses how they could be considered as PUIs (pedagogically useful items), rather than just potential items, in the context of the Japanese syllabus for primary literacy in a foreign language. Syllabus designers should, presumably, include items only if they are pedagogically useful, so the main question to be asked in this chapter is: ‘What are the specific implications of the PPUIs identified in Chapter 7 for primary ELT in Japan?’ By answering this question, I shall attempt to explore a set of criteria for teachers to decide what is pedagogically useful.

Along with the question above, I also address research questions relating to possible similarities and differences across the two corpora (EOCL and JTCL) in order to

contribute to answering RQ 2.2 (i.e. What differences and similarities, if any, can be identified between the sub-corpora of texts written originally in English and those translated into English from Japanese?).

8.2 Implications of PPUIs for primary ELT in Japan

In this section, I discuss the pedagogical implications of PPUIs as identified in Chapter 7 for the teaching of young learners whose first language (L1) is Japanese, in order to demonstrate how the potential items can be claimed to be pedagogically useful. This discussion, in turn, helps to establish certain principles and criteria for identifying pedagogically useful items for specific learners in their particular learning context.

The PPUIs identified in Chapter 7 are largely categorised into four groups, as summarised below:

- PUI-1: Contractions (PPUI #2);
- PUI-2: Semantic groupings (PPUI #5 framework of *X and X and X*);
- PUI-3: Sequences containing articles and prepositions
(PPUI #1 definite article), (PPUI #3 n-grams containing prepositions and articles), and (PPUI #6 colligational frameworks – containing prepositions and articles); and
- PUI-4: Repeated sequences (PPUI #4 full successful extraction of repeated sequences).

Each group is discussed below in relation to its pedagogical usefulness for literacy instruction in primary ELT in Japan.

8.2.1 PUI-1: Contractions

One of the identified PPUIs was contractions (PPUI #2); the contractions ‘-s’ and ‘-t’ were discussed briefly in Chapter 7. The following table shows the contractions found within the 100 frequency rank in the frequency list (see Chapter 7, Table 7.1).

Table 8.1 Frequency of ‘-s’, ‘-t’, ‘-ll’, and ‘-re’ in EOCL and JTCL

Form	EOCL	/million	JTCL	/million
‘-s’	320	966	51	1160
‘-t’	290	875	14	319
‘-ll’	105	317	5	114
‘-re’	56	169	10	228

*normalised frequency per million.

As can be seen in the above table, -s, -t, -ll, -re – are all the possible instances of contractions; these all show relatively high frequency (i.e. ‘s’ ranks 13th, ‘t’ 18th, ‘ll’ 81st, ‘re’ 92nd in the 100 most frequent words). For instance, the following table, Table 8.2, shows the result of a manual analysis and categorisation of ‘-s’ found in EOCL and JTCL with their percentages.

Table 8.2 Different forms of ‘-s’ found in EOCL and JTCL

Contraction	EOCL	%	JTCL	%
Total occurrence ‘-s’	342		60	
Verb contractions	193	57	41	69
Others (possessive -s)	121	35	10	16
Others (irrelevant)	28	8	9	15

As can be seen from the above, more than half the instances of ‘-s’ are identified as contractions, both in EOCL (57%) and JTCL (69%), and others are possessives (35% in EOCL and 16 % in JTCL). Possessive –s is also frequent, suggesting that apostrophe –s denoting possessive is an important item to teach children, as they need to understand the difference in use and meaning between contracted forms with ’s and possessive –’s.

In light of the high frequency of ‘-s’ in contractions as identified above, different forms of contraction were searched in order to examine their frequency. The following table shows the normalised frequency of the contractions identified in each corpus.

Table 8.3 Contractions of different forms

Contraction	EOCL	/million	Contraction	JTCL	/million
<i>I’m</i>	367	1107	<i>I’m</i>	79	1763
<i>we’re</i>	133	401	<i>we’re</i>	25	557
<i>they’re</i>	129	389	<i>they’re</i>	7	156
<i>he’s</i>	148	446	<i>he’s</i>	7	156
<i>she’s</i>	128	386	<i>she’s</i>	6	133
<i>I’ve</i>	179	540	<i>I’ve</i>	18	401
<i>We’ve</i>	75	226	<i>We’ve</i>	3	66
<i>they’ve</i>	29	87	<i>they’ve</i>	3	66
<i>isn’t</i>	132	398	<i>isn’t</i>	13	290
<i>wasn’t</i>	153	461	<i>wasn’t</i>	11	245
<i>weren’t</i>	22	66	<i>weren’t</i>	1	22
<i>don’t</i>	646	1949	<i>don’t</i>	62	1383
<i>didn’t</i>	249	751	<i>didn’t</i>	34	758
<i>can’t</i>	271	817	<i>can’t</i>	19	424
<i>couldn’t</i>	192	579	<i>couldn’t</i>	36	803
<i>won’t</i>	111	334	<i>won’t</i>	18	401
<i>wouldn’t</i>	103	310	<i>wouldn’t</i>	12	267
<i>shouldn’t</i>	9	27	<i>shouldn’t</i>	2	44

As can be seen in the above table (Table 8.3), contractions as a whole are more frequent in EOCL than in JTCL. This could be due to the limited variety of types in JTCL. However, in the case of contraction ‘*I’m*’, the difference in relative frequency seems at first glance to suggest that this contraction is more frequent in JTCL than in EOCL, appearing at 1763 per million in JTCL, as opposed to 1107 per million in EOCL. The chi-square test was applied, to measure whether there was any significance in these differences in frequency. The chi-square result indicated that the difference is at a 0.1% level of significance, which suggests that the difference is highly significant. This result

confirms that *'I'm'* is more frequent in JTCL, while other contractions are more frequent in EOCL than in JTCL.

In Chapter 2, I reported the concern expressed by Butler (2005) who witnessed the phrase *'I'm ~'*, being used preceding a noun without an article in YL classrooms in Japan (e.g. *'I'm apple*'*). The fact that *'I'm'* is included in current teaching is encouraging, but what follows after *'I'm'* in models provided to learners, leaves room for improvement. For instance, in the JTCL corpus, 50% of all instances (39 instances) of *'I'm'* are followed by adjectives or adjectival phrases; 19 instances are present progressives (24%); 19 instances are followed by noun phrases (24%) and 2 instances are a part of phrasal verb 'to be about to' (2.5%). These results indicate that it is more useful and more meaningful for learners to be able to describe feelings (e.g. tired, happy), their state (e.g. busy) with adjectives and/or adjectival phrases; or action (e.g. running, swimming) by using present progressives, rather than 'fictionally' describing themselves as an object (e.g. apple), in a way which can have no communicative purpose in their lives, or anyone else's. This observation shows that the corpus, if it contains the right texts, can contribute to deciding what kind of items can be grouped and taught together.

It has been pointed out by many researchers that one should not rely on frequency alone when deciding what to include in a syllabus, as frequency cannot justify everything. In addition, it is always important to remember that a corpus is always 'naturally restricted' as pointed out by Shin and Nation (2008). However, the comparison of different forms of contraction reveals that high frequency clearly indicates that children are relatively

likely to encounter these forms in their reading. As Sealey and Thompson (2006, p.25) suggest:

[S]tatistics ... cannot directly determine what children should be taught about their language, but if teachers are aware of them they could consider the possibility of making sure that their pupils are familiar both with the most frequent of these contractions and with the reading contexts in which they are most likely to occur.

In addition to the result above, the same phenomena have been observed in the EOCL and JTCL corpora, despite their relatively small scale. It can be claimed that the phenomenon of contraction, along with the words which follow, is worth regarding as a PUI (pedagogically useful item) which merits inclusion in the syllabus.

8.2.2 PUI-2: Semantic grouping

In Chapter 7, it was noted that the lexical items which occupy the ‘X’s, in the framework *X and X and X*, share some semantic meanings. This aspect has pedagogical implications in terms of teaching beginners. For instance, not only can it be used to teach pupils the use of connectors, but it is also useful in teaching them synonyms or the grouping of items according to their meaning. Once those two aspects are mastered, the information about the semantic similarity of Xs, can be used to demonstrate how to process an unknown word in this kind of sequence, making use of the other known word(s). For instance, if learners do not know one of the words (e.g. *bread*) in the following sequence (e.g. *egg and bacon and bread*), given the fact that ‘*bread*’ is likely to be in the same category as ‘*egg*’ and ‘*bacon*’, they can be helped to infer its meaning, or at least guess what sort of category it belongs to (i.e. a category of food), from the two other words (i.e. *egg and bacon = food, breakfast*).

The example above may be too simple, but it can be argued that it could easily contribute to a learner's language awareness. For example, the structure can be used in an exercise where pupils practise the categorisation of vocabulary items according to their meaning, which is usually regarded as an efficient way of memorising lexical items (Kadota, 2002, p.298).

As pointed out above, semantic grouping is useful and has the advantage of helping learners to memorise lexical items efficiently. However, this method should be presented with caution. For instance, Cameron (2003, p. 107) reports a story from a Korean teacher of YL about an incident in her classroom. It took place in a teaching session of the structure 'I like ~/I don't like~', after a few lessons in which the structure was introduced using the topic of food (e.g. 'I like pizza, I don't like hamburgers'). The pupils were terrified when the Korean teacher used the name of a friend with the structure, "e.g. 'Do you like Yong-Hee?'". This was because they had associated the structure with something they could eat, on the basis of the previous teaching sessions around the topic of foods, and had thought the teacher was asking '*Do you like to eat Yong-Hee?*' (ibid, p.107). This case shows the consequences of an over-simplified presentation, where the topic chosen is identified with the structure (in this case, associating the structure with examples of consumables such as food or beverages). It indicates the importance of introducing semantically different lexical items and allowing variations in sequences. In EOCL and JTCL, the following examples are found:

Table 8.4 Frequency of ‘*I like*’; ‘*I don’t like*’ in EOCL and JTCL

Form	Total	EOCL	RF	JTCL	RF
‘ <i>I like</i> ’	17	15	45.2	2	45.5
‘ <i>I don’t like</i> ’	14	14	42.2	0	0

Concordance 8.1 ‘*I like* *’ in EOCL

1	r short arms!" she said to Mr Tickle.	"I like	your beard!" she said to Mr Fussy
2	"In no way," replied the princess.	"I like	exceedingly all that you have dis
3	ughter, which irritated me very much.	I like	my misfortunes to be taken seriou
4	een-Grasshopper cried. 'Good heavens,	I like	that! My dear boy, I am a real v
5	Centipede. 'Mind your own business!'	'I like	that!' 'My dear Earthworm, you'r
6	it came to his turn. 'Hard or soft?'	'I like	it soft, thank you very much,' Ja
7	e beginning?" "Yes please," I said.	"I like	good news." She had finished h
8	ass and roasting the ants one by one.	"I like	watching them burn," he said. "Th
9	Can't you think of something else?'	'I like	blowing things up,' said the Gene
10	rotten ones, where there's no shooting.	I like	the gangsters best. They're
11	r a holiday,' said the Roly-Poly Bird.	'I like	to travel.' He fluffed his marvel
12	sks, Lavender said, 'It's good.	I like	it' 'It's true as well,' a small
13	n the grass. "That was fun," he said.	"I like	picnics!" "You ... do! ... I...
14	delicious dish." He licked his lips.	"I like	food," he explained. Mr Clever's
15	oad. "This hat is your present to me.	I like	it. I will wear it the way it is.

Concordance 8.2 ‘*I don’t like* *’ in EOCL

1	ich he had parked outside the hotel.	"I don't like	driving in the dark!" "Lov
2	Ooooooooooooooh!' the Ladybird said.	'I don't like	this at all!' 'Sssh!' Ja
3	said the Second Umcer. 'I tell you	I don't like	it,' muttered the Captain.
4	great round ball hovering overhead.	'I don't like	it,' the Captain said. 'No
5	up some more silk. Centipede!' 'Oh,	I don't like	this at all,' wailed the Ea
6	not!' 17 Rescue in Minusland	'I don't like	it here at all,' Charlie wh
7	wriggling,' said Grandma Georgina.	'I don't like	things that wriggle. How do
8	rare speeches. 'But whatever it is,	I don't like	it.' 'Aren't you enjoying
9	id Mr Wonka. 'It works by television.	I don't like	television myself. I suppos
10	'So am I! So am I!' cried Mr Twit.	'I don't like	this one little bit!' 'W
11	meant to be clever, too! Well, madam,	I don't like	clever people! They are all
12	es were riveted on the Headmistress.	'I don't like	small people,' she was say
13	and went to work, still sneezing.	"I don't like	this sneezing all the time,
14	c." "Really?" said the wizard. "Well,	I don't like	people who are constantly

Concordance 8.3 ‘*I like* *’ in JTCL

1	wash it all off." said Mom. "Oh, no!	I like	it. I really like it," said the d
2	does. And what's more, I can sing too.	I like	singing." "Sing?" he said. "Blo

There were no occurrences of ‘*I don’t like*’ in JTCL. Nevertheless, the above concordances show that ‘*I like*’ and ‘*I don’t like*’ can occur in the following pattern structures:

Table 8.5 Patterns following ‘*I like*’/‘*I don’t like*’ in EOCL and JTCL

Pattern		EOCL	JTCL
	(pro)noun	✓	✓
I like/I don’t like	ADJ + N	✓	-
	Infinitive	✓	-
	gerund	✓	✓

As can be seen from the table above, ‘*I like/ I don’t like*’ take the following patterns 1) (pro) noun 2) adjective + noun 3) infinitive and 4) gerund. In EOCL, we find all four patterns listed, while in JTCL only 1) (pro)noun and 4) gerund are found. This result could be due to the small size of the JTCL corpus. However, considering the language used in JTCL is relatively simple, this result might have some pedagogical value. For instance, once the pupils master how to express their likes and dislikes about certain objects or events using ‘I like/don’t like + (pro) noun’ (e.g. *I like picnics*) in general, they can be introduced to expressions which in some way modify objects or events, using ‘*I like/don’t like* + ADJ + N’ (e.g. *I like good news, I don’t like hot food*). Moreover, the usage of gerund and to infinitive relating to certain actions ‘I like/don’t like + gerund’ (e.g. *I like/don’t like swimming*) and ‘I like/don’t like + to inf.’ (e.g. *I like/don’t like to travel*) should also be introduced.

While some structures are flexible and admit variations in what can go into the slot, (such as ‘*I like~, I don’t like~*’), others admit only a limited range in Xs. For instance, in the case of sequences identified through the colligational framework method such as: *were X1 of X2 and X3*, when this sequence is preceded by existential ‘*there*’, then X1 is always ‘a quantifier’ describing the nouns in X2 and X3, as these extracts from the examples from concordance lines show:

(8.1) There were lots of snags and complications...

(8.2) There were piles of cakes and buns...

(where the structure is unlike that of the example below)

(8.3) There were hundreds of them, and upon each...

The difference between examples (8.1), (8.2) and (8.3) is that (8.3) has a comma (,) before ‘*and*’. In the sequence “existential ‘there’ + be + ‘quantifier’ + NOUN + *and* + NOUN”, the quantifier modifies the nouns if there is no comma before *and*. Explicitly stating the difference between structures which allow variation and those which do not, such as the examples above, is also useful, in contributing to learners’ awareness of the connection between form and meaning. It may be difficult for learners to spot these aspects by themselves. However for YL teachers, who have control over what they present to their learners, and how they do it, this kind of information is crucial. If teachers take these aspects into account and present them in ways which demonstrate their patterns to the learners, they are likely to contribute to learners’ language awareness. Making learners aware of these aspects, of course, requires choosing an appropriate teaching methodology. This is discussed in Chapter 9 Section 9.3.

8.2.3 PUI-3: Sequences containing articles and prepositions

As mentioned in Chapter 3 (Section 3.3.5), prepositions and articles are two features often considered difficult for Japanese learners of English. This suggests that there is a good reason for focusing on sequences which contain both items. It also indicates that there is still room for more suggestions about the ways in which these elements can be taught.

When children speak in English, their deficient use of articles or prepositions, for example, may not cause many problems, as what they want to convey can be understood in context in speech. Because of this, and because current primary ELT education in Japan generally concentrates on ‘speaking’, rather than reading, one might

argue that in the case of teaching children, we do not need to go into details as long as the activities are enjoyable. By contrast, when it comes to written texts, understanding the right use of prepositions and articles is indeed important for learners' understanding. For instance, the use of articles (indefinite or definite) contributes to the understanding of cohesion in the text (e.g. in terms of anaphoric and cataphoric reference). Moreover, in spoken English, correct use and understanding of prepositions and articles is necessary in expressions such as '*I feel at home*' and '*keep up the good work*'. In addition, it is worth pointing out that there are prepositions which radically change the meaning depending on the combinations: phrasal verbs such as '*pass out = faint*' and '*chill out = relax*'. For these reasons, I would argue that these aspects should not be overlooked even at an early stage of learning.

It has been pointed out that prepositions should not be taught in isolation but in context, using as evidence the co-occurrence of words around prepositions and articles (see Chapter 7 Section 7.2.3.3). For instance, the definite article 'the' sometimes co-occurs with another so-called 'grammatical word' - i.e., 'the' co-occurs with 'of' in a second right position to the node ['the X of'] and with 'to' in the first left position, as in ['to the X']. These can be taught first as a set of phrases, based on evidence from the corpus in terms of variables going into the slot 'X'. These can be treated as prototypical units to master before going on to a colligational framework, which includes more contextual usage in longer sequences. As Stubbs (2007, p.163) suggests, 'these [recurrent phrase] combinations of collocations, colligation...and their meaning often have evaluative communicative functions.' This aspect of the 'evaluative function' of a longer unit is

demonstrated in the example of ‘the most + adjective + noun + in the world’ (cf. Chapter 7 Section 7.3.3).

As Sinclair (1991, p.113) states:

[t]here is a broad general tendency for frequent words, or frequent senses of words, to have less of a clear and independent meaning than less frequent words or senses. These meanings of frequent words are difficult to identify and explain; and, with the very frequent words, we are reduced to talking about uses rather than meanings. The tendency can be seen as a progressive delexicalisation, or reduction of the distinctive contribution made by that word to the meaning.

All the information stated above, which is concerned with processing the meaning of a text by units larger than a word, seems to suggest a shift from a word-based approach - in both analysis and teaching - towards something which gives more attention to phraseology. This suggests that the incorporation of phraseology in the design of a literacy syllabus must thus be crucial, given the support of empirical evidence. This will be addressed in Chapter 9 (Section 9.2) in more detail.

8.2.4 PUI 4: Repeated sequences

One of the significant findings using the method of colligational frameworks was that it proved possible to extract repeated sequences, as demonstrated in Chapter 7. Repetitive features in the texts are typically found in writing for a child audience. Inevitably, there is a link between the repeated sequences identified, and the argument in the previous section about ‘recurrent phrases’ (whether prepositional or not), when considering implications for teaching.

The repeated sequences identified are not always prototypical units throughout the corpus, but are prototypical of a text. It is rare to find the repeated longer sequences identified in Chapter 7 (such as e.g. *I X1 a X2 and I X3 my X4*) across different texts or corpora, as they are usually specific to one text. Other examples of colligational framework could also identify recurrent sequences from particular texts. Such recurrent phrases may be pedagogically useful, even though they are not prototypical features of a whole corpus. The pedagogical implications for these sequences are: 1) they help us to identify recurrent patterns in each text; and 2) they help teachers to identify which particular sequences to focus on when using such texts in reading. However, at the same time, teachers should be aware that this is a sequence from a particular text, which should therefore not be treated in the same way as ‘an idiomatic unit’.

Research on L2 reading, such as that by Kadota and Tada (1992), claims that when people process a text in reading, the unit of process is often greater than a word; what is read is, rather, processed at phrase-level, and they introduce the concept of ‘reading units’. Kadota (1988) finds that word-by-word processing takes much longer than processing by phrase, and also reports a related study by Weaver and Garrison (1977), who studied phrase lists of ‘preposition + noun phrase’ (e.g. *of the dog*) and ‘noun phrase + preposition’ (e.g. *the dog of*). They found that the latter requires more time for people to process in reading (cited in Kadota, 1988). Moreover, Pearson, Roehler, Dole & Duffy (1974) investigated how American children process meaning in text and how they remember sentences. They found that children are more likely to remember the meaning when the sentences are held together as chunks, but not necessarily when the syntax is simple. This finding shows that it is important to make sure that the units

presented are meaningful chunks, and are not to be decided merely by the simplicity of their syntax. This aspect points to an important issue on how to design a syllabus which is not restricted to the traditional grammatical format often found in EFL.

Taking these research findings into account, it can be said that presenting sequences is very relevant to reading instruction. Although the corpus method cannot automatically provide psycholinguistically valid 'reading units' as the research shows (Kadota, 1987; Kadota and Tada, 1992), this previous research pleasingly correlates with what Hunston (2007, p. 264) claims, on the basis of corpus evidence, that readers coming across a sentence, will apply a 'phrasal idiomatic pattern' unless such a pattern does not come to mind, in which case they are obliged to resort to a 'word-for-word interpretation'.

Fries (2008, p.79-80) argues that readers make use of 'redundancy' when reading in order to better comprehend the texts, based on his analysis of miscues produced in reading. Hunston (2007, p. 254) sees the relevance in reading and pattern grammar as follows:

A contention of a functional view of literacy is that readers use redundancy in language when reading ... Redundancy arises as a result of the non-randomness of language: the obvious fact that at very few points in an utterance is it impossible to predict, at least in general terms, what will come next...Grammatical conventions suggest that *the* will be followed by an adjective or noun. Collocational regularities further skew the probabilities of word choice.

As Hunston (2007, p. 260) mentions, '[the] association between pattern and meaning emphasizes the non-random interaction between grammar and lexis.' Taking this into account, if children are taught to process texts as this 'functional view of literacy' recommends, they may be encouraged to infer the meaning of texts as they become

aware of the lexico-grammatical features through reading. There is no doubt that teachers should carefully consider the concepts behind this mode of instruction. The issues associated with the design of such a syllabus for YL reading instruction will be discussed in greater detail in Chapter 9.

8.2.5 Connection between oracy and literacy

The above section concentrated on exploring the implications of data derived from the corpus by these methods for EFL reading instruction. From the perspective of teaching reading to young learners, the following points which emerged from the analysis should be mentioned.

As was pointed out in the previous chapter, the consecutive use of the connector ‘and’ in a sequence (*X and X and X*) is peculiar to writing for children; this is confirmed by a comparison with the occurrence of such sequences in the British books (BrBooks) sub-corpus in the Bank of English (BoE) (see Chapter 7 Section 7.2.3.3). It is noteworthy that this framework reminds us of a characteristic of children’s speech (e.g. using many ‘and’s in between items in reporting an event); thus we can hypothesize that the writers may have reflected this aspect of children’s speech. This hypothesis is apposite to the debate about whether the introduction of literacy should be based on good oral skills. Therefore when children are introduced to reading, and begin to ‘learn to read’, it is important to link oral skills with written forms (Hart et al, 1977 p.34).

8.3 Summary: criteria for deciding what is pedagogically useful for Japanese YLs

The most challenging step for researchers seems to be to draw any conclusions about ‘what is useful’ for learners. In search of pedagogically relevant items, the following steps have been taken:

- Step 1: methods applied and identified candidate PPUIs
(potentially pedagogically useful items).
- Step 2: qualitative analysis using concordance lines (Qualitative analysis led to production of categories of items which can be suggested as potentially useful.)

After the usage is checked through the concordances, it is essential before making recommendations to make use of evidence which can establish what is likely to be pedagogically useful.

The PPUIs (as identified in Chapter 8 according to the criteria of features of English which present challenges to Japanese learners) were reviewed from the perspective of YL English teaching. They were specifically discussed in terms of whether they have direct pedagogical relevance as useful items in teaching young learners.

The discussion can be summarised in terms of the following list of principles for identifying items as PUI (pedagogically useful items).

- 1) Relevance to the criteria identified in SLA research
- 2) Variations (is there enough not to be simplistic, but not so much as to confuse YLs?)
- 3) Frequency (frequent enough, coverage should also be considered)
- 4) Related literature on their use (connection with research in reading and phraseological findings)
- 5) Consideration towards TEYL (specific repetitions, coverage of the repetitive sequences)
- 6) Relationship between form and meaning (e.g. as demonstrated in colligational framework *X and X and X*)
- 7) Consideration of whether the usages are typical, authentic and communicatively useful (Römer, 2005)
- 8) Consideration towards cultural aspects

These can serve as pointers for teachers of young children who are starting to read, who, like teachers in the field of ESP, consider the needs of learners for their specific purposes, and select items according to the needs identified.

The design of the syllabus should be informed by the methodology being used (e.g. if one is in favour of the lexical approach, the teaching syllabus and methodology cannot be primarily grammatical). Moreover, the corpus analysis presented here and the pedagogically useful findings seem to suggest that understanding of these aspects on the part of teachers is necessary. As identified in Chapter 2 and Chapter 5, considering the current scarcity of trained YL English teachers in primary schools, teacher education is another major issue in Japanese ELT for the successful application of a corpus-influenced syllabus. Therefore, practical suggestions for teachers and prospective teachers will be discussed in Chapter 9 (Section 9.4).

CHAPTER NINE

CORPUS-INFLUENCED SYLLABUS DESIGN

9.1 Introduction

As noted in Chapter 3, the first corpus-based syllabus was created by Willis and Willis (1989). They used frequency information from the Bank of English (COBUILD project) as the main basis for the design of a language teaching syllabus (the details can be found in Willis, 1990). While word frequency lists in syllabus design have been used to identify pedagogically useful items (see Chapters 7 and 8), the methods applied made it clear that phraseology plays an important part in the language used in the corpus.

In order to answer RQ 3.2 (i.e. How, if at all, might the data from a corpus of children's literature be used in designing a reading syllabus in an EFL educational setting?), this chapter firstly discusses the relationship between phraseology, SLA and pedagogy, in order to explore how the outcome is relevant to teaching English to young learners. In addition, incorporating phraseological information into the syllabus requires consideration of two aspects. One is how to present the phraseological (lexico-grammatical) features in a syllabus for primary learners. The other is a teaching method and approach for using data derived from the corpus in the teaching of literacy. After considering the above two aspects, this section concludes with some pedagogical implications and some practical examples of Pedagogically Useful Items. These suggestions also have implications for teacher education, which is discussed in the last part of this chapter.

9.2 Corpus-influenced syllabus design for primary ELT in Japan

9.2.1 Implications of phraseological findings, SLA and reading pedagogy

As demonstrated in the analysis in the previous chapters, more and more evidence is available to suggest that meaning is created through the patterning around the words or multi-word units (e.g. Daniellson, 2007, p.1). These findings are relevant to the growing evidence for, and interest in, phraseological phenomena. The results of analysis demonstrated in Chapters 7 and 8 (i.e. the longer sequences being the prominent features) highlight the importance of incorporating these longer sequences into the literacy syllabus for young learners (YLS). Of course, there is no need for learners to know the theory behind this attention to larger units, but the implications of this outcome have pedagogical relevance, and require further discussion. The general implications of this phenomenon in TEYL will be discussed further below.

Research in SLA suggests that phraseological patterns or chunks of the language are sometimes difficult for learners to acquire (Howarth, 1998) and this is one of the reasons that non-native speakers sometimes sound ‘bookish’ (Harwood, 2002) or tend to be ‘ill prepared for casual conversation’ (O’Keeffe et al., 2007, p. 63). The importance of incorporating phraseology and corpus-derived language description in ELT materials has been suggested by many researchers (Bernardini, 2000, p. 21; Howarth, 1998; Nattinger & DeCarrico, 1992), and the implications of this for future pedagogy are considerable. This is not to claim that teaching vocabulary items as individual words is no longer useful, and it is true that, as Sealey (2009, p.40) states, “[i]n any case, in the context of the English L1 primary education syllabus, ‘vocabulary’

features most predominantly as a dimension of literacy, which is consistent with the belief that success in reading and larger vocabularies are mutually reinforcing”. The same can be said for the L2 learners. Word by word interpretation is especially needed when the ‘open-choice’ principle, of attaching meaning to a word rather than to a phrase, is invoked (Sinclair, 1991, 2004). Nevertheless, the growing evidence on phraseology or multiword phenomena suggests that it is worth considering the value of phraseology in relation to the teaching of EFL young learners, from the standpoint of this study.

The potential for taking a phraseological approach lies not only in its usefulness for the learners’ productive fluency but also for receptive skills such as reading. Fries (2008, p. 80) notes that readers make use of what he calls ‘redundancy’ when reading a text. In other words, it is their past experience of encountering recurrent patterns which enables readers to understand the meaning of the text. This suggestion by Fries (2008) is similar to what Darnton (2000) demonstrates, using examples from children’s storybooks. Although Darnton (2000) uses the term ‘repetition’, the main point is that encountering recurrent sequences in their reading helps children learn to read.

