Collocation and Idiom Problems in Secondary Pupils' Writing in Brunei Darussalam

**Alex Henry, Alistair Wood, Adrian Clynes, and
Malai Ayla Surya Malai Hj Abdullah**
Universiti Brunei Darussalam, Brunei

Through his work in corpus linguistics, Sinclair (1991) postulated two types of language production. In the first, the 'open choice' principle, each word produced carries independent meaning. The second principle, the 'idiom principle', is based on the notion that the choice of one particular word is responsible for the choice of other words in the immediate environment. This paper looks at the writing of a variety of text types of a wide range of pupils in the bilingual Bruneian secondary school system, and makes use of Sinclair's principles to describe the transitional stages of their language development from learner English to their target, Standard British English. The results show that language produced by the idiom principle such as phrasal verbs, even common ones, is likely to be problematic for many pupils. Secondly, idioms and strong collocations may be problematic for all pupils including those with very advanced English. Based on these results, the paper offers recommendations for classroom methods and materials based on some of the principles of language awareness.

Key words: corpus, idiomaticity, pupils' writing, collocation, language awareness

INTRODUCTION

The theoretical basis for this paper lies in Sinclair's (1991) work on evidence based language description which makes use of the tools of corpus linguistics to identify natural language use. Based on his analyses of a huge corpus of spoken and written language, Sinclair postulated two types of language production. In the first, the 'open choice' principle, each word produced carries independent meaning. The second principle, the 'idiom principle', is based on the notion that the choice of one particular word is responsible for the choice of other words in the immediate environment, resulting in meaning being carried by 'chunks' of language rather than by individual words (Sinclair, 1991). Although Sinclair was the first to use evidence from real language data to show conclusively the importance of ready to use 'chunks' of language, the first linguist to point this out was Becker in 1975. Up until this time, any form of idiomaticity was considered an aberration and pointedly ignored by linguists, who concentrated on describing the grammar of idiom/collocation free sentences which they obtained not from real instances of language use but constructed themselves for research purposes. In addition to his contribution to the description of real language use, Becker's work also pointed the way to research which showed clearly the importance of pre-formulated chunks of language in first language acquisition (Peters, 1983), second language learning (e.g., Nattinger & DeCarrico, 1992) and native-like fluency (Pawley & Syder, 1983). The implications of the importance of pre-fabricated language for language learning and teaching is that '[e]very word has its own grammar – a set of patterns in which it occurs' (Lewis, 1993, p. 142) and therefore, language learners have to learn prefabricated patterns and the grammar of the pattern. A good example of this is found in phrasal verbs. Sinclair (1991) shows how many of the common verbs in English lose their 'core' meaning when combined with prepositions to form phrasal verbs (e.g., to give up trying). Many of these combinations are idiomatic in nature (e.g., to make up after a fight), very common and, therefore, very important for language learners to

master. The reason for their idiomaticity is the verbs are often de-lexicalised or 'bleached' of their core meaning; compare for example, 'to give money to the poor' with 'to give up on someone'. Another reason for their difficulty is that their use is often linked to specific topics or contexts. Consider these two sentences from the children's book Stories for Children, by Oscar Wilde, published in 1990: 'I thought it was quite the other way, and that we were to be let off in the Prince's honour' (p. 60) and 'He was so damp with crying that he could not go off at all (p. 64). Without sufficient background knowledge of the language associated with fireworks ('let off' and 'go off'), a reader might think the first utterance was from a prisoner expecting amnesty rather than from a conversation between fireworks, while the second utterance would be almost impossible to make sense of.

In Brunei Darussalam English is a very important language within the education system. In order to proceed to the country's main university, Universiti Brunei Darussalam (UBD), all students, even those taking mainly Malay medium courses, must have a credit in the Brunei Cambridge O Level examination taken in Year 5 of secondary school. However, the poor passing rate in this exam has prevented many pupils from entering tertiary education. Given the importance of phrasal verbs in language development, it would seem worthwhile to investigate pupils' use of these verbs in their written output to determine if this aspect of the language was one of the causes for their poor performance in English at school. In addition, there seems to be very little research on phrasal verbs in EFL/ESL school pupils' writing. The main studies in this area have investigated tertiary students' knowledge of phrasal verbs and other multi-word combinations (Liao & Fukuya, 2004; Liu & Shaw, 2001; Siyanova & Schmitt, 2007).

The first aim of this paper was to investigate if pupils in secondary schools in Brunei Darussalam were able to use a common verb in combination with a range of prepositions to create idiomatic expressions. This approach is in line with Liu and Shaw's (2001) approach which investigates learners' qualitative knowledge of a word rather than calculating the quantity of words known. The second aim was to identify examples of specific classroom activities

which can be used to overcome any problems with these verbs. We begin by describing the context in which the study took place.

