

A Data driven learning approach to collocation and colligation

Jonathan Garstang

Masters in Applied Linguistics

Lexis Module

University of Birmingham September 2013

LX/13/04

Title :

Choose a short, authentic, written or spoken text in English (about 200 - 500 words) and select at least twelve examples of collocations, including some fixed expressions. Either discuss how you might use these examples to demonstrate to students the constraints on word-combinations in English, or the extent to which they would present challenges to a translator.

ESSAY COVER & DECLARATION SHEET

IMPORTANT: Please complete ALL sections of this template and PASTE it into the FIRST PAGE of your essay

Student ID number	1304880
Module Number (1-6)	1
Title of Degree Programme:	Masters in applied linguistics
Title of Module:	Lexis
Assessment Task No.	LX/13/04
First or second submission (second submission is used for resubmitted work only)	First
Date Submitted	29/09/13
Name of tutor	R Nickalls
Word Count	4316
	(word count excludes footnotes, references, figures, tables and appendices).

DECLARATION OF AUTHORSHIP

I declare:

- a) that this submission is my own work;
- b) that this is written in my own words; and
- c) that all quotations from published or unpublished work are acknowledged with quotation marks and references to the work in question; and

Date: 29/09/13

Word constraints- a study of collocation and colligation

1.0 Introduction

An aspirational goal of any second language user is that of native like fluency. Such a concept, widely held as secure idiomatic control, could be determined by linguistic output described simplistically as natural. Words learnt as single vocabulary items have been found to be ineffective in the achievement of this goal. The acquisition of prefabricated language- in its many guises- offers a language user a greater capacity to learn a second language, achieving an idiomatic control in excess of that achieved through a primarily structural pedagogical approach.

This study will examine a pedagogical approach that focuses on collocation and colligation to achieve longer and more fluent stretches of communication. It shall be questioned whether learning underlying patterns, resulting in an increase in processing time, can be compensated for by a greater linguistic knowledge of the collocation. The role of a corpus in facilitating this will be analysed alongside a data driven approach that encourages the learner in a more inductive learning style on the journey towards native like fluency.

1.1 Background

If a L2 learner's language output closely matches the phraseology of first language speakers then the resultant communication will be fluent and without hesitation.

Hence, a syllabus or curriculum approach that does not take account of the importance of phraseology will undermine the progression of a second language learner. In the past, language teaching has relied on a structural approach where words filled slots in sentence frames.

Contrary to the structural approach, Nattinger and Decarrio (1992) believe that language is learned through "chunking". Citing Becker (1975:72), they assert that in the process of encoding speech lexis takes prominence over grammar. That is the process of speaking begins with retrieval as a phrase from the lexicon and "then the problem is to stitch these phrases into something roughly grammatical". (ibid:32)

Therefore a possible corollary is that phrases are stored within the lexicon as holistic wholes, and when retrieved the process of syntax is a secondary one. Developing a large lexicon capable of generating this kind of pre-packaged lexis will help a second language learner process communication more effectively.

2.0 Review of the literature: Collocation- a broad perspective

A discussion on word constraints must start with the concept of collocation. J R Firth (1957) is largely regarded as the pioneer of collocations and his definition of "you shall know a word by the company it keeps" illustrates how collocations are created through the habitual reoccurrence of words in specific contexts. In this interpretation, meaning is created through

the use of a word rather than any ideational concept. Or as Firth stated, words in collocations exhibited “mutual expectancies” (ibid 195). Therefore, Firth asserted that one of the meanings of dark is manifested through its collocability with night. Hence, a property of dark is that it will be found with close proximity to night.

Different senses of a polysemous word can be extracted from the differing collocates it forms. Hoey (2005 : 81) asserts that collocations for a word will “systemically differentiate its polysemous senses.” This is illustrated here:

i.	<i>She was not named as a party in the conspiracy</i>
	Sense: a person involved in an enterprise
ii.	<i>He was not a member of a political party</i>
	Sense: Participating in a political organisation
iii.	<i>They celebrated at a cocktail party</i>
	Sense: a social gathering

Figure 1

In figure 1, not only does the various senses of the word “party” form differing collocates, but it can be posited that these different senses determine the selection of the co-text and therefore act as a constraint on the words around them.

The examples above illustrate how collocations need not be adjacent. Collocations that are not adjacent need different pedagogical treatment. This will be looked at later in the study through the analysis of a visual diagram that guides a reader to look at collocations and colligation patterns with surrounding co-text.

Also, a unit of collocation can be calculated by a node word that co-exists with a span of words on either side. Nattinger and Decarrio define collocation as being when:

the node word occurs with a span of particular words at a frequency greater than chance would predict, then the result is a collocation. (1992: 20)

The greater the frequency of the words that occur within the span and the node word then the more idiomatic the collocation.