In addition, the work done by Hart et al. (1977) takes a somewhat phraseological approach to literacy teaching. The Mount Gravatt project established ‘a developmental language reading programme’ in order to help children to read. This programme was created on the basis of a frequency list of children’s spoken utterances and phrases, in the belief that it would be logical to make use of children’s oral proficiency in a language, since this is their first established internalised language system (Hart et al., 1977, p. 14). They go on to say (p.15):

Reading involves identifying, through distinctive features, graphic units which are meaningful because they represent semantic ‘chunks’ of the reader’s organized interpretation of his or her world.

In their report, they discuss the importance of words in sequence, as follows:

Studies of young children’s conception of word boundaries seem to indicate that it is often sequences of words, rather than single words, that are important carriers of meaning in a sentence. It is clear also that when children enter school they do not perceive a word as a separate unit according to printing conventions (ibid, p.17).

In other words, it is normal for children to perceive meaning in a sequence of words rather than in a single word. The work by Hart et al. also shows ‘the minimum oral proficiency which children should have attained before attempting to read’ (ibid, p.38). The situation discussed in Hart et al. (1977) is that of native speakers of English. In the case of EFL, as has been mentioned, it should be taken into account that the process may be slower (Nation, 2009 p. 7), when the students operate in a different environment (as is the case with Japanese students, who rarely come across English outside the classroom). Nevertheless, the acquisition of ‘chunks’ contributes to working memory, and it is certainly useful in learning to read, which involves ‘learning to break the code of print’ (Turbill, 2001, p.274).

Speaking of redundancy, Hunston (2007, p.265) suggests that verbal patterning contributes to the redundancy in a text. The important notion of *grammar patterns* (Hunston and Francis, 2000) is that meaning is often shared when patterns are the same. It should be clear by now that these three aspects (i.e. redundancy, patterns and meaning) are to some extent interrelated and all of them contribute to the understanding of a text. For this reason, it seems important that the notion of phraseology, along with the information from patterns and their meaning, can be incorporated into syllabus

design towards the advancement of reading pedagogy, and the development of learners' literacy skills.

According to Hunston (2002, pp. 147-148), 'the idiom principle and the open-choice principle together provide a theoretical account for two observations'. The two relate to the pervasive nature of phraseology in English, as well as the fact that the make-up of sentences and utterances cannot be accounted for by phraseology alone. This implies that a combination of both principles is needed. Mason (2007) also suggests establishing a hybrid between the idiom and open choice principles for the automatic identification of MWUs. In addition, in relation to pedagogy, if we assume that the open choice principle is related to a word-based approach, since it is concerned with words which have meaning on their own, and that the idiom principle relates to the phraseological approach, then combining these two approaches seems to be a step forward in pedagogy.

Hunston (2007, p.265) suggests that readers only rely on word-for-word interpretation of a sentence as a last resort, that is, when there is a lack of phrasal idiomatic pattern. She stresses that language teaching should be based on grammar supporting this analysis in order to maximise our knowledge of how native speakers process their language. She goes on to discuss the construction and reception of sentences in the same way, seeing the use of pattern grammar to teach reading as a way of changing the nature of language education.

So far, I have discussed how phraseology has been perceived within corpus linguistics, and the possible implications of these insights to TEYL in the EFL context. However, current research available in Japan into ELT for young Japanese learners seems to concentrate on producing wordlists for primary school children (e.g. Chujo & Nishigaki, 2004; Ishikawa, 2007), which recalls the notion of de-contextualised list of words. It seems that the full import of phraseology has yet to be appreciated, and there is undoubtedly a need for materials which incorporate phraseological information. Moreover, if phraseology is to be incorporated into the lexical syllabus, the relevant teaching methodology should also be discussed (Section 9.3). The following section discusses necessary adjustments to syllabus design when incorporating the findings in a language syllabus.

9.2.2 Corpus-influenced syllabus design

According to Richards and Rogers (2001, p.20), there are mainly three decisions which need to be made in applying a methodology to language teaching.

- i) approach: theory of language and language learning
- ii) design: the specification of content, and of the roles of teachers, learners, and materials
- iii) procedure: the variety of pedagogical activities that can be drawn on.
(cited in Harwood, 2010 p. 5)

The methodological aspect of content specification (i.e. syllabus design) is the main focus of this study. In Section 8.3, eight pointers when specifying the PUIs for Japanese YLs were recommended. In the previous section, it was suggested that phraseological information should be incorporated in the syllabus, based on the evidence of the

lexico-grammatical features found in the corpus (Chapter 7 and 8) and their implications to TEYL is discussed (Section 9.2.1).

Hunston (2002) mentions that to design a lexical syllabus, consequently, means to describe a corpus. While Sinclair (2004, p.148) suggests that ‘a word is not the best starting point for a description because meaning arises from words in particular combinations’, Hunston (2002, pp. 147-148) claims that phraseological description alone is not enough for the description of English. In terms of the syllabus design, the remaining question to be addressed is: how can a lexical approach such as that suggested by Sinclair and Renouf (1988) and Willis (1990; 2003) be incorporated into current TEYL in Japan using the findings in Chapters 6, 7 and 8?

9.2.2.1 Selecting and sequencing

Firstly, in terms of the selection, the value of frequency needs to be addressed. As has been suggested by Willis (1990), ‘frequency’ is certainly a crucial criterion in selecting and sequencing lexical items for teaching. For instance, Shin and Nation (2008) use frequency in selecting collocations from the BNC spoken section, for the ‘cost-benefit advantages’ of such information. They further argue that ‘having a list of the most frequent collocations in spoken English to choose from is a useful starting point for syllabus design’ (ibid, p.346). In order for the extraction of PUI (pedagogically useful items) to be implemented in this study, the measure of frequency was essential when using most analytical methods available in corpus linguistics (e.g. n-grams, and colligational approaches, as well as calculating collocates around the node word, all rely on frequency counts).

However, it should be clear by now that word frequency lists should not be the sole guide when one designs a syllabus, though they are useful as starting point in syllabus design. Even at the stage of analysis, as has been pointed out in sections 7.2 and 7.3, both quantitative and qualitative judgements of the findings are necessary, in addition to the frequency information. This is because quantitative analysis contributes to the creation of categories through counting, while qualitative analysis produces categories by detailed interpretation of the data. Although Shin and Nation (2008) used frequency information of collocation as a primary criterion, they also point out the importance of balancing the result with other factors (e.g. 'learner need, range of use, difficulty, teachability, suitability for the age and background of the learners'), bearing in mind the characteristics of the corpus, before applying the result to teaching.

Indeed, the analysis provided in Chapter 6 revealed that some words can have cultural significance (e.g. 'honey' as an endearment usage in EOCL) even if it is infrequent in the CCL. This implies that frequency should not be the only factor, as Hunston (2002, p.193) suggests. Hancioğlu and Eldridge (2007, p.340) also express scepticism about the idea of relying on single-item frequency because this does not tell us 'when and where an item is introduced into the teaching and learning process'. This point is related to how we sequence lexical items in a syllabus or in teaching.

It is clear that syllabus designers or teachers should maintain a balanced attitude towards selecting items for practical application. Along with the frequency information one must bear in mind the practicability of lexical items. Stubbs (2007, p. 181) suggests:

a major criterion is that the results should be comprehensive, down to a given frequency cut-off. ... [T]his restriction is justified both on theoretical grounds (what is typical should receive more prominence) and on practical grounds (e.g. beginning language learners should start with the most frequent and most useful phrases).

In terms of reflecting empirical evidence in teaching and learning practice, Stubbs (2007, p.182) suggests the following: ‘if we are interested primarily in practice, empirical quantitative data about recurrent phrasal constructions can help with designing a new type of dictionary which would list phrasal units with their canonical forms’. Although his concern here is with dictionary design, there are also implications for syllabus design, that is the possibility that phrasal constructions should likewise be listed separately in a syllabus. As Stubbs (2007, p.182) points out, ‘learners might initially require just the most frequent prototypical exponents of the most frequent constructions. However, the best ways of presenting information remain to be investigated’.

In relation to the point raised above, it is often said that applying corpus evidence and phraseology to ELT is far from straightforward (see Chapter 2, Section 2.3.1.). With regard to the application of phraseology to ELT, Hunston (2002, p.152) raises some issues:

To what extent can phraseology be used as a description of what happens in English, in lieu of, say, a grammatical rule or explanation? Will phraseology as a descriptive technique lead to a ‘phrase-book’ approach to English, where learners can produce many phrases fluently but cannot analyse them to attain a more creative mastery of the language?

The incorporation of phraseological information in the syllabus subsequently calls for a re-consideration of the treatment of lexis and grammar presented there.

As the lexical approach focuses on words and their use in context, patterns link the grammatical and lexical aspects of any construction. The question raised at the end of the previous section could be dealt with using the notion of *pattern grammar* (see Hunston and Francis, 2000). In addition, Hunston and Francis (2000), in their extensive research on the relationship between patterns and meanings, discovered that words which have the same pattern also tend to share meaning, both semantically and pragmatically. These principles of pattern grammar unify the traditional division between lexis and grammar. This notion, to my knowledge, has not been taken up in TEYL syllabus design in Japan.

An example from my analysis is that in the Corpus of Children's Literature (EOCL and JTCL), the instances of the verb CRY rarely bore the so-called prototypical sense of CRY (i.e. with the meaning 'shed tears'), but were almost always meant in the sense 'cry out' (i.e. 'shout') (see Chapter 6 Section 6.4 for details). These findings on the sense differentiation according to context and the proportion of each sense, which are very specific to a specialised corpus such as the one in the present study, should be reflected in the syllabus (i.e. by listing them according to the frequency of each sense).

In summary, the suggestion for the incorporation of phraseology into a syllabus is to take the aspect of 'patterns share meaning' and the 'prototypical' sense into consideration. It does not mean that the syllabus should contain every instance of lexical items in the corpus, but rather, as the findings in Chapter 7 and 8 seem to suggest, the proposed syllabus should efficiently incorporate the following:

- 1) grammatical points and lexico-grammatical findings
- 2) perspectives from phraseology (based on colligation, contextual usage, MWUs, longer sequences, strings and phrases).
- 3) particular meaning, if any, attached (e.g. cultural meaning, connotation, metaphors).

Such a suggested syllabus design could group words and phrases as a set, and list them according to their ‘frequency in meaning’, making use of information from the corpus analysis. As Stubbs states, the best way of presenting such empirical information in a syllabus is an unresolved issue, and of course the suggestions made here might not be the complete solution. Nevertheless, it can be argued that they provide a useful and valid development from both theoretical and empirical perspectives.

Nevertheless, one criticism of the lexical syllabus, is that it is merely an ‘itemized syllabi’, Bourke (2006, p.280) describes these as ‘needlessly prescriptive and uni-directional’. It is important to cast aside the ‘false assumption’ of a ‘one to one relationship between teaching and learning’ especially with TEYL (ibid) and the process of teaching and learning should also be considered (Hancioğlu and Eldridge 2007, p.340). This aspect is further discussed in Section 9.3.

9.2.2.2 Implication for the TEYL literacy syllabus

The focus of this thesis is English literacy which has been somewhat neglected in the context of TEYL in Japan. With the issues raised above in mind, I shall now discuss what this suggestion means in relation to TEYL in general and the context of Japanese primary learners in particular.

It is useful to revisit the Japanese local context regarding English literacy teaching before discussing the implications of corpus-influenced syllabus design at primary school level. Children in Japan affected by the issues discussed in this study are at the earlier stage of learning to read in English. They are at the stage of both learning to speak while familiarising themselves with words and phrases in English, just as other children who are learning to read a foreign language need to do (Riley, 2006). In terms of teaching literacy to YLs, there are some tensions surrounding primary English education in Japan, as presented in Chapter 2, between people who are in favour of the introduction of English literacy instruction and those who are against it. Reservations and anxiety have been expressed about teaching English literacy from the point of view of policy makers, who fear it might lead to a delay in Japanese language development. However, researchers such as Butler (2005) provide reassurance that there is no evidence to show that introducing a second language impedes the development of children's mother tongue. Pupils are generally more experienced with literacy in their mother tongue by the time they reach fifth and sixth grade (10-13 years old), which are the grades MEXT stipulates as the start of formal English teaching from 2011 (MEXT 2008).

In relation to the introduction of English literacy, a number of the respondents in the surveys conducted (see Chapter 5 for details) expressed anxiety about teaching literacy. Some reported that 'the pupils asked me how to write the pronunciation of some words or phrases in *katakana* (i.e. a syllabary used for foreign or borrowed words) at the end of the lesson.' This shows that some pupils are curious to know, and also are trying to

make a connection with what they already know (using their existing knowledge of the Japanese language).

Nevertheless, there is certainly a trend in TEYL and research about it to focus on finding methods which further the development of skills in speaking and understanding spoken language (Dlugosz 2000, p.284). As a consequence, reading is often regarded as less relevant in TEYL (Dlugosz, *ibid*) and Japan is no exception (Butler 2005, p.3). MEXT seems to have prohibited the teaching of reading and writing, and there is an almost total absence of written English in the textbook provided (See Appendix 2.1 for a sample of the current 2009 version of the textbook). This may have been part of a compromise in order for English activities to be accepted in the primary curriculum at all, as there has been strong political opposition towards introducing primary English education. Nevertheless, as we have seen from evidence in the surveys with primary teachers in this study (see Chapter 5), concentrating on only ‘sound’, without presenting written words visually, can be problematic for pupils in an EFL setting. Moreover, Dlugosz (2000, p. 284) argues that ‘early students of English, even those who have not yet learnt to read in their native tongue, will benefit from learning to read from the beginning of their contact with a foreign language’. Considering the discussion in this section, one might conclude that teaching without any recourse to written language is misleading and unconventional for the pupils’ learning.

In reading instruction, two ways of processing are often mentioned: bottom-up processing and top-down processing. Bottom up (or data-driven) processing involves prior knowledge of the language system (phonology, grammar, vocabulary) (Kirsch,

2008: 102). Top-down (concept-driven or knowledge-driven) processes involve activation of background information about the topic, an understanding of how discourse is organized with respect to different genres, purposes or topics and a knowledge of the particular context, situation and speaker (Kirsch, 2008, p.103). Teachers have different views which will be reflected in their teaching approach. As Kirsch (2008, p.121) suggests:

Educationalists working in a structural tradition tend to suggest a linear progression from reading sounds, single words and simple phrases to short texts [i.e. bottom-up approach]. ... Teachers who favour a holistic approach tend to opt for a different direction: they start with a text, the meaningful whole, and work 'down' to the language items [i.e. top-down approach]. They work on text comprehension and shift their focus to form wherever appropriate.

Emphasising a holistic approach to teaching, Cameron (2003, p.107) claims that 'children see the foreign language "from the inside".' She stresses the importance of making sense of language through action, interaction, and intention, as opposed to studying system and form which she describes as seeing language 'from the outside'. She further claims that 'even if the syllabus they are taught is structural, children's learning will be communicative, in the most basic sense of "communicative", as being used with meaning and for action'. On the other hand, Cameron (2003, p.107) alerts us to the fact that 'if teaching or materials do not enable children to find meaning in new language, learning will be stultified'.

Whichever stance teachers take, it is necessary for them to be trained to introduce YLs to exercises which focus on form. I argue that this is where the corpus evidence is useful in presenting some focus on form, while maintaining the meaning of the text. In order for children to understand the use of words or phrases at a deeper level (i.e. to be

accustomed to English patterns and rhythms), they need to be given many opportunities for exposure to words or groups of words which are specific to certain situations and contexts. Gregory (1996; 2008) sums up this kind of language as ‘chunks of language’. These ‘chunks of language’ often refer to meaningful segments of sentences, which are thus likely to contribute to the child’s understanding of how the sentence is constructed in order to create meaning. In order for children to be able to read and understand the meaning, they need some understanding of how the sentence is constructed and how to decode texts into meaningful segments. This is important because there is a danger that if the decoding of sentences becomes overemphasised, the comprehension of the overall meaning of the text could be neglected (Riley, 2006 p. 216).

Riley (2006, p.217) emphasises the importance of verse and rhymes (both orally and through reading texts) as representing ‘the exact way the words are placed together’. He goes on to stress the necessity of ‘chunks’ in initiating the important strategy of ‘syntactic prediction cueing systems’, meaning an ability to know whether or not a word fits in a given situation. Suggestions from a corpus linguistics perspective (i.e. a hybrid between the idiom principle and open choice principle, and its implications to reading methodology) have some relevance to what is recommended for TEYL in Japan.

The approach seen in Halliday’s *Breakthrough to Literacy* (1970), seems potentially applicable in primary English literacy teaching in Japan. Riley (2006 pp. 217-218) champions the approach used in *Breakthrough to Literacy*, using ‘look-alike materials where key words are written on strips of card, which then are read, made into sentences, rearranged and read again, and learned.’ He feels that later review of such sentences can

help learners to internalize the language. This approach may also help learners to develop the ‘syntactic predicting cueing system’ (Riley, 2006 p. 217) described above.

The type of pattern practice seen in *Breakthrough to Literacy* (op cit) is especially important for EFL learners at the earlier stages, on account of their relatively limited knowledge of the structure of English, meaning that the prediction of a text is difficult. As Dlugosz (2000, p.288) also suggests, ‘[d]iscovering that sentences are constructed according to different patterns will help children to speed up the development of their listening comprehension and language production skills’.

Not only pattern practice, but other activities should be provided in the long run. Dlugosz (2000, p.288) emphasises the need for maintaining frequent contact with English texts along with recorded materials, which she believes have direct influence on ‘the [pupils]’ retention of words and grammatical constructions in long-term memory’.

Kirsch (2008, p.118) suggests that ‘[r]esearch studies have shown that eight-year-olds tend to have an ability to think analytically’, suggesting that there is a greater likelihood of pupils finding lessons enjoyable if they have ‘specific linguistic outcomes’, rather than merely ones which are interesting or fun (Sauer, 1997 cited in Wolff and Rueschoff 2000, p.112). She further points out that ‘it is completely artificial to separate oral and written language’ implying that it is beneficial for YLs around or above that age to use tasks which require children’s analytical skills in their foreign language learning.

Making use of L1 literacy knowledge is another aspect which can be considered. Riley (2006, p. 216) alerts us to the awareness of the ‘code aspect of written language’ among EFL young learners. He hypothesises that this may be a result of their knowledge of ‘another written language, even if they are unlikely to be bi-literate... when they operate in the early stages of reading in English in school.’ Sassoon (1995) mentions that pupils who have a prior knowledge of L1 literacy in more complex scripts, such as Chinese or Japanese, tend to find the orthography of English relatively straightforward. However this does not necessarily mean that they can understand the meaning. Furthermore, we should be sensitive to the huge demand we may place on children in expecting them to decode English sentences.

One of the strengths of corpus linguistics methods is in identifying ‘recommendable items’ to teach. In other words, they can tell us how ‘frequent’ some items are (Hancioğlu and Eldridge 2007, p.340), which in turn might contribute to deciding ‘what’ items to teach, but not ‘how’ to teach them. In designing a language syllabus, it is crucial to consider methodological and procedural information (White, 1988). These are naturally inter-related issues, which need to be addressed. In light of the teaching of literacy to YLs, the next section discusses how corpus evidence can be employed in teaching children who are learning to read in English as a foreign language. The next section examines pedagogical considerations needed in teaching YLs, including methodology which can maximise the usefulness of PUI (pedagogically useful items), and practical examples of activities are presented.

9.3 Pedagogical considerations and recommendations

9.3.1 Teaching methodology

As pointed out in previous section, it is necessary to discuss teaching methodology with the local context of TEYL in Japan in mind. Willis and Willis (1988) proposed TBL (Task-Based Learning) as the methodology most suited for teaching based on a Lexical Syllabus, and TBL has also been used and recommended as appropriate for young learners (Carless, 2004, Nunan, 1994; Slattery & Willis, 2001). However, the original model for TBL presupposes a certain working knowledge in L2 and the ability to communicate in order to complete the task (Bourke 2006, p.284). Kirsch (2008, p.63) also recognises shortcomings of Task-based instruction, pointing out that ‘beginner learners are likely to find it difficult or even impossible to use the target language’ in negotiating around the tasks. In the case of Japanese primary ELT, if we were to strictly follow the sequence of TBL as suggested by J. Willis (1996), it could be too demanding for pupils, because of the linguistic and cognitive demands the communicative task requires in the target language (i.e. J. Willis’s recommendation is to give focus-on-form instruction after communicative tasks).

Bourke (2006, p.284) suggests that how and what we teach should depend on the ‘learners’ readiness’. If the learners have some English skills, it is possible to conduct communicative tasks with them as the original TBL suggests. However, for those who know little English, ‘some time will need to be spent on building a linguistic platform’ (ibid). This point is relevant in the discussion of whether phraseology should be dismissed as a ‘phrasebook approach’ (see Section 7.3.2.1). It is desirable for learners to

acquire the creative use of language, but it can also be argued that initially building the foundation of the language is necessary in the context of teaching young learners.

In Section 9.2.1, the importance of repetitive expressions, redundancy of texts (as also expressed by Darnton, 2000, Fries, 2008, Hunston, 2007) and their relevance for the development of literacy skills in YLs was discussed. Considering this discussion, the outcome for this approach is therefore to make pupils aware of the lexico-grammatical aspect of language by presenting them with sequences of language on the basis of corpus findings.

In Chapter 2, there was a discussion of TBL in terms of its practicality in an Asian context. As has also been discussed above, the order of task-based approach instruction can be flexible in TEYL, in consideration of learners' 'readiness' (Bourke 2006, p. 284). Since YLs of English as a foreign language need more support, and unless they are presented with some form of models which should be used in the Task Stage, it is almost impossible to expect them to carry out the tasks in the target language without any support. As suggested by J. Willis (2009) in her modified framework of TBLT for beginners (e.g. 'brainstorming session for words and phrases'), I suggest adding more support in a 'pre-task'. This can be done through introducing the target language to be used in the task stage as a means of linking new language with what pupils already know (e.g. language or concepts) for the context of TEYL in Japan (and the same might be applied in other EFL countries).

In the previous section, it was suggested that the approach seen in *Breakthrough to Literacy* can be used in YL teaching in Japan. The main concept of *Breakthrough to Literacy* is related to teaching situations where pupils practise producing sentences by rearranging cards printed with phrases and words (see Appendix 9.1). At a very primary stage, the techniques of this approach can also be used at the ‘pre-task’ stage (i.e. presenting language in order to prepare pupils for the task to be followed).

It is also important to remember to include aspects of recycling in the task-cycle. For instance, if teachers are asking children to retell stories, they could produce a set of cards which contain phrases (or sequences) which are of high-frequency in the corpus, so that the language is recycled, and at the same time, it introduces the new set of phrases or expressions.

Nevertheless, it is important to remember that ‘[e]ach teacher needs to make informed decisions based on the learners’ needs about what methods work best with particular pupils in a particular context at a particular moment’ (Kirsch, 2008, p.48). In addition, Kirsch (2008, p. 121) stresses that it is necessary for teachers to be able to decide where they incorporate the focus on form during their lessons. Thus, for instance, even if teachers are conducting the lesson using a holistic approach, some focus on form might be used where appropriate.

In addition, it should be noted that the suggested activities have a focus primarily on literacy using the findings from CCL. However, it is important to incorporate this aspect

with speaking and listening activities, so that the tasks are not too demanding for pupils and incorporate the main speaking activities which pupils are accustomed to.

9.3.2 Practical examples

In the survey conducted with current primary teachers in Japan as a part of this study (see Chapter 5), some respondents expressed anxiety about teaching English literacy. It is important to consider the needs from the local context when designing materials (Harwood, 2010 p. 15). Thus, some examples of how the findings from the CCL can be used in the TEYL classroom are presented in this section. It should be stressed that, as discussed in Section 9.3.1, the evidence from the corpus is particularly of use to teachers or designers when they devise a focus on form within their lesson. Needless to say, any task in pedagogy should be planned with consideration to pupils' readiness (Bourke, 2006). Each section is divided for the purpose of presentation, however, presented examples may contain some overlapping aspects of language.

9.3.2.1 PUI 1 – contractions

In Section 8.2.1, it was discussed that '*I'm*' is often followed by adjectives or adjectival phrases, and noun phrases or present progressives. The pedagogical implication of the finding from the corpus analysis is useful in choosing items which follow the contraction '*I'm*'. For instance, *I'm* followed by adjectives (Concordance 9.1):

Concordance 9.1 Some examples of *I'm* + *ADJ*

1	h dear," he cried, leaping out of bed. "I'm late !" And he rushed into his bat
2	ful and I can look at the robe whenever I'm lonely ." Holding the robe very caref
3	once. Earthworm is right!' 'Of course I'm right ,' the Earthworm said. 'And we
4	The doggy goes to sleep all by himself. I'm sleepy . Good night. Meow! The kitty
5	voice, 'The fact of the matter is that I'm shy , so if I don't do this by comma
6	here seems to be no end to it. "Oh, but I'm tired !" the big bird said and starte

In the pre-task stage, for example, the presentation of a sequence '*I'm hungry*', can be divided into two cards: '*I'm*' and '*hungry*'. Different variations can be brainstormed by interacting with students. Retaining the meaning is one of the key principles in teaching children (Cameron, 2003). This brainstorming stage, in which pupils are actively engaged by demonstrating what they know, helps pupils focus on the language while meaning is also retained. Moreover, from the viewpoint of pupils' readiness, using '*I'm*' is a good starting point, as pupils are used to using it in spoken form (e.g. in response to the phrase 'How are you?' at the beginning of lessons).

Using the contracted form and adjective, both the contracted subject '*I'm*' the adjective part '*hungry*' can be swapped over with other variations of contractions (*She's*, *He's*, *They're*) and adjectives of state (e.g. *hungry*, *sleepy*, *tired*) in order to clarify and give further examples of the structures. This method of presentation is particularly useful when presenting a sequence where some words in the sequence allow variations.

It is often pointed out that literacy makes demands on pupils, especially at the beginning. In preparation for writing, pupils can be encouraged to express meaning with pictures and word or phrases cards.

Using the findings from the corpus, teachers can prepare a set of pictures in advance, which corresponds with the different sets of lexical items and combinations. The task is to find a set of word cards which best describes the picture (e.g. *he's a magician*. *he's walking*). If pupils are able, this task could also be extended to make it a little more challenging. For instance, teachers could prepare some irrelevant cards which do not fit in the groups described above (e.g. *he's finished...* *he's invented a way of making*

chocolate) asking pupils to work out why some cards are not included in a certain group.

This stage of the task should be carefully introduced so that the pupils do not get discouraged.

The post-task can include a focus on form activity, in which they recycle the first set of focus in the task (i.e. in this case, what follows after *I'm*, identified adjectives and nouns/phrases). For example, as can be seen in the following concordances (Concordance 9.2), the usage of *very/too/so* which appear before adjectives to put an emphasis on what is described.

Concordance 9.2 *I'm + very/so/too + ADJ*

very	re. "Please help me," he said to them. "I'm very, very hungry." "Oh! what a poor
very	, if it's not inconvenient, perhaps, as I'm very thirsty, provide me with, do y
very	. 'We're in the Chocolate Factory!' 'I'm very glad to hear it,' said Mrs Buck
so	shaking Mr Bucket warmly by the hand, 'I'm so very glad to meet you. You mustn'
so	with tears streaming down his face. "Oh I'm so unhappy," he keeps sobbing, over
so	e newborn baby. "Ahhh ..." yawned Ken. "I'm so tired of sitting around here and
so	"To heck with the vases!" she shouted. "I'm so thrilled I don't care if I break
so	ing to pull it off again. 'Do you think I'm so stupid I'd glue this thing to my
so	s beard. 'I didn't mean it. Forgive me. I'm so sorry. Good-bye, Mrs Gloop! And
so	question. 'Oh, what will happen? Help! I'm so nervous! What if I can't behave
so	to be happy," said Mr Miserable. "But I'm so miserable I don't think I could e
so	between two tall trees! "Oh. Mr Brave, I'm so lonely," she sobbed. "Nobody wil
so	ack into the bushes to hide. "I'm so hungry now," he said t
so	g. "I thought I'd nearer see you again! I'm so happy you got away!"
so	etty you look in that lovely mink coat! I'm so glad you could come! Dear me, thi
so	ng's papers! Just in time, my dear boy! I'm so glad! So happy for you! And this?
so	ng almost exactly like you used to! Oh, I'm so glad!' 'Just as long as it all
so	they managed to de-juice her after all. I'm so glad. And how healthy she looks!
so	do beg your pardon,' said Muggle-Wump. 'I'm so excited I hardly know what I'm sa
too	good at jobs," he confessed, "because I'm too quiet." "Ah," smiled Mr Happy.
too	'You finish it,' said Grandpa Joe. 'I'm too nervous.' 'No, Grandpa.
too	aid to me. "No, thank you," I said. "I'm too excited to eat. And I've got to

As this type of task tends to be more challenging for pupils, it is important for YL teachers to make sure that they have completed the groundwork necessary (e.g. by grouping the concordance examples in advance) for pupils to be able to carry out the task which leads to developing language awareness.

