

## Essay Cover Sheet

Student ID number	972700
Module Number (1-6)	4
Title of Degree Programme:	MA Applied Linguistics
Title of Module:	Corpus Linguistics
Assessment Task No.	CL/09/02
Submission	1
Date Submitted	31/07/09
Name of tutor	Christine Mackie
Word Count	4490 (not including long quotations and figures/tables)

### Essay contents

<b>Section</b>	<b>Title</b>	<b>Page</b>
1	Introduction.....	1
2	The idiom principle and phraseology.....	1
3	Phrases and their identification.....	2
	3.1 Types of 'semi-preconstructed phrase'.....	2
	3.2 Identifying phrases.....	3
4	Previous studies on the idiom principle.....	3
5	Choice of text and corpus.....	4
6	Using a text to distinguish between the idiom principle and the open choice principle.....	4
	6.1 The methodology used for identifying semi-preconstructed phrases in the text.....	4
	6.2 The identification of phrases and instances of open choice.....	5
	6.3 Calculating the idiom and open choice principles for a text.....	13
	6.4 Discussion of findings.....	14
7	Pedagogical applications.....	15
8	Conclusion.....	17

CL/09/02

Select a short text extract (about six lines) and use it to demonstrate the idiom principle and the open-choice principle (Sinclair 1991). You should raise questions of how the two may be defined and distinguished, and how the proportion of each may be quantified. Begin with an account of Sinclair's theory, and end with a short discussion of applications to teaching or to translation.

## 1. Introduction

Sinclair's idiom principle (1991) holds that most text is constructed from ready made phrases; it contrasts with the traditional view that most text production involves 'free' choices. This paper attempts to use a short text to distinguish between the idiom and open choice principles, applying the principles of phraseology derived from Sinclair's work. Section 2 provides a brief overview of the idiom and open choice principles and their relationship to phraseology. Section 3 overviews types of 'semi-preconstructed phrase' (Sinclair, 1991) and how they may be identified. Section 4 gives a brief overview of previous studies that have attempted to distinguish between the idiom principle and the open choice principle and estimate their respective contributions to text. Section 5 outlines the reasons for the choice of text and reference corpus. Section 6 offers an analysis of the text and its constituent phrases, as well as a means of calculating the 'extent of the phraseological tendency' (Cheng et al, 2008) and a brief overview of the limitations of the approach taken. Finally, section 7 discusses pedagogical applications.

## 2. The idiom principle and phraseology

Sinclair's idiom principle states that

'A language user has available to him or her a large number of semi-preconstructed phrases that constitute single choices, even though they might appear to be analysable into segments' (Sinclair, 1991: 110)

This principle is perhaps best understood in contrast with what Sinclair termed the 'open choice principle' (ibid.), or the lack of acknowledgement of the interrelationships between lexis and grammar in many traditional models of language. According to such models, grammar and lexis are largely independent of each other (Francis, 1993), implying that at almost any point in a clause or utterance almost any member of an appropriate word class might be chosen freely. Theorists who take this approach therefore consider the description and explanation of idioms and phrases peripheral to models of language (Ellis, 2008). This is a view of language that Sinclair (1991 et seq) indicated was misleading in most cases, holding that the default principle in interpreting text is the

idiom principle (Sinclair, 1991). Other theorists who have noted the importance of phraseology to descriptions of language and native-speaker competence include Becker (1975), Pawley & Syder (1983), and Nattinger & De Carrico (1992).

### **3 Phrases and their identification**

#### **3.1 Types of ‘semi-preconstructed phrase’**

Sinclair (2008) postulates two main types of phrase, the ‘phraseological item’ and the ‘lexical item’.

The phraseological item is largely based on Sinclair and Renouf’s (1991) ‘collocational framework’ where two or more closed-class items enclose a gap to make ‘skeletons of full-bodied phrases’ (Sinclair, 2008; 408), an example of which would be *a \* of*. As Sinclair and Renouf (ibid.) indicate, the items that occur in the gap of *a \* of* can be categorized to give evidence of phraseological tendencies in the manner of Pattern Grammar (Hunston & Francis, 2000).

Sinclair’s ‘lexical item’ (1996; 1998) is the hypothesis he created to account for the variation that occurs within the semi-preconstructed phrase while also indicating its regularities; to identify it, the researcher needs to ‘exercise [their] intuition in the presence of as much relevant data as can be assembled’ (Tognini-Bonelli, 2001: 178). The item contains a lexical core which is built up by considering the frequent co-selection of the core and specific words, or collocation; the co-selection of the core and members of a grammatical class or a grammatical form, or colligation; and the association between the core of the phrase and a semantic set, or ‘semantic preference’ (Sinclair, 1996, 1998). At the level of the phrase Sinclair theorizes that there is a ‘semantic prosody’ (ibid.), which is the attitudinal, pragmatic meaning which is revealed when one looks at the lexical item as a whole in its co-text and, where possible, its context (Sinclair, 1996, 1998; Hunston, 2007; Stubbs, 2009).

The important thread that unites all these phraseological phenomena is the ‘co-selection of lexis and grammar’ (Stubbs, 2009; 120) to express particular meanings in particular social situations (Hunston, 2007; Sinclair 2008; Stubbs, 2009). This allows us to view phraseology and idiomaticity in their full variety; while it may be possible to identify a ‘canonical form’ of a phrase or collocation, one has to accept a certain amount of ‘turbulence’ to which this canonical form is subject in the course of everyday use (Cheng et al, 2008; Sinclair, 1991; Francis, 1993; Philip, 2008).