2.1 - Collocations and the spectrum of idiomaticity

Individual words that occur within collocations can vary along both the paradigmatic and syntagmatic axis. This is shown below

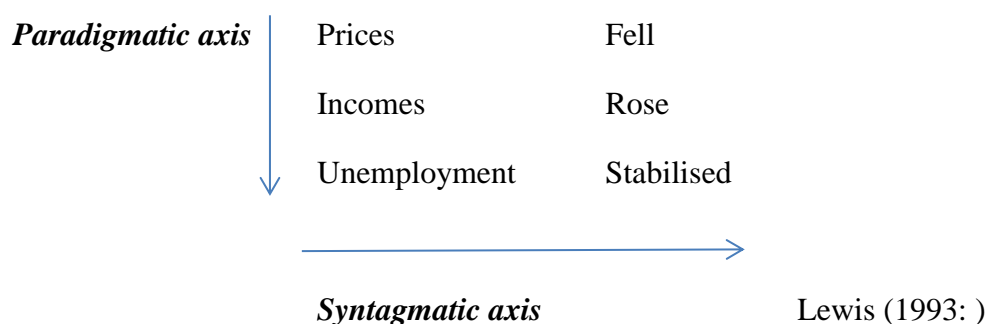


Figure 2

Here each of the words is substitutable with another on each axis to form a collocation.

Rather than existing as a dichotomy, collocations can be plotted along a spectrum of fixed to free. Wood(cited in Nattinger and Decarrio) argues that productivity and compositionality of form exists along a:

continuum from complete frozenness to full freedom of combination (ibid: 34)

Where a collocation can be located along this continuum will provide an indication of its predictability and transparency of meaning. At the fixed end of the spectrum, with idioms say, ossification has occurred along both the paradigmatic and syntagmatic axis.

Hill(2003) believes that medium strength collocations have significant pedagogical value- a way of recording these will be explored in section 4.

2.2- Colligation

The underlying premise of colligation is the co-occurrence of a lexical item with grammatical functions. A major premise of “The lexical Approach” is that language is formed by grammaticalised lexis. Hoey (2005) asserts that colligation needs to be examined beyond normal grammatical parameters. In this view, colligation of a lexical item is examined not only through the co-occurrence with a grammatical function, but also with regard to its sentential position.

The gist of my argument is that understanding the colligation patterns of a collocation will create opportunities for the language user to form longer stretches of communication. Nevertheless, colligation patterns can be intricate and are a good example of investigative work best undertaken through data driven learning.

2.3-Formulaic sequences and the idiom principle

The apparent diversity of terms for phraseological language means gaining a definition is problematic. Wray (2000) tabled more than forty different categorisations for this phenomenon, asserting that as labels differed it is wise to assume that there is some differentiation of underlying meaning. Wray (2000) arrives at a definition, classifying formulaic sequences as:

a sequence, continuous or discontinuous, of words or other meanings elements, which is, or appears to be prefabricated: that is stored and retrieved whole from meaning at the time of use, rather than being subject to the generation or analysis by the language grammar. ”

(ibid:465)

Establishing the importance of these formulaic sequences, one need only examine the ubiquity of this language phenomenon. Estimates vary on the prevalence, with Wray (2000: 466) estimating that 80% of spoken communication is of this form, whilst Schmitt (1996 : 1) cites Erman and Warren estimate of 58.6% of “spoken discourse”. Even if the figure for formulaic language lies somewhere in between the two, it suggests that the majority of language output is achieved through this phenomenon.

A corollary of this is that the pervasiveness of formulaic sequences demonstrates how communicative meaning is created in chunks of language. This exemplifies Sinclair’s principle of idiom, in which he stated:

a language user has available to him or her a large number of semi-preconstructed phrases that constitute single choices, even though they appear to be analysable into segments

(Sinclair 1991: 110)

L2 learners, whilst decoding a text, rely too heavily on the open choice principle. A teacher needs to engage in strategies to present the text in such a way that promotes the engagement of the idiom principle.

2.4 A review of teaching approaches of collocations and formulaic sequences

Wray (2000) examines the two approaches used in this study that of Nattinger and Decarrio (1992) and Lewis (1993). The main focus was whether their approach to analysis of formulaic sequences was conducive to a mastery of native like fluency. In their curriculum approach, Nattinger and Decarrio focus on the pragmatic influence of the lexical phrases. On the other hand, Lewis asserts the need to expose the learners to a comprehensive vocabulary, focusing on lexical phrases rather than individual words , and selection of items that maintain a balance of statistically rare words with high content meaning and high frequency words with low content meaning.

Both approaches downplay the role of grammar. In the approach of Nattinger and Decarrio, the assertion is that knowledge of a language is derived from the estimation of the possible variations in any of the sequences, and this analysis comes from experience rather than from rules. Lewis (1993: 100) postulates that sentences “chosen for archetypicality” will enable the learner to infer rules from the limitations and generalisations from this language. Wray summarises that these two assumptions contradict corpus data that suggests that learners find it difficult to make grammatical generalisations from linguistic input. Wray concludes by

stating that analysis is possible through generalisation but it needs to be accepted that this is due to a process of non- native processing.

These findings are important for this study. Firstly, it will be shown how Lewis' approach is excellent in developing competency of two word collocations. However, in terms of making grammatical generalisations, it will be shown how an inductive approach to corpus based learning is better at extracting underlying patterns. Collocations will be examined in relation to their co-text and this analysis will show that the importance of colligation patterns in forming longer stretches of communication.

2.5 Data driven learning- A corpus based approach

The process of Johns' data driven learning changed the emphasis from deductive learning to one of induction. Thus, the role of the teacher becomes a facilitator, allowing the learner to form observations from their own evaluation of data.