THE BRUNEI CONTEXT

Negara Brunei Darussalam (Brunei) is a small, oil-rich, Malay, Islamic sultanate situated on the north coast of Borneo with a population of about 380,000. Although Bahasa Melayu (Malay) is the official language, a wide range of other languages are found within the country. Martin (1995) lists seven indigenous languages: Malay, Kedayan, Murut, Belait, Dusun, Tutong and Bisaya. In addition, there are speakers of Iban and Penan which are not considered indigenous as they originated from areas of Borneo outside of Brunei (Martin & Sercombe, 1992). To these languages we can add Mandarin, Hokkien, Hakka or other Chinese languages spoken by the Chinese community which makes up about 11% of the population (Jones, 2008).

Before 1984, the brightest pupils in Brunei received their education in the medium of English, while the less able were educated in Malay. In 1984, the government introduced a Malay-English bilingual education for all pupils entering the education system. Twenty-five years since its introduction, the results of the policy appear to be mixed. Larkings (1996) contended that the implementation of the bilingual policy had not been particularly successful in the areas of the syllabus, teaching materials, test and examinations, the inspectorate, teacher supply and teacher education. A study by Henry, Rohaniah, and Metussin (1999) suggested that many of the students in Form 1 did not have sufficient vocabulary in English to cope with the demands of their English-medium subjects. However, Jones (2003) claims that by the year 2000, 15 years after the implementation, the country had an 'embarrassment of riches' of well-educated bilinguals. Jones does however, point out that not all pupils are learning English equally well, and that at the other end of the spectrum there are pupils who are 'abject failures', and this may in part be

due to the special demands made on pupils in a bilingual system. These failures may be caused by English being a foreign language in some parts of Brunei where the situational context plays little part in the learning of English (Beardsmore, 1996). Clearly then there is a need to investigate in detail what is actually happening in the education system in terms of language development and to determine what scope there is for improvement.

METHODOLOGY

The Participants

To obtain a representative sample of pupil writing, four secondary schools were selected based primarily on their geographical locations. The 'city' school, situated in the capital Bandar Seri Begawan (BSB), was chosen because it is one of the top schools in the country: pupils who have excelled academically in any of the primary schools in the country may be chosen for enrollment. The 'rural' school, situated in the most rural eastern district of Brunei, Temburong, was included as it is the only secondary school in that area. Two 'town' schools from the semi-rural Tutong district were included as they were single sex schools, which provided opportunities to compare language development between genders. In the Temburong and Tutong schools pupils were streamed according to their academic abilities, so data were collected from the top, middle and bottom streams in Forms 3, 4 and 5. In the selective city school pupils are not streamed, so classes from each of these years were selected based on convenience for the teaching staff.

Data Collection Instruments

In order to collect a varied sample of writing pupils were given tasks (with instructions in English and Malay) which elicited five text types: narrative, exposition, argument, description and instruction. The prompts used are

shown in Appendix A.

The Corpus

The data used in this paper were extracted from a corpus of writing collected during a University of Brunei Darussalam sponsored research project. The aim of the project was to create a corpus of writing from pupils in the Brunei bilingual education system and to analyse the data to provide a description of the pupils' writing. A total of 764 scripts were collected and stored in an Excel spreadsheet. Details of the composition of the corpus are shown in Table 1.

TABLE 1
The Composition of the Corpus

Text Type	No. of Scripts	No. of Words	Ave. Length of Script
Argument	144	23,170	160.90
Description	160	29,713	185.71
Exposition	144	32,446	225.32
Instruction	147	25,421	172.93
Narrative	169	38,930	230.36
Whole corpus	764	149,680	196.00

As Table 1 shows, the longest scripts were exposition followed by narrative. The text type with the fewest number of words was argument. The differences in the average length of the text types are probably explained by the difficulty of the tasks which elicited the texts.

Data Analysis

The corpus was analysed both as a whole and as various sub-corpora from different years in the schools, both manually and using software, including *Concapp* (available free online at the time). By creating a cross-sectional corpus of pupils' writing we were able to investigate the stages of development of particular lexico-grammatical structures in the education

system. We thus describe the pupils' language development in its own right, rather than simply comparing it to the target language (Ellis & Barkhuizen, 2005). This further allows us to identify the sequence of language acquisition as pupils move through a series of 'transitional structures' (Dulay, Burt & Krashen, 1982) on their path to the target structure. Identifying the paths of development should help determine when and how teacher intervention might be effective.

Sub Corpora

The corpus described above was subdivided into three sub-corpora, BSB, Tutong and Temburong, and each of these sub-copora was then further divided into Forms 3, 4 and 5, thus forming a total of nine sub-corpora. Unlike other studies on phrasal verbs which used multiple choice, translation or gap fill, the data collection instruments were not designed to specifically elicit phrasal verbs. We thus had to find a suitable verb for investigation from the corpus. 'Look' was chosen for investigation because it is a common verb in English and, as such, is used in various combinations with prepositions to form a variety of phrasal verbs which are, in Sinclair's sense, idiomatic in nature. In addition, this verb was well represented in the corpus.