9.3.2.2 PUI 2 - semantic grouping

Let us consider the potential of the findings further using some examples from my analysis. For instance, in Chapter 8, it was observed that in some colligational structures, the lexical items slotted into the Xs share some semantic meaning (e.g. *X and X and X*). It was suggested that this aspect was useful from a pedagogical point of view. It was suggested that, for instance, if learners cannot understand one of the words in this sequence, given the fact that it is likely that the meanings are from the same semantic domain, it becomes easier for the learners to infer its meaning from the other two words. The following activity involves grouping words together according to the patterns and phraseological information obtained from the corpus analysis. For instance, as discussed in Chapter 7, in the PUI example of *X and X and X*, the nouns are related semantically.

Figure 9.1 Task sequence: PUI-2 semantic grouping

[Pre-task]

In the pre-task, pupils listen to a story (either read by the teacher or from a CD), and they are encouraged to identify what kind of items or characters which/who appear in the story. (supplement with pictures where necessary)

[Task]

Teachers prepare cards for items/characters appearing in the story and the conjunction 'and' card. Pupils group items (Xs) according to their semantic similarities, and produce sequences using vocabulary cards. The pupils then categorise items listed in the pre-task stage, and link those items using 'and'. (Supplement with pictures where necessary)

[Planning/Presenting]

Pupils present the sequences to the class, or within small groups, after which the teacher gives feedback to the whole class. Some pupils may create mixed categories, not by prototypical categorisation. However, it is also possible that they might be trying out their creativity in producing the sequence.

[Post-task]

The story can be read to them again exemplifying the pattern. The teacher may involve pupils in a storytelling activity by setting a rule that pupils need to raise their card when they hear the items they used in their sequence. This raises pupils' attention to the story as a listening activity, as well as making them participate in the storytelling activity.

By demonstrating the sequences on the board at the Presentation Stage, this procedure will enable pupils to use the techniques and try them out by themselves when they are ready to compose written sentences. In addition, by using the semantically related items for Xs, the task can be extended asking pupils to group items of Xs according to their word families.

In relation to verbal patterns, an awareness-raising activity which encourages pupils to make a connection between meaning and patterns can also be suggested. For instance, GO also occurs with home (occurring L2 position). The following table shows the number of instances which occur with the ‘verb + *home*’, ‘verb + *back*’.

Table 9.1 Frequency of Verbs (GO/COME/RUN/WALK) + *home/back* in CCL

	GO	COME	RUN	WALK	<i>total</i>
+ <i>home</i>	84	19	9	8	120
+ <i>back</i>	40	58	16	6	120

Based on this evidence, it is possible to present learners with groups of words which share patterns. For example, the verbs GO/ COME/ RUN/ WALK can be grouped to illustrate ‘moves/direction’ followed by words which indicate ‘place or direction’ such as *back*, *home*. It is not necessary to show the numbers to pupils such as above, but it is possible to highlight the verbal patterning and shared meaning among those verbs.

Moreover, by selecting sequences which are relevant to pupils’ life (‘GO to school’ (*school* occurs as one of the collocates in R2 position) and ‘GO home’ ‘GO back to + place’), it is possible to expose them to the basic words that they usually encounter with lexico-grammatical features (e.g. Eigo Note 2, Lesson 7 includes a schedule of a typical

pupil's day). For instance, the verb GO has *to* as one of the frequent collocates in R1 position. These verbs can be followed by prepositions such as, for example, *go to*, *go back to school*. The concordance results confirm that *go to* + place is a common construction, but *go home* always occur together without a preposition in between (*go to home**). The same applies to verbs such as RUN, COME and WALK. These can be highlighted for learners by the use of simplified concordance results. The concordance lines can be prepared so that learners can easily recognise the patterns. In a pre-task, teachers may instruct pupils to find words (which are key to the patterns) and colour the instances (e.g. *go* in red, *to* in green, *home* in blue). As reported by Sealey and Thompson (2004 and 2007), presentation with coloured concordances is useful in the direct use of concordance lines in the L1 primary literacy classroom.

Different activities are also possible. Considering the nature of this grouping, an activity such as 'Karuta' can be employed. *Karuta* is a Japanese traditional card game, which was originally used to challenge the memory by recalling the lines of traditional *waka* (Japanese) poetry (百人一首 *hyakunin isshu*). The nature of the card game itself is familiar to Japanese learners, who are taught it as a means of learning traditional *waka* poetry.

The activity can be used with the aim of encouraging YLs to recognise patterns and meanings. For instance, the teacher reads out each verb card 'run/come/walk/go' and YLs find the card bearing a word or words to follow the verb, in order to complete a pattern. This may be supplemented with a work-sheet which has a simplified version of concordance results, so that pupils have support (access to the language usage in written

form). The fastest pupil to touch the correct card keeps it, and they continue this procedure until there are no cards left on the table or the floor. The game can be repeated after shuffling the cards. Once the simple combinations is remembered by pupils, teachers might ask them to add some words they know which can follow. The advantage of this approach is that teachers can test pupils' understanding without them realising they are being evaluated. Throughout the game, teachers are able to observe the pupils' needs and identify possible support and extension for subsequent lessons.

After pupils get used to this approach, it can be done in small groups. Especially in a Japanese primary school, it is common to have a large number of pupils (usually 30 – 40 pupils) in class, so the teacher may set it as group-work, in which pupils select two cards in turn, one from the first set of cards (containing verbs) and one from the other (containing prepositions and nouns/noun phrases which frequently appear in the corpus). Pupils are in charge of checking whether the pattern presented by their peers is possible, using the simplified (paper-based) concordance list supplied to them for checking. Giving responsibility to pupils in this way may also encourage them to learn autonomously. Likewise, the teacher can monitor and check the understanding of pupils at this stage. Once the patterns or combinations are created by each group, teachers might ask some pupils to come to the front and demonstrate the matching of the cards on the board. As a post-task, categorisation of the lexical items can be done so that pupils think about the form and meaning one more time.

9.3.2.3 PUI 3 – phrases containing preposition and articles

As discussed, the use of definite and indefinite articles is a challenge for EFL learners, even those at advanced level. This is especially the case for learners whose L1, like Japanese, does not have articles. As reported in Chapter 7 and 8, the analysis of the corpus of children's literature revealed, as is the case with other corpora, that the indefinite/definite articles often co-occur with prepositions in sequences (i.e. identified as PUI-3 in Chapter 8). From this finding, it was suggested that the use of prepositions and articles can be introduced in a sequence rather than as separate entities of grammatical items.

For instance, there are a number of instances of *a X1 of X2* structure found in both EOCL and JTCL (709 instances, 37 variations in X1). Moreover, *the X1 of X2* is also very frequent, (1745 instances, having 73 variations of X1). It is not necessary to teach all of the instances to pupils, but teachers may select words which are common and useful in practice, such as the following: *a bag of, a packet of, a pair of*

Awareness-raising by means of focus on form can be provided as well, by using some examples from the corpus. For instance, the lexis that goes into X1 can influence the form of X2, as the following concordances exemplify (e.g. *a pair of X2, a loaf of X2*).

Concordance 9.3 *a pair of X2*

1 decided he'd better go home. He passed a pair of worms who were having a chat.
2 ed jaws started chewing away on it like a pair of tongs. 'Don't!' said M
3 ly, her old wrinkled lips shut tight as a pair of tongs and the hand that held t
4 t to his medicine cabinet and took out a pair of tiny red boots. -This should
5 outs, but it was still a shock to meet a pair of them in the flesh. 'Matilda's
6 you like to know what the present was? A pair of socks. Well. Not quite a PAIR
7 he lady behind the counter. "You mean a pair of socks," she smiled, and showe
8 t was? A pair of socks. Well. Not quite a PAIR of socks. One sock was white! And
9 inish up as a bundle of old clothes and a pair of shoes?' Mr Twit put on a ve
10 in the end there's nothing left except a pair of shoes and a bundle of old clot
11 it, however, by a little dwarf who had a pair of seven-league boots, which are
12 She became so angry that she grabbed up a pair of scissors and cut the sparrow's
13 ouldn't slip off. So Mrs Wormwood took a pair of scissors and cut the thing of
14 othing but a small suitcase containing a pair of pyjamas and a toothbrush, was
15 ncing upon a girl of about ten who had a pair of plaited golden pigtailed hangin
16 a sweater with holes, some torn shirts, a pair of pants, a curtain and some old
17 an inch from my nose. I noticed also a pair of little furry front paws restin
18 the Ladybird, 'our Centipede, who has a pair of jaws as sharp as razors, is u
19 r to the door of my grandmother's room. A pair of her shoes was standing outside
20 et with jewels. This done, she gave her a pair of glass slippers, the prettiest
21 nose-holes were waving in and out like a pair of fish-tails. "It's getting stro
22 !" One boy grabbed up a broom; another, a pair of fire tongs; and the third, a d
23 aw, to their astonishment, nothing but a pair of feet resting on the pillow. A
24 from out of the smock were encased in a pair of extra-ordinary breeches, bot
25 r. And where his arms had been, he had a pair of duck's wings instead! 'But..

Concordance 9.4 *a loaf of X2*

1 ho might come along.â€ So the fox left a loaf of freshly baked bread in the bas
2 he birds live in holes! And if you want a loaf of bread you don't go to the bake
3 . And then he said, quietly, "I'd like a loaf of bread, please, Mrs Crumb." M
4 mb, pretending not to hear. "I'D LIKE A LOAF OF BREAD!!" Mr Noisy shouted. "S
5 went into Mrs Crumb's shop. "I'D LIKE A LOAF OF BREAD," he boomed. -Sorry! Wh
6 went into Mrs Crumb's shop. "I'D LIKE A LOAF OF BREAD," he boomed. "A what?"
7 : A bacon rind, some rancid lard, A loaf of bread gone stale and hard,

In 'Eigo Note 1' (Lesson 6) includes a sentence 'What do you want?' as a key sentence in the context of ordering meals at restaurants. In a similar manner, the examples above can be used in the following task sequence.

Figure 9.2 Task sequence (*I'd like a X1 of X2, please.*)

[Pre-task]

Presentation: sequences (*a X1 of X2*) are presented using pictures and word cards.

Pair-work activity:

The teacher presents a worksheet with a set of pictures for each pair, and pupils need to find the matching cards for the picture, and prepare a shopping list. (*a pair of _____, a loaf of _____*)

[Task]

Shopping game (*I'd like a pair of socks, please. I'd like a loaf of bread, please.*)

[Planning/Presenting]

In pairs, pupils prepare items they got in the shopping task, and prepare for presentation.

[Post-task]

Present what they have prepared to the class, after which the teacher gives feedback to the whole class.

It is necessary, as far as possible, for the lexical items used in this task to be ones which have already been introduced to the pupils, as the task is to use the extended sequence (in this case, *a X1 of X2*). As a part of the task, pupils may be asked to draw pictures, illustrating the sequence in order to prepare for the task (i.e. creating a shopping list).

As the main goal for this activity is to make pupils aware of the lexico-grammatical aspect of language, it is best to avoid increasing the risk of discouraging learners by doubling the demands on them by introducing too many items at once. As mentioned before, it would be easier for YLs if they know the content words.

The presentation stage is also important as teachers can measure the pupils' understanding of the connection between oral and written language. If pupils experience any difficulty with a particular item, this item should be remembered and recycled.

As an extended activity, or for more experienced learners, outside the context of shopping, the following sequence can be presented, to show that the sequence is not restricted to one noun (or object) but can be combined with the use of ‘*and*’.

Concordance 9.5 Colligational framework - *of X1 and X2*

than anything else she took a slice of bread and margarine and started to "and nibble through all the packets of raisins and cornflakes and chocolate Through they were made out of a mixture of cotton-wool and candy floss and thin at Mr Grumpy lying a mid a jumble of sausages and cakes and newspapers the greengrocer with his window full of apples and oranges , and the butcher
--

9.3.2.4 PUI 4 - repeated sequence

In Chapter 7 (Section 7.2.3.3), the following concordance (Concordance 7.4) was presented as a result of the colligational framework. As mentioned, the sequences were identified from a single text. Children’s storybooks often contain repetitive sequences, which teachers reported as being useful in preparing classes (see Survey 1). Making use of repetitive features of the texts, the following task can be suggested.

Figure 9.3 Task sequence: PUI-4 repeated sequences

[Pre-task]

Listen and sort activity

A story which contains repeated sequences is read to the pupils. Teacher initiates questions about characters (e.g. animals appearing in the text) to pupils.

[Task]

Pair activity - Running dictation

[Planning/ Presenting]

In pairs, pupils fill the every gap. Pupils then read out the story which they have completed

[Post-task]

Storytelling again but pupils participate in the reading where possible and contrast the outcome they have heard with what they have completed.

As an introduction to the task, it is possible to conduct a pre-task of a 'listen and sort' activity in which pupils listen to what the teacher reads out and make sequences by rearranging the cards as they talk about the characters as they appear in the text, stressing the key phrases.

A further task could involve techniques used in the running dictation. In preparation for a running dictation activity, a story is cut into parts (or some parts of the story are missing in a text), and stuck randomly around the room. This is a pair activity and a worksheet is given to each pair (Student A and B) with missing information (e.g. an incomplete story). The task for pupils is to work out the sequence of the story by finding the appropriate order of the missing sequence/items. Student A firstly goes around the room and finds a slip, remembers it, and goes back to his/her partner Student B. Student A dictates to Student B, and Student B writes down what he/she hears. They take turns until they complete the story on their worksheet.

As for post-task, the role of the teacher can be decreased, depending on their participation and levels, and they can take turns to read out the sequence when they are used to it. Once the sequences being introduced to the pupils have become automatic, the teacher might want to read out the story one more time, or read with pupils as a post task (e.g. the book '*Rosie's Walk*' (Pat Hutchens, 1987) is full of such sentences).

For more experienced learners, the Task Stage can be related to writing, e.g. composing a sentence with the cards they have learnt. This is where the original approach used in

Breakthrough to Literacy is particularly useful for pupils to try out combinations of words using the boards.

9.3.2.5 Making use of L1 knowledge

In relation to making use of YLs' L1 literacy knowledge, it is necessary to make sure that this prior knowledge is distinguished carefully from what is introduced. For instance, Japanese has two types of phonological scripts (*hiragana* and *katakana*) and Roman script is also used in some contexts. This way of writing involves a direct one-to-one correspondence of spoken form and written form in Japanese (combination of consonants and vowels). It is often reported that this Romanised way of writing Japanese (called *romaji*) tends to create confusion when it comes to the phonological aspects of English and thus may be one of the causes of particular problems with pronunciation and spelling by many Japanese learners (Okada, 2005). As has been pointed out in Chapter 5 and by teachers and prospective teachers, it is necessary to present learners the written language, rather than restricting them, so that they can be aware of the difference between Roman script in Japanese and English writing, even though they use the same alphabet.

As discussed in Chapter 6, the spellings of proper nouns are different in Japanese original texts and the translated versions. In this respect, if teachers have the Japanese original texts, it would be interesting to compare the original texts with their English translated versions as a teaching activity. Moreover, an activity based on spelling can be proposed, using proper nouns. For example, the teacher prepares cards which have different spellings of the names in English (e.g. Ken-Kon, Ellie-Eri, Amy-Aki) and have

pupils firstly identify using a game of ‘I spy’ (Wyse, 2007, p.73-74) and ask them to classify them into a Japanese name group and an English name group or both. When pupils are confident about the spelling, a modified activity such as asking pupils to make up names using alphabet cards or panels can be conducted. This might contribute to raising awareness among YLs of the spelling differences between English and *Romaji* which they learn at primary level. Therefore highlighting the differences might contribute to raising awareness of it and may help children to realise that *Romaji* is not English despite its superficial similarity.

9.3.2.6 Pedagogical implication of cultural keywords

In Chapter 6, it was suggested that cultural keywords have some implication for reading instruction (see Chapter 6, Section 6.3). When children are learning to read in English, the following awareness and knowledge need to be developed: ‘awareness and knowledge about print’, ‘graph-phonetic knowledge’, ‘lexical knowledge’, ‘syntactic knowledge’ and ‘semantic knowledge’ (Brewster et al, 2002, p.110). In this section, some activities are suggested using the cultural keywords identified in Chapter 6. The activities suggested below could contribute to build pupils’ awareness of different types of knowledge and make links between them.

In order to understand the story, it is also important to be able to relate one event with another stressing the ‘meaning’ before focusing on decoding written words (Brewster et al, 2002, p.112). Taking the ‘*monkey*’ instances as an example, the following four stages of an activity can be suggested.

Figure 9.4 Example of an activity using keywords

Stage 1	Teacher reads out the story, showing the key characters, items and places using pictures
Stage 2	Teacher asks about the monkey character: (e.g. what he does at each stage of story, what happens to him, and what kind of person he is, by asking pupils questions related to the content.)
Stage 3	Teacher would know what verbs to highlight in the story (for instance ‘ <i>saw</i> ’, ‘ <i>watch</i> ’). Pupils learn the connections (cause and effect) between what the monkey did and what happened to him.
Stage 4:	After a connection is established in the above stage, and the children’s understanding has been checked, the teacher asks children to act out the stories (employing a drama method). This would make the activity more interactive and meaning can be consolidated with the lexical items or phrases they learn with the story.

A similar activity can be planned using the stories which include instances of ‘*rice*’. As mentioned, ‘*rice*’ is used to make mainly *mochi* in the case of JTCL. Using something related to learners’ culture and life experience is very important as it helps them understand the content easily, enabling them to focus on language. The associations between the rice-cakes (*mochi*) and ‘*rice*’ can be highlighted using a sentence from the text (e.g. *And we don't have any rice at all. So we won't be able to make rice-cakes.*). After establishing the link between the rice-cakes and the ingredient ‘*rice*’, teachers can also point out that other rice recipes are available in other countries (e.g. rice pudding made from ‘*rice*’ as found in EOCL) to raise awareness of other food cultures and make links between syntactic and semantic knowledge.

Instances of ‘*says*’, as has been observed in the analyses are mostly within speech marks. Taking this into account, a ‘speech bubbles activity’ (Brewster et al, 2002, p.114) might be useful where pupils try matching the dialogues with characters by making

connections between the story, what they read, and what they hear. For more able students, practising the dialogues by means of a role play can also be useful as it can integrate other skills such as listening and speaking.

As Cameron (2001, p. 95) points out, making pupils 'notice links' with what they already know contributes to children's learning of new vocabulary items, and this is necessary for developing their conceptual knowledge. She also suggests that lexis is best learnt by means of interaction with discourse, rather than as separate items. This is also true in reading, as it is necessary for children to understand the relationships between lexical items such as characters and places, and how they are represented in the texts as well as how the information relates to the main theme.

Nesi (1995) studied animal metaphors and suggested that their use may 'facilitate cross-cultural communication' in the EFL classroom (Nesi, 1995 p. 276-277). As was found in the instance of similes in EOCL and JTCL, a task can be introduced to describe certain animals or their peers using similes found in texts.

Overall, we can conclude that this comparative analysis of EOCL and JTCL also served as a good starting point for providing some insights into the pedagogical considerations necessary for teachers when designing a syllabus, materials and activities suitable for TEYL.

9.3.2.7 Summary

Up to now, the possibility of practical applications of PUIs has been discussed, specifically in terms of how teachers can use the identified PUI in classroom. I have explained the activities in the framework mainly of task-based learning sequences, but adding more support at pre-task stage for L2 young learners.

The approach taken by *Breakthrough to Literacy* is designed to integrate receptive (reading) skills and productive (writing) skills (Coles and Hall, 1998 p, 4). In addition, it was suggested that teachers should be ready to adopt a different sequence according to the pupils' level of 'readiness'. It is also necessary for a teacher to create a situation where reading or writing is purposeful in teaching foreign language literacy (Griva, Tsakiridou & Nihoritou, 2009, p. 147), and I would argue that the suggested examples of task sequence are far from a so-called 'drill-based approach' which is often held in contempt for TEYL classrooms (Bourke 2006, p.286). Kirsch (2008, p.118) suggests that children around the age of eight or above can deal with activities which call upon their analytical skills, and it contributes to their motivation if they recognise their achievements of 'specific linguistic outcomes'. As children tend to be distracted easily if the practice and activities are repetitive and boring, it is necessary for teachers to find ways to keep them motivated. It is also suggested to use activity games (from the YLs' point of view) at the Task Stage while teachers monitor their language. These activities can also be incorporated within the existing lesson for key sequences or phrases of a

story to do some work on focus on form, even when teachers conduct their classes using a holistic approach¹³.

The points discussed here might be useful for Japanese primary teachers who have received little training in teaching English, especially with English literacy. In relation to teacher training, the next section discusses further implications of the corpus-influenced syllabus for YL reading instruction. In particular, I discuss how the introduction of a corpus-influenced syllabus may affect primary English teacher training in Japan, and may contribute to raise teachers' own awareness of English language in general.

9.4 Implications of corpus-based research for TEYL teacher education

So far, the relevance of phraseological findings, and of the importance of phraseology in teaching literacy, has been discussed in relation to the lexico-grammatical features found to be prominent in the analysis of the corpus (see Chapters 7 and 8). In relation to EFL language syllabuses, Mindt (1996, p. 294) states that 'most language-teaching syllabuses are structural syllabus... [and the] use of grammatical syllabus has a long tradition'. In Japanese schools also, the grammatical syllabus has a long tradition, and is still found in most textbooks, where decisions about the sequence of the courses are based on carefully graded grammatical information. The influence of this tradition can be seen in the 'Teaching English to Young Learners' course for prospective teachers of

¹³ Kirsch (2008, p 132) explains 'a holistic approach' as follows: 'a holistic approach considers language as an entity rather than an assemblage of rules and patterns, integrates work on the four language skills (listening, speaking, reading, writing) and encourages interaction with text and language, collaboration with other learners and communication for a given purpose within a meaningful context.'

YL, which I taught at my own workplace. For instance, when my students had a task which involved planning a lesson for TEYL, most students concentrated on choosing language which they could use with the target structure but with little regard for appropriacy when designing a lesson for young learners (e.g. using ‘Who are you?’ in a conversation asking about jobs). Students were around the age of 21 in 2009, and had received their English language education via the so called ‘grammar translation method’ when they were in junior-high and high schools. This suggests that prospective teachers of YLs need new ways of looking at the English language.

Nevertheless as reviewed in Chapter 2, there is a movement towards not explicitly teaching grammar at all and it was pointed out that this movement may have caused the misleading oversimplification of sample English sentences. However, this oversimplification can cause confusion when pupils encounter real usage in written form and in their later learning. Because of this concern, and also because of the tension over current educational policy in teaching YLs at state schools in Japan, I have argued that combining a lexical syllabus (Sinclair and Renouf, 1988; Willis, 1991) with some lexico-grammatical information may help YLs in beginning to read (see Chapter 2).

I mentioned that young learners do not need to know the theory behind the phenomenon. However, in the case of preparing prospective teachers to teach young learners to read, it is necessary for them to be aware of the importance of phraseological phenomena. This seems to be one of the direct implications of this research and one which is useful for future teacher education in TEYL. For instance, teachers using a holistic approach to literacy, often employ ‘story-telling activities’ (Kirsch, 2008; Slattery and Willis, 2001).

It is important for teachers to recognise the units of meaning in each sentence, and demonstrate this when reading the stories out to their YLs, instead of reading the story word by word.

More importantly, as J. Moon (2008, p.400) highlights, there is ‘a general lack of systematic and informed focus on early L2 literacy in most YL coursebooks’. She points out that it is difficult to deal with L2 literacy in globally available textbooks, as there are differences in each teaching context, such as the aim of curriculum, levels and the writing systems in learners’ L1. This points towards the importance for primary teachers to be equipped to produce materials locally for their own pupils.

In Chapter 3, I explained that DDL (data-driven learning) requires a lot of inductive tasks on the part of learners, and therefore it is perhaps too demanding for children, if the DDL approach were to be directly applied in TEYL in Japan. However, as has been argued by Moudraia (2001, p.2), it is the ‘teachers’ mind-set’ which needs to be changed in order to implement a lexical approach. Furthermore, there is a possibility of using corpora in the field of teacher training, as well as for syllabus design, as has been discussed up to now in this chapter. For instance, prospective teachers can be encouraged to appreciate the value of corpora, through DDL. The concordance lines visually represent patterns, which are useful in raising awareness of the phraseological features of English.

For instance, in Chapter 5, it was reported that teachers were concerned about the usage of plural forms of nouns appearing in the construction of ‘Do you like + N?’ (e.g. children or some teachers say ‘Do you like orange?, instead of ‘Do you like oranges?’).

Indeed, the following concordance results from CCL suggests that there are uncountable nouns and countable nouns following the sequence ‘*do you like*’. More interestingly, the nouns often refer to a particular object with the use of a possessive pronoun (e.g. *my*, *your*) or demonstrative pronoun (e.g. *this*).

Concordance 9.6 *do you like* + *X*

1	ded down three magnificent hams. And do you like bacon, Badger?
2	ay to do it properly! The only way! And do you like my trees?' he cried, pointin
3	something different and delicious! And do you like my meadows? Do you like my g
4	ace she loved to read most was in bed. Do you like reading in bed? Lots of peop
5	ge. "Hello, Postman," he said. "I say, do you like my new hat, have you ever s
6	delicious! And do you like my meadows? Do you like my grass and my buttercups?
7	But, porridge on toast! Really! "Why do you like porridge on toast?" Mr Nosey
8	feels so good!" said Amy. "So, Ken how do you like your first bath?" asked Gran
9	longer one for the Earthworm. 'And how do you like your bed?' she said to Jame
10	his head. "Now!" roared the bear, "how do you like this. If you don't give up,

As was also mentioned by Cameron (2003) as an example of illustrating how children try to understand the meaning, the construction of ‘do you like ~?’ is one of the most common structures used in the TEYL classroom. These concordance results suggest that learners may encounter different usages of ‘do you like + N’ in reading. As suggested in Chapter 8, section 8.2.2, oversimplification of what follows (in the case of above examples, noun phrase) in presentation to learners may cause confusion such as the one reported by Cameron (2003).

The corpus evidence suggests that even with an apparently simple construction such as ‘do you like + N?’, the usage is not as straightforward as the teacher (as above, in Chapter 5) had assumed, but instead includes different usages, though it is still limited.

Nevertheless, as Sealey and Thompson (2006, p.27) suggest, corpus evidence can serve as a resource for teachers in deciding the language to be presented to learners.

Furthermore, using corpora directly in the educational programme for prospective teachers has further benefits. Not only can the notion of lexico-grammatical aspects of language be introduced and serve as an awareness-raising device for YL teachers, but also teachers have external access to a range of expressions through the use of corpora, which cannot be covered by reading paper-based textbooks. Knowing how to access and make use of corpora can certainly broaden their horizons, and this will lead to developing their qualities as language teachers. If teachers know how to make use of corpora, their awareness of language use will increase, which in turn can promote teaching and materials which are not limited to what they come across in textbooks. Such a quality among teachers is necessary for the success of primary English education.

In September 2009, the new DPJ (The Democratic Party of Japan) government¹⁴ announced a plan to change the whole teacher education system (currently 4 years of university education) to a 6-year system in the near future. In this plan of the fundamental review of the teacher education system, there is no mention of primary ELT teachers. However, if 6-year teacher education is to become a reality, then it is possible to offer more highly specialised education to prospective teachers. This is a source of hope for future education in Japan. In the meantime, for teachers currently

¹⁴ <http://www.dpj.or.jp/special/manifesto2009/index.html>

working at primary school, it may be necessary to make use of the licence renewal workshops¹⁵ to supplement training sessions specially designed for TEYL.

9.5 Limitations

One obvious limitation of this study is the degree of generalisation of the results obtained from the corpus. If we were to use only the corpus of children's literature, it might not be possible to generalise the results. As there exist limitations with the nature of the corpus being a collection of written texts, it may be difficult to find suitable examples in terms of spoken English, such as those found in the YL teaching practice session. In this session, students were preparing for their teaching practice, and the interaction below was suggested by one group of my students. The aim of the class was to introduce lexical items relating to 'occupation'. The following is the conversation model which my students suggested in the first draft of a lesson plan.