Bernadini (2004 : 17) suggests that concordancing can prove an invaluable tool in the process of inductive learning. What is asserted here is that concordancing facilitates learning that synthesises meaning and form. A manifestation of this in terms of collocation is that the learner will be guided towards recognition and understanding of the collocation, whilst developing an understanding of the prevalent underlying patterns.

Willis (90:40) in his book in developing a lexical syllabus states that one of the problems facing selection of items in the classroom is "proliferation". Hence, recognising the sheer volume of linguistic input needed to develop a lexicon similar to a L1 speaker, a curriculum designer must recognise that it is not possible to replicate this time in a classroom situation. Benardini posits that a concordancer may reduce the learning load of a L2 learner by focusing on specific examples on language output that elucidates both meanings and patterns. It will be shown in section four how a corpora based approach can be used to derive the colligation patterns of the selected collocations.

3.0- Selection of Text

The chosen text is from the W.J.E.C board and forms part of a GCSE English Language reading paper. As such, it not only demonstrates usage of phraseological language and collocations, but in doing so, it presents the challenges that this type of examination might pose for students learning English as a second language. The original article was published in a magazine called "Time Out" whose primary function is to inform readers of ways of spending recreation time in the city of London.

3.1 Categorisations for analysis

Items were selected by intuition and items chosen reflect the differing types of collocation needed to develop fluency. In appendix one, there is the typology of lexical collocation and grammatical collocation from the BBI dictionary. As such, their definitions of collocations

will serve best for this study. A grammatical definition of collocation is a phrase where a lexical word co-occurs with a grammatical structure, whereas a lexical collocation consists of a combination of lexical class words. A full typology used to analyse is included in appendix 1. Also forming part of the analysis will be a type of fixed lexical phrase classified as institutionalised expressions. This can be defined as those which:

allow the language user to manage aspects of the interaction; they are pragmatic in nature

Lewis (93 : 94)

Allowing for no paradigmatic or syntagmatic substitution, their advantages in textual cohesion will be illustrated, and their appropriacy for either spoken or written discourse will be explored. The Cobuild Bank of English corpora was used where appropriate to determine frequency using “T” scores.

3.2 Items under analysis- Collocations

The following are the items under analysis with their respective coding and the T scores to show their statistical frequency.

Lexical Collocation	BBI Classification	T score
Wonderful way	L3	14.5
Traffic jam	L5	23.5
Torrential rain	L3	22.9
Less polluted	L3	3.74
Reassuring presence	L3	5.9
Do errands	L1	3.4
Grammatical collocation	BBI Classification	T score
take up	G8D	88
notion of	G1	74
sodden with	G5	5.5
pick up	G8D	120

Figure 3

Although the items were selected by intuition, the use of the Bank of English corpora demonstrates they all have a T score that illustrates their statistical reliability as a collocation.

3.2 Items under analysis- Fixed term expressions

These were previously defined as institutionalised expressions and they can be according to Lewis employed in either written or spoken discourse. Nattinger and Decarrio(1992:38) classify these as polywords and assert that this type of lexical phrase must be pragmatic in nature: such as shifting the topic and summarising. Most importantly for classroom practice, it is necessary to ascertain what type of discourse they are most likely to be found in. Figure 4 below shows The T scores from two different sub corpora of the Bank of English:

Institutionalised expression	UK spoken corpora	Times corpora
Believe me	3.3	7.4
That said	0	17.9
Worse still	1.73	10.3

Uk spoken corpora over 200 million words

Figure 4

Times corpora over 51 million words

After using different corpora to establish what type of discourse a lexical item is from can inform what pedagogical treatment is most appropriate.

4.0 Collocation dictionaries and L1 translation

As a means of introduction to the text, student should be instructed to find and highlight the nouns in the text. Walker (2008) asserts that one of the ways of extracting underlying meaning of collocations is by a breakdown of their semantic components. Establishing meaning of these nouns will assist with subsequent understanding of their collocates. The BBI dictionary is an example of a collocation dictionary, which can aid learning. Below is an example of a visual tool that used interactively with the collocation dictionary can record collocations in this type of exercise. Known as a 5-1 box, it has been adapted to include the L1 translation of the collocates.

L2	L1 equivalent	
drenching	rain	
pouring		
heavy		
torrential		

Collocates from BBI Dictionary

Figure 5

The process of investigation using a dictionary not only encourages learners to correctly identify collocates but the translation from L2 will deepen underlying semantic understanding of the collocate. Thus it might reasonably be posited that it will assist in long term storage of the collocation. Lewis (1997) suggests that after the identification of the key noun, the learner searches for collocating verbs and adjectives. Although the end product may not be a collocation in the L1, the use of analogy will assist with understanding.

4.1 Fixed term expressions

In a study of the highest frequency phrasal verbs, Gardner and Davies (2007) examined this phenomenon in order to inform better pedagogical approaches. One feature of their study was to examine the polysemous nature of phrasal verbs. Included were the two phrasal verbs under analysis for this study. It was found that “take up” had 16 different senses while “pick up” had thirteen. Further investigation using the Bank of English corpora of these phrasal verbs highlighted that they only formed meaningful collocations for a limited number of senses. These collocations are shown in appendix 2. It is suggested that a corpora can be used to assist with the selection of language that is most probable.