Each of the nine sub-corpora was then searched for instances of 'look', 'looks', 'looking' and 'looked', using the search facility of *Concapp*. The results from each search, in the form of a Key Word in Context (KWIC) data set, were then pasted into a Microsoft Word document, analysed and sorted into two categories: Standard Use and Non-standard Use. We then investigated the non-standard occurrences to identify the various transition stages. As an example, an extract from a sorted KWIC from the corpus is shown in Figure 1 below.

FIGURE 1
Extract from the KWIC 'Look' in Combination with 'at'

because it was very funny to ***look*** at a fat man's stomach that we
long the path, some villagers ***look*** at him and laugh.   John knows
gossip talking and a pet dog, ***look*** at him sharply.   After he pa
 walk to the mirror to take a ***look*** himself, then he found that
 of home, Smalle took another ***look*** at himself.   His knee was graz
go infront of the mirrow. He ***look*** at his body and suprise because
tant in our life.   That men ***look*** his body at the mirror.   He
an very slowly.   A boy said, ***look*** at his stomach!   He is pregnan
en he look at the mirror.   He ***look*** his tummy, his arm and face
 old book from the stores and ***look*** at it. The title of book are _

RESULTS

We start with an overview of the different patterns of use of 'look' found in the corpus.

Although it might be possible to distinguish between a collocation and an idiom using the possibility of substitution as a criterion, the exact difference between a strong collocation and an out and out idiom is problematic. Therefore, in this paper, the terms 'idiomatic' and 'idiomaticity' will refer to both. Thus the expression 'take a bath', which could be considered a collocation because 'take' can be substituted with 'have' without changing the meaning, and 'kick the bucket', which allows no such substitution, will be both be considered idiomatic.

Overview of the Use of 'Look'

Table 2 below provides an overview of the different patterns of 'look' in the corpus.

TABLE 2
Syntactic Patterns of 'Look' Found in the Corpus

	Syntactic Pattern (meaning)	Example	No. of Occurrences in Corpus
1	Look + AdjP	You look tired.	55
2	Look like (+ NP)	His stomach looks like a balloon.	10
3	Look + preposition of direction (to look in a particular direction rather than at a particular object or person)	He looked under the bed.	6
4	Look for + NP (to search)	I am also looking for a penfriend.	10
5	Look + forward to (to anticipate in a positive way)	I'm looking forward to the party.	4
6	Look + through (to browse or sort through)	Mr Brown looked through the book.	7
7	Look + at' to look at a particular person or object	He looked at the deep wound made by the dog.	129
8	Look + in the mirror (to look at a reflection in the mirror)	But on that day, he looked in the mirror and was disgusted.	53

As Table 2 shows, eight different syntactic patterns of 'look' were found in the corpus. All of these patterns occur frequently in English and are therefore worthy of investigation. However, as this paper is concerned only with 'look' in combination with prepositions, we will only look at patterns 3 to 8.

Look with a Preposition of Direction Meaning to Look in a Particular Direction Rather than at a Particular Object or Person

The pupils used a variety of prepositions including 'around', 'back', 'up' and 'inside' with 'look', to form seven standard phrases (e.g., he looked around his room . . .). Perhaps the reason why the pupils' were able to use the standard form was due to the fact that the verb-preposition combinations were non-idiomatic with the meaning of the prepositions corresponding to

their core meaning as in the phrase 'He looked back and couldn't see them'.

'Look for + NP' (to Search)

The pupils had few problems using this combination. All ten occurrences were used in standard form and were evenly spread in the corpus with five instances being found in the BSB corpus, three in Temburong and two in Tutong. This result is unexpected from a contrastive analysis perspective because in Malay the notion of 'looking for' is realized by a verb such as 'mencari' without any preposition and is not idiomatic. The pupils who used this expression would appear to have learned the verb and preposition combination as a unit and were able to use it in an appropriate context. The reason for the low occurrence of this combination in the corpus is probably due to the data elicitation tasks which did not encourage its use.

'Look Forward to + v-ing' (to Anticipate in a Positive way)

This structure was attempted only four times in the corpus, probably because the prompts did not encourage its use rather than because of avoidance. None of these occurrences were completely in standard English. In two instances, pupils selected the preposition 'for' instead of 'to'. Of the remaining two sentences with the correct proposition, one was non-standard in terms of tense and aspect as in 'will be looking forward'. It is likely that the pupil did not know that the present tense form carried a future meaning. In the second occurrence the pupil made the classic syntactic error associated with this phrase of not using the -ing form of the verb. It would seem that this idiomatic phrase, although not found widely in the corpus, is problematic. This may be because in Malay the corresponding verb phrase is not realized by a verb and preposition combination. In terms of transition stages, the pupils who attempted to use this lexical item have learned that it is a multi-word phrase, are aware of a context in which it can be used, but have not fully mastered its grammar.

'Look Through +NP' (to Browse)

Given that this combination is idiomatic in nature, we might have expected the pupils to have some problems. However, of the seven occurrences, six were in standard form. The seventh instance (He **_looked_** through a book title keep fit) may be standard if we consider the non-standard form to be in the choice of the word 'title' instead of 'entitled. Of the seven instances, four occurred in the elite school sub-corpus of BSB and three were found in the semi-rural schools in Tutong. This encouraging result cannot be explained by the existence of a parallel structure in Malay as 'to look through' is realized through a verb such as 'meneliti' rather than with a verb and preposition combination. We can assume that the pupils who used this item have learned that it is a multi-word combination, have learned its grammar and are aware of how to use it appropriately in this context.