S1: I'm a doctor. Who are you? *

S2: I'm a photographer.

The question 'Who are you?' is not grammatically incorrect, but as any competent speaker of English knows, using the question form 'Who are you?' in order to ask people's occupation is unacceptable. It can be explained that these students prioritised the use of simple structures over the naturalness of the question, not only ignoring the most typical form of asking someone's occupation, but also failing to recognise the different meanings attached to 'who' and 'what'.

¹⁵ MEXT requires a compulsory workshop for teacher licence renewal for teachers who have been in service more than 10 years since their acquisition of teaching licence. For details see http://www.mext.go.jp/a_menu/shotou/koushin/

As stated above, a limitation of the corpus used in the study is the impossibility of finding every example which teachers or prospective teachers would like to look for. Moreover, while the use of a concordance helps analyse instances in a vertical manner to identify patterns, it cannot tell us how to teach them. As some of the following instances from CCL (9.1-9.5) exemplify, it is possible to explain the usage of (9.1), (9.2) and (9.4) as a ‘spatial scene’ (Tylor and Evans, 2003 p. 27), in the case of (9.3) and (9.5), it may be easier to treat them as a chunk of language. These decisions have to be made by teachers when extracting examples from any corpus.

(9.1) I’m going on a lovely white horse.

(9.2) I’m going on a lion.

(9.3) I’m going on a picnic.

(9.4) I’m going into town.

(9.5) I’m going for a walk.

Nevertheless, with these limitations in mind, it is still justifiable to claim that using a corpus of children’s literature in teacher training is a useful tool. This is especially the case in making teachers aware of the lexico-grammatical features of English in the texts which have been chosen for Japanese YLs for pedagogical application.

As MEXT suggests, the aim of ELT in Japan is to foster English ability which can be used in the real world. Taking this aim into account, if we are to train teachers in ‘communicative competence’, it is important for them to know what is natural and what is not. Not only the corpus of the current study, but also other available corpora could be used for teacher training.

9.6 Conclusion

This chapter has discussed corpus-influenced syllabus design and the direct implications of features identified in the corpus, mainly in relation to their relevance to YL literacy syllabus design. In section 9.2 it was suggested that the notion of pattern grammar (Hunston and Francis, 2000) can be a bridge between lexis and grammar, and can be effectively incorporated into a syllabus to be designed for TEYL in Japan. Based on the suggestions made by Fries (2008, p.80), that language is ‘pervasively redundant’, and furthermore ‘readers use these redundancies as they make sense of the text they are reading’ as well as those by other researchers (Darnton 2000, Hunston 2007), it was argued that a substantial input of lexico-grammatical context has to be given to learners at an early stage. Thus it was argued that incorporating this view of language into teaching literacy is advantageous for young learners.

Nevertheless, one might argue that some of the approach and suggestions outlined in Section 9.3.3 includes something somewhat similar to a ‘phrasebook approach’ in which learners are encouraged to simply produce phrases, which may prevent them from achieving creative use of language (Hunston, 2002). However, it was argued that for primary learners in the context of EFL, it is necessary to lay the groundwork in order to enable them to read and write. At the same time, the importance of introducing variation where appropriate was stressed (i.e. presenting lexical items in a sequence where variation is allowed, but not oversimplifying the set of lexical items found with certain grammatical constructions).

Moreover, in applying the findings of corpus analysis, I have mentioned the importance of shifting the perception of language among current and prospective YL teachers (i.e. raising their awareness of the lexico-grammatical features of language rather than the traditional view of lexis and grammar as separate entities). The direct relevance of this suggestion to teacher education was discussed, and it was pointed out that it is necessary to formalise teacher training if this notion is to be reflected in the teaching of literacy to young learners. The DDL approach using the corpus for teacher education is proposed, in order to make them aware of the lexico-grammatical aspects of language. This can effectively contribute to the comprehension of texts and consequently to developing their ability to teach literacy to YLs. Overall, a specific corpus which is compiled with an ethnographic approach (Flowerdew, 2002, p.111) based on the teacher's knowledge of TEYL was proved to be useful in describing what kind of features predominantly occur and hence are likely to be encountered by learners, in order to inform corpus-influenced syllabus design.

CHAPTER TEN

CONCLUSION

10.1 Introduction

This chapter firstly readdresses the three main questions posed in this thesis in an attempt to examine the extent to which it has met its objectives. The aims of the study, provided in Chapter 1, were as follows:

I aim to find out, in relation to the primary educational sector in Japan, and, with reference to the teaching of English literacy skills to learners aged around 10-12:

- 1) What government and policy-makers think the learners should learn; and what the practitioners' views are on:
 - a) introducing pupils to English literacy;
 - b) the texts used to introduce English at primary level in practice;
 - c) what kinds of teacher training are needed;
- 2) In what ways culture can be explored in texts;
- 3) what role a corpus-based approach might have in the teaching of literacy skills in English.

The subsequent sections summarise the contribution to knowledge in the field made by this study, with reference to these research questions. Section 10.2 presents a prototype syllabus for the Japanese TEYL context, using the results obtained from the specific corpus used in this study. Section 10.3 summarises the extent to which the thesis has met its aims referring to the answers to research questions. Section 10.3.1 addresses the cultural context of Japanese ELT which is relevant to the RQs 1.1; 1.2; 1.3; 1.4 and 1.5. Section 10.3.2 summarises the contribution to the pedagogic corpus initiatives and the investigation of culture, which are relevant to RQs 2.1; 2.2; and 2.3. Section 10.3.3 considers the methodological contribution of the pedagogic corpus analysis (RQs 3.1

and 3.2), and the implications for the literacy instruction of young learners. A summary of the contribution of this study is given in Section 10.4. Limitations of this research are discussed in Section 10.5, followed by Section 10.6 which suggests potential for further development of this research.

10.2 A prototype syllabus for Japanese primary ELT

In Chapter 2, it was pointed out that the current Japanese TEYL policy makes no reference to literacy teaching in state-funded schools. Considering the needs of literacy skills from the pupils' side, as revealed from the surveys conducted in this study and by other researchers (e.g. Allen-Tamai, 2010; Butler 2005), there appears to be evidence of the need for some literacy instruction to young learners in Japan. This study provides support for this initiative by exploring the possibility of a corpus-influenced syllabus for literacy teaching to young learners in Japan.

In Chapter 3, different types of syllabuses were reviewed in terms of their benefits and drawbacks, and it was argued that it is hardly feasible to produce a single type of syllabus which universally suits any teaching context. Rather, it was pointed out that it is more common to find integrated syllabuses which draw on the principles of different approaches (e.g. topics, functions, grammar etc.) as all of them play a useful part in language teaching (Brown, 1995 pp.12-14; O'Dell, 1997 p. 272; White 1988, p.111). In Section 3.2.5, it was also stressed that syllabus designers need to consider the practicality of the syllabus, rather than focusing on specifying a single type of syllabus to employ for a particular situation.

Given that current primary school teachers receive little support with regard to English instruction, and bearing in mind the limited time available for ‘Foreign Language Activities’, it is useful to suggest a prototype syllabus, which provides guidance on what kind of literacy teaching is feasible in accordance with the current official material ‘Eigo Note’.

It is worth recalling, as stated in Chapter 3, that syllabuses can be defined as ‘ways of organising courses and materials’ (Brown, 1995 p. 14), and can therefore be regarded as sub-divisions of curricula. The following figure (10.1) shows the overall contents of ‘Eigo Note 1’ (MEXT, 2009c pp. 2-3).

Figure 10.1 Syllabus of Eigo Note 1 (MEXT, 2009c pp.2-3)

Lesson	Title	Contents to be taught			
		Session 1	Session 2	Session 3	Session 4
1	Say "hello" in different languages	Learn that there are different greetings in the world.	Learn manners of greetings, to do greetings with a positive attitude and be able to say your own name.	Greet your friends. Name card exchange game.	
	expressions	Hello. / What's your name?/ My name is Ken./ Nice to meet you. etc			
2	Making gestures	Learn about feelings and gestures	Learn that gestures are important. Communicate using gestures.	Using gestures to communicate with partners.	Using gestures to communicate feelings.
	expressions	How are you? / I'm happy. etc			
3	Playing with numbers	Learn that there are several ways of counting in the world. Numbers from 0 to 10.	Understand different ways of counting from around the world. Say numbers 0 to 20.	Asking "how many?"	Game using numbers.
	expressions	How many?/ Five. etc			
4	Self-introduction	Listen to and understand what others like and dislike	Express likes and dislikes	Ask others about their likes and dislikes	Introduce oneself including one's own likes and dislikes
	expressions/ words	Do you like apples?/ Yes, I do. No, I don't. I like apples./ Thank you. etc			
5	Getting to know the different costumes in the world	Learn about the different costumes in the world and different clothes.	Understand the importance of expressing one's own opinions and learn about expressions used in shopping	Express one's own likes and dislikes. Talk to others so that they can shop with pleasure.	Introduce the purchased items to peers.
	expressions	I don't like blue. etc			
6	Using foreign loan words	Understand the difference between loan words and English pronunciation.	When asked, express what you would like.	Order what you would like and make your own fruit parfait	Present your fruit parfait to the class
	expressions	What do you want?/ Melon please. etc			
7	General knowledge quiz	Notice how interesting it is that the same thing has a different name in English and Japanese	Understand "What's this?" and answer.	Learn names of animals in English	Enjoy a "quiz" with your friends
	expressions	What's this?/ It's a pencil. etc			
8	Daily schedules	Learn what kids in foreign countries learn in school. Learn names of subjects in English.	With a positive attitude, play a game using days of the week and subject names	Tell your friends about your original timetable that you have made.	Present the timetable you made with your group members.
	expressions	I study Japanese. etc			
9	Making a lunch menu	Learn that foreign countries' breakfasts are different to those common in Japan.	Learn food and cooking phrases	Using polite language ask for things you'd like.	Ask your friends what they like and make a "special lunch"
	expressions	What would you like? I'd like juice. etc			

In this section, suggestions for Lessons 4 and 5 will be highlighted in order to demonstrate how corpus findings can contribute to syllabus design as a means of supplementing the existing guidelines with literacy teaching. It is not feasible to address all of the lessons due to limitations of space.

In these two lessons, pupils learn how to express their likes and dislikes. However, under the current ‘Eigo Note’ guideline, the objects of likes and dislikes are limited to the pattern *I like + Noun*. In Chapter 9, it was pointed out there is a need to avoid oversimplification of the set of lexical items typically found with certain grammatical constructions. In this respect, Lesson 4 of ‘Eigo Note 1’ has already provided a wide variety of items (e.g. food, drinks, animals, and sports). However, the nouns introduced in Lesson 5 are limited to colours (e.g. *I like blue. I don’t like red.*) and names of clothes (e.g. *pants, skirt, hat, cap*). Moreover, the plural forms of the noun are not included when expressing likes and dislikes in the workbook for pupils, leading to the production of unacceptable output such as ‘**I like apple.*’ instead of ‘*I like apples.*’ In addition, it should be noted that a large proportion of the lexis introduced in MEXT-mandated lessons consists of words (such as names of fruits) which are similar to their Japanese equivalents, but often differ in pronunciation. As there is not enough attention given to pronunciation, this is a potential problem of which teachers of English need to be made aware.

The following figure (10.2) shows part of a prototype syllabus suggested for Lessons 4 and 5 in ‘Eigo Note 1’ (MEXT, 2009a). The underlined parts highlight the suggestions based on corpus analysis conducted in this study. Suggestions are made in terms of

lexical items and their use, adding structural information, functions and tasks, as well as the modification of topics which I consider suitable for the situations employed in those lessons.

Figure 10.2 Suggested prototype syllabus accompanying the material ‘Eigo Note’

Lesson Title	Lesson 4 Self-introduction	Lesson 5 Getting to know the different costumes in the world
Topics	<u>Favourite things</u>	Clothes and shopping
Functions	Expressing likes <u>Complimenting others</u>	Expressing opinions about likes and dislikes
Structures	I like ~.	I like ~. I don't like ~.
Lexical items appearing in 'Eigo Note'	banana, strawberry, pineapple, ice-cream, milk, orange juice, dog, rabbit, bird, fish, cat swimming, ski, basketball, soccer	orange, red, black, blue, white, yellow shirt, cap, shoes, socks, T-shirt
Suggested items based on corpus analysis	<p>1) I like/ I don't like + (pro)noun; 2) I like/ I don't like + ADJ + N; 3) I like/ I don't like + gerund; and 4) I like/ I don't like + infinitive; and</p> <p><i>NB) If pupils are able, infinitive and gerund forms can also be introduced. (e.g. 'I like soccer' may be replaced by 'I like playing soccer' 'I like to play soccer.' or 'I like watching soccer').</i></p> <p><i>NB) Majority of food items introduced in Eigo Note tend to be used in plural forms after 'I like/ I don't like ____', therefore plural forms should be introduced.</i></p>	
Suggested function based on corpus analysis	<p><u>Expressing likes/ Complimenting others</u> <u>I like ~'s + (ADJ) + N; I like your/his/her + (ADJ) + N</u></p> <p><i>NB) Pupils need to learn the difference between when they talk about a specific item, and the items in general.</i></p>	
Suggested Literacy Tasks	<p>A) 1. Interview friends about things they like in a group and make quizzes (e.g. Guess who?) for the rest of the groups in class. 2. Write down the items on cards for each student (Name of each pupil should be written on one side, and the items on the other side, one card per person). 3. Using the cards, items side face-up, each group reads out the quizzes they prepared. Take turns in groups to guess, and the pupil ending with the most cards wins.</p> <p>B) 1. Complimenting others in groups and writing memos (see Appendix 10.1) 2. Write down the list, in the form of 'I like ____'s ____.' 3. Presentation by groups. Others say 'I like it (that) too' if they agree.</p>	

As identified in Chapter 8 (Table 8.5), the patterns following ‘*I like*’/‘*I don’t like*’ in the corpus used for this study include *I like/ I don’t like* + 1) (*pro*)*noun*; 2) *ADJ* + *N*; 3) *infinitive*; and 4) *gerund*. These results indicate which structures could be introduced for expressing likes and dislikes.

In addition, based on the results of corpus output (see section 9.4), further tasks can be suggested with the function of complimenting others by introducing ‘*I like your/her/his/NAME’s ~.*’ (e.g. *I like your shirt. I like your hair.*) As seen in Chapter 8 (section 8.2.1), the second most frequent occurrence of ~’s was with possessives, but this use of possessives does not appear in ‘Eigo Note’. This is one of the examples in which the current material, ‘Eigo Note’, tends to omit certain features which frequently occur in texts written for children. Pupils need to be exposed to these items in order for them to be able to identify and make sense of them. As discussed in Chapter 3, this point is crucial in order for learners to encounter words a number of times in different contexts in order to fully acquire vocabulary (Cameron, 2001; Schmitt, 2000). Corpus findings are useful in identifying features which are likely to be encountered by learners, and this information is useful in constructing a corpus-influenced syllabus. Moreover, the construction ‘*I like _____*’ is often followed by a pronoun in the corpus (e.g. *I like it. I like that.*). Taking these findings into account, a task can be suggested as shown in Figure 10.1. In this task, pupils find something they like about their classmates and compliment each other. It is designed so that pupils will also have opportunities to use ‘*I like + pronoun*’ when they agree with a classmate’s idea.

In this proposal, it may appear that vocabulary learning is placed at the centre of the teaching cycle, rather than grammar. However, as important grammatical features are tied into words, it can be said that by being presented with these examples based on corpus output, pupils are inevitably exposed to grammar at the same time, as suggested by researchers supporting the idea of corpus-oriented approaches to syllabus design (Renouf & Sinclair, 1988; Willis 1990; Sealey & Thompson, 2006).

As stated in Chapter 3, topics and situations are important in motivating learners, as long as they relate to their own interests. However, the situation introduced in Lesson 5 of ‘Eigo Note’ (e.g. shopping for clothes) has little connection with what most pupils do in their daily lives. Furthermore, the main structure introduced here (i.e. *I like/I don’t like* ____) seems inappropriate for the situation. It is important to select lexical items which are suitable for a particular situation, considering the needs and interest of the learners (see section 3.3.3) when designing a syllabus. There is a mismatch between the situation provided by the activities and the structures designated to be introduced in Lessons 4 and 5 of ‘Eigo Note’.

As was pointed out in Chapter 9, some expressions in ‘Eigo Note’ are uncommunicative, so they need to be supplemented. In Lesson 4, the section title is specified as ‘self-introduction’, but the structure appearing here is only ‘*I like ~*’ introducing pupils’ favorite things. In reality, it is unlikely for people to start their self-introduction with ‘*Hello, I like ~*’. The section title of Lesson 4 may need revision to ‘Introduce things you like’.

The example prototype syllabus presented above (Figure 10.2) has been created in accordance with course guidelines suggested by the Japanese Ministry of Education (MEXT, 2008; 2009a), for practical applications. This study makes a case for the inclusion of more realistic and practical language in official syllabus design in the field of TEYL in Japan, based on corpus analysis such as the one described here.

10.3 Reflections on how the thesis has met its aims

10.3.1 The cultural context of Japanese primary ELT

First of all, I would like to submit that the thesis has met the first objective in identifying current issues in primary ELT especially in relation to policy makers' views on what should be taught, and teachers' views on literacy teaching and teacher education in order to discover the contextual considerations necessary for designing a corpus-influenced syllabus. This study by no means aimed to give a comprehensive review of current primary ELT in Japan, but the challenges in primary ELT have been discussed in relation to the policy issued by MEXT. It was suggested that there still exist some ideological conflicts in the Action Plan and its implementation in general (e.g. classroom contexts), even after the revised course of study in 2008. In Chapter 5, the original focus of Survey I was primarily on teachers' attitudes towards the introduction of literacy. However, the discussion of the findings highlighted problems in teacher training for English at primary level in Japan, which led to the further investigation of challenges perceived by current teachers and prospective teachers (Surveys II and III).

In the case of English literacy, the Survey I carried out in 2007 indicated some discrepancies in ideas about English literacy instruction among teachers. Some teachers expressed opinions which are less enthusiastic about introducing English literacy to primary level education. This reflects the current climate in primary ELT in Japan in which there is strong opposition to the introduction of English literacy education (e.g. Torikai 2006; Kageura, 2006). The main reasons for opposition often provided are the beliefs that 1) the introduction of written material is the main cause of pupils disliking English and can lead to a poor start for learners and 2) it is better to wait until pupils have fully established Japanese language skills which form part of their Japanese identity.

Although some teachers were resistant to accepting the teaching of English literacy for fear of not being able to develop children's Japanese language skills, some teachers who had received teacher training tended to think that introducing some form of English literacy is necessary. In addition, results of the survey indicated that some practitioners feel the need to use some kind of written form despite the present situation (i.e. the fact that MEXT discourages the use of the written form in teaching), and expressed their support for the introduction of English literacy. Moreover, some examples from respondents in the surveys indicated that pupils at the higher grades, aged around ten to twelve (5th and 6th grades), needed (or wanted) to see the written form of new language presented to them, judging from the fact that pupils asked teachers to write the new language in *katakana* (as found in three of the surveys reported in Chapter 5).

While there is little research available in terms of SLA with Japanese children, some research indicates that it is necessary to make young learners aware of the written aspect of English (e.g. Dolouz, 2000; Gregory 1996, 2008). It is certainly not a good idea to force pupils to read when they are not ready (Cameron, 2001) and it should be noted that my suggestion here does not mean that TEYL in Japan needs to start English literacy education at the very start of schooling. However, as the survey and other research have shown, there is a need as perceived by pupils. It could be perhaps explained that pupils around the age of 10-12 know from experience that language has a written form, and therefore it is quite natural for them to want to know how language practised in spoken form can be represented in written form. It could be argued that prohibiting access to written language could mean limiting children's potential for language learning.

Despite the suggestion by MEXT, texts are often used in practice, especially in pilot schools and private schools examined in Survey I (where the Course of Study issued by MEXT is regarded as the minimum requirement and therefore deviation from it is not a serious matter). In Chapter 5, texts which are used by and are familiar to teachers at primary level in Japan were identified based on results of Survey I and served to answer RQ 1.2 for the compilation of the corpus. The result of Survey II indicated that teachers working at primary schools which are not pilot schools nor private schools, receive little training.

The results obtained from three surveys indicated the need for gradual literacy teaching, and two further challenges are identified. Firstly, in the light of the predominant focus

on ‘listening and speaking’ in the policy (MEXT 2003; 2008), there is a need to consider a syllabus and appropriate teaching methodologies which balance the skills. Secondly, it is necessary to improve the quality of the teacher training system to support current and prospective teachers.

It has been claimed that local contexts, such as practitioners' opinions and real classroom challenges for teaching English literacy should be taken into account as well as the policy, in designing the syllabus. Judging from the results of the three surveys, the following local and contextual considerations are relevant for the planning of a corpus-influenced syllabus:

- intercultural education maintaining the Japanese identity
- incorporation of non-explicit teaching of grammar
- a literacy syllabus which uses materials and content familiar to pupils
- (though more teacher training is required), the syllabus needs to be simple enough and accompanied by procedural elements so that it would be easily understood even for teachers with limited training or experience in teaching English literacy.

10.3.2 Pedagogic corpus initiatives and aspects of culture

The thesis has succeeded in its second objective of identifying corpus initiatives and some cultural similarities and differences.

The compiled corpus of texts written for children is designed to be as close as possible to what learners might encounter in the future. The design of the corpus is ensured by the ‘ethnographic method’ (L. Flowerdew, 2002) by means of the survey conducted

with current primary teachers (Survey I). The information gained through the ethnographic method was deemed useful in creating a ‘specialised’ corpus, as can also be seen in ESP studies (for the list of texts included in the corpus, see Appendix 4.6).

In relation to the differences and similarities between the sub-corpora of EOCL and JTCL, these are identified in the comparative analysis conducted in Chapters 6 and 8. The differences and similarities identified are divided into two aspects: 1) linguistic and 2) cultural. One of the objectives of ‘Foreign Language Activities’ is international understanding including knowledge of different cultures. For this reason, a comparative analysis of EOCL and JTCL was conducted by means of the *Keyword* function in Wordsmith ver 4.0 in order to identify cultural differences between the two sub-corpora (see Chapter 6). We should bear in mind that WordSmith Keyword function facilitates the initial search criteria, which assisted to identify culturally-related keywords for further analysis.

It was stressed that any differences or similarities identified in this analysis could be limited to the sets of texts included in the corpora, and the results cannot be generalised easily. Some aspects of culture were identified through the analysis, though the corpora were not created in order to represent any particular culture as such. It should be mentioned that the results obtained were not significant, and they largely depend on the texts chosen in the corpus for each type. It was concluded that the comparative analysis can also be a good starting point for considering pedagogical implications for literacy instruction (e.g. encouraging pupils to notice links with what they already know, using

the culturally related words). Such considerations are necessary for teachers when designing a reading syllabus, materials and activities suitable for TEYL.

10.3.3 Pedagogic corpus analysis methods and implications

The third main objective was to find ways in which to identify pedagogically useful items (PUIs) to be included in a corpus-influenced syllabus (see Chapters 7 and 8). Three methods were used for the analysis of the corpus: frequency, n-grams, and colligational frameworks. The rationale for this analysis was to evaluate these three methods as a means of identifying pedagogically useful items for Japanese young learners. These analyses were divided into two stages: Stage 1 and Stage 2.

Stage 1 involved the statistical and computational automatic extractions of items. It was soon clear that qualitative interpretation of the data was needed, because automatic identification relies on frequency or statistical methods alone, and items found using only these methods are sometimes not meaningful. Furthermore, these automatic extractions of n-grams do not lead to the extraction of variations within a similar framework, such as a *bit of*, *a lot of* – the “*a * of*” framework etc. Additionally, statistically identified n-grams sometimes failed to produce meaningful units, requiring further analysis by means of concordances. A program devised by Danielsson (2001; 2003; 2007) was used to extract candidate colligational frameworks, as a way of avoiding the omission of less frequent structures and variations within the meaningful multi-word units. The analysis shows that this procedure produces a large number of candidates, including the low frequent items.

The Stage 2 analysis involved a subsequent search in concordances by means of the co-selection of a collocate and a node. This stage circumvented the difficulty of determining the sequence length and identifying variation within the frameworks. The Stage 1 method combined with the subsequent search for co-selection with collocates described in Stage 2, helped to highlight the potential for pedagogically useful items.

In deciding whether the items were pedagogically useful, a more qualitative analysis of concordance lines was needed, as well as drawing insights from relevant research findings, as statistical data alone ‘do not themselves carry any guarantee of pedagogic relevance’ (Widdowson 1991, p. 20-21). The findings of SLA research were also considered in discussing the pedagogically useful items. In discussing the pedagogical implications, the researcher’s intervention is necessary in all application and interpretation of the methods. This is the most challenging step for researchers in drawing conclusions about ‘what is useful’ for learners. Moreover, opinions and insights gained from current teachers in the surveys were also an essential source of supporting evidence in order to consider the relevance of this research to a specific pedagogical situation. The identified PPUIs were reviewed from the perspective of YL English teaching. Eight sets of considerations were suggested as principles for identifying items as PUI (see Chapter 8, Section 8.3) in order to serve as pointers for teachers of young children who are starting to read.

The employed methods can be claimed as being useful in specifying the potentially pedagogically useful items for the teaching of literacy to young learners. As Sinclair (1991, p. 39) recommends, it is important for material and syllabus designers to

investigate what to teach in the corpus (i.e. attested language data) at the beginning. Consequently, the same applies to practitioners as they are in the position of putting the research findings into practice in actual classrooms, thus teacher training in using a corpus is also proposed.

In Chapter 9, the implications regarding teaching have been suggested along with the consideration of teaching methodology (cf. Chapter 9, Sections 9.2 and 9.3). In terms of the implications regarding teaching methodology, some researchers (e.g. Hunston 2007, Fries 2008) point out that the lexico-grammatical features found in corpora suggest a rather different approach to reading featuring a ‘functional view to literacy’ which is somewhere between the traditional top-down and bottom-up approaches, as readers make use of their past experience in encountering ‘redundancy’ in understanding texts (Fries, 2008). Taking this point into account, it was suggested to incorporate the ‘functional view of literacy’ as a means of teaching children how to process texts, so that they can be trained to receive meaning through redundancy of units (i.e. sequences, patterns), rather than through a word-by-word process. As Hunston (2007, p. 254) suggests, native speaker readers only resort to word-for-word interpretation when faced with an absence of phrasal idiomatic patterns. It is important for the learners to recognise ‘words as units’ in learning to read as ‘this is built firmly into our general model of language’ (Sinclair, 2004 p. 24). This view has some relevance to the study of ‘reading units’ and the concept of ‘reading units’ is based on a claim that when people process a text in reading, the unit of processing is often greater than a word; what is read is, rather, processed at phrase-level (Kadota and Tada, 1992). In addition, research suggests that children are more likely to remember the meaning when the sentences are

held together as chunks, but not necessarily when the syntax is simple (Pearson, 1974). These insights show the importance of making sure that the units presented are meaningful chunks, rather than being decided by syntactical simplicity. This points to an important issue on how to design a syllabus which is not restricted to the traditional grammatical format often found in EFL.

Considering the identification of ‘units of meaning’ is especially pertinent to the relationship between reading process and reading units, it can be argued that the employed methods of analysis (i.e. the combination of Stage 1 and 2) in this study were appropriate, coinciding with the purpose of identifying the features which are beneficial to learners’ development of literacy.

It is clear that the teaching approach should be considered according to the children’s progress in learning and examples of tasks employing a modified TBL approach were presented (see Chapter 9, section 9.3). I believe this investigation provides useful implications for the teaching of English literacy in order to make learners aware of certain linguistic features which are useful for them when starting to read.

10.4 Contribution of this thesis

The main contribution of this thesis is to have suggested ways in which a corpus-influenced syllabus for the teaching of young learners in Japan could be designed using the concept of the lexical syllabus (Sinclair and Renouf, 1988, Willis, 1990) and the notion of a ‘pedagogic corpus’ (Willis, 1990 and 2003) as well as

employing an ‘ethnographic approach’ to corpus compilation for TEYL. In particular, the methodological investigations demonstrated the identification of pedagogically relevant features of the texts which can be used for literacy development, drawing relevance from the relationship between reading and ‘redundancy’ which is a product of patterns (Hunston, 2007). Thus, the contribution of the thesis can be summarised in the following statements.

- i) The ethnographic method of compiling a corpus is effective for identifying the pedagogically useful features of a corpus, relevant to the target young learners.
- ii) The combinations of objective methods used in this study deemed appropriate for identifying potentially PUIs for literacy instruction.
- iii) The methodological procedure followed in this thesis highlighted the relevance of the ‘functional view of literacy’.
- iv) The analysis and findings of the PUIs lead to an appropriate teaching approach for basic literacy for YLs.