This leaves a definition of a phrase that is relatively broad: a semi-preconstructed phrase is a combination of items that is frequently used to express a particular meaning in a particular context but which allows for a degree of flexibility ‘perhaps with some discontinuity thrown in’ (Sinclair, 2004b: 5). This means that the ‘open choice’ principle can be said to be operating either outside these semi-preconstructed phrases or in the slots in such phrases that can be filled by a wide range of items.

### **3.2 Identifying phrases**

Accepting a broad definition of a semi-preconstructed phrase based on Sinclair’s theories carries with it certain methodological and theoretical implications. The first of these is that one takes a distributional approach to identifying phrases, focusing on those that are repeated relatively frequently – there are varying views as to what level of frequency is significant (Gries, 2008). An important though limited tool in this process is the intuition of the researcher; it is accepted that, while intuition is not reliable when estimating word frequencies or explaining grammatical ‘rules’ (Sinclair, 1991), a native speaker can be relied upon to say whether an utterance is idiomatic or not (Pawley & Syder, 1983; Owen, 1996). This means that intuition acts as an initial guide when formulating corpus searches as well as in identifying patterns in the data. A second implication is that a researcher needs to employ very flexible search techniques to identify semi-preconstructed phrases since some semi-preconstructed phrases do not even share a lexical item as a common core (Francis, 1993; Philip 2008).

### **4. Previous studies on the idiom principle**

A number of researchers have estimated the proportion of text consisting of phrases; examples include Pawley and Syder (1983) and Mel’čuk (1998), who estimate a ratio of phraseological units to individual words in a language to be approximately 10 to 1. Wray (2002; 28) reports on several studies that have tried to establish the percentage of particular corpora that consist of prefabricated language (‘prefabs’) and whose findings range from 4% to 80%, depending on what analysts include within the scope of phraseology. One important example is Erman and Warren’s (2000) study, which estimates that ‘prefabs’ make up around 55% of the texts that they studied. They also present the method they use to calculate this figure which involves counting the number of ‘slots’, or orthographic words in the text, identifying which of them are part of phrases and thereby establishing the percentage of text which forms part of what they term ‘prefabs’ (ibid).

## **5. Choice of text and corpus**

The text extract chosen for this assignment (Gardner, 1991; see Appendix 1) forms the introduction to an authentic text reprinted in the in-house produced EAP textbook used at the university where I teach. The text is written by Prof. Howard Gardner, an American academic, for the online journal ‘New Horizons for Learning’, which is aimed at education professionals. Since the text is typical in terms of register and style of those used in the textbook as a whole, it was thought that such an analysis might provide useful insights into the phraseology typically used in this context, as well as providing useful teaching material both for productive and receptive use; examples of such material are provided in Section 7 below.

The methodology used requires access to a large corpus to check whether potential phrases, or those identified intuitively, are frequently attested; as Sinclair (1991) and Hunston (2002) point out, the larger the corpus, the more reliable the results obtained. The Bank of English, with 450 million tokens, is a suitably large corpus. Since the corpus access software provides information as to the incidences per million for a search for each of the subcorpora of which it is composed, it also allows the researcher to check whether a particular phrase is more prevalent in the register being investigated and whether it is worth focusing particular attention on (Flowerdew, 2004).

## **6. Using a text to distinguish between the idiom principle and the open choice principle**

### **6.1 the methodology used for identifying semi-preconstructed phrases in the text**

The method used to search for semi-preconstructed phrases in the text is informed by those outlined by Francis (1993), Groom (2007), Philip (2008), and Sinclair (1991 et seq). First of all, possible semi-preconstructed phrases were identified in the source text; it was decided where possible to focus on instances of closed class items in collocational frames (Sinclair & Renouf, 1991) since excluding lexical words broadens the search and tends to be more revealing in terms of phraseology (Groom, 2007).

Frames and phrases identified were then checked for attestations in the Bank of English. Extensive use was made of the ‘picture’ tool available to Bank of English users, which provides information about the most significant collocates of a node at each position for up to 6 items either side, and can therefore suggest further lines of investigation, such as members of a similar semantic set that may replace a particular item in a phrase; concordances were also searched for repeating patterns. The

‘picture’ data was based on t-scores since the calculation of t-scores by its nature provides information on ‘upward’ collocation, or co-selection with more frequent items (Sinclair, 1991; Barnbrook, 1996), meaning that a ‘picture’ of the most common t-score collocates of a particular node can give an indication of phraseological patterning (Hunston, 2002).

## 6.2 The identification of phrases and instances of open choice in the text

The foregoing analysis will examine the text (see Appendix 1) one sentence at a time, giving information about the semi-preconstructed phrases and instances of open choice tentatively identified, the means of identification and the corpus evidence for such findings. All data provided is from the Bank of English unless otherwise stated; dates of searches are given on the recommendation of Walker (2009). Specific word forms are given in italics, but where a lemma rather than a word form is considered, it is given in capital letters. Semantic sets are indicated by putting the descriptor in brackets; where search formulas are given, they are enclosed by inverted commas.

[1] The concept of intelligence, a very old one, has been employed in the most varied ways over the centuries.
--

Since the frame *the \* of* is so common, with 81,222 hits (29/07/09) in the Bank of English’s ‘usacad’ corpus, which consists of academic textbooks (Oahey, personal communication) and hence reflects closely the register of interest in this assignment, it was possible to focus on instances occurring solely in this subcorpus. Examination of the items that have the highest t-scores for the intervening slot revealed that ‘concept’, at twenty-first in the list, is amongst the most significant collocates, suggesting that *the concept of \** is a semi-preconstructed phrase in academic textbooks. The paucity of instances of *the concept of intelligence*, with nine in the whole Bank of English, precludes considering it as a stand-alone phrase; it was therefore seen as an instance of ‘open choice’.