What is interesting about this particular lexical item is that the data collection instrument used in the study did encourage its production as the pupils were asked to describe a picture which showed a man reading a keep-fit book, or maybe browsing through it, in order to find a way of losing weight. This means that the pupils must have found another way of describing the picture in question. By using the search term 'book' in the corpus, we were able to investigate the alternatives they choose. It was found that 36% of the pupils were not able to use appropriate verbs to describe the activity in the picture as shown in Table 3.

TABLE 3
Verbs Used to Describe a Man Reading a Book

Standard Use (no. of occurrences)	Non-standard Use (no. of occurrences)
look through the book (7)	look the book (13)
read the book (29)	look at the book (13)
opened the book (13)	looked out the book (1)
learned from the book (2)	analysed the book (1)
found from the book (1)	examined the book (1)
perused the book (1)	have the book (1)
referred to the book (1)	

One thing is clear from these data: the pupils' are not trying to avoid the use of multi-word expressions as has been postulated in other studies of phrasal verb use, or lack of use (e.g., Siyanova & Schmitt, 2007). The pupils who wrote 'look the book' seemed not to know that 'look' needs to be followed by a preposition. Those that produced 'look at the book' seemed not to know the difference in meaning between 'looking at a book' and 'looking through' or 'reading' a book. They did seem aware though that 'look' needed a preposition but chose the wrong one as did the pupil who wrote 'looked out the book'. Those pupils who used 'analysed' and 'examined' may have been aware of the general lexical set of verbs available for the context but made the wrong choice.

In terms of transition stages, we can see that many of the pupils had not passed through the first stage of identifying the appropriate lexical item to describe the activity in the picture. Even those who used the seemingly appropriate 'read the book' and 'opened the book' may have been avoiding describing the activity of browsing through the book to find specific information.

Pupils' Use of 'Look + at'

This pattern was attempted 129 times in the whole corpus. Of these 67 were used in a standard way while the remaining 62 were non-standard. By far the most frequent non- standard form was for pupils to omit 'at', while other pupils replaced 'at' with an unsuitable preposition. Considering that these data were produced by pupils in Forms 3 to 5, an optimistic interpretation might be that they represent a positive movement from non-standard English in the lower forms to standard English in the upper forms. However, a more detailed analysis, as shown in Table 4, reveals that most of the non-standard forms were produced by pupils in the town and rural schools.

TABLE 4
Pupil's Use of 'Look' in the Whole Corpus

School	Standard use	Non-standard use
City (selective)	17	1
Town (Tutong)	36	42
Rural (Temburong)	14	19
Total	67	62

From Table 4, it is clear that pupils in the selective school have no real problems with the most common syntactic pattern involving 'look'. However, for pupils in the other two schools this structure appears problematic as they do not appear to have moved from the transition structure 'look + NP' to the fully formed structure 'look + at + NP'. We now look at a breakdown of the instances of 'look + at' in these two schools in more detail to determine if there are any transitional changes occurring in use between Forms 3, 4 and 5.

Table 5 shows quite conclusively that in Temburong, pupils in Forms 4 and 5 are more likely to use the non-standard form of this structure than the standard form.

TABLE 5
Use of 'Look + at' in the Temburong Corpus

Form	Instances of Standard Use	Instances of Non-standard Use
3	5	4
4	8	10
5	1	5
Total	14	19

We find only 14 acceptable instances of 'Look + at' and 19 occurrences where 'at' is missing or has been replaced by an unacceptable preposition. Thus there is no real evidence of pupils moving through a transition stage towards the target structure.

Table 6 shows a similar trend in the town schools in Tutong with regard to Forms 3 and 4, but by Form 5 they used the standard and non-standard forms almost equally.

TABLE 6
Use of 'Look + at' in the Temburong Corpus

Form	Instances of Standard Use	Instances of Non-standard Use
3	13	17
4	14	17
5	9	8
Total	36	42

These data indicate that in the town and rural areas, there has not been a general movement within the education system from the transition form (look + NP) to standard English (look + at +NP). It would seem that for many of the pupils the meaning of 'look at' is contained in the verb itself and not in the combination of the verb and the preposition. This would imply that pupils have either not had enough exposure to the target form, either at school or out of school, or they have simply not noticed the verb and preposition combination. The data may be an indication that fossilization (Selinker, 1992) has taken place in the rural schools. Given the bilingual education system in Brunei, it is hard to imagine pupils not being sufficiently exposed to the standard form. It is likely that they simply have not noticed the correct form when reading, nor have they had their attention drawn to it. A third possibility is the error correction procedure in the schools has not been effective. If learners are unaware that their intended meaning is not getting across, they will not be under any pressure to learn the correct form (Liu & Shaw, 2001).