Considering the current state of TEYL in Japan, there is a need to find ways to raise learners’ awareness about grammatical construction without explicitly teaching grammar. The corpus analysis helped to confirm that some phrases are not neatly explained in traditional grammar as proposed in Sinclair’s idiom principle (Sinclair, 1991), and the lexico-grammatical aspect can be utilised in promoting learners’ general awareness of language, without explicit grammatical instruction. I believe the procedure presented in this study sheds light on the way in which an ‘ethnographically designed

corpus' can serve to identify features which learners need to assimilate for the development of their English literacy as well their cultural understanding.

10.5 Limitations

There are a number of limitations to be mentioned relating to three areas: 1) the data 2) method of analysis and 3) research implications.

1) The data

First, this research is limited in terms of its generalisability. The degree of generalisability of the respondents' view towards primary ELT is limited by the number of respondents. The degree of generalisation of the results obtained from the corpus is also limited, related to the nature of the corpus used in this study. In particular, the corpus is relatively small, due to the restricted nature of the kinds of texts which could be included. It should be noted that there were constraints in the creation of the corpus, mainly due to copyright issues involved in storing literacy texts in a digitised format. As a result, the collection of texts slightly differs from those which are likely to be read by or to pupils in Japan.

The influence of the nature of the corpus was noticed in some of the results. For instance, with regard to the cultural aspects explored in Chapter 6, the examples are limited in certain ways. Therefore, there is a possibility that, if different texts were used, the items identified may be different. The findings observed in EOCL and JTCL could provide a basis for extracting keywords in a more quantitative way, and these could be

explored further in order to obtain more reliable results, if there were similar but much larger corpora available.

In addition, a full investigation of phraseology is limited by the small size of the corpus used in this study. Although some redundancy in the children's literature under investigation (i.e. repetitive expressions) was extracted successfully using the methods suggested in this study, the potential for generalisation of the identified items is limited.

Without the copyright constraints mentioned above, it would have been possible to compile a larger corpus, which in turn would have made it possible to extract items with a wider relevance for the teaching of literacy to young learners. In addition, although the use of this or other corpora in teacher training is suggested in Chapter 9, there are inevitable limitations in any corpus. Therefore it is impossible to find every example that teachers or prospective teachers would like to search for. As Sealey and Thompson (2006, p. 23) put it:

Most advocates of the approach [corpus-based approach] are quite willing to accept that [the corpus] cannot be allowed to dictate every aspect of what is taught.

2) Methods of analysis

The second area of limitation can be also seen in the methods of corpus analysis used for this study. The corpus linguistic approach failed to identify any significant cultural differences between the two sets of texts. There are several possible explanations for this. It may be that some aspect of the methods used was at fault, or it could be due to

the size of the two sub-corpora. Another explanation, of course, is that the supposed contrasts were simply not there to be found.

In addition, the methods described in Chapter 7 are inter-related or overlapping (e.g. frequency is involved in every search), but the differences between methods arises when items are viewed as ‘word-based’, or as longer sequences or frameworks which allow variations in slots. Moreover, though the main purpose was to show methodological means of identifying pedagogically useful items for literacy teaching, it must also be pointed out that not all the examples could be explained in depth due to the constraints of time and space.

3) Research implications

Finally, there are some limitations in the implications of this research. Although tasks and procedures are suggested in Chapter 9 for pedagogical exploitation of the results, this is intended as a starting point for future avenues for exploration. The implications for literacy teaching for young learners have to be evaluated in actual classrooms and a stronger case can be made if the items identified by analysis of the corpus are found to coincide with items and sequences demonstrated to be of value to learners.

10.6 Further research

Despite the limitations suggested above, the following points are in need of further study:

Expansion of the corpus

It was identified that some of the results are limited due to the size of the corpus. Therefore, in terms of the corpus design, it can be said that further expansion is necessary for the purpose of creating a more reliable pedagogic corpus. With the expansion of the current corpus, similar studies could be carried out in order to extract phraseological features and test the methods suggested in this study.

Teacher training

Considering the current scarcity of trained English teachers in primary schools, teacher education is another major issue in Japanese ELT, and clearly relevant to this research. It was identified in Chapter 5 that those teachers surveyed were not confident with their English abilities. The criteria for appointing and training English teachers at primary level are in need of review. Teachers should be aware of the concepts behind the mode of instruction suggested in this study for the successful application of a corpus-derived syllabus. As discussed in Chapter 8, it is necessary for teachers to realise the lexico-grammatical aspects of language when teaching children who are starting to read. Moreover, as MEXT states the purpose of English language teaching in Japan is to promote the development of real-world English ability, therefore for teachers to understand the meaning of communicative competence they need to know what is natural and what is not. Corpora such as the one used in the current study could be a useful tool for teacher training. Incorporation of a DDL approach into a teacher training course can be proposed as a means towards an increase in awareness about language.

Experimental study with pupils

Taking into consideration the limited research evidence available for TEYL and English literacy education in general, experimental studies with Japanese pupils need to be conducted in order to observe their progress in literacy skills (e.g. raising-awareness of the lexico-grammatical features of the text, a hybrid of top-down and bottom-up approaches).

A true evaluation of the implications of this study will only become apparent when we see its contribution to children's needs and progress. Nevertheless, it is hoped that the insights gained from the empirical analysis of the corpus in this study may be applied to the teaching of primary ELT in Japan.

REFERENCES

- Adams, R. & Newton, J. (2009). TBLT in Asia: constraints and opportunities. *Asian Journal of English Language Teaching* 19: 1-17.
- Adolphs, S. (2006). *Introducing Electronic Text Analysis: A practical guide for language and literary studies*. London: Routledge.
- Aijmer, K. (1996). *Conversational routines in English: Convention and creativity*. London: Longman.
- Aijmer, K. (1998). On the phraseology of spoken English: The evidence of recurrent word combinations. In A. P. Cowie (ed.). *Phraseology: Theory, analysis, and applications*. Oxford: Clarendon Press, pp. 101-122.
- Aijmer, K. (ed.). (2009). *Corpora and language teaching*. Amsterdam: John Benjamins.
- Aitchison, J. (2003). *Words in the mind. Third edition*. Oxford: Blackwell.
- Asahi Newspaper (2004) "Chikai shorai ni zengok-de" shogattko-no eigokyoiku-de monkasho [Minister of Education states that "English will be at all elementary schools in the near future"] Available from: <http://www.asahi.com/national/update/0113/038.html> [29th May, 2007].
- Aston, G. (2000). Corpora and language teaching. In L. Burnard & T. McEnery (eds.). *Rethinking language pedagogy from a corpus perspective: Papers from the third international conference on teaching and language corpora*. Frankfurt: Peter Lang, pp. 7-17.
- Allen-Tamai, M. (2010). *Shogakko eigo no kyoikuhou: riron to jittsusen [Educational methodology for primary ELT: Theory and practice]*. Tokyo: Taishukan-shoten.
- Ano, K. (2005). 'Japanese English: Fluency and Accuracy in the Spoken English of High School Learners'. *Proceedings of the 8th Conference of Pan-Pacific Association of Applied Linguistics*: 9-18. Available from: <http://www.paaljapan.org/resources/proceedings/PAAL8/>
- Anthony, L. (2007a). Automatic evaluation and simplification of vocabulary with applications in course materials development. *Proceedings of the JACET 46th Annual Convention*, pp. 116-117. 7th September, 2007. Hiroshima, Japan.
- Anthony, L. (2007b). *AntConc (version 3.2.1 w)* [Computer Software]. Tokyo: Waseda University. Available from: http://www.antlab.sci.waseda.ac.jp/antconc_index.html
- August, D. & Shanahan, T. (2008). Introduction and Methodology. In D. August, T. Shanahan (eds.). *Developing reading and writing in second-language learners: Lessons from the report of the National Literacy Panel on language-minority children and youth*. New York: Routledge, pp.1-18.
- Baker, C., & Freebody, P. (1989). *Children's first school books*. Oxford: Blackwell.
- Baker, N.D. & Nelson, K. E. (1984). Recasting and related conversational techniques for triggering syntactic advances by young children. *First Language* 5: 3-21.
- Baker, P., Hardie, A., & McEnery, T. (2006). *A Glossary of Corpus Linguistics*. Edinburgh: Edinburgh University Press.
- Barlow, M. (2004). *Collocate 1.0. Collocation extraction software*. [Computer Software]. Houston: Athelstan.
- Bauer, L. & Nation, I. S. P. (1993). Word families. *International Journal of Lexicography* 6 (3): 1-27.

- Barnbrook, G. (1996). *Language and Computers: A practical introduction to the computer analysis of language*. Edinburgh: Edinburgh University Press.
- Barnbrook, G. and Sinclair, J. M. (2001). Specialised corpus, local and functional grammars. In M. Ghadessy, A. Henry & R. Roseberry (eds.). *Small corpus studies and ELT*. Amsterdam: John Benjamins, pp. 237-276.
- Beaugrande, R. (2001). Text grammar revisited. *Lagos and Languages* Special issues. Available from: <http://www.beaugrande.com/Revisited.htm> [Accessed on 7 December 2006].
- Benesse. (2007). *A survey regarding the primary English education*. Tokyo: Benesse.
- Bernerdingi, S. (2000). Systematising serendipity: Proposals for concordancing large corpora with language learners. In L. Burnard and T. McEnery (eds.). *Rethinking Language Pedagogy from a Corpus Perspective: Papers from the third international conference on Teaching and Language Corpora*. Frankfurt am Main: Peter Lang, pp. 225-234.
- Bernerdingi, S. (2004). Corpora in the classroom: An overview and some reflections on future developments. In J. Sinclair (ed.). *How to Use Corpora in Language Teaching*. Amsterdam: John Benjamins, pp. 15-36.
- Biber, D. (1993). Representativeness in corpus design. *Literary and Linguistic Computing* 8 (4): 243-257.
- Biber, D., Conrad, S., & Reppen, R. (1994). Corpus-based approaches to issues in applied linguistics. *Applied Linguistics* 16 (2): 169-189.
- Biber, D., Conrad, S. & Reppen, R. (1998). *Corpus linguistics: Investigating language structure and use*. Cambridge: Cambridge University Press.
- Biber, D., & Finegan, E. (2001). Intra-textual variation within medical research articles. . In S. Conrad & D. Biber (eds.). *Variation in English: Multi-Dimensional Studies* Harlow: Longman, pp. 108-123.
- Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Grammar of Spoken and Written English*. London: Longman.
- Block, D. (1991). Some thoughts on DIY material design. *ELT Journal* 45(3): 211-217.
- Bolander, M. (1989). Prefabs, patterns and rules in interaction? Formulaic speech in adult learners' L2 Swedish. In K. Hyltenstam & L. K. Obler (eds.). *Bilingualism across the Lifespan*. Cambridge: Cambridge University Press, pp. 73-86.
- Bourke, J. M. (2006). Designing a topic-based syllabus for young learners. *ELT Journal* 60(3): 279-286.
- Bowker, L. & Pearson, J. (2002). *Working with specialized language: a practical guide to using corpora*. Oxon: Routledge.
- Brace, I. (2004). *Questionnaire Design: How to plan, structure and write survey material for effective market research*. London: Market Research in Practice Series.
- Breen, M. (1987). Learner contributions to task design. In C. N. Candlin & D. Murphy (eds.). *Lancaster Practical Papers in English Language Education: Vol. 7. Language learning tasks*. Englewood Cliffs, NJ: Prentice Hall, pp. 23-46.
- Brewster, J., Ellis, G., & Girard, D. (1991). *The Primary English Teacher's Guide*. London: Penguin.
- Brewster, J., Ellis, G., & Girard, D. (2002). *The primary English teacher's guide: new edition*. Essex: Pearson Education.
- Brown, J. D. (1995). *The elements of language curriculum: A systematic approach to program development*. Boston: Heinle & Heinle Publishers.

- Brown, D. (2000). *Principles of language learning and teaching*. Fourth edition. London: Longman.
- Brumfit, C. (1980). Notional syllabuses: A reassessment. In C.J. Brumfit (ed.). *Problems and principles in English teaching*. Oxford: Pergamon Press Ltd, pp.98-106.
- Brumfit, C.J. & Johnson, K. (eds.). (1979). *The communicative approach to language teaching*. Oxford: Oxford University Press.
- Bryman, A. (2001). *Social Research Methods*. Oxford: Oxford University Press.
- Bygate, M., Skehan, P., & Swain, M. (eds.). (2001). *Researching pedagogic task, second language learning and testing*. Harlow: Longman.
- Butler, Y. G. (2005). *Nihon no shogakkou-eigo wo kangaeru: Asia no shiten karano kenshou to teigen. [English language education in Japanese elementary schools: Analysis and suggestions based on East Asian perspectives]*. Tokyo: Sanseido.
- Butler, Y. G., & Iino, M. (2005). Current Japanese reforms in English language education: The 2003 "Action Plan". *Language Policy* 4: 25-45.
- Cameron, L. (2001). *Teaching languages to young learners*. Cambridge: Cambridge University Press.
- Cameron, L. (2003). Challenges for ELT from the expansion of teaching children. *ELT Journal* 57(2): 105-112.
- Candlin, C. N. & Mercer, N. (2001). Introduction. In C. N. Candlin & N. Mercer (eds.). *English language teaching in its social context: A reader*. Oxon: Routledge, p.1-8.
- Candlin, C. N. & Mercer, N. (eds.). (2001). *English language teaching in its social context: A reader*. Oxon: Routledge.
- Carless, D. (2002). Implementing task-based learning with young learners. *ELT Journal* 56(4): 389-396.
- Carless, D. (2003). Factors in the implementation of task-based teaching in primary schools. *System* 31(2): 485-500.
- Carless, D. (2004). Issues in teachers' reinterpretation of a task-based innovation in primary schools. *TESOL Quarterly* 38(4): 639-662.
- Carless, D. (2009). Revisiting the TBLT versus P-P-P debate: Voices from Hong Kong. *Asian Journal of English Language Teaching* 19: 49-66.
- Carter, R. (1998a). *Vocabulary: applied linguistics perspective. Second Edition*. London: Routledge.
- Carter, R. (1998b). Orders of reality: CANCODE, communication, and culture. *ELT Journal* 52(1): 43-56.
- Carter, R., & McCarthy, M. (eds.). (1988). *Vocabulary and Language Teaching*. Essex: Longman.
- Celce-Murcia, M. & Olshtain, E. (2000). *Discourse and Context in Language Teaching: a guide for language teachers*. Cambridge: Cambridge University Press.
- Carter, R., & McCarthy, M. (eds.). (1988). *Vocabulary and Language Teaching*. Essex: Longman.
- Carroll, B. J., Davis, P. & Richman, B. (1971). *The American Heritage Word Frequency Book*. New York: American Heritage Publishing Co. Inc.
- Cheng, W. & Warren, M. (1999). Inexplicitness: what is it and should we be teaching it? *Applied Linguistics* 20(3): 293-315.

- Chujo, K., & Nishigaki, C. (2004). Creating e-learning material to teach essential vocabulary for young EFL learners. *IWLeL: An Interactive Workshop on Language e-learning*: 35-44.
- Chujo, K., Nishigaki, C., Hasegawa, S., & Uchiyama, M. (2008). 'Yurori Kyoiku' jidai no koukoukyoukasho goi wo kangaeru: 1980 nendai to 2000 nendai no koukoukyoukasho goi no hikakubunseki kara no kousatsu.' [The impact of Yutori Kyoiku: A comparative study of 1988 and 2006 High school textbook vocabulary]. *English Corpus Studies* 15: 57-79.
- Clarke, F. D. (1991). The negotiated syllabus: What is it and how is it likely to work? *Applied Linguistics* 12 (1): 13-28.
- Clear, J. (2000). Do you believe in Grammar? In L. Burnard & T. McEnery (eds.). *Rethinking language pedagogy from a corpus perspective: Papers from the third international conference on teaching and language corpora*. Frankfurt am Main: Peter Lang, pp. 19-30.
- Coles, M. & Hall, C. (1998). *Breakthrough to Literacy: Teacher's Guide*. Harlow: Longman.
- Coniam, D. (2004) Concordancing yourself: a personal exploration of academic writing. *Language Awareness* 13(1): 49-55.
- Cook, G. (1998). The uses of reality: reply to Roald Carter. *ELT Journal* 52(1): 57-63.
- Collins, H. (2000). Materials design and language corpora: a report in the context of Distance Education. In L. Burnard & T. McEnery (eds.). *Rethinking Language Pedagogy from a Corpus Perspective: Papers from the third international conference on Teaching and Language Corpora*. Frankfurt am Main: Peter Lang, pp. 51-64.
- Conrad, S. (2001). Variation among disciplinary texts: A comparison of textbooks and journal articles in biology and history. In S. Conrad & D. Biber (eds.). *Variation in English: Multi-Dimensional Studies*. Harlow: Longman, pp. 94-107.
- Cooper, L. (1989). *Language planning and social change*. Cambridge: Cambridge University Press.
- Cowie, A. P. (ed.). (1998). *Phraseology: Theory, analysis applications*. Oxford: Oxford University Press.
- Coxhead, A. (2000). A new academic word list. *TESOL Quarterly* 34(2): 213-238.
- Crystal, D. (2003). *English as a global language*. Cambridge: Cambridge University Press.
- Danielsson, P. (2001). *The Automatic Identification of Meaningful Units in Language*. Unpublished PhD Thesis. Göteborg University, Gothenburg, Sweden.
- Danielsson, P. (2003). Automatic extraction of meaningful units from corpora: a corpus-driven approach using the word *stroke*. *International Journal of Corpus Linguistics* 8(1): 109-127.
- Danielsson, P. (2007). What constitutes a unit of analysis in language? *Linguistik online* 31(2). Available from: http://www.linguistik-online.de/31_07/danielsson.html [1st August, 2008].
- Darnton, A. (2000). Repeat after me: The role of repetition in the life of an emergent reader. In M. Scott & G. Thompson (eds.). *Patterns of text*. Amsterdam: John Benjamins, pp. 193-212.
- Day, R. R., Omura, C., & Hiramatsu, M. (1991). Incidental EFL vocabulary learning and reading. *Reading in a Foreign Language* 7: 541-551.

- Day, R. R., & Bamford, J. (1998). *Extensive Reading in the Second Language classroom*. Cambridge: Cambridge University Press.
- Day, R. R., & Bamford, J. (2002). Top ten principles for teaching extensive reading. *Reading in a Foreign Language* 14(2): 136-141.
- DeCock, S. (1998). A recurrent word combination approach to the study of formulae in the speech of native and non-native speakers of English. *International Journal of Corpus Linguistics* 3(1): 59-80.
- DeCock, S., Granger, S., Leech, G., & McEnery, T. (1998). An automated approach to the phrasicon of EFL learners. In S. Granger (ed.). *Learner English on computer*. London: Longman, pp. 67-79.
- Deller, H., & Hocking, D. (2000). *Innovations: an intermediate/upper intermediate course*. Hove: Language Teaching Publications.
- Deng, C. & Carless, D. (2009). The communicativeness of the activities in a task-based innovation in Guangdong, China. *Asian Journal of English Language Teaching* 19: 113-134.
- Di Pietro, R. J. (1981) Discourse and real-life roles in the ESL classroom. *TESOL Quarterly* 15(1): 27-33.
- Dlugosz, D. W. (2000). Rethinking the role of reading in teaching a foreign language to young learners. *ELT Journal* 54(3): 284-290.
- Dodd, T. (2001). *SARA: Technical Manual*. Oxford University Computing Services. Available from: <http://info.ox.ac.uk/bnc/sara/TechMan/>.
- Dogancay-Aktuna, S. (2005). Intercultural communication in English language teacher education. *ELT Journal* 59(2): 99-107.
- Doughty, C. & Williams, J. (1998). Issues and terminology. In C. Doughty & J. Williams. (eds.), *Focus on form in classroom second language acquisition*. Cambridge: Cambridge University Press, pp. 1-11.
- Doughty, C. & Williams, J. (eds.). (1998). *Focus on form in classroom second language acquisition*. Cambridge: Cambridge University Press.
- Dörnyei, Z. (2007). *Research Methods in Applied Linguistics*. Oxford: Oxford University Press.
- Dublin, F. & Olsthtain, E. (1986). *Course Design*. Cambridge: Cambridge University Press.
- Dulay, H. and Burt, M. (1973). Should we teach children syntax? *Language Learning* 23: 95-123.
- Dulay, H. and Burt, M. (1974). Natural Sequences in child second language acquisition. *Language Learning* 24: 37-53.
- Dupuy, B., & Krashen, S. D. (1993). Incidental vocabulary acquisition in French as a foreign language. *Applied Language Learning* 4: 55-63.
- Dunning, T. (1993) Accurate methods for the statistics of surprise and coincidence. *Computational Linguistics* 19(1): 61-74.
- Elley, W. (1989). Vocabulary acquisition from listening to stories. *Reading Research Quarterly*, XXIV (2): 174-187.
- Ellis, R. (1993). Second language acquisition and the structural syllabus. *TESOL Quarterly* 27: 91-113.
- Ellis, R. (1995). Interpretation tasks for grammar teaching. *TESOL Quarterly* 29(1): 87-105.
- Ellis, R. (2003). *Task-based language learning and teaching*. Oxford: Oxford University Press.

- Ellis, R. (2006). Current issues in the teaching of grammar: An SLA perspective. *TESOL Quarterly* 40(1): 36-42.
- Ellis, G., & Brewster, J. (1991). *The Storytelling handbook for primary teachers*. London: Penguin.
- Engwall, G. (1994). Not Chance, but Choice: Criteria in Corpus Creation. In B.T.S. Atkins & A. Zampolli (eds.). *Computational Approaches to the Lexicon*. Oxford: Oxford University Press, pp. 49-82.
- Erman, B. & Warren, B. (2000). The idiom principle and the open choice principle. *Text* 20(1): 29-62.
- Faucett, L., West, M., Palmer, H. E., & Thorndike, E. L. (1936). *The Interim Report on Vocabulary Selection for the Teaching of English as a Foreign Language*. London: P.S. King.
- Farrar, M. J. (1992). Negative evidence and grammatical morpheme acquisition. *Developmental Psychology* 28: 90-98.
- Fein, D. & Baldwin, R. (1986). Content-based curriculum design in advanced levels of an intensive ESL program. *Newsletter, English for Foreign Students in English-speaking Countries Interest Section, TESOL* 4 (1): 1-3.
- Firth, J. R. (1957). *Papers in Linguistics 1934-1951*. Oxford: Oxford University Press.
- Flowerdew, J. (1993). An educational or process approach to the teaching of professional genres. *ELT Journal* 47(4): 305-316.
- Flowerdew, J. (1996). Concordancing in language learning. In M. Pennington (ed.). *The Power of CALL*. Houston, TX: Athelstan, pp. 97-114.
- Flowerdew, J. (2001). Concordancing as a tool in course design. In M. Ghadsessy, A. Henry & R. L. Roseberry (eds.). *Small corpus studies in ELT: theory and practice*. Amsterdam: John Benjamins, pp. 71-92.
- Flowerdew, L. (2002). Corpus-based analyses in EAP. In J. Flowerdew (ed.). *Academic Discourse*. Harlow: Pearson Education, pp. 95-114.
- Flurkey, A. D., Paulson, E. J., & Goodman, K. S. (eds.). (2008). *Scientific realism in studies of reading*. New York and Oxon: Lawrence Erlbaum Associates.
- Fransis, W. N. & Kucera, H. (1982). *Frequency Analysis of English Usage*. Boston: Houghton Mufflin Company.
- Freyd, P., & Baron, J. (1982). Individual differences in acquisition of derivational morphology. *Journal of Verbal Learning and Verbal Behavior* 21: 282-295.
- Fries, C. C. (1945). *Teaching and learning English as a foreign language*. Ann Arbor: University of Michigan Press.
- Fries, P. H. (2008). Words, context and meaning in reading. In A. D. Flurkey, E. J. Paulson, & K. S. Goodman (eds.). *Scientific realism in studies of reading*. New York and Oxon: Lawrence Erlbaum Associates, pp. 53-82.
- Funabashi, Y. (2000). *Aete eigo kyogoron. [Nevertheless, English as an official language]*. Tokyo: Bungei Shunju.
- Gaies, S. (1977). The nature of linguistic input in formal second language learning: Linguistic and communicative strategies in ESL teachers' classroom language. In H.D. Brown, C.A. Yorio, & R. H. Crymes (eds). *On TESOL '77. Teaching and learning English as a second language: Trends in research and practice*. Washington, D.C: TESOL, pp. 201-212.
- Gamble, N. & Yates, S. (2002). *Exploring Children's Literature: teaching language and reading fiction*. London: Sage Publishing.

- Garton, A. & Pratt, C. (1998). *Learning to be literate: The development of spoken and written language. Second edition*. Oxford: Blackwell.
- Ghosn, I. K. (2002). Four good reasons to use literature in primary school ELT. *ELT Journal* 56(2): 172-179.
- Gillham, R. (2000). *Developing a Questionnaire*. London: Continuum.
- Glaser, R. (1998). The stylistic potential of phraseological units in the light of genre analysis. In A. P. Cowie (ed.). *Phraseology: theory, analysis and applications*. Oxford: Clarendon Press, pp.125-143.
- Gleason, H. A. (1961). *An Introduction to Descriptive Linguistics*. New York: Holt & Rinehart.
- Gledhill, C. (1995). *Scientific innovation and the phraseology of rhetoric: posture, reformulation and collocation in cancer research articles*. Unpublished Ph.D. thesis. Aston University, Birmingham, UK.
- Gledhill, C. (1996). Science as a collocation. Phraseology in cancer research articles. In S. Botley, J. Glass, T. McEnery & A. Wilson (eds.). *Proceedings of Teaching and Language Corpora 1996*. The University of Lancaster: UCREL Technical Papers Volume 9, pp.108-126.
- Gledhill, C. (2000). The discourse function of collocation in research article introductions. *English for Specific Purposes* 19: 115-135.
- Goodman, K. (1986). *What's whole in whole language?* Portsmouth: N.H. Heinemann.
- Goulden, R, Nation, P. & Read, J. (1990). How large can a receptive vocabulary be? *Applied Linguistics* 11(4): 341-363.
- Granger, S. (ed.). (1998). *Learner English on computer*. London: Longman.
- Granger, S. (2005). Pushing back the limits of phraseology: How far can we go?
Available from:
<http://sites-test.uclouvain.be/cecl/archives/Granger%202005b.doc>. [4th November 2008].
- Granger, S., & Tribble, C. (1998). Learner corpus data in the foreign language classroom: Form-focused instruction and data-driven learning. In S. Granger (ed.). *Learner English on Computer*. Harlow: Longman, pp.109-209.
- Gregory, E. (1996). *Making sense of a new world: Learning to read in a second language*. London: Paul Chapman Publishing.
- Gregory, E. (2008). *Learning to read in a new language: Making sense of words and worlds*. London: Sage Publications.
- Griva, E., Tsakiridou, H., & Nihoritou, I. (2009). A study of FL composing process and writing strategies employed by young learners. In M. Nikolov. (ed.). *Early Learning of Modern Foreign Languages: Process and outcomes*. Bristol: Multilingual Matters, pp. 132-148.
- Groom, N. (2005). Patterns and meaning across genres and disciplines: An exploratory study. *Journal of English for Academic Purposes* 4(3): 257-277.
- Halliday, M.A.K. (1978). *Language as social semiotic: The social interpretation of language and meaning*. London: Edward Arnold.
- Hancioğlu, N., & Eldridge, J. (2007). Texts and frequency lists: Some implications for practicing teachers. *ELT Journal* 61(4): 330-340.
- Harklau, L. (2005). Ethnography and ethnographic research on second language teaching and learning. In E. Hinkel (ed.). *Handbook of research in second language teaching and learning*. Mahwah, N. J.: Laurence Erlbaum, pp. 179-194.