A search for the frame *a very \* one* gives 1261 hits in the Bank of English, indicating it is a relatively common ‘phraseological item’, though instances by subcorpora indicate this is far more commonly used in spoken language, with around 4 times as many instances per million, than in academic registers. A number of evaluative adjectives can complete this phrase, one of the most frequent of which is *old* (see Figure 1), suggesting that *a very old one* is a semi-preconstructed phrase in its own right, or at least a relatively common exponent of the lexical item (BE) *a very [evaluative adjective] one*.

<b>'a+very+1,1one': t-score picture from Bank of English (retrieved on 17/07/09)</b>				
is	a	NODE	good	one
was		NODE	important	
not		NODE	difficult	
and		NODE	old	
be		NODE	real	
s		NODE	long	
but		NODE	happy	
been		NODE	serious	
albeit		NODE	small	
though		NODE	different	

Figure 1

A search to investigate whether *the concept of* is part of a phrase completed by EMPLOY provided little evidence of this (see Appendix 2). However, there appears to be a strong relationship of colligation between *over the centuries* and the 'present perfect' form of the verb (see Figure 2), suggesting that the use of this phrase at the end of this sentence influences the choice of this form of the lemma EMPLOY. Figure 2, however, indicates that *over the centuries* predominantly co-occurs with verbs that express the meaning 'change' rather than 'use', casting doubt on the phraseological status of *has been employed ... over the centuries*.

<b>'over+the+centuries': t-score picture first ten rows, left-hand side (retrieved on 20/07/09)</b>						
has	have	have	developed	over	the	NODE
have	has	has	up			NODE
that	been	been	evolved			NODE
of	had	built	<p>			NODE
had	many	and	changed			NODE
been	by	many	little			NODE
how	church	changed	which			NODE
attracted	of	of	accumulate			NODE
built	changed	its	altered			NODE
place	which	grown	europe			NODE

Figure 2

While the phrase *in the most varied ways* is not attested in the corpus, searches of the strings *in \* \* \* ways*, *in \* \* ways* and *in \* ways*, shown in Figure 3, indicate that it is a variant of a common lexical item which is used to talk about differences in how certain things are done or in how people behave. Evidence for this can be seen in the presence of words relating to or suggestive of difference or variety. Interestingly, USE appears to collocate with this item, suggesting that the choice of EMPLOY in this text was a stylistic one.

**'in+3,3ways' t-score picture, left-hand side; first 10 rows (retrieved on 20/07/09)**

be	used	in	a	variety	of	NODE
is	them		one	number	two	NODE
are	itself		all	of	different	NODE
been	done		their	lot	many	NODE
their	up		any	sorts	three	NODE
can	me		so	and	some	NODE
do	meeting		new	in	other	NODE
was	together		lots	many	discuss	NODE
your	behave		fact	kinds	various	NODE
to	work		other	least	several	NODE

**'in+2,2ways' t-score picture, left-hand side; first 10 rows (retrieved on 20/07/09)**

can	be	used	in	many	different	NODE
to	are	behave		their	many	NODE
has	their	but		so	other	NODE
and	been	work		two	own	NODE
may	is	me		very	important	NODE
with	were	them		more	following	NODE
both	can	different		several	main	NODE
could	each	expressed		lots	subtle	NODE
are	use	both		all	same	NODE
we	do	us		three	specific	NODE

**'in+1,1ways' t-score picture, left-hand side; first 10 rows (retrieved on 20/07/09)**

can	can	be	is	in	many	NODE
they	this	are	but		some	NODE
is	to	is	and		different	NODE
to	different	their	<p>		other	NODE
that	has	been	set		two	NODE
this	could	different	was		various	NODE
different	may	very	think		several	NODE
it	they	this	them		more	NODE
think	are	can	itself		what	NODE
but	use	i	behave		certain	NODE

Figure 3

[2] During the past century, there has been considerable movement on the "intelligence front," and this trend shows no sign of stopping.

The string *during the past century* is not frequent, with just 26 hits in the corpus. However, *during the past*, which occurs 3434 times, is commonly followed by a period of time (see Figure 4), confirming that this is a semi-preconstructed phrase:

*during the past* [period of time]

which, like *over the centuries* seems to strongly attract the present perfect form.

**‘during+the+past’ t-score picture, right half (retrieved on 20/07/09)**

NODE	the	past	two	years	mr	has
NODE			year	months	years	have
NODE			few	days	<p>	been
NODE			five	weeks	has	unruh
NODE			decade	decades	have	chamberlai
NODE			three	fiscal	the	stork
NODE			week	century	there	years
NODE			six	seasons	he	jessup
NODE			four	hours	but	were
NODE			12	year	according	had

Figure 4

Evidence derived from the usacad corpus (see Figure 5), chosen here because of apparent differences from general Bank of English findings for this phrase, supports the intuition that *there has been* is a phraseological item, suggesting further that it forms part of a complex lexical item

*there has been/occurred/emerged/developed* [extent] [change]

frequently used with periods of time. The fact that the items *considerable* and *movement* do not co-occur in the corpus merely reflects the variability available; they are compatible with the whole phrase in terms of meaning and function.