Analysis of 'Look + in the Mirror'

One striking finding we noticed when we looked through the initial data was the difficulty pupils had with the phrase 'look + in the mirror'. This common and seemingly uncomplicated lexical phrase was problematic for the pupils in all schools. The data in Table 7 were obtained from an analysis of the whole corpus.

TABLE 7
Pupils' Use of 'Look' with 'the Mirror'

Syntactic pattern	No. of Occurrences
1. Look + the mirror	1
2. Look at + mirror	1
3. Look at + the/a mirror	33
4. Look + to + the mirror	3
5. Look + on + the mirror	1
6. Look + into + the mirror	7
7. Look + in + the mirror (Standard)	7
Total	53

There appears to be at least two possible explanations for the data. The first is that there is a gradual progression, as reflected in the seven syntactic patterns, from learner English to standard English with few pupils actually going through all the transition stages to the target form. In the first transition stage, the pupil has omitted the preposition. In the second stage the pupil is aware that a preposition is necessary but has chosen 'at' and the article 'the' or 'a' is omitted. In the third stage, beyond which the majority of the pupils have not progressed, the pupils assume that the most common pattern of 'look + at' applies to 'the mirror'. In the syntactic patterns 4, 5 and 6, the pupils are aware that 'look' needs to be followed by a preposition other than 'at' but have chosen the wrong alternative. In pattern 6, the pupils are very close to the acceptable syntactic patterns but in standard English when someone 'looks **into** the mirror' the person is going beyond looking simply at his or her reflection. Only seven pupils have been able to use the target form.

A second possible explanation is that the phrase 'look at the mirror' is used so frequently that it is a candidate to be considered Brunei English. When the issue of 'look in the mirror' was discussed with one of the Bruneian participants in this project, she pointed out that she had only learned the 'correct' form during her second year at UBD studying TESL. According to our informant, 'look at the mirror' meaning 'look in the mirror' is widely used in Brunei and she considers it be Bruneian English. The question this raises is has this lexico-syntactic feature become part of Brunei English? The

answer is probably 'no', since this structure would not seem to satisfy any of the generally agreed criteria for inclusion: it does not contribute to a marker of Bruneian solidarity (Svalberg, 1998); its use in the standard form would not be seen as being pedantic or pretentious as Martin (1996a, p. 33) claims would be the case if a Bruneian used Standard Malay rather than Brunei Malay; it does not express a non-standard meaning; it is not innovative; and it does not allow Bruneians to express their culture more accurately nor does it help them avoid expressions that would somehow be inappropriate to their culture (Cane, 1996). Thus it would seem that the reason why pupils have not learned such as common phrase is that they have simply not noticed the form of the structure themselves or have not had their attention drawn to the standard form resulting in fossilization.

CONCLUSIONS

After a detailed investigation of the syntactic patterns commonly associated with 'look', we have found that pupils were successful in producing the correct forms of 'look for' and 'look through' although these combinations were not well represented in the corpus. On the other hand, pupils in the rural and semi-rural areas had difficulty producing the combinations 'look forward to' and the most common form 'look at'. It would seem that in these schools a certain degree of fossilization is occurring which will inhibit these pupils' ability to communicate accurately. All pupils, including those in the elite school, had difficulty with 'look in the mirror'.

From these findings we can draw two main conclusions. The first is that idiomaticity involving verb and preposition combinations is likely to be problematic for pupils outside the 'elite' selective schools in Brunei Darussalam. Secondly, unusual idioms and strong collocations may be problematic for all learners in the secondary system including those in the elite schools. These results are similar to other studies on phrasal verbs such as Nesselhauf's (2003) study of German speakers learning advanced English

and Siyanova and Schmitt's (2007) study of undergraduates and young professional using English on a daily basis. Nesselhauf found that students have more trouble learning collocations than other kinds of chunks and that congruence with the first language was an important factor in the learning of chunks of language – especially collocations, while Siyanova and Schmitt (2007, p. 132) found that even advanced learners of English, from a variety of L1 backgrounds, did not use multi-word verbs in a 'in a manner congruent to native speakers'. Our data would also suggest that pupils in rural and semi-rural schools are not avoiding using phrasal verbs but have not noticed the meanings associated with these verb and preposition combinations which are so common and therefore so important in English. It is likely that they see the words 'look' and 'through' and because they are aware of their base meanings, wrongfully assume they understand the meaning of their combination. By not noticing these meanings they have been unable to acquire these vocabulary items receptively or productively. Thus the data from this study would suggest that the problems of multi-word verbs are mainly due to interlanguage development as suggested by Liao and Fukuya (2004) rather then through deliberate avoidance as has been found in other studies (e.g., Siyanova & Schmitt, 2007).

This study also suggests that in the rural and semi-rural areas, pupils' interlanguage in terms of phrasal verbs may be fossilized. This may have serious implication for, as Liu and Shaw (2001, p. 188) point out, '[o]nce the profile becomes fossilized, a potentially negative result is a serious loss of precision in meaning'. We now look at how these problems might be addressed using some of the principles of language awareness.