Pedagogical treatment of this phenomenon can be difficult. One reason which Condon (2008) alludes to is the unmethodical relationship between the particle and the verb. Therefore, a classroom approach must, as Hill asserts, focus “on the phrase as a whole.” (2003:51).

Such a pedagogical approach for introducing phrasal verbs would be gap filling activities. An example is below:

Fill in the gap with right phrasal verb

- 1) *I would like to residence in a foreign country*
- 2) *The train..... speed when it left the station*
- 3) *I would like to some information on **cooking***
- 4) *The centre forward will a position behind the defender*

Take up
Pick up

Figure 6

This concept of introducing selected items as holistic units would also be best applied to the institutionalised expressions. Gap filling exercises can be formulated by extracting key word in concordance lines as evidenced below:

- a) This series showed they are really no more bizarre than the rest of us.1)the second series, which starts tonight,
- b) that is taken for granted in British life today. 2), the situation seems to deteriorate by the
- c) Life just gets unbearable if you go everywhere with the entire contents of your bathroom.3)....., I know exactly how bad it could be

Fill in which fixed term expression you think matches each sentence

That said

worse still

believe me

Figure 7

What should be noted here is that the examples were specifically selected to show the discourse functions of these fixed phrases. Students should be directed to discuss the cohesive features exhibited by each fixed phrase by close examination of the co-text.

4.2Using grids

Throughout the literature, delexicalised language is alluded to as being important in the development of native like control. As Lewis (1997) asserts they have little lexical meaning but the collocations formed are idiomatic because the meaning of the resultant phrase cannot be worked out by its individual parts. Sinclair (1991) stressed the lack of independent meaning of the delexicalised word by itself. A corollary from this is the propensity for delexicalised verbs to take their meaning from the nouns that follow them.

Channel (1981: 120) stressed the importance to record collocations on first appraisal. A collocation grid was, she posited, a means to encourage this.

	errands	a break	dinner	mischief	money	mischief	amends	homework
Do								
Have								
Make								

Adapted from Channel (81)

Figure 8

The grid will allow learners to learn the collocates of common delexicalised verbs, negating the common errors associated with this high frequency language. Wray (2002: 202) cites Grainger (1998), that non-natives treated collocations as separate units, and as such they were generated through the rules of grammar rather than from the phrasal lexicon. Thus, the implementation of a grid will facilitate that the delexicalised verb and its collocate is stored as a holistic unit.

4.3 Collocation and colligation radar

An invaluable investigative tool to sift data in a corpora is creating a “T” score profile. Using this function a picture tree is created including six words sorted to the left or right of the node. These six words are ranked in terms of their T score frequency. Using the example of traffic jam, the following words were produced from its profile:

in	a	traffic	NODE	on	the	way
a	the		NODE	in	90	some
an	massive		NODE	and	were	m25
avoid	hour		NODE	at	london	hour
through	huge		NODE	there	is	car
into	long		NODE	<p>	up	two
mile	mile		NODE	delayed	stretched	had
caused	worst		NODE	as	route	autobahn

Example of T score picture tree

Figure 9

Using an inductive learning style, students can be taught how to create a T score picture profile and then extract any words that have a T score of 2 or above. The teacher can supply them with a collocation radar grid that allows the students to plot words according to their distance from the node. An example here is provided here:

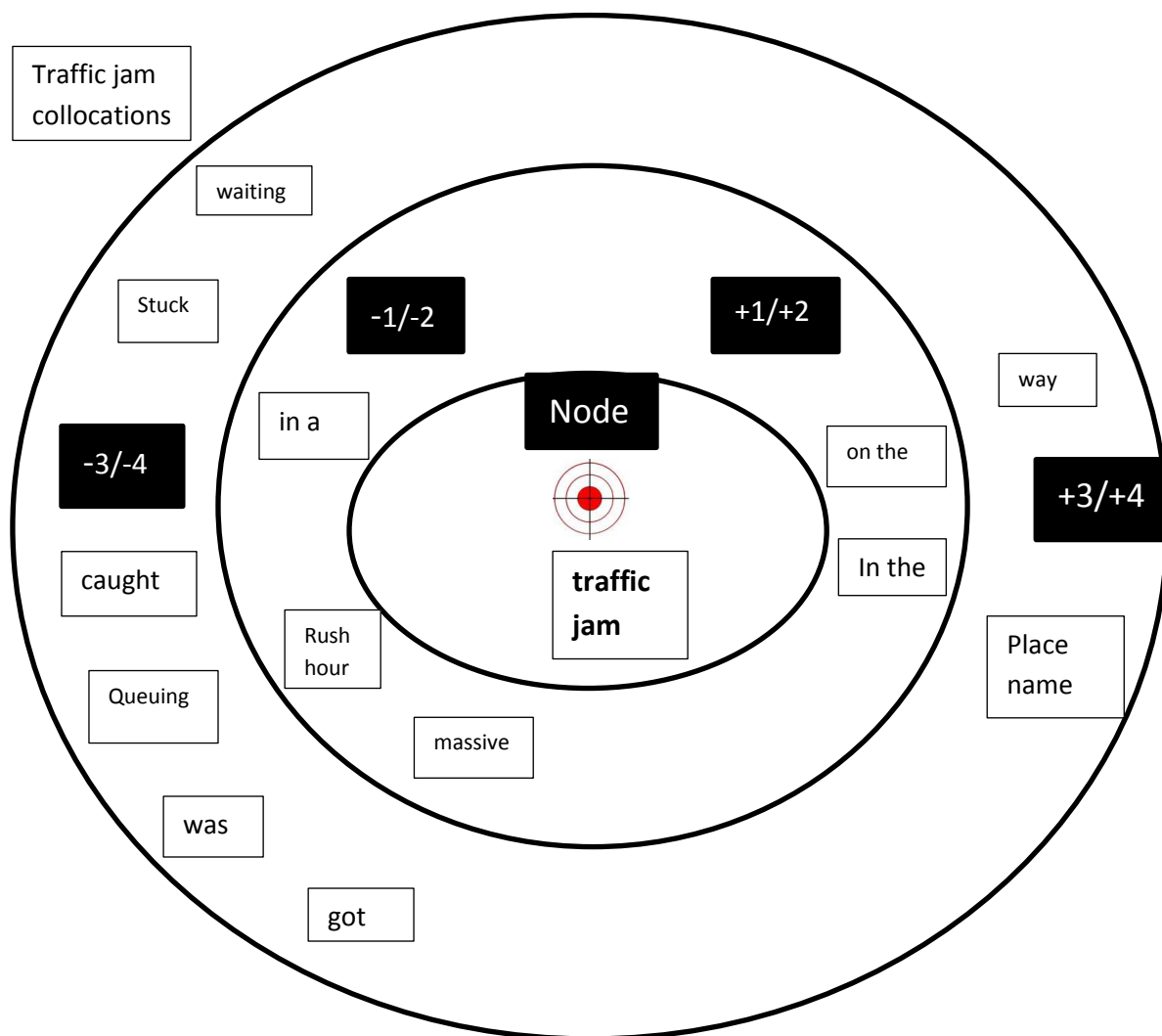


Figure 10

This type of resource can be used to promote Hoey's (2005) psycholinguistic concept of lexical priming by illustrating how the co-text of traffic jam "nest" to former longer stretches of communication. In this example, it can be seen how traffic collocates with jam and subsequently collocates with other words to form greater collocations.

Through this pedagogical device, learners can be directed to examine the sentential position that the collocation is found in. This can be demonstrated in the diagram above by analysing the frequency of the collocations to the right and left of the node. Here, it would appear from the data that the collocation "traffic jam" has a preference for avoiding sentence initial

position. A learner would also be directed to the observation that “traffic jam” collocates with certain verbs in the past tense; the significant pattern before the node is the prepositional phrase “in the” and after the node it is “on the”. Once students have completed their collocation radar, they can be encouraged to form as many different sentences as possible.

4.4 Key word in concordance lines

A corpora can be used to extract concordance lines which can then be used as a pedagogical tool to explore co-text. As Woolard asserts:

Concordances provide much richer sources of co-textual information than dictionaries, and they can lead to a more efficient exploration of the collocates of a word.

(2000: 40)

Using the data as a means of inducting learning, students can be given a variety of tasks to perform:

displays of stars unimaginable above our own light-polluted island
northern town on the estuary of a sprawling, much-polluted river.

we perched on the roof of this foul, polluted place

of Methil, whose inhabitants still drank foul polluted water

circulation of fresh air and filtration of stale, polluted air will again usually
be issues

carrot tops, and fresh dandelion leaves from a non-polluted source.

they smoke and non-smokers, breathing the same polluted air suffer losses

mode of transport and has led to the congested, polluted transport network we
have today.

. Dirt will precipitate from the thin, polluted air of summer

are increasing. This is partly due to a more polluted atmosphere

The fishing village is now a noisy, polluted city of over 1 million people
the changes from which will make it a less polluted atmosphere

Key word in concordance lines from adjective + polluted

Figure 11

Here, learners can extract the collocate modifiers of polluted, and in turn the nouns that polluted modifies. This will lead to an understanding of the complex noun phrases that are formed with this collocation. An example is provided of a means to analyse this information:

Noun phrase analyser

Determiner	adjective/ noun modifier	headword	noun	of	noun
<i>the</i>	<i>same</i>	<i>polluted</i>	<i>air</i>		
<i>the</i>	<i>congested</i>	<i>polluted</i>	<i>transport network</i>		
<i>a</i>	<i>more</i>	<i>polluted</i>	<i>atmosphere</i>		
<i>a</i>	<i>sprawling/ much</i>	<i>polluted</i>	<i>river</i>		
<i>a</i>	<i>less</i>	<i>polluted</i>	<i>atmosphere</i>		
<i>our/own</i>	<i>light</i>	<i>polluted</i>	<i>island</i>		
<i>the</i>	<i>thin</i>	<i>polluted</i>	<i>air</i>	<i>of</i>	<i>summer</i>
<i>a</i>	<i>noisy</i>	<i>polluted</i>	<i>city</i>	<i>of</i>	<i>over one million people</i>

Figure 12

This resource would allow learners to explore the key word in concordance lines to extrapolate this information..