**‘there+0,1has’ usacad corpus t-score picture, right half, top twelve rows\* (retrieved on 19/06/09)**

NODE	has	been	a	a	in	in
NODE	also	also	been	progress	increase	understand
NODE		not	no	decline	of	of
NODE		has	an	too	trend	decline
NODE		never	much	steady	decline	increase
NODE		occurred	some	tendency	little	shift
NODE		however	little	overall	change	difference
NODE		emerged	substantia	revolution	among	interest
NODE		recently	extensive	little	about	on
NODE		indeed	progress	major	narrowing	the
NODE		developed	increasing	study	empirical	indisputab
NODE		always	significan	long	reduction	toni

**\*to include all items in column 3**

Figure 5

*On the intelligence front* also appears to be part of a discontinuous phrase *on the \* front*, where the symbol \* can represent one to three words. Evidence is provided by searching for ‘on+the+1,3front’, which gives over 2800 hits in the Bank of English. Certain forms of this phrase are very common and might be considered as one choice, but *intelligence* is not one of these; this is an example of an open choice. It is also interesting to note that this phraseological item also attracts the present perfect form and overlaps to some extent with the previous item in what Hunston and

Francis (1998: 68) term ‘pattern flow’ in that both *progress* and *development* are among the more significant collocates (see Figure 6).

<b>‘on+the+1,3front’ t-score picture, first ten items, left of node (retrieved on 20/07/09)</b>					
on	on	on	the	western	NODE
all	<p>	the	national	home	NODE
good	quiet	wider	internatio	eastern	NODE
german	news	political	human	domestic	NODE
bad	forces	bed	interest	economic	NODE
little	war	double	wider	political	NODE
been	progress	military	foreign	diplomatic	NODE
quiet	fighting	lithograph	mountain	central	NODE
more	action	headquarte	arab	internatio	NODE
news	developmen	mkiii	israeli	work	NODE
<b>‘on+the+1,3front’ t-score picture detail (retrieved on 20/07/09)</b>					
NODE	the	has	has	have	war
NODE	there	is	also	front	a
NODE	but	are	been	has	of
NODE	and	s	is	new	series
NODE	he	the	were	to	agreed
NODE	in	well	are	that	germany
NODE	line	was	s	world	bad
NODE	<p>	have	war	been	week
NODE	bench	will	not	were	left
NODE	it	also	had	war	nato

Figure 6

*This trend* accounts for around 10% of the instances of *trend* in the corpus, suggesting that they form one choice. The t-score picture for *this trend* (see Figure 7) further indicates it frequently co-occurs with CONTINUE or its instancial antonym REVERSE; when writers use the word *trend*, it is often to say whether it will continue or not.

<b>‘this+trend’ t-score picture detail (retrieved on 20/07/09)</b>						
this	NODE	is	continue	to	continue	the
	NODE	has	be	be	by	local
	NODE	continues	not	by	the	years
	NODE	will	likely	facilitate	reverse	least
	NODE	was	been	continue	is	far
	NODE	may	expected	will	women	american
	NODE	towards	set	reversed	pronounced	more
	NODE	toward	reversed	as	obvious	reversed
	NODE	but	also	developed	on	next
	NODE	continued	already	recently	rates	itself

Figure 7

A search for ‘no+sign|signs+of’ (see Figure 8) revealed evidence of the lexical item

SHOW *no sign(s) of* [slowing/finishing].

Other relevant items appearing further down the second column to the right of the node include *ending, easing, waning, diminishing* and *stopping* (see Appendix 2). Corpus evidence suggests that the choice of item preceding SHOW, while relatively open, constrains the choice of item that follows *no sign(s) of*, leading to the conclusion that at the level of phraseology two choices are involved here, the core – SHOW *no sign(s) of*– and the items surrounding this core, which attract each other, a phenomenon pointed out by Hunston (2009).

<b>‘no+sign signs+of’ t-score picture, first ten items (retrieved on 20/07/09)</b>						
there	was	no	NODE	of	a	or
he	shows		NODE		any	down
has	showed		NODE		abating	up
was	is		NODE		life	in
had	show		NODE		the	yet
is	showing		NODE		slowing	break
they	with		NODE		an	entry
but	still		NODE		him	breakthrou
still	were		NODE		being	<p>
she	shown		NODE		letting	so

Figure 8

[3] In this essay I briefly describe seven historical steps, or phases, in the development of thinking about intelligence, focusing in particular on work inspired by the theory of multiple intelligence.

The results of a search for *in this \* I* (see Figure 9), which was limited to the usacad corpus in an attempt to minimize irrelevant concordance lines, led to the hypothesis of the following lexical item, which in academic writing has the specific function of telling the reader exactly what the forthcoming text will be about:

*In this* [piece of writing] *I* (will/intend to) [talk about]

There also seems to be a closely related item which seems to be used to summarise the main idea or ideas – for example ‘*In this book, I have tried to describe the ...*’.

**'in+this+1,i' usacad corpus; concordance lines (retrieved on 19/06/09)**

ips founded in feminist practice. In this book, I will explore the means by  
 minist social change goals. Later in this chapter I will explore at greater  
 m and ethno-racial particularism. In this chapter, I want to explain the appeal  
 ing David Tracy Andrew M. Greeley In this paper I propose to outline my efforts  
 oth biological and psychological. In this section, I'll briefly review some of  
 Chapter 5 (see pp. <bx> and <bx> In this section, I will list some of the more  
 is, as `postsocialist" societies. In this chapter I am concerned only with those  
 ng featured in the movie. <p> 33. In this paragraph I am responding to Charles  
 to be seen. Have you noticed that in this section I have tried to focus on the  
 elves in such a generic portrait. In this book, I have tried to describe the  
 r with sufficient conviction. <p> In this book, I have tried not to play this  
 tituted centuries before. Earlier in this book I have discussed the hypothesis  
 re in this territory anyway." <p> In this volume I have chosen to refer to  
 arguments of the kind I am making in this chapter. I believe that proponents of  
 ," depending on your experience. (In this experiment, I won't allow you the

Figure 9

The hypothesis that, since many academic textbooks have more than one author, *we* was a likely substitute for *I* prompted the search 'in+this+1,i|we'. This confirmed the initial hypothesis of a lexical item, with 203 hits for the phrase *in this* [piece of writing] *I/we*; some editing was required to remove strings such as *in this way we*. The t-score picture for this phrase (see Figure 10) indicates, moreover, that, the position following the reporting verb is commonly filled either by an explicit number or a vague expression indicating that there is more than one entity or phenomenon that will be described, explored or outlined. This leads to the framing of the longer lexical item:

*In this* [piece of writing] *I/we* (will/intend to) [talk about] [an explicit or vague number of] entities/phenomena

**'in+this+1,i|we' t-score picture (retrieved on 08/08/09)**

in	this	chapter	NODE	will	consider	some	of
		section	NODE	ll	how	how	some
		topic	NODE	have	review	the	few
		book	NODE	examine	see	two	with
		appendix	NODE	focus	briefly	three	these
		paragraph	NODE	develop	examine	review	major
		volume	NODE	consider	explore	at	important
		paper	NODE	look	reviewed	a	overview
		discussion	NODE	explore	focus	firms	examining
		analysis	NODE	ve	discuss	on	describe
		study	NODE	discuss	look	only	on

Figure 10

There is insufficient evidence to show that *steps* or *phases* are part of this lexical item; they are just two of a large number of possible items that follow it. However, there is evidence to show that *step(s)*, *phase(s)* and the related *stage(s)* commonly precede the strings *in the development of* and the related *in the evolution of* (see Figure 11); this is hardly surprising since it is normal to

conceptualize development or evolution as happening in steps or stages. It can also be seen from Figure 11 that the phrase attracts adjectives with the general meaning of ‘important’ or ‘significant’, of which *historical* could be said to be a member, although such adjectives tend to co-occur with the singular forms *stage*, *step*, and *phase*. Once again pattern flow seems to link this phrase and the preceding one through their attraction of adjectives meaning ‘important’ and also numbers, indicating that the presence of the item *seven* here is not an instance of open choice, while making it harder to say which phrase it ‘belongs to’.

One might propose the lexical item:

stage(s)  
 [important] step(s)                    in the development/evolution of  
 phase(s)

<b>‘stage step phase+in+the+development evolution+of’ Bank of English t-score picture</b>						
important	stage	in	the	NODE	of	the
next	step			NODE		a
new	phase			NODE		our
another				NODE		human
first				NODE		psychoanal
early				NODE		israeli
crucial				NODE		industrial
latest				NODE		culture
key				NODE		american
major				NODE		life

*Figure 11*

The choice of items to follow *the development|evolution of* is a relatively open one judging from the evidence available. However, the following two words, *thinking about*, are clearly closely associated, which is indicated by the fact that *about* is the most frequent item immediately after *thinking* in the usacad corpus, occurring in about 15% of all instances; this presents strong evidence that they are a semi-preconstructed phrase, one confined to two words since there is no apparent pattern in the words that follow (see Appendix 2), making the choice of *intelligence* an open one.

The t-score picture derived from the search ‘focus+0,2on’ (see Figure 12) provides evidence that *in particular* is one of several words or phrases following *focusing* that serves to narrow the focus, other options including *only*, *solely*, *exclusively* and *particularly*.

**‘focusing+0,2on’ t-score picture: first 10 items at R1, R2 and R3 (retrieved 29/07/09)**

NODE	on	the	on
NODE	attention	on	historical
NODE	instead	attention	issues
NODE	more	a	and
NODE	their	how	particular
NODE	only	what	core
NODE	too	one	own
NODE	solely	particular	rather
NODE	exclusivel	much	business
NODE	particular	specific	problems

*Figure 12*

While it is possible to find evidence that the string *focusing in particular on* forms one variant of a semi-preconstructed phrase, there is no evidence that *work* or even its close contextual synonym, *research*, co-occurs with this phrase, suggesting this is a more open choice.

It is clear that *inspired by* is a common combination accounting for more than 35% of the occurrences of *inspired* in the Bank of English. Searching ‘work+0,4inspired+by’ gives 62 hits in the corpus (see Appendix 2), a finding which suggests that *work* may be co-selected with *inspired by* in this instance, although it is acknowledged that there is not strong evidence for this; this phrase does not have an incidence rate higher than 0.5 per million in any of the Bank of English subcorpora.

Regarding the string *the theory of multiple intelligence*, this is clearly what Wray (2002, 2008) would term a ‘formula’ for Gardner, since he uses it four times in the text alone (see Appendix 3). There is also evidence that this phrase is becoming current, with 31 hits for the search ‘multiple+intelligence|intelligences’ in the Bank of English, of which 12 occur with *theory of*, *theories of*, or *concept of* (see Appendix 2).

### **6.3 Calculating the extent of the idiom and open choice principles for a text**

Based on the findings of the analysis in section 6.2, it is possible to divide the text up into putative semi-preconstructed phrases and calculate how many of the available slots in the text form part of these phrases and how many of these slots are ‘open choice’ ones. The term ‘slots’ is used instead of ‘words’ since even working according to the ‘open choice’ principle some slots are not open; the ‘choice’ involved, for example, after ‘has’ when forming the passive is at a different ‘stratum’ (Sinclair, 2008: 408) from that involved in the choice of the lemma EMPLOY. To calculate the

percentage of the text that is covered by the idiom principle, one then divides the number of slots which are part of semi-preconstructed phrases by the total number of slots in the text and multiplies this figure by 100 (see Figure 13). In the case of the text analyzed in this paper, the figure thus gained is 92.3%.