RECOMMENDATIONS

The second aim of this paper is to explore how such problems can be overcome using a language awareness approach to teaching and learning. Two teaching and learning objectives emerge from the findings of the above

investigation. The first is that the most common uses of verbs in combination with prepositions should be given more consideration in the Bruneian classrooms, perhaps at the expense of teaching the typical or core meanings of verbs on their own. Secondly, pupils need to be taught to identify and use idiomatic verb combinations. As such, these combinations should be taught as whole units rather than as single items. Based on these objectives, the following sample language awareness and conscious-raising activities may be useful in helping pupils recognize and learn idiomatic chunks of language.

Dictation

Fanselow (n.d.) points out that while dictation is a very worthwhile learning activity, word for word dictation, as often practiced in language classrooms, is mere copying and so does not necessarily involve the mental processing necessary for language learning. To combat this he outlines the following steps to encourage mental processing rather than copying:

1. Learners put down their pens and listen
2. Teacher dictates a sentence (learners are not allowed to write, only listen)
3. Teacher allows a 10 to 15 second pause (to allow for processing)
4. Learners write what they heard (learners are not allowed to use erasers)

In order to develop accuracy when listening Fanselow recommends:

1. Asking learners to write 1a, 1b and 1c on separate lines of their paper
2. Telling learners to write their first rendition on line 1a, their second rendition on line 1b, and their third rendition on line 1c.
3. Asking learners to then compare their different renditions (by comparing what they hear with what they think they heard over and over they build up their ability to listen accurately)

In language awareness terms the 'gap' (or 'space between') is created

when learners have the opportunity to compare what they heard with what they thought they had heard. This type of dictation would appear ideal for encouraging learners to notice the difference between 'One day when Mr Ong was looking **at** the mirror', a common non-standard use found in our corpus, with 'One day when Mr Ong was looking **in** the mirror'.

Comparing Learner Output with Acceptable Models

As mentioned previously, one of the principles of language awareness exercises is that learners should be encouraged to 'notice' differences in language data. A second principle is that teaching should begin with what learners already know. With this in mind, one teaching method would be to:

1. Have learners respond to a writing task (e.g., describe a man looking in a mirror)
2. Provide them with an acceptable version in the target language, or encourage them to make use of dictionaries, reference books and corpora to find for themselves the acceptable version
3. Ask them to compare their writing with the acceptable version

In this case the 'gap' to be noticed is created when there are differences between the learners' writing and the acceptable versions. In the next section, we look in more detail at how learners can access acceptable output through the use of a relatively new source, language corpora.

Using Corpora

The main aim of corpus linguistics is to identify lexico-grammatical features and their meanings through the compilation and analysis of naturally occurring language using specialized software. In terms of language teaching, most research and development concerning the use of corpora has been carried out in tertiary education where it has been shown to be effective (e.g.,

Varley, 2009). Although there has been insufficient research to allow us to make strong claims with regard to the benefits of using corpora with young learners, a recent study by Sealey and Thompson (2007) has shown that using corpora with this age group may be beneficial. Their study with 8-10 year olds in the L1 classroom showed that the use of corpora can develop their metalinguistic awareness to the point where they could classify words based on common features, particularly grammatical class. There is also some evidence that using corpus tools might be acceptable to teachers (Farr, 2008). This would suggest that there is a possibility of using corpora effectively in Bruneian classrooms with pupils in Forms 3 to 5. We might expect them to be able to at least identify and perhaps classify verb/preposition combinations as 'corpus based teaching facilitates exploration of language patterns which are less evident in other kinds of materials' (Sealey & Thompson, 2007, p. 221). Exercises designed to make learners aware of the difference between 'look + preposition of direction' and the idiomatic 'look through' could make use of data such as those in Figure 2 below.

FIGURE 2
KWIC of 'Look through' and 'Look + Preposition of Direction

how to keep fit.   Mr Brown **looked** through the book and found a

if day, Andrew woke up. He **looked** around his white plain room a

rold woke up from sleep.   He **looked** up at the ceiling and stared at

desperate for a partner.   He **looked** through the magazine but

talking something.   He was **looking** through his mail and didn't see the dog

dropped his fork. He **looked** under his chair but couldn't find it

In addition, there are a number of other exercises in which learners' consultation of corpora would create a language/awareness learning environment. One example would be traditional gap filling and sentence jumbling exercises. The main aim of exercises based on gap filling and jumbled sentences in relation to idiomatic phrases is to make learners aware of multi-word expressions and to teach their meanings. For example, a normal cloze passage gaps every n^{th} word but a cloze aimed at teaching

idiomatic phrases deletes parts of the multi-word target expression such as 'I look . . . . . . to seeing you' and 'I look forward to . . . . . . you'. Jumbled sentences can be made to raise awareness of a particular phrase as in the example 'looked / in the mirror./ He' which highlights the collocation 'look + in the mirror'. For more advanced learners, gap filling with complete multi-word expressions can take place in short conversation as shown in Appendix A.