Further investigative work could then be planned such as the following:

- What type of nouns does polluted come before
- Using the information can you devise your own “polluted + noun +of noun” phrase
- What type of modifier comes before polluted

Here the analysis work undertaken, first by students, and then is directed by the teacher can deepen understanding of the collocations of polluted, but also encourage the correct use of polluted in the future.

4.5 Lexical Notebooks

Recording collocations is an important process in their acquisition. Revisiting a word or its collocates, and subsequently reproducing it with the correct context and meaning is pivotal in the acquisition phase. Establishing this protocol is critical because:

a single encounter with a word is not enough to ensure its acquisition, and that subsequent encounters- research suggests a minimum of perhaps seven- are essential.

Woolard (2003: 43)

To ensure the information is recorded in the appropriate manner, Woolard suggests the use of a vocabulary notebook. Specifically, these notebooks can be essential in the acquisition of two word collocates: adjective and noun collocates; and verb and noun collocates. Reflecting a learner's own lexicon, their function can be dual purpose: a way of decoding communication, as well as a means of enabling encoding during production of writing. Woolard coined the term "Lexical notebooks" which will encourage the learner to perceive their role in developing a holistic lexicon.

Using an adaption of this idea, it is possible to format versions which allow for collocations to be presented with the fullest possible information.

4.6 The lexico- grammatical vocabulary notebook

Investigation using the Bank of English corpora gave the most common verb and adjective collocates for the target words of "presence" and "notion"(see appendix 4). One of the issues with teaching collocations is that of explanation. Rote learning collocations from their linguistic environment, perhaps, does not contribute to the most stimulating and motivating lessons. Postulating the need for a deeper underlying understanding of collocations, Walker attempts to account for the phraseological patterns underlying these phenomenon:

It is obvious that words can be associated with various types of phraseological unit, and that the tendency to form such units has to be taken into account in describing their collocational behaviour

(2008: 306)

Through the use of the B.O.E. it is possible to extract the grammatical patterning from a given word and this can then be recorded in the notebook, along with a student's favourite example of each pattern.

Word	Translation	Verb collocate	Adjective collocate	Grammatical Pattern	Favourite
Presence (noun)	presencia	establish	strong	<i>verb+ det+adj+noun</i>	to increase its global presence
		maintain	global	<i>verb+prep+det+noun</i>	maintain in his presence
		feel	military		
notion	noción	support	whole	<i>verb(inf)+det+noun+that</i>	to support the notion that
		reject	very	<i>verb(inf)+det+noun+prep</i>	to accept the notion of
		accept	popular		
		dispel			

Two of the nouns under analysis “presence” and “notion” with a Spanish translation

Figure 13

Wray (2000: 463) in analysing the treatment of grammar in a lexical language teaching claims that learning structure in the way above promotes “unnative like processing behaviour.” As collocations are stored and retrieved as holistic wholes, they circumvent the need for analysis.

However, recording patterns as in figure 8, the assumption is that the emphasis is on what Woolard terms word grammar, the advantages of which are that this approach:

complements the traditional approach to grammar by directing the students’ attention to the syntactic constraints on the use of lexis.

Woolard (2003: 45)

Such focus on word grammar emphasises that lexis is already grammaticalised and the patterns recorded above are a further example of how the learner is exposed to lexis that is seen in its natural linguistic environment.

In appendix 2, there is a further example of how a lexical notebook can be adapted. In this case, the adaption allows for learner to record and learn collocates of different senses of a word.

4.7 Colligation- a data driven approach

Teaching learners how to use a corpora will lead to more independent learning under the guidance of the teacher. Such a project was undertaken by the university of Oslo, and its founder Ebeling states that enabling students to discover through problem based activities facilitates outcomes commensurate with ability and motivation. This project was devised for undergraduates, but it will be evident that it can be adapted for learners of varying ability- although there could be a strong argument that it might provide an invaluable tool for learners to overcome the intermediate plateau.

The focus of the project was on three key learning activities: multiple choice activities, gap filling activities and open exploratory activities. Having established in our project the statistical viability of the selected collocations, the next process to show to the learner is what patterns the collocation forms with surrounding co-text.

Below is an example of how this type of investigation might be introduced in the classroom, examining the patterns prevalent with the collocation “torrential rain”. It is necessary for the teacher to research on underlying patterns for the collocation first. The first lesson here directs learners to the knowledge that a significant pattern of this collocation is that of a passive voice construction.

Lesson one:Task- Understanding the patterns for “ torrential rain”

- 1) Using the BOE undertake a search for torrential+ rain
- 2) Now create a “T” score picture tree
- 3) Establish what is the node word and then look at node -1
- 4) Now fill in the following table

Rank by T-score	Node -1 Pattern	Collocation
1	<i>by</i>	torrential rain
2	<i>in</i>	torrential rain
3	<i>of</i>	torrential rain

Figure 14

The second lesson expands the knowledge of this colligation pattern by investigating the statistical collocates of this pattern:

Lesson two: Investigation of pattern 1- multiple choice

Which of the following is the correct:

- By the resort torrential rain pounded
- Torrential rain swept by the hills
- The resort was pounded by torrential rain

Can you explain why ?