- Hart, N., Walker, R. & Gray, B. (1977). *The language of children: A key to literacy*. London: Addison Wesley.
- Harwood, N. (2002). Taking lexical approach to teaching: principles and problems. *International Journal of Applied Linguistics* 12(2): 139-155.
- Harwood, N. (2005a). What do we want EAP teaching materials for? *Journal of English for Academic Purposes* 4 (2): 149-161.
- Harwood, N. (2005b). *I hoped to counteract the memory problem, but I made no impact whatsoever*: discussing methods in computing science using *I. English for Specific Purposes* 24 (3): 243-267.
- Harwood, N. (2005c). *'Nowhere has anyone attempted...In this article I am to do just that'*. A corpus-based study of self-proportional I and WE in academic writing across disciplines. *Journal of Pragmatics* 37(8): 1207-1231.
- Harwood, N. (2005d). *'We do not seem to have a theory...The theory I present here attempts to fill this gap'*: inclusive and exclusive pronouns in academic writing. *Applied Linguistics* 26 (3): 343-375.
- Harwood, N. (2010). Issues in materials development and design. In N. Harwood. (ed). *English Language Teaching Materials: Theory and practice*. Cambridge: Cambridge University Press, pp. 3-30.
- Harwood, N. (ed). (2010). *English Language Teaching Materials: Theory and practice*. Cambridge: Cambridge University Press.
- Harwood, N. & Hadley, G. (2004). Demystifying institutional practices: critical pragmatism and the teaching of academic writing. *English for Specific Purposes* 23 (4): 355-377.
- Hawkins, B. (1991). Teaching children to read in a second language. In M. Celce-Murcia. (ed.). *Teaching English as a second or foreign language*. New York: Newbury House, pp. 169-184.
- Haynes, M. & Baker, L. (1993). American and Chinese readers learning from lexical familiarization in English text. In T. Huckling, M. Haynes, & J. Coady (eds.). *Second language reading and vocabulary learning*. Norwood, N.J.: Ablex Publishing Cooperation, pp. 130-152.
- Hinkel, E. (ed.). (2005). *Handbook of research in second language teaching and learning*. Mahwah, N. J.: Laurence Erlbaum.
- Hinkel, E. & Fotos, S. (eds.). (2002). *New perspectives on grammar teaching in second language classrooms*. London: Routledge.
- Hoey, M. (1991). *Patterns of Lexis in Text*. Oxford: Oxford University Press.
- Hoey, M. (2000). The hidden lexical clues of textual organisation: a preliminary investigation into an unusual text from a corpus perspective. In L. Burnard & T. McEnery (eds.). *Rethinking language pedagogy from a corpus perspective: Papers from the third international conference on teaching and language corpora*. Frankfurt am Main Peter Lang, pp. 31-42.
- Hoey, M. (2005). *Lexical Priming: a new theory of words and language*. London: Routledge.
- Hoey, M., Mahlberg, M., Stubbs, M., & Teubert, W. (2007). *Text, discourse and corpora: Theory and analysis*. London: Continuum.
- Hofstede, G. (2001). *Cultures' consequences: Comparing values, behaviours, institutions, and organizations across nations*. California: Sage Publications.
- Hofstede, G. & Hofstede, G. J. (2005). *Cultures and organizations: software of the mind. Second edition*. New York: McGraw-Hill.

- Holliday, A., Hyde, M., & Kullman, J. (2004). *Intercultural communication*. London: Routledge.
- Holmes, J. (1988). Doubt and certainty in ESL textbooks. *Applied Linguistics* 9(1): 21-44.
- House, J. (1996). Developing pragmatic fluency in English as a foreign language: routines and metapragmatic awareness. *Studies in Second Language Acquisition* 18: 225-252.
- Howarth, P. (1996). *Phraseology in English academic writing: some implications for language learning and dictionary-making*. Tübingen: Niemeyer.
- Howarth, P. (1998). Phraseology and second language proficiency. *Applied Linguistics* 19(1): 24-44.
- Hulstijn, J. H. (1992). Retention of inferred and given word meanings: Experiments in incidental vocabulary learning. In P. Arnaud & H. Bejoint (eds.). *Vocabulary and applied linguistics*. London: Macmillan, pp. 113-125.
- Hulstijn, J. H. (1993). When do foreign language readers look up the meaning of unfamiliar words? The influence of task and learner variables. *Modern Language Journal* 77: 139-147.
- Hunston, S. (2002). *Corpora in applied Linguistics*. Cambridge: Cambridge University Press.
- Hunston, S. (2007). Grammar patterns and literacy. In A. McCabe, M. O'Donnell & R. Whittaker (eds.). *Advances in language and education*. London: Continuum, pp. 254-267.
- Hunston, S. (2009). The usefulness of corpus-based descriptions of English for learners: the case of relative frequency. In K. Aijmer (ed.). *Corpora and Language Teaching*. Amsterdam: John Benjamins, pp.141-156.
- Hunston, S., & Francis, G. (1998). Verbs Observed: A corpus-driven pedagogic grammar. *Applied Linguistics* 19: 45-72.
- Hunston, S., & Francis, G. (1999). *Pattern grammar: A corpus-driven approach to the lexical grammar of English*. Amsterdam: John Benjamins.
- Hunston, S., Francis, G., & Manning, E. (1997). Grammar and vocabulary: showing the connections. *ELT Journal* 51(3): 208-216.
- Hunston, S. & Thompson, G. (ed.). (2000). *Evaluation in text: authorial stance and the construction of discourse*. Oxford: Oxford University Press.
- Hunt, P. (ed.). (1990). *Children's literature: the development of criticism*. London: Routledge.
- Hunt, P. (ed.). (1992). *Literature for children: contemporary criticism*. London: Routledge.
- Hunt, P. (1994). *An Introduction to children's literature*. Oxford: Oxford University Press.
- Hunt, P. (ed.). (2005). *Understanding children's literature. Second edition*. London: Routledge.
- Hutchinson, T., & Torres, E. (1994). The textbook as agent of change. *ELT Journal* 48(4): 315-328.
- Hwang, K. & Nation, P. (1989). Reducing the vocabulary load and encouraging vocabulary learning through reading newspapers. *Reading in a Foreign Language* 6(1): 232-335.
- Hyland, K. (1994). Hedging in academic writing and EAP textbooks. *English for Specific Purposes* 13(3): 239-266.

- Hyland, K. (1998). Talking to students: metadiscourse in introductory coursebooks. *English for Specific Purposes* 18(1): 3-26.
- Hyland, K. (1999). Disciplinary discourses: writer stance in research articles. In C. N. Candlin & K. Hyland (eds.). *Writing: texts, processes and practices*. Harlow: Longman, pp. 99-121.
- Hyland, K. (2000) *Disciplinary discourses: social interactions in academic writing*. Harlow: Pearson Education.
- Hyland, K. (2004). Disciplinary interactions: metadiscourse in L2 postgraduate writing. *Journal of Second Language Writing* 13: 133-151.
- Hyland, K. (2009). Metadiscourse: mapping interactions in academic writing. *Nordic journal of English studies. Special Issue on Metadiscourse* 9(2): 125-143.
- Inoue, F. (2001). *Nohongo-wa ikinokoreruka [Can the Japanese language survive?]*. Tokyo: PHP shinsho.
- Iryjo, S. (1986). Don't put your leg in your mouth: transfer in the acquisition of idioms in a second language. *TESOL Quarterly* 20(2): 287-304.
- Ishikawa, S. (2007). L1/L2 corpus kaiseki nimotozoku jidoueigokyouiku no tameno goi material chushutsu shisutemu no kaihatu [Choosing basic English words for Japanese elementary school students: An analytical approach based on L1/L2 corpora]. *Journal of English Language Teaching Conference in central Japan* 36: 317-332.
- Izumi, E. & Isahara, H (2004). Investigation into language learners' acquisition order based on the error analysis of the learner corpus. *Proceedings of IWLeL 2004: An Interactive Workshop on Language e-Learning*: 63-71.
- Jabbour, G. (1998). *Corpus Linguistics, contextual collocation and ESP syllabus creation: A text analysis approach to the study of medical research articles*. Unpublished PhD Thesis, University of Birmingham, Birmingham, UK.
- Jenkins, J., Stein, M., & Wysocki, K. (1984). Learning vocabulary through reading. *American Educational Research Journal* 21: 767-787.
- Jenkins, J. (2003). *World Englishes*. Abingdon: Routledge.
- Johns, T. (1989). Whence and whither classroom concordancing? In T. Bongaerts, P. de-Haan, S. Lobbe & H. Wekker (eds.). *Computer applications in language learning*. Dordrecht: Foris Publications, pp. 9-33.
- Johns, T. (1991). Should you be persuaded: Two samples of data-driven learning materials. *ELR Journal* 4: 1-16.
- Jones, B. (1995). *Exploring otherness: an approach to cultural awareness. Pathfinder Series for Language Teachers* 24. London: CILT.
- Kachru, B. B. (ed.). (1992). *The other tongue: English across cultures. Second edition*. Chicago: University of Illinois Press.
- Kadota, S. (1988). *Dokkaini okeru shikakuninnshiki no mekanizumu: Gendai no eigo kenkyu [Mechanism of sight-recognition in comprehension: Current English language research]*. Tokyo: Kinseido.
- Kadota, S. (2002) *Hanashi kotoba to kakikotoba wa ikani kankeishiteiruka [The relationship between spoken and written language]*. Tokyo: Kuroshio Shuppan.
- Kadota, S. & Tada, M. (1992). *Eibun oyobi nihonbun no dokkai to shoritani. [Comprehension and process units of English and Japanese texts]*. Tokyo: Kuroshio Shuppan.
- Kamiya, N. Hasebe, I, Nishina, Y. & Hirata, E. (2010). Jidoueigo kyouiku ni okeru teikei hyougen: eigo-note to kokunai jidoueigo kyoikuyoukyozai no hikaku wo

- chushin ni. [Fixed expressions in TEYL textbooks: a comparative study of Eigo-Note and other TEYL materials in Japan]. Presentation at *The third joint conference on English Vocabulary and Lexicography*, Tokyo: Waseda University. 12th December 2010.
- Kaneko, T. (2006). Semantic network makes things more problematic: Use of vertical axis prepositions by Japanese learners of English. *Gakuen Communication Journal* 786: 1-16.
- Kaprowski, M. (2005). Investigating the usefulness of lexical phrases in contemporary coursebooks. *ELT Journal* 59(4): 322-332.
- Kelly, P. (1990). Guessing: No substitute for systematic learning of lexis. *System* 18(2): 199-207.
- Kennedy, C. (1986). Innovation for a change: teacher development and innovation. *ELT Journal* 41 (3): 163-170.
- Kennedy, G. (1998) *An introduction to corpus linguistics*. Harlow: Longman.
- Kirkpatrick, A. (2007). *World Englishes: Implications for international communication and English language teaching*. Cambridge: Cambridge University Press.
- Kirsch, C. (2008). *Teaching foreign languages in the primary school*. London: Continuum.
- Kjellmer, G. (1987). Aspects of English collocations. In W. Meijs (ed.). *Corpus linguistics and beyond*. Amsterdam: Rodopi, pp. 133-140.
- Koda, K. (2005). *Insights into Second Language Reading: a cross-linguistics approach*. Cambridge: Cambridge University Press.
- Koda, A. (2000). Shogakko kara chugakko eigo he no kyouzaikaihatsu niyoru paipu zukuri: Authentic na kyouzai de fukai rikai he michibiku tameno ichikousatsu. [Successful transition from elementary school English to junior high school English: Use of previously studied and authentic teaching materials]. *JASTEC Journal* 19: 59-71.
- Koizumi, M. (2011). Shogakko de eigo wo gakushu surukotode naniga kwaruka [What will change with the introduction of primary English education]. *The English Teachers's Magazine* 59 (11): 31-33.
- Knowles, M., & Malmkjær, K. (1996). *Language and control in children's literature*. London: Routledge.
- Kramsch, C. (1993). *Context and culture in language teaching*. Oxford: Oxford University Press.
- Krashen, S. (1978). Is the 'natural order' an artifact of the Bilingual Syntax measure? *Language Learning* 28: 175-183.
- Krashen, S. (1980). The theoretical and practical relevance of simple codes in second language acquisition. In R. Scarcella and S. Krashen (Eds.) *Research in Second Language Acquisition*. Rowley, Ma.: Newbury House, pp. 7-18.
- Krashen, S., Seliger, H., & Hartnett, D. (1974). Two studies in adult second language learning. *Kritikon Litterarum* 2 (3): 220-228.
- Kubaneck-German, A. (2000). Emerging intercultural awareness in a young learner context: a checklist for research In J. Moon & M. Nikolov (eds.). *Research into teaching English to young learners*. Pecs: University of Pecs Press, pp. 49-65.
- Kubota, R. (2002). The impact of globalization on language teaching in Japan. In D. Block & D. Cameron (eds.). *Globalization and language teaching*. London: Routledge, pp. 13-28.

- Kuno, Y. (1999). *Konna fuu ni hajimete mitaha? Shogakko eigo [English for children: A new perspective]*. Tokyo: Sanseido.
- Kuno, Y., Sato, R., Nagai, J. Komeya, K. (2006). *Kokoga pointo! Shogakkoeigo [Key points for primary ELT]*. Tokyo: Sanseido.
- Laufer, B. (1991). The development of L2 Lexis in the expression of the advanced learner. *The Modern Language Journal* 75: 440-448.
- Laufer, B. (1997). What's in a word that makes it hard or easy: some intralexical factors that affect the learning of words. In N. Schmitt & M. McCarthy (eds.), *Vocabulary: Description, acquisition, and pedagogy*. Cambridge: Cambridge University Press, pp.140-155.
- Laufer, B. (2001). Reading, word-focused activities and incidental vocabulary acquisition in a second language. *Prospect* 16: 44-54.
- Laufer, B., & Paribakht, T. S. (1998). The relationship between passive and active vocabularies: Effects of language learning context. *Language Learning* 48: 365–391.
- Lee, H. (2006). *Using English children's literature in particular Arthur Ransome's Swallows and Amazons for teaching EFL in Taiwan*. Unpublished PhD Thesis University of Birmingham, Birmingham, UK.
- Lee, J. (2000). *Tasks and communicating in language classrooms*. Boston: McGraw-Hill.
- Lewis, M. (1993). *The Lexical Approach*. Hove: Language Teaching Publications.
- Lewis, M. (1997). *Implementing the lexical approach: putting theory into practice*. Hove: Language Teaching Publications.
- Lewis, M. (ed.). (2000). *Teaching collocation: Further developments in the lexical approach*. Boston: Thomson Heinle.
- Lewis, M. (1997). *Implementing the lexical approach*. Hove: Language Teaching Publications.
- Lewis, M., & Ellis, S. (ed.). (2006). *Phonics: practice, research and policy*. London: Sage Publications.
- Lightbown, M. P., & Spada, N. (1999). *How languages are learned. Second edition*. Oxford: Oxford University Press.
- Lightbown, M. P. & Spada, N. (2006). *How languages are learned. Third edition*. Oxford: Oxford University Press.
- Linse, C. (2007). Predictable books in the children's EFL classroom. *ELT Journal* 61(1): 46-54.
- Long, M. H. (1988). Instructed interlanguage development. In L. M. Beebe (ed.). *Issues in second language acquisition: Multiple perspective*. Cambridge, MA: Newbury House/Harper and Row, pp.115-141.
- Long, M. H. (1991). Focus on form: A design feature in language teaching methodology. In K. de Bot, R. Ginsberg, & C. Kramsch (eds.). *Foreign language research in cross-cultural perspective*. Amsterdam: John Benjamins, pp. 39-52.
- Long, M.H. & Crookes, G. (1992). Three approaches to task-based syllabus design. *TESOL Quarterly* 26 (1): 27-56.
- Long, M. H. & Robinson, P. (1998). Focus on form: Theory, research and practice. In C. Doughty & J. Williams. (eds.). *Focus on form in classroom second language acquisition*. Cambridge: Cambridge University Press, pp.12-41.
- Louw, B. (1991). Classroom concordancing of de-lexical forms and the case for integrating language and literature. *ELR Journal* 4: 151-178.

- Louw, B. (1997). The role of corpora in critical literary appreciation. In A. Wichman, S. Fligelstone, T. McEnery & G. Knowles (eds.). *Teaching and language corpora*. Harlow: Addison Wesley Longman, pp. 240-251.
- Macalister, J. & Nation, I.S.P. (eds.). (2011). *Case studies in language curriculum design: concepts and approaches in action around the world*. London: Routledge.
- Mahlberg, M. (2007a). Lexical items in discourse: identifying local textual functions of *sustainable development*. In M. Hoey, M. Mahlberg, M. Stubbs and W. Teubert (eds.). *Text, discourse and corpora: Theory and analysis*. London: Continuum, pp. 191-218.
- Mahlberg, M. (2007b). Corpus stylistics: bridging the gap between linguistic and literary studies. In M. Hoey, M. Mahlberg, M. Stubbs and W. Teubert (eds.). *Text, discourse and corpora: Theory and analysis*. London: Continuum, pp. 219-246.
- Mason, O. (2006). *The automatic extraction of linguistic information from text corpora*. Unpublished PhD thesis. University of Birmingham, Birmingham, UK.
- Mason, O. (2007). Multi-word units as a model of grammar. *Proceedings of Corpus Linguistics 2007*. Available from: <http://ucrel.lancs.ac.uk/publications/CL2007/>.
- Mason, O. & Hunston, S. (2004). The automatic recognition of verb patterns: A feasibility study. *International Journal of Corpus Linguistics* 9 (2): 253-270.
- Marco, M. J. L. (2000). Collocational frameworks in medical research papers: A genre-based study. *English for Specific Purposes* 19(1): 63-86.
- Matsuka, Y. (2001). 'Uta chantsu wo riyō shita kyōzai to sono shidōhō [Teaching methods and materials using songs and chants.]' In K. Nakayama (ed.). *Jidō eigo kyōiku wo manabu hito no tameni [Handbook for prospective teachers of English to young learners]*. Tokyo: Sekaishisōsha, pp. 170-171.
- Matsukawa, R. (2007). Shōgattō eigo kōiku-no kyōkaka-no kanōsei. [Prospects for implementing primary English education]. *The English Teacher's Magazine* 49: 14-16.
- McCabe, A., O'Donnell, M., & Whittaker, R. (eds.). (2007). *Advances in language and education*. London: Continuum.
- McCarthy, M. (1990). *Vocabulary*. Oxford: Oxford University Press.
- McDonough, J., & Shaw, C. (2003). *Materials and methods in ELT: Second edition*. Oxford: Blackwell.
- McEnery, T. & Wilson, A. (2001). *Corpus linguistics. Second edition*. Edinburgh: Edinburgh University Press.
- McEnery, T., Xiao, R., & Tono, Y. (2006). *Corpus-based language studies: An advanced resource book*. London: Routledge.
- McKay, W. F. (1965). *Language teaching analysis*. London: Longman.
- McKay, S. (1978). Syllabuses: Structural, situational, notional. *TESOL Newsletter*, 12 (5), 11.
- MEXT. (1947). Gakushū shidō youryō 1947 [The Course of Study 1947].
- MEXT. (1951). Gakushū shidō youryō 1951 [The Course of Study 1951].
- MEXT. (2002a). Developing a strategic plan to cultivate "Japanese with English Abilities". Available from: <http://www.mext.go.jp/english/news/2002/07/020901.htm> [22nd January, 2007].
- MEXT. (2002b). *Practical handbook for elementary school English activities*. Tokyo Kairyudo.

- MEXT. (2003). Regarding the establishment of an Action Plan to Cultivate “Japanese with English Abilities”. Available from:
<http://www.mext.go.jp/english/topics/03072801.htm> [22nd January 2007].
- MEXT. (2008). Shogakko gakushu shido yoryo [The Course of Study for primary education 2008]. Available from:
http://www.mext.go.jp/a_menu/shotou/new-cs/youryou/syo/syo.pdf
- MEXT. (2009a). *Eigo Note 1: Student's book*. Tokyo: MEXT
- MEXT. (2009b). *Eigo Note 2: Student's book*. Tokyo: MEXT.
- MEXT. (2009c). *Eigo Note 1: Teacher's book*. Tokyo: MEXT.
- MEXT. (2009d). *Eigo Note 2: Teacher's book*. Tokyo: MEXT.
- MEXT. (2010). Chapter 4 Foreign Language Activities (translation of Chapter 4, Course of Study 2008, updated on the 7th October, 2010). Available from:
http://www.mext.go.jp/component/a_menu/education/micro_detail/_icsFiles/afildfile/2010/10/20/1261037_12.pdf [21 November, 2010].
- Mindt, D. (1996). English corpus linguistics and the foreign language teaching syllabus. In T. J & M. Short (eds.). *Using corpora for language research: studies in the honour of Geoffrey Leech*. London: Longman, pp. 232-247.
- Milton, J. & Meara, P. M. (1995). How periods abroad affect vocabulary growth in a foreign language. *ITL 107-108*: 17-34.
- Mitchell, R. & Myles, F. (1998). *Second language learning theories. Second edition*. London: Arnold.
- Mondria, J-A. & Wit de-Boer, M. (1991). The effects of contextual richness on the guessability and the retention of words in a foreign language. *Applied Linguistics* 12(3): 249-267.
- Moon, J. (2008). L2 children and writing: a neglected skill? *ELT Journal* 62(4): 398-400.
- Moon, R. (1997). Vocabulary connections: multi-word items in English. In N. Schmitt & M. McCarthy (eds.). *Vocabulary: description, acquisition, and pedagogy*. Cambridge: Cambridge University Press, pp. 40-63.
- Moon, R. (1998). *Fixed expressions and idioms in English: a corpus-based approach*. Oxford: Oxford University Press.
- Morizumi, M. (2009). Japanese English for EIAL: What it should be like and how much has been introduced. In K. Murata. & J. Jenkins (eds.). *Global Englishes in Asian contexts: Current and future debates*. Hampshire: Palgrave Macmillan, pp. 73-93.
- Morrow, K., & Schocker, M. (1987). Using texts in a communicative approach. *ELT Journal* 41(4): 248-256.
- Moudraia, O. (2001). Lexical approach to second language teaching. Available from:
<http://www.cal.org/resources/digest/0102lexical.html> [3rd November, 2007].
- Mumby, J. (1978). *Communicative Syllabus Design: A sociolinguistic model for defining the content of purpose-specific language programmes*. Cambridge: Cambridge University Press.
- Murata, K. & Jenkins, J. (eds.). (2009). *Global Englishes in Asian contexts: Current and future debates*. Hampshire: Palgrave Macmillan.
- Nagy, W. E. & Anderson, R.C. (1984). How many words are there in printed school English? *Reading Research Quarterly* 19: 304-330.
- Nagy, W., Herman, P., & Anderson, R. (1985). Learning words from context. *Reading Research Quarterly*, 20, 233– 253.

- Nagy, W., Anderson, R., & Herman, P. (1987). Learning word meanings from context during normal reading. *American Educational Research Journal*, 24, 237–270.
- Nasselhauf, N. (1996). *Collocations in a learner corpus*. Amsterdam: John Benjamins.
- Nasselhauf, N. (2004). Learner corpora and their potential for language teaching. In J. Sinclair (ed.). *How to use corpora in language teaching*. Amsterdam: John Benjamins, pp. 125-156.
- Nation, I.S.P. (1990). *Teaching and learning vocabulary*. Boston: Heinle & Heinle Publishers.
- Nation, I.S.P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.
- Nation, I.S.P. (2009). *Teaching ESL/EFL reading and writing*. New York: Routledge.
- Nation, I.S.P. & Macalister, J. (2010). *Language curriculum design*. London: Routledge.
- Nation, I.S.P. & Macalister, J. (2011). Introduction. In M. Macalister & I.S.P, Nation (eds.). *Case studies in language curriculum design: concepts and approaches in action around the world*. London: Routledge, pp.1-7.
- Nation, I.S.P. & Newton, J. (2009). *Teaching ESL/EFL listening and speaking*. New York: Routledge.
- Nation, I.S.P. & Waring, R. (1997). Vocabulary size, text coverage and word lists. In Schmitt, N. and M. McCarthy (eds.). *Vocabulary: Description, Acquisition and Pedagogy*. Cambridge: Cambridge University Press, pp. 6-19.
- Nattinger, J. (1980). A lexical phrase grammar for ESL. *TESOL Quarterly* 14: 337-344.
- Nattinger, J., & DeCarrico, J. (1992). *Lexical phrases and language teaching*. Oxford: Oxford University Press.
- Nazari, A. (2007). EFL teacher's perception of the concept of communicative competence. *ELT Journal* 61(3): 202-210.
- Nesi, H. (1995). A modern bestiary: a contrastive study of the figurative meanings of animal terms. *ELT Journal* 49(3): 272-278.
- Noro, T. (2007). Shouchuu renkei wo meguru kadai [Issues in collaboration between primary and secondary ELT]. In R. Matsukawa & K. Ohshita (eds.). *Shoggakko eigo to chugakkou eigo wo musubu [Linking primary ELT and secondary ELT: Collaboration between primary schools and junior high schools]*. Tokyo: Kouryosha shoten, pp. 102-118.
- Nunan, D. (1988). *Syllabus design*. Oxford: Oxford University Press.
- Nunan, D. (1989). *Designing task for the communicative classroom*. Cambridge: Cambridge University Press.
- Nunan, N. (1994). *Task-based language teaching*. Cambridge: Cambridge University Press.
- Nunan, D. (2004). *Task-based language teaching*. Cambridge: Cambridge University Press.
- O'Dell, F. (1997). Incorporating vocabulary into the syllabus. In N. Schmitt & M. McCarthy (eds.). *Vocabulary: description, acquisition, and pedagogy*. Cambridge: Cambridge University Press, pp. 258-278.
- Ohtsu, Y. (2006a). Koritsugakko niokeru eigokyoiku: giron no genjyo to kongo no kadai [English education at public primary schools: current debate and future problems]. In Y. Ohtsu (ed.). *Nihon no kyoiku ni hitsuyonakoto: shogakko eigo to eigokyoikuseisaku [What is needed for English education in Japan: Primary*

- school English and English education policy*]. Tokyo: Keio University Press, pp. 207-226.
- Ohtsu, Y. (ed.). (2006b). *Nihon no kyoiku ni hitsuyonakoto: shogakko eigo to eigokyoikuseisaku* [What is needed for English education in Japan: Primary school English and English education policy]. Tokyo: Keio University Press.
- Okada, T. (2005) Remarks on poor spellers in English. Available from: www.intcul.tohoku.ac.jp/okada/PDFs/PoorSpellers.pdf [22nd May 2008].
- O'Keeffe, A., McCarthy, M., & Carter, R. (2007). *From corpus to classroom: Language use and language teaching*. Cambridge: Cambridge University Press.
- Oliver, R. (1995). Negative feedback in Child NS/NNS conversation. *Studies in Second Language Acquisition* 17: 459-482.
- Ortega, L., & Long, M. H. (1997). The effects of models and recasts on the acquisition of object topicalization and adverb placement in L2 Spanish. *Spanish Applied Linguistics* 1(1): 65-86.
- Owen, C. (1993). Corpus-based grammar and the Heineken effect: lexico-grammatica description for language learners. *Applied Linguistics* 14(2): 167-187.
- Owen, C. (1996). Do concordancers require to be consulted? *ELT Journal* 50(3): 219-224.
- Pérez-Paredes, P. (2003). Small corpora as assisting tools in the teaching of English news language: A preliminary tokens-based examination of Michael Swan's Practical English Usage news language wordlist. *English for Specific Purposes World* 3(6), Volume 2, 2003 Corpus linguistics. Available from: http://http://www.esp-world.info/articles_6/issue_6.htm [Accessed on 7th January, 2008].
- Paul, D. (2003). *Teaching English to children in Asia*. Hong Kong: Longman Asia ELT.
- Parry, K. (1993). Too many words: learning the vocabulary of an academic subject. In T. Huckin, M. Haynes and J. Coady (eds.). *Second language reading and vocabulary learning*. Norwood, N.J.: Ablex Publishing Cooperation, pp. 109-129.
- Pawley, A., & Syder, F. H. (1983). Two puzzles for linguistic theory: Nativelike selection and nativelike fluency. In J. C. Richards & R. W. Schimdt (eds.). *Language and communication*. London: Longman, pp. 191-225.
- Pearce, M. (2008). Investigating the collocational behaviour of MAN and WOMAN in the BNC using Sketch Engine. *Corpora* 3(1): 1-29.
- Pearson, P.D., Roehler, L.R., Dole, J.A., & Duffy, G.G (1992). Developing expertise in reading comprehension. In J. Samuels and A. E. Farstrup (eds.). *What research has to say about reading instruction*. Second edition. International Reading Association, pp. 145-199.
- Perera, N. (2001). The role of role of prefabricated language in young children's second language acquisition. *Bilingual Research Journal* 25: 327-356.
- Peters, A. (1983). *The units of language acquisition* Cambridge: Cambridge University Press.
- Pitts, M., White, H., & Krashen, S. (1989). Acquiring second language vocabulary through reading: A replication of the clockwork orange study using second language acquirers. *Reading in a Foreign Language* 5: 271-275.
- Prabhu, N.S. (1987). *Second language pedagogy: a perspective*. London: Oxford University Press.