Language Awareness Using the First Language

One founding feature of language awareness is to exploit the gap between the learners' first language and the target language (Hawkins, 1984). In the complex sociolinguistic Bruneian context, where many of the pupils first language is not Malay, some of the pupils would be comparing the two languages used in the education system rather than English and their mother tongue.

A typical scenario for this approach is as follows:

1. Provide learners with data in their L1 and in the target language.
2. Ask them to notice any similarities/differences in the lexical items

For example, in Malay 'Dia ada *masa depan* yang **cerah**' could be presented to learners for comparison and contrast with the English equivalent 'He has a **bright** *future*'.

In this way, learners have the opportunity to see similarities and differences between their LI and the target language. In the above example, Malay and English are both similarly idiomatic in that 'bright' collocates with 'future'. In this respect, the two languages are congruent. If we look at an example from our study, the phrases 'look forward to', we find that English and Malay do not share the same lexical patterning. In Malay 'I **look forward to** seeing you' is not realized by any of the 'seeing/looking' verbs but instead by non-idiomatic expressions like 'Saya **harap** kita *akan*

berjumpa (lagi)' which translates literally as 'I **hope** we *will meet* (again)'. In this case, and in the case of all the phrasal verbs investigated in this study, Malay and English are non-congruent. By presenting such data, learners may be able to work out for themselves the differences between the two languages. However, if this proves too demanding, Lewis (1993, p. 121) recommends learners be taught, presumably explicitly, appropriate translations of these expressions to prevent them from translating word for word and should also be taught to recognise the boundaries between idiomatic phrases.

Finally, as has been pointed out earlier in the paper whether learners find idiomatic phrases in English difficult or not depends on whether a particular phrase is congruent or non-congruent with their own language. In order to determine which exercise type might be most appropriate, Nesselhauf (2003) recommends that consideration be given to whether a particular lexical item is congruent or non-congruent with their counterparts in the first language. For this to occur, teachers, syllabus designers and materials writers should have a working knowledge of the first language of their students. With regard to the Bruneian context, it has been shown that although Malay, particularly Brunei Malay, is used informally by teachers in the classroom, especially in the primary classrooms (Martin, 1996b), the present set of textbooks used in the English classes does not have any Malay and English contrastive exercises.

FURTHER RESEARCH

This paper has shown that idiomaticity is a major problem in the majority of the schools in Brunei. An examination of some of the principles of language awareness and the types of exercises these principles generate has identified some possible solutions to the problem. Research is now needed to determine how effective such classroom activities would be in the Bruneian and other contexts.

THE AUTHOR

Alex Henry is Senior Lecturer and Programme Leader for Professional Communication and the Media in the Faculty of Arts and Social Sciences, University of Brunei Darussalam. His research interests include genre analysis, small corpus linguistics and English for specific purposes.
Email: alexander.henry@ubd.edu.bn

Alistair Wood is Senior Lecturer and Programme Leader for Graduate Studies in the Faculty of Arts and Social Sciences, University of Brunei Darussalam. His research interests are in the fields of Bruneian English, scientific English, and Science, Technology and Society. His current research projects are on the Asian Corpus of English, public understanding of science in Brunei and university-industry-state research linkages in Brunei.
Email: alistair.wood@ubd.edu.bn

Adrian Clynes is Senior Lecturer in the English Language and Linguistics programme of Universiti Brunei Darussalam in Brunei. His current research interests include descriptive linguistics and minority language documentation. His recent publications include the Influence from dominant languages on minority language documentation projects (2010) and Bound roots in Balinese and Malay (2010).
Email: aclynes@ubd.edu.bn

Malai Ayla Surya Malai Hj Abdullah is Lecturer at the Faculty of Arts and Social Sciences, Universiti Brunei Darussalam. Her current research interests cover Autism, Brunei English, Language and Media and Youth and Media. Her current research projects are The Asian Corpus of English, Public Understanding of Science and The Impact of the Internet on the Social Behaviour of Youth in Brunei.
Email: ayla.abdullah@ubd.edu.bn

REFERENCES

Beardsmore, H. B. (1996). Reconciling content acquisition and language acquisition in bilingual classrooms. *Journal of Multilingual and Multicultural Development, 17*(2-4), 114-122.

Becker, J. D. (1975). The phrasal lexicon. In R. Schank, & B. Nash-Webber (Eds.), *Theoretical issues in natural language processing* (pp. 70-73). Boston: Bolt, Beranek and Newman.

Cane, G. (1996). Syntactic simplification and creativity in spoken Brunei English. In P. Martin, C. Ozog & G. Poedjosoedarmo (Eds.), *Language use and language change in Brunei Darussalam* (pp. 209-222). Ohio: Ohio University Press.

Dulay, H., Burt, M., & Krashen, S. (1982). *Language two*. New York: Oxford University Press.

Ellis, R., & Barkhuizen, G. (2005). *Analysing learner language*. Oxford: Oxford University Press.

Farr, F. (2008). Evaluating the use of corpus-based instruction in a language teacher education context: Perspectives from the users. *Language Awareness*, *17*(1), 25-43.