<u>Semi-preconstructed phrase / open choice item</u>	<b>number of slots</b>	<b>Number of slots forming parts of semi-preconstructed phrases</b>
<u>The concept of intelligence</u>	4	3
<u>a very old one</u>	4	4
<u>Has been employed</u>	1	-
<u>in the most varied ways</u>	5	5
<u>over the centuries</u>	3	3
<u>During the past century</u>	4	4
<u>there has been considerable movement</u>	4	4
<u>on the "intelligence front,"</u>	4	3
<u>And</u>	1	-
<u>this trend ... [not] stopping</u>	3	3
<u>shows no sign of</u>	4	4
<u>In this essay I briefly describe seven</u>	7	7
<u>(seven) historical steps [or phases] in the development of</u>	6	6
<u>thinking about intelligence</u>	3	2
<u>focusing in particular on</u>	4	4
<u>work inspired by</u>	3	3
<u>the theory of multiple intelligence.</u>	5	5
<b>Total</b>	65	60 (92.3%)

Figure 13: Calculation of the extent of the idiom principle for the text extract analyzed

#### 6.4 Discussion of findings

The concept of choice is a very subtle and complex one and could be interpreted in different ways; the decision taken here was to interpret Sinclair's reference to 'single choices' (1991: 10) as being the choice of a particular semi-preconstructed phrase rather than, for example, the 'choice' involved in selecting one particular exponent of a 'lexical item' over another. As mentioned above, this is an issue referred to by Sinclair (2008) who suggests looking at phraseology in terms of different levels or 'strata' of choice.

The implication of this methodological decision, in broadening the types of searches to allow for greater variation within the phrase, can be seen in the high percentage of the text attributed to the idiom principle. The fact that a higher proportion of language is thus 'captured' is felt to be a

benefit to both teachers and learners. While the phrases thus revealed tend to be at one remove from direct observation through concordance searches, which explains the need for ‘picture’ data, they can be seen as generalizing more about how language is used and help learners to identify patterns where they might not otherwise be evident.

Hoey (2005) also refers to other links between items, such as colligation or association with semantic concepts such as ‘specificity’ which, while frequently not falling under the heading of phraseology still affect choices in subtle and apparently unconscious ways. Also, as Mason (2007) points out, the extent to which the idiom principle will apply may vary considerably from text to text. Obtaining widespread agreement on the extent of idiomaticity in terms of a precise figure is therefore an impossible task, not least because, as Teubert (2009) points out, once one goes beyond linguistic data and into interpretation, there are always going to be different opinions, for example as to what parameters might be used to define a phrase and how they are applied (Gries, 2008). The approach to identifying lexical and phraseological items presented here allows for a wide range of phrases, some of which might not be recognized as such by other analysts.

## **7. Pedagogical applications**

The acknowledgement that a very high proportion of language consists of semi-preconstructed phrases which are often very variable and frequently discontinuous has many pedagogical implications. The most obvious of these is the need to focus more on language at a phraseological level instead of at the word level, which necessitates an integration of grammar and lexis teaching to enable investigation and elaboration of the types of constraints that act upon the phrase. Words should thus not be presented in lists of separate decontextualized items, but as parts of phrases that they commonly occur in. Equally, resources such as the Collins Cobuild Grammar Patterns series (Francis et al., 1996; 1998) may be used to point out similarities in phraseology that tend to apply to words belonging to the same semantic sets.

Also, bearing in mind the vast number of phrases available to a native speaker (Pawley & Syder, 1983; Mel’cuk, 1998), it is important to identify and focus on those phrases which are deemed to be most useful to a particular group of learners and to distinguish between those that are useful for decoding and those more likely to be used for encoding text. This may be achieved by focusing on the phrases which most commonly realize those functions which are frequently needed in the particular register or field relevant to the learners, and by adhering to Sinclair’s (1991) recommendation to start by investigating the most common words in the language.

In terms of learner training and independence, students need to be encouraged to develop the skills to be able to identify patterns or items for themselves and record them, noting variations that occur (Johns, 1991). This is where approaches such as ‘chunking’ (Lewis, 1993) can provide some useful training, and where, as Johns (1991: 3) puts it, the teacher becomes more of a ‘director and coordinator of student-initiated research’. An example of Johns’ ‘Data Driven Learning’ approach based on some patterns surrounding *there has been* and containing data derived from the usacad corpus is presented in Figure 15. This example provides students with substantial guidance as to the patterns to look for; as students become more adept at identifying phrases and patterns, this kind of guidance can gradually be reduced to encourage learner independence. Where materials such as this make use of concordance lines, such lines should be selected with reference to ‘picture’ findings where available to ensure that the examples provided are indicative of the phraseology.

Further, students can be encouraged to consult online sites such as ‘Just the Word’ (Sharp, 2008), which offers mediated access to its corpus data by organizing the results of searches for particular items into semantic sets and grammatical patterns, thus making the phraseology more salient.



using their intuition and knowledge of the register in question to identify which items in a text belong to either 'phaseological items' or 'lexical items' (Sinclair's 2008) and which are 'free'. It is then possible to calculate the proportion of the text in question that forms part of phrases and what proportion is attributable to open choice; in this instance the proportions were found to be 92.3% and 7.7% respectively. The phraseological information thus gathered can then form the basis for pedagogical applications and help to motivate a move away from an artificial separation of lexis and grammar (Sinclair, 1991; Francis, 1993).