- Pressley, M., Levin, J. R., McDaniel, M. A. (1987). Remembering versus inferring what a word means: Mnemonic and contextual approaches. In M. McKeown and M. Curtis (eds.). *The nature of vocabulary acquisition*. Hillsdale, N.J.: Lawrence Erlbaum, pp. 107-123.
- Prodromou, L. (1988). English as cultural action. *ELT Journal* 42(2): 73-83.
- Renouf, A. & Sinclair, J.M. (1991). Collocational frameworks in English. In K. Aijmer and B. Altenberg (eds.). *English Corpus Linguistics*. London: Longman, pp. 128-144.
- Ricento, K., & Hornberger, H. (1996). Unpeeling the onion: Language planning and policy and the ELT professional. *TESOL Quarterly* 30(3): 401-427.
- Richards, J. (1976). The role of vocabulary teaching. *TESOL Quarterly* 10(1): 77-89.
- Richards, J., & Rodgers, T. (2001). *Approaches and Methods in Language Teaching*. Second edition. Cambridge: Cambridge University Press.
- Rigg, P. (1991). Whole language in TESOL. *TESOL Quarterly* 25(3): 521-542.
- Riley, J. (2006). *Language and literacy: 3-7*. London: Sage Publications.
- Römer, U. (2005). *Progressives, patterns pedagogy: A corpus-driven approach to English progressive forms, functions, contexts and didactics*. Amsterdam: John Benjamins.
- Römer, U. (2009). Corpus research and practice: What help do teachers need and what can we offer? In Aijmer, K. (ed.). *Corpora and language teaching*. Amsterdam: John Benjamins, pp. 83-98.
- Ruppenhofer, J. Baker, C.F. and Fillmore, C.J. (2002). Collocational information in the FrameNet database. In A. Braasch & C. Povlsen (eds.). *Proceedings of the Tenth Euralex International Congress*. Copenhagen, Denmark. Vol. I, pp. 359-369.
- Rayson, P. and Garside, R. (2000). 'Comparing corpora using frequency profiling.' Available from:
http://www.comp.lancs.ac.uk/computing/users/paul/publications/rg_acl2000.pdf
 [Accessed on 27th July, 2009].
- Sampson, G. and McCarthy, D. (eds.). (2004). *Corpus linguistics: Readings in widening discipline*. London: Continuum.
- Sandra, D. (1993). The use of lexical morphology as a natural aid in learning foreign language vocabulary. In J. Chapelle & M. T. Claes (eds.). *Memory and memorization in acquiring and learning languages*. Centre de Langues a Louvain-la-neuve et en Woluwe.
- Sasson, R. (1995). *The acquisition of a second writing system*. Oxford: Intellect.
- Samuda, V. (2005). Expertise in pedagogic task design. In K. Johnson (ed.). *Expertise in second language learning and teaching*. Basingstoke: Palgrave Macmillan, pp. 230-254.
- Scherfer, P. (1993). Indirect L2-vocabulary learning. *Linguistics* 31: 1141-1153.
- Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics* 11: 129-158.
- Schmidt, R. (1993). Consciousness, learning and interlanguage pragmatics. In G. Kasper & S. Blum-Kulka (eds.). *Interlanguage pragmatics*. Oxford: Oxford University Press, pp. 21-42.
- Schmitt, N. (1997). Vocabulary learning strategies. In N. Schmitt & M. McCarthy (eds.). *Vocabulary: Description, acquisition and pedagogy*. Cambridge: Cambridge University Press, pp. 199-227.

- Schmitt, N. (2000). *Vocabulary in language teaching*. Cambridge: Cambridge University Press.
- Schmitt, N. (ed.) (2004). *Formulaic sequences: Acquisition, processing and use*. Amsterdam: John Benjamins.
- Schmitt, N., & Meara, P. (1997). Researching vocabulary through a word knowledge framework: Word associations and verbal suffixes. *Studies in Second Language Acquisition* 19(1): 17-36.
- Schmitt, N., & McCarthy, M. (eds.). (1997). *Vocabulary: Description, acquisition, and pedagogy*. Cambridge: Cambridge University Press.
- Scollon, R. S., & Scollon, W. (1995). *Intercultural communication*. Oxford: Basil Blackwell.
- Scott, M. (2004). *Oxford Wordsmith Tools version 4.0*. [Computer software]. Oxford: Oxford University Press.
- Scott, M., & Tribble, C. (2006). *Textual patterns: Key words and corpus analysis in language education*. Amsterdam: John Benjamins.
- Sealey, A. (1996). *Learning about language: Issues for primary teachers*. Buckingham: Open University Press.
- Sealey, A. (2000). *Childly language: children, language, and the social world*. Harlow: Pearson Education.
- Sealey, A. (2009). Exploring vocabulary with young L1 learners: the contribution of a corpus. In B. Richards, M. H. Daller, D. D. Malvern, P. Meara, J. Milton & J. Treffers-Daller (eds.). *Vocabulary studies in first and second language acquisition: The interface between theory and application*. Basingstoke: Palgrave Macmillan, pp. 39-58.
- Sealey, A., & Thompson, P. (2004). "What do you call the dull words?": Primary school children using corpus-based approaches to learn about language. *English in Education* 38(1): 80-91.
- Sealey, A., & Thompson, P. (2006). "Nice things get said": Corpus evidence and the National Literacy Strategy. *Literacy* 40(1): 22-28.
- Sealey, A., & Thompson, P. (2007). Corpus, concordance, classification: Young learners in the L1 classroom. *Language Awareness* 16 (3): 208-223.
- Seidlhofer, B. (2000). Operationalising intertextuality: Using learner corpora for learning. In L. Burnard & T. McEnery (eds.). *Rethinking language pedagogy from a corpus perspective: Papers from the third international conference on Teaching and Language Corpora*. Frankfurt am Main: Peter Lang, pp. 207-224.
- Shin, D. & Nation, I.S.P. (2008). Beyond single words: The most frequent word in English. *ELT Journal* 62 (4): 339-348.
- Shiozawa, T. (2005). Kanrenbunya no chicken to shogakko eigokyoiku eno shisa [Knowledge from related subjects and suggestions for primary school English education]. In T. Higuchi, T. Kanamori & T. Kunikata (eds.). *Korekarano shogakko eigokyoiku: riron to jittsusen [The future of primary English education: Theory and practice]*. Tokyo: Kenkyusha, pp. 50-73.
- Shu, H., Anderson, R., & Zhang, H. (1995). Incidental learning of word meanings while reading: A Chinese and American cross-cultural study. *Reading Research Quarterly* 30: 76-95.
- Simpson, P. (1993). *Ideology and Point of View*. London: Routledge.

- Sinclair, J. (1985). Selected Issues. In R. Quirk & H. Widdowson (eds.). *English in the world: teaching and learning the languages and literatures*. Cambridge: Cambridge University Press, pp. 248-254.
- Sinclair, J. (1991). *Corpus concordance collocation*. Oxford: Oxford University Press.
- Sinclair, J. (2004). *Trust the text: Language corpus and discourse*. London: Routledge.
- Sinclair, J., Daley, R., & Jones, S. (1970). *English lexical studies*. Report no. 5060 London: OSTI.
- Sinclair, J., & Renouf, A. (1988). A lexical syllabus for language learning. In R. Carter (ed.). *Vocabulary and language teaching*. Essex: Longman, pp. 140-160.
- Singleton, D. (1997). Language and processing vocabulary. *Language Teaching* 30: 213-225.
- Singleton, D. (1999). *Exploring the second language mental lexicon*. Cambridge: Cambridge University Press.
- Skehan, P. (1998). Task-based instruction. *Annual Review of Applied Linguistics* 18: 268-286.
- Skehan, P. (2003). Task-based instruction. *Language Teaching* 36: 1-14.
- Slattery, M., & Willis, J. (2001). *English for Primary Teachers*. Oxford: Oxford University Press.
- Sökmen, A. J. (1997). Current trends in teaching second language vocabulary. In N. Schmitt & M. McCarthy (eds.). *Vocabulary: Description, acquisition, and pedagogy*. Cambridge: Cambridge University Press, pp. 237-257.
- Sripicharn, P. (2002). *Evaluating Data-driven Learning: the use of classroom concordancing by Thai learners of English*. Unpublished PhD Thesis. University of Birmingham, Birmingham, UK.
- Stahl, S. A. & Fairbanks, M. M. (1986). The effects of vocabulary instruction: A model-based meta-analysis. *Review of Educational Research* 56: 72-110.
- Steedman, C. (1982). *The Tidy House: Little girls writing*. London: Virago Press.
- Stephens, J. (1992). *Language and ideology in children's fiction*. Harlow: Addison Wesley Longman.
- Sternberg, R. J. (1987). Most vocabulary is learned from context. In M. G. McKeown & M. E. Curtis (eds.). *The nature of vocabulary acquisition*. Hillsdale, NJ: Erlbaum, pp. 89-105.
- Stevens, V. (1991). Classroom concordancing: Vocabulary materials derived from relevant, authentic text. *English for Specific Purposes* 10: 35-46.
- Stubbs, M. (1996). *Text and corpus analysis*. Oxford: Blackwell.
- Stubbs, M. (2002). *Words and phrases: corpus studies of lexical semantics*. Oxford: Blackwell.
- Stubbs, M. (2007). Quantitative data on multi-word sequences in English: the case of the word *world*. In M. Hoey, M. Mahlberg, M. Stubbs, & W. Teubert (eds.). *Text, discourse and corpora: Theory and analysis*. London: Continuum, pp. 163-189.
- Summers, D. (1996). Computer lexicography: The importance of representativeness in relation to frequency. In J. Thomas & M. Short (eds.). *Using corpora for language research*. London: Longman, p. 260-266.
- Suzuki, T. (1999). *Ninonjin-wa naze eigo-ga dekinaiika. [Why aren't Japanese good at English?]*. Tokyo: Shonchosa.
- Swan, M. (1997). The influence of the mother tongue on second language vocabulary acquisition and use. In N. Schmitt & M. McCarthy (eds.). *Vocabulary:*

- Description, acquisition, and pedagogy*. Cambridge: Cambridge University Press, pp. 156-180.
- Swan, M. (2005a). Legislation by hypothesis: The case of task-based instruction. *Applied Linguistics* 26(3): 376-401.
- Swan, M. (2005b). Rod Ellis, 2003, task-based language learning and teaching (book review). *International Journal of Applied Linguistics* 15(2): 251-256.
- Swaffar, J. (1992). Written texts and cultural readings. In C. Kramsch & S. McConnell-Ginet (eds.). *Text and context: Cross-disciplinary perspectives on language study*. Lexington, MA: D.C. Heath, pp. 238-250.
- Takahashi, M. (2000). *Eigo kyoikugaku gairon*. [Introduction to English language education]. Tokyo: Kinseido.
- Takahashi, K. (2010). 'Eigo Note' wo dou katsuyousuruka: sono chousho to mondaiten wo fumaete. [How can we make use of 'Eigo Note': Benefits and problems of 'Eigo Note']. *The English Teachers' Magazine* 59 (9): 10-12.
- Tanko, G. (2004). The use of adverbial connectors in Hungarian university students' argumentative essays. In J. Sinclair (ed.). *How to use corpora in language teaching*. Amsterdam: John Benjamins, pp. 157-184.
- Taylor, B. P. (1982). In search of real reality. *TESOL Quarterly* 16 (1): 29-42.
- Terauchi, M. (1994). Keitaiso no shutoku [Acquisition of morphemes]. In I. Koike (ed.). *Daini gengo syutoku kenkyu ni motodoku saishin no eigokyouiku* [Current English language teaching based on SLA research]. Tokyo: Taisyukan, pp.24-39.
- Teubert, W. (2001). Corpus Linguistics and Lexicography. *International Journal of Corpus Linguistics* 6: 125-153.
- Thompson, P. (2000). Citation practices in PhD theses. In L. Burnard & T. McEnery (eds.). *Rethinking language pedagogy from a corpus perspective: Papers from the third international conference on Teaching and Language Corpora*. Frankfurt am Main: Peter Lang, pp. 91-102.
- Thompson, P., & Sealey, A. (2007). Through children's eyes? Corpus evidence of the features of children's literature. *International Journal of Corpus Linguistics* 12(1): 1-23.
- Thorndike, E. L., & Lorge, I. (1944). *The teacher's word book of 30,000 words*. New York Teacher's College: Columbia University.
- Tomlinson, B. (2003). Materials development courses. In B. Tomlinson (ed.). *Developing materials for language teaching*. London: Continuum, pp. 445-461.
- Tono, Y. (2002). *The Role of Learner Corpora in SLA Research and Foreign Language Teaching: The Multiple Comparison Approach*. Unpublished PhD thesis. Lancaster University, Lancaster, UK.
- Torikai, K. (2006a). *Ayaushi! Shogako Eigo* [English education in primary school at risk]. Tokyo: Bungeishunju.
- Torikai, K. (2006b). Jizokukanouna miraihen communication kyoiku [Feasibility of practical and continuous communication education]. In Y. Ohtsu (ed.). *Nihon no kyoiku ni hitsuyonakoto: shogakko eigo to eigokyoikuseisaku* [What is needed for English education in Japan: Primary school English and English education policy]. Tokyo: Keio University Press, pp. 135-151.
- Tribble, C. (2001). Small corpora and teaching writing. In M. Ghadsessy, A. Henry & R. L. Roseberry (eds.). *Small corpus studies in ELT: theory and practice*. Amsterdam: John Benjamins, pp. 71-92.

- Tribble, C. (2000). Genres, keywords, teaching: Towards a pedagogic account of the language of project proposals. In L. Burnard & T. McEnery (eds.). *Rethinking language pedagogy from a corpus perspective: Papers from the third international conference on Teaching and Language Corpora*. Frankfurt am Main: Peter Lang, pp. 75-90.
- Trousdale, A. (2007). The role of literature in children's spiritual development. In M. d. Souza, G. Durka, K. Engebretson, R. Jackson & A. McGrady (eds.). *International Handbooks of Religion and Education*. Amsterdam: Springer Netherlands, pp. 1225-1235.
- Tsui, A. M. (2004). What teachers always wanted to know- how corpora can help. In Sinclair, J. (ed.). *How to use corpora in language teaching*. Amsterdam: John Benjamins, pp. 39-61.
- Tsui, A. M. (2003). *Understanding expertise in teaching: Case studies of second language teachers*. Cambridge: Cambridge University Press.
- Tyler, A. & Evans, V. (2003). *The semantics of English prepositions: Spatial scenes, embodied meaning and cognition*. Cambridge: Cambridge University Press.
- Van Ek, J.A. (1975). *The Threshold Level in a European Unit: Credit system for modern language learning by adults*. Systems Development in Adult Language Learning. Strasbourg: Council of Europe.
- Van Ek, J.A., & Alexander, L.G. (1977). *Waystage*. Strasbourg: Council of Europe.
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological process*. Cambridge, MA: Harvard University Press.
- Wade, B. (ed.). (1990). *Reading for real*. Buckingham: Open University Press.
- Weinert, R. (1995). The role of formulaic language in second language acquisition: A review. *Applied Linguistics* 16: 180-205.
- Wall, B. (1991). *The Narrator's Voice: the dilemma of children's fiction*. London: Macmillan.
- Wang, C. and Dowker, A. (2008). Interpretation of animal metaphors: Evidence from Chinese and English children and adults. *Proceedings of Child Language Seminar 2007*: 226-235.
- Waring, R., & Takaki, M. (2003). At what rate do learners learn and retain new vocabulary from reading a graded reader? *Reading in a Foreign Language* 15: 130-163.
- Webb, S. (2008). The effects of context on incidental vocabulary learning. *Reading in a Foreign Language* 20: 232-245.
- West, M. (1930). Speaking-vocabulary in a foreign language. *Modern Language Journal* 14: 509-521.
- West, M. (1953). *A general service list of English words*. London: Longman.
- White, R. (1988). *The ELT curriculum: Design, innovation and management*. Oxford: Blackwell.
- Widdowson (1979). The communicative approach and its applications. In H. G. Widdowson (ed.). *Explorations in Applied Linguistics*. Oxford: Oxford University Press, pp. 251-264.
- Widdowson, H. (2000). On the limitation of linguistics applied. *Applied Linguistics* 21(1): 3-25.
- Widdowson, H. (2003). *Defining issues in English language teaching*. Oxford: Oxford University Press.
- Wilkins, D. A. (1972). *Linguistics in language teaching*. London: Edward Arnold.

- Wilkins, D. A. (1976). *Notional Syllabuses*. Oxford: Oxford University Press.
- Williams, M. (1988). Language taught for meetings and language used in meetings: is there anything in common? *Applied Linguistics* 9 (1): 45-58.
- Williams, G. C. (1998). Collocational networks: Interlocking patterns of lexis in a corpus of plant biology research articles. *International Journal of Corpus Linguistics* 3: 151-171.
- Willis, D. (1990). *The lexical syllabus: A new approach to language teaching*. London: Collins ELT.
- Willis, D. (1993). Grammar and lexis: some pedagogical implications. In J. M. Sinclair, M. Hoey & F. Gwyneth (eds.). *Techniques of Description: spoken and written discourse. A festschrift for Malcolm Coulthard*. London: Routledge, pp. 83-93.
- Willis, D. (1994). A lexical approach. In M. Bygate, A. Tonkyn & E. Williams (eds.). *Grammar and the Language Teacher*. Hemel Hempstead: Prentice Hall, pp. 55-66.
- Willis, D. (2003). *Rules, patterns and words: Grammar and lexis in English language teaching*. Cambridge: Cambridge University Press.
- Willis, J. (1996). *A framework for task-based learning*. Harlow: Longman.
- Willis, J. (2009). The TBL Framework. In C. Van den Branden, M. Bygate, & J. Norris (eds.). *Task-Based language teaching: A reader*. Amsterdam: John Benjamins, pp. 227-242.
- Willis, J., & Willis, D. (1987-8). *Collins COBUILD Course*. Collins.
- Willis, D. & Willis, D. (2007). *Doing task-based teaching*. Oxford: Oxford University Press.
- Wray, A. (2000). Formulaic sequences in second language teaching: practice and principle. *Applied Linguistics* 21(4): 463-489.
- Wray, A. (2002) *Formulaic language and the lexicon*. Cambridge: Cambridge University Press.
- Wray, A., & Perkins, M. R. (2000). The functions of formulaic language: an integrated model. *Language and Communication* 20(1): 1-28.
- Wynne, M (ed.). (2005). *Developing linguistic corpora: a guide to good practice*. Oxford: Oxbow Books. Available from: <http://ahds.ac.uk/linguistic-corpora/> [1st May 2008].
- Wyse, D. (2007). *How to help your child read and write*. Harlow: Pearson Education.
- Yanagise, Y. (2006). Eigo kyoiku no genri ni tsuite [About the principles of English Language education] In Y. Ohtsu (ed.). *Nihon no kyoiku ni hitsuyonakoto: shogakko eigo to eigokyoikuseisaku [What is needed for English education in Japan: Primary school English and English education policy]*. Tokyo: Keio University Press, pp. 33-52.
- Yoshida, K. (2003). *Atarashii eigo kyouiku heno challenge: shogakusei kara eigo wo oshieru tameni [A new challenge in English education: Introducing English in primary schools]*. Tokyo: Kumon.
- Zimmerman, C. B. (1997). Historical trends in second language vocabulary instruction. In J. Coady & T. Huckin (eds.). *Second language vocabulary acquisition: A rationale for pedagogy*. Cambridge: Cambridge University Press, pp. 5-19.


APPENDICES

Appendix 2.1 Sample of 'Eigo Note'

Lesson 6


What do you want?

外来語を知ろう



Let's Listen


CDを聞いて、どの
絵かさがそう。




Activity

身近な外来語を
さがそう。


①




②




③




④




⑤



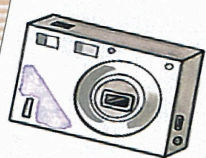
⑥



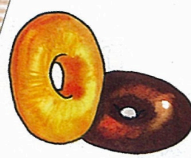
⑦




⑧



⑨






Let's Play ①

キー・ワード・ゲーム

ペアでやってみよう。




Let's Play ②

おはじきゲーム

おはじきを使って
やってみよう。

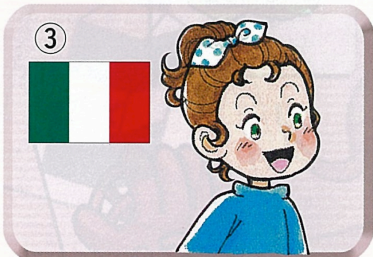
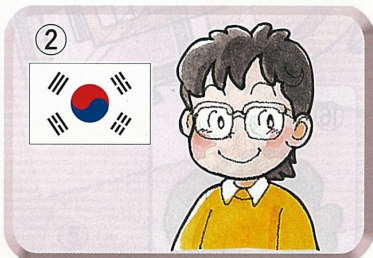
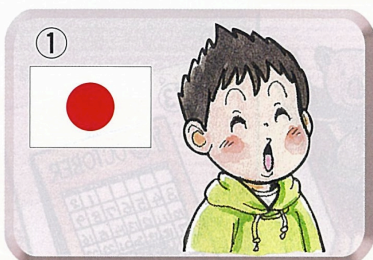
36



From: 'Eigo Note' Book 1 Lesson 6 (MEXT, 2009a p.36)

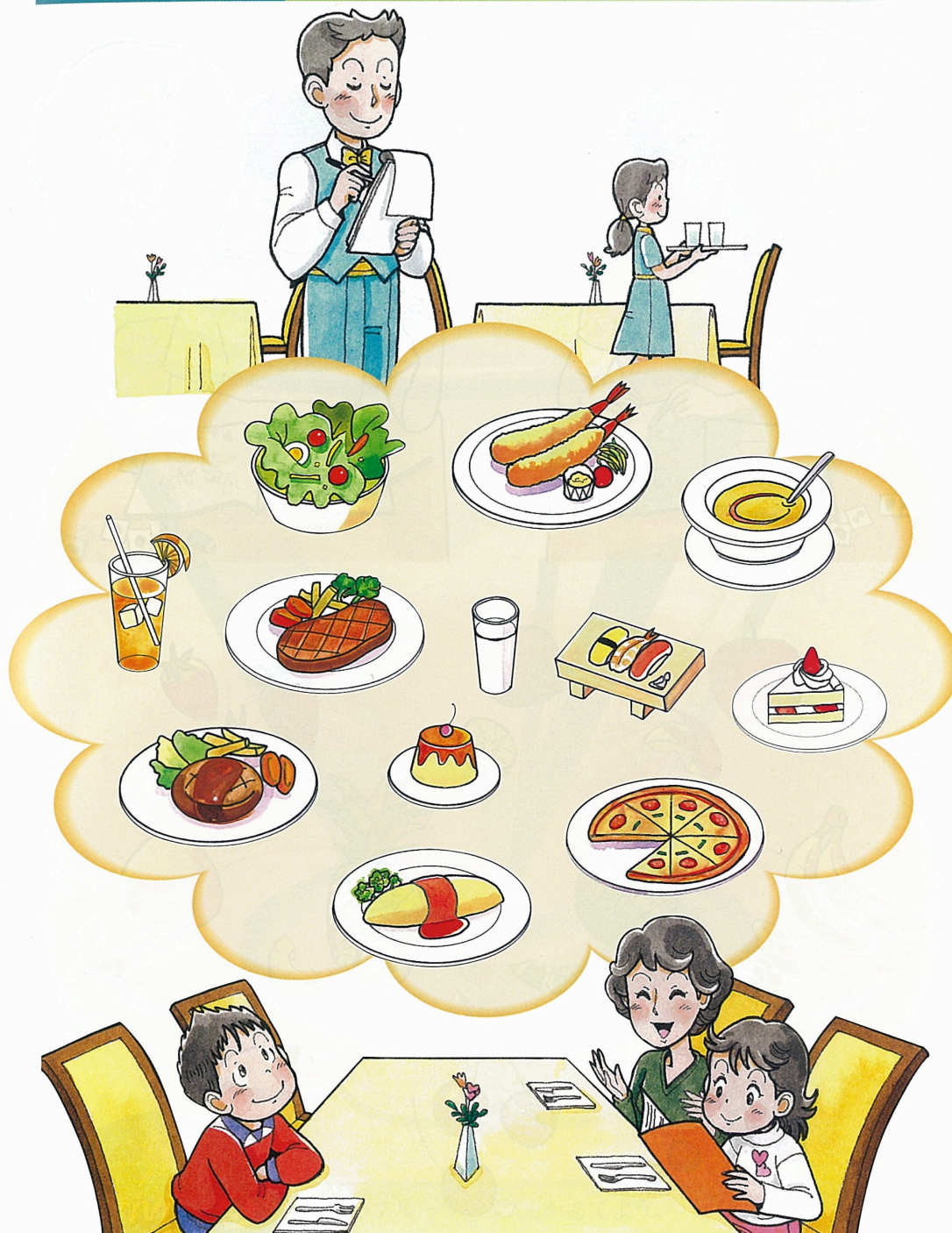
Let's Listen 1

CDを聞いて、どこの国からきた食べ物か考えて、線で結ぼう。



Let's Listen 2

CDを聞いて、健が注文したものを○で、
麻衣が注文したものを□で囲もう。



From: 'Eigo Note' Book 1 Lesson 6 (MEXT, 2009a p.39)

Appendix 4.1 Translated questionnaire Survey I



UNIVERSITY OF
BIRMINGHAM

Eri Hirata
Department of English
University of Birmingham
Edgbaston, Birmingham, UK
B15 2TT
Email:

Date

Dear Participant

Thank you for agreeing to co-operate with my research into the characteristics of children's literature.

I am interested in finding out what kind of books (storybooks, picture books, reading book series) you use in the practical education setting for literacy work in the classroom, and the reasons for your choice of materials. If you would kindly complete the attached questionnaire, and e-mail it to the above address I would be most grateful.

I would like to reassure you that by answering the questionnaire, the resulting answers will be used for research purposes only. The research will be anonymous and no one identified by name. Any personal information you give, such as your name or email address, will remain confidential.

If you have any questions regarding this research, please feel free to contact me. Your participation in the study is most appreciated and I would be more than happy to let you know the results after finishing the study.

Yours sincerely,

Eri Hirata

Questionnaire

Q1: In what kind of institutions/ schools do you teach? Please select (put an x in all boxes that apply).

<input type="checkbox"/>	Public	<input type="checkbox"/>	Private	<input type="checkbox"/>	Juku	<input type="checkbox"/>	Other ()
--------------------------	---------------	--------------------------	----------------	--------------------------	-------------	--------------------------	-----------------------------

Q2: Have you ever attended any training programmes before starting to teach English at primary level?

<input type="checkbox"/>	YES	<input type="checkbox"/>	NO
--------------------------	------------	--------------------------	-----------

Q3: If you answered “Yes” to Q2, please name the institution or organization which gave you training.

Name of the institution /organization.	
---	--

Q4: Which age group of students do you teach? Please select all that apply.

<input type="checkbox"/>	a. 0-3	<input type="checkbox"/>	b. 4-5	<input type="checkbox"/>	c. 6-7	<input type="checkbox"/>	d. 8-10	<input type="checkbox"/>	e. 11-13
--------------------------	---------------	--------------------------	---------------	--------------------------	---------------	--------------------------	----------------	--------------------------	-----------------

Q5: Do you use storybooks in your classroom?