Fanselow, J. (n.d.) Ke wave. Retrieved 16th February 2003, from www.ipc.ac.nz.

Hawkins, E. (1984). *Awareness of language: An introduction*. Cambridge: Cambridge University Press.

Henry, A., Rohaniah, D. H., & Metussin, P. H. (1999). An investigation into the levels of difficulty of certain semantic word classes in a bilingual setting. *International Journal of Bilingual Education and Bilingualism*, *2*(1), 13-29.

Jones, G. (2003). Bilingual education equals a bilingual population? The case of Brunei Darussalam. In W. C. Daniel, & G. Jones (Eds.). *Education and society in plurilingual contexts* (pp. 129-142). Brussels: VUB Brussels University Press.

Jones, G. M. (2008, February). *The evolution of language-in-education policies in Brunei Darussalam*. Paper presented at the Invitational Workshop, SEAMEO and World Bank, Bangkok, Thailand.

Larkings, L. G. (1996). Bilingualism through the classroom: Strategies and practices in Brunei Darussalam. *Journal of Multilingual and Multicultural Development*, *17*(2), 296-311.

Lewis, M. (1993). *The lexical approach: The state of ELT and a way forward*. Boston: Heinle and Heinle.

Lewis, M. (1997). *Implementing the lexical approach: Putting theory into practice.*

Boston: Thompson/ Heinle.

Liau, Y., & Fukuya, Y. J. (2004). Avoidance of phrasal verbs: The case of Chinese learners of English. *Language Learning, 54*(2), 193-226.

Liu, E. T. K., & Shaw, P. M. (2001). Investigating learner vocabulary: A possible approach to looking at EFL/ESL learners' qualitative knowledge of the word. *IRAL 39*(2001), 171-194.

Martin, P. W. (1995). Whither the indigenous languages of Brunei Darussalam? *Oceanic Linguistics, 34*(1), 44-60.

Martin, P. W. (1996a). Brunei Malay and Bahasa Melayu. In P. Martin, C. Ozog & G. Poedjosoedarmo (Eds.), *Language use and language change in Brunei Darussalam* (pp. 27-36). Ohio: Ohio University Press.

Martin, P. W. (1996b). Code-switching in the primary classroom. *Journal of Multilingual and Multicultural Development, 17*(2-4), 128-144.

Martin, P. W., & Sercombe, P. G. (1992). An update of the Penan of Brunei. *Borneo Research Bulletin 24*, 86-92.

Nattinger, J., & DeCarrico, J. (1992). *Lexical phrases in language teaching*. Oxford: Oxford University Press.

Nesselhauf, N. (2003). The use of collocations by advanced learners of English and some implications for teaching. *Applied Linguistics, 24*(2), 223-242.

Pawley, A., & Syder, F. H. (1983). Two puzzles for linguistic theory: Nativelike selection and nativelike fluency. In J. Richards & R. Schmidt (Eds.), *Language and communication* (pp. 191-226). London: Longman.

Peters, A. M. (1983). *The units of language acquisition*. Cambridge: Cambridge University Press.

Sealey, A., & Thompson, P. (2007). Corpus, concordance, classification: Young learners in the L1 classroom. *Language Awareness 16*(3), 208-223.

Selinker, L. (1992). *Rediscovering interlanguage*. New York: Longman.

Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford: Oxford University Press.

Siyanova, A., & Schmitt, N. (2007). Native and nonnative use of multi-word vs. one-word verbs. *IRAL, 45*(2007), 119-139.

Svalberg, A. M-L. (1998). Nativization in Brunei English: Deviation vs. standard. *World Englishes, 17*(3), 325-344.

Varley, S. (2009). I'll just look that up in the concordancer: integrating corpus consultation into the language learning environment. *Computer Assisted Language Learning, 22*(2), 133-152.

Wilde, O. (1990). *Stories for children*. New York: Macmillan.

APPENDIX A
Prompts Used to Elicit Data

Argumentation

Your school is organising an awareness campaign on the importance of a healthy diet. Write an essay that persuades people to stop eating junk food and to start eating healthily. Write at least 200 words.

Exposition

Write a letter introducing yourself to a new friend in London, describing yourself, your family, your country, your hobbies, your school, and any other information you think is interesting. Write at least 200 words.

Instruction

Explain how to make something you like to eat or drink. Write at least 200 words.

Description

Describe your house in as much detail as you can. Write at least 200 words.

Narrative

Write a short story based on the set of pictures below. Write at least 200 words.

(A series of pictures which describe a man attempting to lose weight).

APPENDIX B

Complete short conversations/text with appropriate phrases containing look. (**Answers in bold**)

1. A. I asked for a pay rise and the boss said he would ……….. (**look into it**)
   B. Doesn't sound too encouraging

2. A. My girlfriend dumped me.
   B. …... you won't have to go shopping every Saturday. (**Look on the bright side**)

3. A. hey …… (**look at you**)!! What a change!
   B. So you like my new hairstyle??

4. A. Excuse me I think you sold me a lemon!
   B. …., it was a perfectly good car! (**Now look here**).