Now use the following table to record collocates of this pattern

Generators	Pattern	Collocation
	by	torrential rain

Figure 15

What is important here is the use of terminology. It is suggested that the term generator is employed instead of that of verb or noun. Through this term learners can draw the conclusion that certain lexis generates specific patterns. A discussion of the use of the word generator would facilitate knowledge of the constraints that words placed on forming patterns.

Conclusion

The importance of word constraints in learning a second language must be stressed. The prevalence of formulaic sequences and in particular of collocations means that a learner cannot achieve mastery of a language without competence in this area. The Lexical approach has this premise at the heart of its curriculum and its strength lies in the acquisition of two or three word collocations. It has been shown, however, in this study that examining colligation patterns underlying collocations can lead a learner to formulate longer and more fluent stretches of communication. Using a corpora is essential in establishing these statistically reliable colligation patterns. Then by using a data driven learning approach a learner can examine and extrapolate these important patterns. The loss of the advantage of processing time in the analysis of the composition of the phrases will be compensated for by the ability of the learner to form longer and more fluent stretches of communication.

Bibliography

- Benson, M., Benson, E. & Ilson, R. (1986) "The BBI Combinatory Dictionary of English." John Benjamins Publishing Company, Amsterdam
- Boers F. Lindstromberg S. (2008) "Cognitive linguistic approaches to teaching vocabulary and phraseology"
- Bernedini S. (2004) "Corpora in the classroom." In Sinclair J. "How to use corpora in language teaching." John Benjamins Publishing Company, Amsterdam
- Channell, J. (1981) 'Applying semantic theory to vocabulary teaching'. *ELT Journal* 2:115-122.
- Condon N. (2008) "How cognitive linguistic motivations influence the learning of phrasal verbs." In Boers F. Lindstromberg S. (2008) "Cognitive linguistic approaches to teaching vocabulary and phraseology"
- Ebeing S. (2004) "Oslo Interactive English." In Sinclair J. "How to use corpora in language teaching." John Benjamins Publishing Company, Amsterdam
- Firth, J.R. (1957) *Papers in Linguistics: 1934-1951*. London: Oxford University Press.
- Gardner D. and Davies M. (2007) 'Pointing out Frequent Phrasal Verbs: A Corpus-Based Analysis.' *TESOL Quarterly*, Vol. 41, No. 2 pp. 339-35
- Hill, J. (2000) 'Revising priorities: from grammatical failure to collocational success'. In Lewis, M. (ed.) (2000) *Teaching Collocation: Further Developments in the Lexical Approach*. London: Heinle Cengage Learning.
- Hoey, M. (2005). *Lexical Priming*. Oxon: Routledge.
- Hunston, S. & Francis, G. (2000). *Pattern Grammar*. Amsterdam/Philadelphia: John Benjamins.
- Lewis, M. 1993. *The Lexical Approach*. Hove: Teacher Training Publications
- Lewis, Michael (1997) *Implementing The Lexical Approach: Putting theory into practice*, Heinle, Boston.
- Lewis, M. (ed.) (2000) *Teaching Collocation: Further Developments in the Lexical Approach*. London: Heinle Cengage Learning.
- Mackin, Ronald. "On collocations: Words shall be known by the company they keep." *Honour of AS Hornby* (1978): 149-165.
- Nattinger, J.R. and DeCarrico, J.S. (1992) *Lexical Phrases and Language Teaching*. Oxford: Oxford University Press.

Schmitt and Carter (2004) "Formulaic sequences in action." P1- 22 In Formulaic Sequences : Acquisition, Processing, and Use. John Benjamins Publishing Company

Sinclair J. " How to use corpora in language teaching." John Benjamins Publishing Company, Amsterdam

Sinclair J.M. (1991) Corpus, Concordance, Collocation Oxford: Oxford University Press

Walker C. (2008) "Factors which influence the process of collocation." Boers F.

Lindstromberg S. (2008) "Cognitive linguistic approaches to teaching vocabulary and phraseology"

Willis, D. 1990. The Lexical Syllabus. London: HarperCollins

Woolard, G. (2000) 'Collocation – encouraging learner independence'. In Lewis, M. (ed.)(2000) Teaching Collocation: Further Developments in the Lexical Approach. London: Heinle Cengage Learning.

Wray, A. 2000. 'Formulaic Sequences in Second Language Teaching: Principle and Practice'. Applied Linguistics 21/4: 463-489

Wray A. 2002 Formulaic Language and the Lexicon: "Formulaic Sequences in the Second Language Acquisition Process: A Model"

Appendix

Appendix 1

BBI Typology of collocations- Grammatical collocations

Notation	Classification
G1	Noun + preposition
G2	Noun + to infinitive
G3	Noun + that clause
G4	Preposition + noun
G5	Adjective + preposition
G6	Adjective + infinitive
G7	Adjective + that clause
G8 A	Verb + indirect object + direct object
G8 B	(Trans) Verb + direct object + indirect object
G8 C	(Trans) Verb + direct object + for + indirect object
G8 D	Verb + preposition
G8 E	Verb + to infinitive
G8 F	Verb + infinitive
G8 G	Verb + verb-ing
G8 H	(Trans) Verb + direct object + to infinitive
G8 I	(Trans) Verb + direct object + infinitive
G8 J	Verb + object + verb-ing
G8 K	Verb + possessive(pronoun/noun) + gerund
G8 L	Verb + that clause
G8 M	(Trans) Verb + direct object + to be + adjective/ past participle/noun/pronoun
G8 N	(Trans) Verb + direct object + adjective/ past participle/ noun/ pronoun