## Appendices

	<b>page</b>
<b>Contents:</b> Appendix 1 – Text extract with numbered sentences	<b>19</b>
Appendix 2 – Searches referred to in assignment	<b>19</b>
Appendix 3 – Gardner’s text as presented in the textbook	<b>23</b>
Appendix 4 – Appendix 4: composition of the usacad corpus (Oakey, personal communication)	<b>26</b>

### Appendix 1: Source text extract with numbered lines

[1] The concept of intelligence, a very old one, has been employed in the most varied ways over the centuries. [2] During the past century, there has been considerable movement on the "intelligence front," and this trend shows no sign of stopping. [3] In this essay I briefly describe seven historical steps, or phases, in the development of thinking about intelligence, focusing in particular on work inspired by the theory of multiple intelligence. (Gardner, 1991)

### Appendix 2: Other searches referred to in assignment

Unless specified otherwise, all the data presented below comes from the Bank of English

#### **1. Co-occurrences of *the concept of* and EMPLOY found using the searches ‘concept+0,4employ@’ and ‘employ@+0,4concept’**

or consistently employ the concept of mode of production, and his victimologists have employed the concept of `victim precipitation" as a way interpretations usually employ the concept of totalitarianism as a unifying Victor Hugo employed the Christian concept of the Wandering Jew to describe psychoanalysts have employed the concept of repression to explain what critically on the way in which the concept of `corporatism" had been employed

#### **2. ‘no+sign|signs+of’ t-score picture, first 25 items (retrieved on 20/07/09)**

there	was	no	NODE	of	a	or
he	shows		NODE		any	down
has	showed		NODE		abating	up
was	is		NODE		life	in
had	show		NODE		the	yet
is	showing		NODE		slowing	break
they	with		NODE		an	entry
but	still		NODE		him	breakthrou
still	were		NODE		being	<p>
she	shown		NODE		letting	so
and	s		NODE		wanting	end
have	been		NODE		it	happening
could	are		NODE		her	let
far	saw		NODE		ending	no
iraq	found		NODE		them	on
which	see		NODE		easing	struggle
who	gave		NODE		nerves	recovery
economy	seen		NODE		anyone	early
government	sees		NODE		having	life
yesterday	yet		NODE		forced	from
tests	absolutely		NODE		giving	at
s	but		NODE		either	anywhere
certainly	detected		NODE		waning	despite
hussein	betraying		NODE		diminishin	its
however	certainly		NODE		stopping	compromise

### 3. 'thinking+about' usacad search; concordance lines including 'thinking' as a noun.

conclusions clashed with much of the 1960s thinking about thinking. Recall that  
n based on race, class, and culture. <p> Thinking about the struggles of these  
produced significant changes in American thinking about religious and  
eristics, habits, ways of perceiving and thinking about life processes, and  
treatments that combine all of the best thinking about how to improve mental  
essence of nineteenth-century bourgeois thinking about gender. Marital oneness  
an's work reveal about sixteenth-century thinking about the role of women in  
) . Summary A mix of modern and classical thinking about concepts was represented  
higher-order idea occurs, it constrains thinking about the details of the  
is not alive would reflect conventional thinking about life. In a number of  
, we have introduced some of the current thinking about the sense of pain. It is  
truggle forced Americans to do some deep thinking about both American and  
h questioning probably stimulates deeper thinking about mathematics in Asian  
nited States, where it tends to dominate thinking about the nature of  
ch on imagery, some of the most exciting thinking about imagery pertains to its  
d that this distinction is essential for thinking about biological evolution,  
reviously, conflicts that stimulate hard thinking about moral choices and  
tandard [in a shorter, wider glass]. His thinking about this problem is said to  
t the analogy used makes a difference in thinking about a target concept was  
d. Piaget's ideas caused a revolution in thinking about how thinking develops,  
once dominated - and still influences - thinking about women and men. The  
In these two arenas, Piagetian-inspired thinking about education is enjoying a  
er emotional life? While all intelligent thinking about eating problems mentions  
and how teaching behaviors interrelate-- thinking about teaching is more complex  
xclusion of other problems of knowledge. Thinking about motivation is  
re inspired by Piagetian and Kohlbergian thinking about the cognitive growth  
ligence tests, they have stimulated much thinking about the many processes  
y (1986), which formed the nucleus of my thinking about feminist therapy theory  
il War, an event that generated bold new thinking about republican government  
period was characterized by a rebirth of thinking about the nature of the world  
In the 1960s, there was a great deal of thinking about how to use principles of  
ng an abortion. The third was a study of thinking about rights and  
al judgment missed the rich diversity of thinking about moral issues, something  
ollege in stimulating the development of thinking about values comes through in  
lowed 'politics by men' to structure our thinking about other societies leading  
ands of years and continues to guide our thinking about 'a proper meal.' Sahlins  
models of education as well as Piagetian thinking about the stimulation of  
he national security state has processed thinking about Communism comes under  
econdary schools are designed to promote thinking about alternative positions,  
ivity. <p> Thus, feminist psychoanalytic thinking about the culture's impact on  
he middle decades of the century, recent thinking about 'community' has  
ther." Chomsky's ideas so revolutionized thinking about language and human  
eenth century gradually changed people's thinking about the appropriate  
only a few attempts to evaluate Piaget's thinking about adulthood, although  
p. 640) An interview tapping a subject's thinking about the dilemma then  
90s offering data challenging Kohlberg's thinking about universality, however.  
stency of Kohlberg, Erikson, and Perry's thinking about the role of college in  
recently dominated American sociological thinking about education. It attempts  
ents' social-class positions shape their thinking about American society. While  
e the party label to help organize their thinking about politics: to guide them  
et for thinking and then to expand their thinking about the target followed by a  
readers, and &hellip; readers use their thinking about authors to enhance their  
ution require that students reveal their thinking about the problem. Closing  
ay off as well as male ways of thinking. Thinking about how to create more  
either a universalized mode of thought - thinking about thinking - or one that  
here who are doing it came at it through thinking about feminist theory. We have  
cted nor triggered a revolution in white thinking about the race question. Its  
A revolutionary shift occurred in white thinking about using black men as  
ip to murder rates, see James Q. Wilson, Thinking about Crime (New York: Basic