<input type="checkbox"/>	YES	<input type="checkbox"/>	NO
--------------------------	------------	--------------------------	-----------

Q6: a) If you answered “YES” to Q5, please look at the list of titles below, and indicate which books you are familiar with by putting “X” in the first left hand column.

b) Please also indicate which books you have used by putting “X” in the second left hand column.

a) Familiar	b) Used	Title of Graded Readers
		Oxford Reading Tree (Stage 1, 1+, 2-16)
		Oxford Classic Tales
		Oxford Bookworms Library (Starters – Stage6)
		Oxford Bookworms Factfiles (Stage 1-5)
		Oxford Dominoes (Starter – Level 3)



		Oxford Start with English Readers (Grade 1-6)
		Longman Literacy Land (Level 1-13)
		Longman Chatterbox
		Penguin Young Readers (Level 1-4)
		Penguin Readers (Easy starts-Level 6)
		Puffin Easy-to-Read (Level 1-3)
		All Aboard Reading (Level 1-3)
		I Can Read Book (My First, Level 1-4)
		Step into Reading (Level 1-5)
		Scholastic Reader (Level 1-4)
		McMillan New Wave Readers (Level 1-5)
		McMillan Guided Readers (Starter – Upper)
		Cambridge English Readers (Level 1-6)
		Houghton Mifflin Leveled Readers (Grade 1-6)
a) Familiar	b) Used	Title of English Book Series
		Longman Shared Reading
		Powerpuff Girls Series
		Time-to-Discover Readers
		100 English
		Sight Word Readers
		Dr. Seuss Series
		Gaspard and Lisa Series
		Big Red Reader (Clifford series) penguin p
		Goosebumps series
		Catwings series
		Cobble Street Cousins
		Let's-Read-And-Find-Out-Science
		Moonbear series
		Nate the Great series
		A-Z Mysteries
		Marvin Redpost series
		MAXimum Boy series
		The Zack Files series
		The Boxcar Children series



		Full House Michelle series
		Jake Drake series
		Judy Moody series
		The Julian Stories series
		Magic Tree House
		My Father's Dragon
		Nancy Drew Notebooks
		The Worst Witch series
		Akiko series
		Baby-Sitters Club series
		Captain Underpants series
		The Famous Five series
		Horrid Henry series
		Ramona series
		The Secrets of Droon series
		Animorphs series
		Deltora Quest series
		A Series of Unfortunate Events
		The Saga of Darren Shan
		Rowan of Rin
		Rookie Readers
		TIME for Kids (newspaper for children)
		Mr Men series (Roger Hargreaves)
		Little Miss series (Roger Hargreaves)
a) Familiar	b) Used	Title of English individual books
		Curious George (Short stories)
		Clifford Board Books (Norman Bridwell)
		It's Mine! (Leo Lionni)
		The Giving Tree (Shel Silverstein)
		Dustbin Baby (Jacqueline Wilson)
		Lizzie Zipmouth (Jacqueline Wilson)
		Frindle (Andrew Clements)
		Little House in the Big Woods (Laura Ingalls Wilder)
		Number the Stars (Louis Lowry)



		Pippi Longstocking (Astrid Lindgren)
		There's a Boy in the Girl's Bathroom (Louis Sachar)
		Tales of a Fourth Grade Nothing (Judy Blume)
		The Velveteen Rabbit (Margery Williams)
		Charlotte's Web (E.B. White)
		Emil and the Detectives (Erich Kastner)
		The Enormous Crocodile (Roald Dahl)
		The Magic Finger (Roald Dahl)
		Charlie and the Chocolate Factory (Roald Dahl)
		Matilda (Roald Dahl)
		James and the Giant Peach (Roald Dahl)
		The Witches (Roald Dahl)
		The Big Friendly Giant (Roald Dahl)
		Harry Potter
		The Lord of the Rings
a) Familiar	b) Used	OTHERS (If there are any books you are familiar with or you have used that are not on the list, please give details below.)

Q7: Do you use Japanese stories translated into English in your classroom?

	YES		NO
--	-----	--	----

Q8: If you answered "YES" to Q7, please list these.

	Translated Japanese books or texts
1	
2	



3	
4	
5	
6	
7	

Q9: If you answered “NO” to Q7, would you use translated Japanese storybooks or texts if they were available?

	YES		NO
--	-----	--	----

Q10: Please would you clarify the reasons why you chose these specific reading materials, and what other factors influence your choices generally.

--

Q11: Please give your opinion regarding the introduction of English literacy work to children at primary school level.

--



UNIVERSITY OF
BIRMINGHAM

Thank you for your co-operation, your feedback will really help with my research. I would be grateful if you could supply the following information for research purposes and future reference. I can confirm again that the research will be anonymous and no one will be identified by name. Any information you give below will remain confidential.

Name:	
Nationality:	
Country of workplace:	
English teaching experience	Primary school, Junior-high school, High school, Other ()
Email:	
Would you like to receive my research results via email? (Yes/ No)	

FUKUOKA JO GAKUIN UNIVERSITY Faculty of Humanities	
Dear Teachers	
<i>Thank you for agreeing to answer the questionnaire. It should take 10-15 minutes to complete. Please rest reassured that the results obtained through this survey will be used only for the purpose of research and the improvement of the training course, and no individual or school will be identified.</i>	
A. First of all, please provide answers to the following information about yourself.	
Q1:	How many years of experience do you have (a) in teaching at primary schools in general; and b) in conducting English lessons?
(a).	(years)
(b).	(years)
Q2:	To what extent are you confident in using English?
	1. not confident at all, 2. not very confident, 3. cannot say, 4. fairly confident, 5. very confident
Q3:	In terms of your English ability, which of the following do you feel you need to improve? Please select all that apply.
a.	speaking skills
b.	listening skills
c.	pronunciation skills
d.	reading skills
e.	writing skills
f.	vocabulary
g.	grammatical knowledge
h.	others ()
B. Please answer the following questions regarding teacher training.	
Q4:	Have you attended any training workshops for teaching English?
a.	Yes
b.	No
Q5:	So far, what kind of training have you received? Please select all that apply.
a.	material design for teaching young learners
b.	teaching methods & techniques
c.	theories about second language acquisition
d.	educational and language policy
e.	lesson planning
f.	observation at schools
g.	others (Please specify below)
Q6:	How many times did you attend the training sessions?
a.	once
b.	twice
c.	three times or more

TEYL course



Q7:	What kind of frustrations or problems do you have in conducting English lessons? Please provide details below.
Q8:	What are the most important things that prospective teachers need to learn at university before starting to teach English? Please select all that apply.
	<ul style="list-style-type: none">a. development of children's psychologyb. knowledge about policyc. lesson planning and teaching demonstrationsd. teaching methods and techniquese. general English skillsf. syllabus design and material designg. others (please specify below)

Thank you for your cooperation.

Eri Hirata
Department of English
Faculty of Humanities
Fukuoka Jo Gakuin University

Appendix 4.3 Translated questionnaire Survey III (with prospective teachers)

FUKUOKA JO GAKUIN UNIVERSITY
Faculty of Humanities



A. Reflect on your experience in TEYL course and self-evaluate.	
While taking the TEYL training course, I have...	<i>unsatisfactory</i> <i>satisfactory</i>
1. worked hard on all the assignments and teaching	1-----2-----3-----4-----5
2. cooperated with other prospective teachers	1-----2-----3-----4-----5
3. worked hard to improve my English skills	1-----2-----3-----4-----5
B. Please answer the following about teaching experience	
Q1:	What were the most difficult aspects of teaching pupils?
	a. communication with pupils b. time management c. the use of activities which suit pupils d. keeping pupils' attention. e. dealing with unexpected circumstances f. using English in class g. others (please provide the details)
Q2:	What were the most useful aspects of training course when you taught pupils? (Multiple choice permitted)
	a. examples of activities used in classes at primary schools b. children's development c. foreign cultures d. use of classroom English e. current situations and issues in TEYL in Japan and other countries f. teaching practice at primary schools g. teaching methods suitable for pupils h. teaching methods suitable for different skills i. observations at primary schools j. lesson planning & teaching demonstrations k. theories about second language acquisition l. material design & syllabus design m. educational and language policy n. phonology o. others (Please provide the details)
Q3:	What are the aspects of knowledge or skills you felt were lacking when you did teaching practice? Please provide the details below.

TEYL course



Q4:	What do you think prospective teachers, such as you, should learn before going on to do teaching practice?
	<ul style="list-style-type: none">a. development of children's psychologyb. knowledge about policyc. lesson planning & teaching demonstrationsd. teaching methods and techniquese. general English skillsf. syllabus design & material designg. others (please specify below)
Q5:	Are there any aspects which you think positively changed through taking this course? What are they? And why do you think the positive change happened?

Thank you for your cooperation & I hope you enjoyed the course!

Appendix 4.4 Reflection assignment given to prospective teachers

Name:

ID:

Group:

FUKUOKA JO GAKUIN UNIVERSITY



Last Week, I was a Teacher...

Write about your experience of teaching practice.

Submission Deadline:

What was the most enjoyable part of teaching practice?	
What was the most difficult part of teaching practice?	
Describe your relationship with the pupils in your school.	
Describe your relationship with the teachers in your school.	
What was the most unexpected thing that happened?	
What was the most successful thing that you did?	
Did anything go wrong? Write about it.	

Name:
ID:
Group:

FUKUOKA JO GAKUIN UNIVERSITY



What kind of teaching <u>methods</u> did you use?	
Was (were) the chosen method(s) appropriate for the class you taught?	
In what ways has your attitude towards teaching changed?	
What have you learnt from this teaching experience?	
What advice would you give to somebody about to start teaching practice?	



Who do you think had the best experience during teaching practice?

Why do you think so?

Appendix 4.5 Letter sent requesting permission



UNIVERSITY OF
BIRMINGHAM

Eri Hirata
English Department
University of Birmingham
Edgbaston, Birmingham
B15 2TT
Email

Dates

Dear Sir/ Madam

I am a research student in the English Department at the University of Birmingham, where I am researching the teaching of English to young learners in Japan. I am planning to construct a corpus of current children's literature (including reading books and series), which would involve storing the texts, or parts of texts, in an electronic format, and all text will be deleted after I conclude my research. I would very much appreciate it if you would grant me permission to include the texts listed in the attached sheet in my corpus.

I should stress that my interest is solely in research and therefore non-commercial. No other person will have access to these stored texts. If you have any questions about my research and what is involved using the texts in this way, please do not hesitate to get in touch. My supervisor at the university is Dr. Alison Sealey, who can be contacted at the above address.

I shall be happy to share the results of my research with you due course. I hope it will be useful to teachers, children's authors and others involved in texts for young learners.

I look forward to hearing from you.

Best wishes,

Eri Hirata

Enc.

Contact address: Eri Hirata,
Email:

Appendix 5.1 List of texts identified as ‘used’ or ‘familiar’ by the respondents

a) Familiar	b) Used	Title of books/ series/ graded readers
45	1	Oxford Reading Tree (Stage 1, 1+, 2-16)
32		Longman Literacy Land (Level 1-13)
4	1	Penguin Young Readers (Level 1-4)
65		Penguin Readers (Easy starts-Level 6)
12		Puffin Easy-to-Read (Level 1-3)
32		I Can Read Book (My First, Level 1-4)
3		Scholastic Reader (Level 1-4)
9		Dr. Seuss Series
56	2	Gaspard and Lisa Series
27	6	Big Red Reader (Clifford series)
12	3	Nate the Great series
29		Full House Michelle series
5		The Famous Five series
31	2	Mr Men series
27	2	Little Miss series
62	3	Curious George (Short stories)
12	4	Clifford Board Books
27	2	It's Mine!
12	2	The Giving Tree
68		Little House in the Big Woods
54		Pippi Longstocking
14	1	Charlotte's Web
8	1	The Enormous Crocodile
6		The Magic Finger
75	1	Charlie and the Chocolate Factory
18		Matilda
7	1	James and the Giant Peach
12		The Witches
5		The Big Friendly Giant
92		Harry Potter
89		The Lord of the Rings

Appendix 5.2 List of texts additionally mentioned by the respondents

Recommended texts	No.	Recommended texts	No.
Guri and Gura	8	Grandma Baba Wants Sunshine!	3
Guri and Gura's Seaside Adventure	7	Grandma Baba's Big clean-up!	3
Guri and Gura's Surprise Visitor	8	Peach boy	11
Guri and Gura's Spring Cleaning	8	Little Red Riding Hood	5
Guri and Gura's Playtime Book of Seasons	6	Little Prince	1
Guri and Gura's Song's of the seasons	6	The Giving Tree	1
Guri and Gura's Special Gift	8	Harry, the Dirty Dog	1
Big Beanie and the Lost Fish	3	The Happy Day	1
Blackie, the Crayon	6	The Rabbit's Wedding	1
Chugging all the Way	8	Go away Big Green Monster	1
Little Mouse's Red Vest	14	Cinderella	1
Naomi's Special Gift	9	Fantastic Mr. Fox	1
The Ainu and the Fox	2	Balloon Cat	4
Katie and Boo	3	Chug-chug Train	2
Jeeper the Fire Engine	7	Good Evening, Dear Moon	1
Groompa's Kindergarten	2	Don't Want to Go to Bed?	1
Amy and Ken Visit Grandma	12	Suho's White Horse: A Mongolian legend	19
Sid the Signal	2	Japanese Children's Favorite Stories 1 **	2
Where's the Fish?	13	Japanese Children's Favorite Stories 2 **	2
Blue Mouse, Yellow Mouse	3	The Snowman	3
Elephee's Walk	2	Elmer and Rose (Elmer Series)	4
Sudsy Mom's Washing Spree	4	The Bad-Tempered Ladybird	6
Miki's First Errand	5	The Very Hungry Caterpillar	26
Peek-a-Boo	4	From Head to Toe	25
The Giving Chair	16	Brown Bear, Brown Bear, What Do You see?	23
Barney and the Kitten	2	We're Going on a Bear Hunt	1
Barney's First Snow	2	Days With Frog and Toad	5
The Rainy Trip Surprise	1	Frog and Toad Are Friends	12
Sleepy Time	2	Frog and Toad Together	8

* no. indicates the number of times mentioned by the respondents. ** these books include several stories.

Appendix 6.1 List of keywords extracted from the comparison of JTCL and EOCL

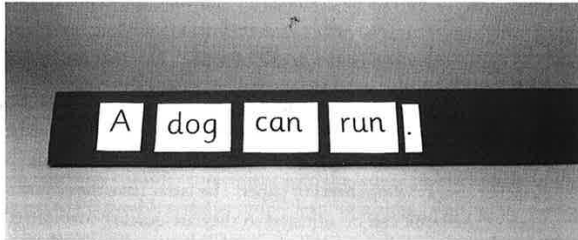
N	Key word	JTCL as TC		EOCL as RC			
		Freq.	%	Freq.	RC. %	LL	P
1	GURI	87	0.19	0		370.5	2.1E-17
2	GURA	79	0.17	0		336.4	2.8E-17
3	SAYS	108	0.24	37	0.01	304.6	4.03E-17
4	RICE	72	0.16	4		276.3	5.62E-17
5	KEN	57	0.12	0		242.7	8.8E-17
6	MONKEY	90	0.20	42	0.01	228.8	1.08E-16
7	CHUG	52	0.11	0		221.4	1.21E-16
8	MAN	156	0.34	217	0.06	212.4	1.41E-16
9	KINTARO	44	0.09	0		187.3	2.22E-16
10	THE	3071	6.86	17531	5.29	176.6	2.75E-16
11	CLOTH	50	0.11	6		176.3	2.78E-16
12	FINALLY	62	0.13	23		170.6	3.14E-16
13	OLD	216	0.48	513	0.15	164.0	3.64E-16
14	AMY	38	0.08	0		161.8	3.83E-16
15	BARNEY	35	0.07	0		149.0	5.25E-16
16	SEA	60	0.13	35	0.01	139.3	6.84E-16
17	SUHO	32	0.07	0		136.2	7.48E-16
18	SO	376	0.84	1345	0.40	135.3	7.68E-16
19	DOG	49	0.10	19		132.9	8.27E-16
20	SABURO	30	0.06	0		127.7	9.71E-16
21	MOUNTAIN	46	0.10	18		124.3	1.08E-15
22	ELF	29	0.06	0		123.4	1.12E-15
23	HE	884	1.97	4293	1.29	120.5	1.23E-15
24	BABA	28	0.06	0		119.2	1.29E-15
25	TURTLE	40	0.08	12		117.1	1.39E-15
26	GROOMPA	27	0.06	0		114.9	1.51E-15
27	SPARROW	30	0.06	2		113.2	1.6E-15
28	VERY	312	0.69	1124	0.33	110.3	1.8E-15
29	SQUASH	31	0.06	4		108.1	1.96E-15
30	MOMOTARO	25	0.05	0		106.4	2.1E-15
31	MIKI	25	0.05	0		106.4	2.1E-15
32	SKY	64	0.14	71	0.02	103.7	2.36E-15
33	LET'S	62	0.13	69	0.02	100.2	2.75E-15
34	BEAR	44	0.09	28		98.2	3.02E-15
35	KITTY	25	0.05	1		98.2	3.02E-15
36	TARO	23	0.05	0		97.9	3.06E-15
37	NAOMI	23	0.05	0		97.9	3.06E-15
38	MORTAR	23	0.05	0		97.9	3.06E-15
39	YELLOW	51	0.11	44	0.01	97.1	3.18E-15
40	FOREST	38	0.08	20		92.1	4.07E-15
41	CAKES	37	0.08	19		90.6	4.41E-15
42	EEL	21	0.04	0		89.4	4.7E-15
43	MOM	21	0.04	0		89.4	4.7E-15
44	JEEPER	21	0.04	0		89.4	4.7E-15
45	PRINCESS	52	0.11	54	0.01	88.2	5.02E-15
46	FIREWOOD	22	0.04	1		85.7	5.79E-15

47	SING	34	0.07	17		84.1	6.34E-15
48	BROTHER	37	0.08	24		81.8	7.3E-15
49	FIRE	58	0.12	80	0.02	79.4	8.51E-15
50	BLUE	65	0.14	102	0.03	79.3	8.55E-15
51	BOAR	18	0.04	0		76.6	1.03E-14
52	HOME	113	0.25	298	0.08	73.2	1.32E-14
53	JUNKMAN	17	0.03	0		72.3	1.41E-14
54	BIG	108	0.24	280	0.08	71.9	1.47E-14
55	SID	16	0.03	0		68.1	2.01E-14
56	YOSAKU	16	0.03	0		68.1	2.01E-14
57	SHIRO	16	0.03	0		68.1	2.01E-14
58	WOMAN	59	0.13	100	0.03	66.8	2.26E-14
59	WONDERFUL	50	0.11	71	0.02	66.8	2.26E-14
60	CLOUDS	22	0.04	6		66.0	2.42E-14
61	KETTLE	22	0.04	6		66.0	2.42E-14
62	ISLAND	25	0.05	11		64.9	2.71E-14
63	STRAW	23	0.05	8		64.5	2.8E-14
64	WEAVING	18	0.04	2		64.1	2.92E-14
65	RAZOR	15	0.03	0		63.8	3E-14
66	FOOTPRINTS	15	0.03	0		63.8	3E-14
67	INCH	24	0.05	10		63.5	3.11E-14
68	HIS	438	0.97	2099	0.63	62.7	3.36E-14
69	MEW	14	0.03	0		59.6	4.76E-14
70	ELLIE	14	0.03	0		59.6	4.76E-14
71	HORSE	44	0.09	63	0.01	58.3	5.52E-14
72	RED	72	0.16	164	0.04	57.8	5.93E-14
73	WHITE	72	0.16	165	0.04	57.3	6.29E-14
74	DANCE	30	0.06	26		56.9	6.59E-14
75	HARD	51	0.11	88	0.02	56.7	6.81E-14
76	JAPAN	15	0.03	1		56.6	6.89E-14
77	CHARCOAL	15	0.03	1		56.6	6.89E-14
78	BEAUTIFUL	59	0.13	117	0.03	56.3	7.15E-14
79	RABBIT	24	0.05	14		55.7	7.81E-14
80	TEMPLE	13	0.02	0		55.3	8.22E-14
81	SANDALS	13	0.02	0		55.3	8.22E-14
82	PRIEST	13	0.02	0		55.3	8.22E-14
83	AND	1769	3.95	10860	3.27	52.9	1.18E-13
84	WEAVE	14	0.03	1		52.5	1.26E-13
85	WINE	16	0.03	3		52.3	1.3E-13
86	SHOBEI	12	0.02	0		51.0	1.6E-13
87	MOMMY	12	0.02	0		51.0	1.6E-13
88	LOBSTER	12	0.02	0		51.0	1.6E-13
89	BEANIE	12	0.02	0		51.0	1.6E-13
90	CHUBBY	12	0.02	0		51.0	1.6E-13
91	BUMBUKU	12	0.02	0		51.0	1.6E-13
92	CASEY	12	0.02	0		51.0	1.6E-13
93	FISHY	12	0.02	0		51.0	1.6E-13
94	BLACKIE	12	0.02	0		51.0	1.6E-13
95	ROBE	13	0.02	1		48.3	2.68E-13

96	CRAB	13	0.02	1		48.3	2.68E-13
97	JELLYFISH	13	0.02	1		48.3	2.68E-13
98	MOUNTAINS	22	0.04	15		47.5	3.24E-13
99	BOO	17	0.03	6		47.5	3.25E-13
100	CARRYING	33	0.07	43	0.01	47.3	3.35E-13
101	SUDSY	11	0.02	0		46.8	3.78E-13
102	KITTEN	11	0.02	0		46.8	3.78E-13
103	ZOOF	11	0.02	0		46.8	3.78E-13
104	GOVERNOR	11	0.02	0		46.8	3.78E-13
105	MISUZU	11	0.02	0		46.8	3.78E-13
106	FRIENDS	42	0.09	74	0.02	45.7	4.97E-13
107	DAY	139	0.31	516	0.15	45.5	5.23E-13
108	BEGAN	87	0.19	261	0.07	45.2	5.69E-13
109	SURPRISED	27	0.06	29		44.7	6.44E-13
110	BABY	28	0.06	33		43.4	9.54E-13
111	I'LL	72	0.16	199	0.06	43.2	1.02E-12
112	PINE	17	0.03	8		43.0	1.07E-12
113	NEW	47	0.10	97	0.02	42.8	1.17E-12
114	ASKS	15	0.03	5		42.6	1.24E-12
115	WARRIORS	10	0.02	0		42.5	1.27E-12
116	CHUNG	10	0.02	0		42.5	1.27E-12
117	BASKET	22	0.04	19		41.8	1.64E-12
118	NAMED	13	0.02	3		40.6	2.7E-12
119	PA	11	0.02	1		40.2	3.35E-12
120	FISHERMAN	11	0.02	1		40.2	3.35E-12
121	OGRES	12	0.02	2		40.1	3.5E-12
122	PHEASANT	12	0.02	2		40.1	3.5E-12
123	SAW	87	0.19	280	0.08	39.4	5.05E-12
124	JIZO	9	0.02	0		38.3	9.93E-12
125	CRAYONS	9	0.02	0		38.3	9.93E-12
126	KATIE	9	0.02	0		38.3	9.93E-12
127	WEN	9	0.02	0		38.3	9.93E-12
128	ELEPHEE	9	0.02	0		38.3	9.93E-12
129	PEARL	9	0.02	0		38.3	9.93E-12
130	DRAGON	23	0.05	25		37.8	1.45E-11
131	VILLAGE	27	0.06	37	0.01	37.1	2.48E-11
132	BECAME	47	0.10	109	0.03	36.8	3.59E-11
133	SOME	112	0.25	416	0.12	36.6	4.49E-11
134	DEMON	11	0.02	2		36.1	7.73E-11
135	FAN	10	0.02	1		36.1	8.29E-11
136	THEN	254	0.56	1224	0.36	35.4	2.74E-10
137	COOKIES	18	0.04	15		34.9	4.29E-10
138	SELL	14	0.03	7		34.6	1.02E-09
139	WOODCUTTER	15	0.03	9		34.3	1.58E-09
140	STEPPES	8	0.01	0		34.0	2.41E-09
141	STOREKEEPER	8	0.01	0		34.0	2.41E-09
142	ZIPPY	8	0.01	0		34.0	2.41E-09
143	CRANE	8	0.01	0		34.0	2.41E-09
144	OCTOPUS	8	0.01	0		34.0	2.41E-09

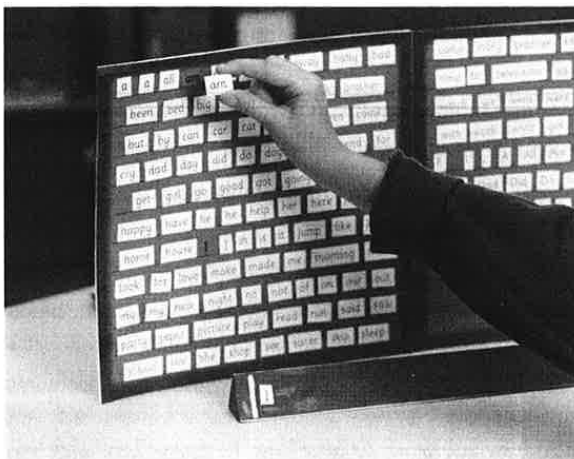
145	CHUGGA	8	0.01	0		34.0	2.41E-09
146	LANKY	8	0.01	0		34.0	2.41E-09
147	YAMS	8	0.01	0		34.0	2.41E-09
148	DEVIL'S	8	0.01	0		34.0	2.41E-09
149	DOUSER	8	0.01	0		34.0	2.41E-09
150	CHIRP	8	0.01	0		34.0	2.41E-09
151	TENGU	8	0.01	0		34.0	2.41E-09
152	CARVED	8	0.01	0		34.0	2.41E-09
153	TOOTHPICKS	8	0.01	0		34.0	2.41E-09
154	NOSE	46	0.10	114	0.03	32.7	7.39E-09
155	PAT	10	0.02	2		32.2	1.05E-08
156	OGRE	24	0.05	36	0.01	30.5	2.96E-08
157	WOLF	23	0.05	33		30.4	3.13E-08
158	ROLLING	18	0.04	19		30.1	3.63E-08
159	ANIMALS	18	0.04	19		30.1	3.63E-08
160	CAT	30	0.06	57	0.01	30.0	3.83E-08
161	YEAH	7	0.01	0		29.8	4.49E-08
162	RAINMAKERS	7	0.01	0		29.8	4.49E-08
163	LOOM	7	0.01	0		29.8	4.49E-08
164	HERDBOY	7	0.01	0		29.8	4.49E-08
165	COOKIE	12	0.02	6		29.6	4.75E-08
166	SWIM	17	0.03	17		29.5	5.14E-08
167	COTTON	14	0.03	10		29.5	5.18E-08
168	TREASURE	13	0.02	8		29.4	5.39E-08
169	MANY	52	0.11	149	0.04	29.3	5.78E-08
170	HITS	10	0.02	3		29.2	5.93E-08
171	OUR	53	0.11	154	0.04	29.1	6.32E-08
172	HOSPITAL	9	0.02	2		28.3	9.6E-08
173	PRINCE	5	0.01	223	0.06	-29.7	4.73E-08
174	IN	449	1.00	4313	1.30	-29.9	4.22E-08
175	KNOW	30	0.06	572	0.17	-34.2	1.88E-09
176	MY	131	0.29	1593	0.48	-34.4	1.42E-09
177	SHOUTED	16	0.03	407	0.12	-35.0	6.24E-10
178	QUITE	8	0.01	299	0.09	-35.6	1.67E-10
179	GRANDMOTHER	5	0.01	255	0.07	-36.4	5.15E-11
180	ROUND	5	0.01	286	0.08	-43.2	1.02E-12
181	HER	189	0.42	2264	0.68	-46.5	4.04E-13
182	OF	603	1.34	5914	1.78	-47.4	3.3E-13
183	CRIED	35	0.07	782	0.23	-58.2	5.62E-14
184	YOU	442	0.98	4776	1.44	-65.0	2.68E-14
185	HONEY	3		380	0.11	-74.0	1.25E-14
186	SAID	339	0.75	4027	1.21	-80.7	7.83E-15
187	SHE	251	0.56	3314	1.00	-93.0	3.88E-15
188	MRS	4		484	0.14	-93.3	3.83E-15
189	NOT	57	0.12	1344	0.40	-106.8	2.07E-15
190	I	338	0.75	4850	1.46	-170.1	3.17E-16
191	MISS	3		1175	0.35	-269.1	6.16E-17
192	MR	39	0.08	3316	1.00	-584.4	4.8E-18

Appendix 9.1 Sample of pictures of ‘Breakthrough to Literacy’



From:
Coles, M. and Hall, C. (1998). *Breakthrough to Literacy: Teacher's Guide*. Harlow: Longman. (p. 28).

The Sentence Maker in use



A child begins making a sentence.

From:
Coles, M. and Hall, C. (1998). *Breakthrough to Literacy: Teacher's Guide*. Harlow: Longman. (p. 32).

Appendix 10.1 A sample worksheet

Task 1:

グループでお友達の持っているものを褒め合ってみよう。

下にお友達の名前と好きなものをメモしておこう。

(Praise your friends' belongings in a group. Write a memo below.)

Friend's name	Things you like
<i>Ken</i>	<i>cap</i>



Task 2:

メモをもとに、好きなものについてまとめてみよう。

(Based on the memo, summarise what you like below.)

<i>I like</i>	Friend's name	Things you like
	<i>Ken's</i>	<i>cap.</i>
	<i>Eri 's</i>	<i>glasses.</i>
	<i>'s</i>	
	<i>'s</i>	
	<i>'s</i>	
	<i>'s</i>	
	<i>'s</i>	



Task 3:

メモをもとに他のグループのみんなに発表しよう。どんなものがあるかな？

お友達の意見と同じ人は、'I like it too!' または 'I like that too' と言ってみよう。

(Share your ideas with other members of the class.

If you like the same thing, please say 'I like it (that), too'.)