G8 O	(Trans) Verb + object + object (not allowing “to” or “for” prep phrases)
G8 P	Verb + adverbial
G8 Q	Verb + interrogative
G8 R	(Trans) It + verb + to/that clauses
G8 S	(Intrans) Verb + noun/adjective

Lexical collocations

Notation	Classification
L1	Verb + noun/pronoun
L2	Verb + article + noun
L3	Adjective + noun
L4	Noun + verb
L5	Noun + of + noun 2
L6	Adverb + adjective
L7	Verb + adverb

Appendix 2-

Statistical collocations of the senses of phrasal verbs – “take up” and “pick up”

Take up

Sense	Collocation	T-score
to occupy	residence	11.2
to occupy	space	4.9
to develop an interest in	golf	4.7
to develop an interest in	exercise	3.1
to accept	work	3.9
to accept	employment	3.8
to accept	position	3.4
to use up	time	3.0

Pick up

Sense	Collocate	T-score
To collect or gather	information	3.78
To collect or gather	litter	3.74
To gain	speed	6.23
To gain	business	3.16
To gain	momentum	2.8
To gain	support	2.5

Appendix 3

Lexical notebook- adapted to disambiguate differing senses of a word

Word	Translation (Spanish)	Sense 1	Sense 2	Sense 3	Sense 4
Presence	presencia	<i>The state or fact of being present</i>	<i>Immediate proximity in time or space</i>	<i>The diplomatic, political, or military influence of a nation</i>	<i>A person who is present</i>
Collocates		strong mere constant reassuring	global	military international naval	human female

The word “presence” and a Spanish translation

How a lexical notebook can be used to disambiguate the differing senses of a word

Appendix 4

The chosen text

The Pleasures of London Cycling *Deborah Moggach*

1)I've biked in London all my life. My children bike; everyone
2) I know, with any sense, bikes. How else does one get
3)around the place? Nothing beats weaving through the
4)rush-hour traffic or whizzing past the eternal gridlock that is
5)the Strand. Cycling is the only way to free ourselves from
6)the misery of the tube, the wall-to-wall buses that line
7)Oxford Street, the hopelessness of even thinking about
8)driving. It's the only way you can get where you want to go,
9)on time, fast.

10)We all know the advantages of cycling. It's free, it's non-
11)polluting and it's a wonderful way to keep fit. Isn't it weird
12)that people take the tube to work and then spend their lunch
13)hour sweating away on an exercise bike? What is less
14)recognised, however, is its efficiency. You know exactly
15)when you'll arrive at your destination. It takes me 25
16)minutes from my front door to the West End. Any other form
17)of transport entails allowing for delays – endless waits,
18)traffic jams. What a waste of life!

19)Not only that. I can do errands en route. Drop off a parcel,
20)do a bit of shopping here and there, pick up some tickets,
21)stop for a coffee. And I do these things on impulse. Once
22)you start cycling, the city opens up for you. No longer are
23)you fighting it, hot and frustrated. No longer are you at the
24)mercy of bus drivers, roadworks, decisions made by others
25)and over which you have no control. Believe me, once
26)you've tasted this freedom, you're hooked.

27)More and more of us are taking up cycling. Side streets are
28)less polluted, and more interesting anyway. Besides, the
29)more of us who cycle the safer it is. Traffic has to
30)accommodate us. We also make the streets safer. For a
31)lone woman, walking home at night, a passing cyclist is a
32)reassuring presence. We give the city a human scale.

33)Some people find the notion of confronting London's
34)aggressive car drivers scary, especially in places like Hyde

35)Park Corner, but if you're spooked you can always get off
36)and walk. And as long as you treat all drivers as idiots who
37)are just about to fling open the door of their parked car, or
38)cut you up with their horrible 4x4, or squash you with their
39)ghastly bendy bus, you'll be okay. Of course there are other
40)disadvantages, like arriving at your destination sodden with
41)sweat and with hideous 'helmet hair'. Sexy clothes and
42)biking don't go together, so one has to pop into an alleyway
43)to change, hopping about on one foot and looking a bit of a
44)twit. And of course, cycling is no fun in torrential rain.
45)Drivers can be aggressive, but I've met with a lot of
46)politeness. Perhaps it's because I look harmless tootling
47)along with my basket of shopping.

48) That said, how friendly is London to cyclists? There is still a
49)scandalous lack of facilities. Ever noticed how few cycle
50)racks adorn our streets? So we have to lock our bikes to a
51)parking meter or a lamppost. The bike then falls over and
52)gets damaged, or trips up a pedestrian. Worse still, many
53)railings have a sign saying 'Bikes will be removed' – railings
54)of offices whose bosses arrive each day in chauffeur-driven
55)cars and, best of all, railings at a heart hospital! Isn't that
56)lunatic?

57)Still, if you love London there's no better way to see it. On a
58)bike, it's you who owns the city, instead of the city owning
you.

<http://www.wjec.co.uk/uploads/papers/w12-4171-02a.pdf>

Downloaded on 22/04/2013