#### 4. 'work+0,4inspired+by'search; concordance lines (retrieved on 29/07/09)

work on first language acquisition inspired by Chomsky's ideas, see T.  
work from The Australian Ballet, inspired by Malouf's novel Fly Away Peter,  
Doctrine, a work she claimed to be inspired by the Secret Masters. It was the  
up artist whose work appeared to be inspired by a beach holiday was Ellie  
that Mary Daly's work has been inspired by obscure debates within academic  
to see this work as having been inspired by the legend of St George and the  
Scott said her work had been inspired by the opulence of the early 19th  
audiences to see work so clearly inspired by a culture almost totally  
work by three craftspeople inspired by sea and sun. Sian Tucker's  
a fine three-movement work directly inspired by Roberts's playing on the  
work, "Modern Masterpiece," is inspired by his predecessor, Vaslav  
on the foundation work that is inspired by the Spirit, work done with a  
by Nicholas Mosse, whose work is inspired by the rural beauty of  
Kwara State, Nigeria, his work is inspired by the oral tradition of the  
by Bonhams, Denison's work is inspired by "the things I most enjoy in my  
<p> Georgina's provocative work is inspired by her own African and English  
in his native country. His work is inspired by examples from the past-such as  
in the rest of the work which is inspired by the life of the 13th century  
<p> While much of his work is inspired by the environment, urban  
literary" novelist, whose work is inspired by ideals of pacifism and  
the work, though abstract, is inspired by the convenience stores" which  
stories. <p> Elegy, a new work, is inspired by Dostoevsky's novel The Idiot,  
Jennings says his work is often inspired by religion - but not the  
Atlette Chacock, whose work seems inspired by boldly colored geometric and  
work, Lucid in the Sky, inspired by Scottish sunlight, will be in a  
as work on mathematical thinking inspired by traditions such as  
Each and every work was variously inspired by the landscape around Broadheath  
Pochner says their work was inspired by events in the upper atmosphere  
in Oslo, says that the work was inspired by an experience the artist had  
a trace mineral. The work was inspired by the observation that in parts  
output. Much of his work was inspired by jazz, by American folk music,  
says. "I used to do work which was inspired by classical figurative art, and  
the late 1950s. Erikson's work was inspired by his personal experience and his  
body of work he produced that was inspired by the creepy equipment you find  
trends. In 1931, he began a work inspired by the threat modernism posed to  
Canadian artist Joe David's work -- inspired by a dream about the Princess --  
company, SEITA, to produce a work inspired by the famous gypsy silhouette  
s birthplace Vitebsk and work inspired by the constructivists; a second  
include Suky Best, with work inspired by teen-magazines photo love  
to produce five pieces of work inspired by a particular aspect of life for  
exquisite, bejewelled work inspired by myth, fairy story and history  
4. <ZZ1> PHOTOS <ZZ0> <p> Nice Work Inspired by its history and setting, Don  
in Sandwich includes work inspired by the sea, views of interiors and  
cultural community producing work inspired by tragedy. What does he see as  
was a reference to a Klein work inspired by Greek and Roman heroic  
vintage examples and work inspired by them on sale at Artissimo, 7378  
and change the human world. Work inspired by Marxist thought tends to centre  
perform Appearances, a new work inspired by Metheny and created especially  
students will perform work inspired by war at the Queensland Art  
in an abstract improvisational work inspired by the poems of Vincent Woods. <p>  
Stucky and ends with another work inspired by the United States, Dvorak's  
work, it is the first AB work inspired by a novel and Graeme Koehne's  
choreographers. An abstract work inspired by Matthias Claudius' poem, which  
in dance for Rites, a work inspired by Rite of Spring and created for  
new media. They have exhibited work inspired by astronomy at the Photographers'  
Persephone, part of a larger work inspired by Eliot's The Waste Land, on  
and for some neat approach work inspired by the energetic and skilful  
culminating in the new work inspired by Auster's novel. Through it all,  
it. Tippett finished the work, inspired by the light on a lake in Senegal,  
of Christopher Bruce's work inspired by the life of the boxer Rubin  
director, presents a new work inspired by Arvo Part's Fratres for violin  
She wanted to display a work inspired by the fate of two workers who,

## 5. 'multiple+intelligence|intelligences'search; concordance lines (retrieved on 29/07/09)

thus giving rise to the theory of multiple intelligences (verbal, Frames of Mind: The Theory of Multiple Intelligences (New York: Basic is Gardner's (1983) theory of multiple intelligences. Gardner believes consistent with Gardner's theory of multiple intelligences. In short, there is for considering the theory of Multiple Intelligences and its implications what it might be. <o> The theory of Multiple Intelligences allows him to do has developed a theory of multiple intelligence, arguing that each Gardner, favours a theory of 'multiple intelligences". It is the speed pedestals. The theory of multiple intelligences (MI) has been out of the several theories of multiple intelligence that we have already just one net. Mirroring theories of multiple intelligences, it is made up of argues instead for the concept of 'multiple intelligences" - practical, in the latest academic theory about multiple intelligence. Teachers should we are certain that Gardner's multiple intelligences theory is going to University and the author of Multiple Intelligences. <p> Henry Louis Determined Modules in the Mind and Multiple Intelligences Concluding Comments Determined Modules in the Mind and Multiple Intelligences The idea that mind make clear that if people inherit multiple intelligences, what they inherit the discussion of Gardner's multiple intelligences (Chapter 6). A to rule. <o> So I think of these multiple intelligences as a little group factorial approach, (d) Gardner's 'multiple" intelligences and (e) the Vernon, 1969, p. 10). d) Gardner's multiple intelligences Guilford's view of social scientist. This is the 'multiple intelligences" approach of Howard to set up standardised testing of 'multiple intelligences", he does suggest processing. The Centurian Multiple-Intelligence Test below consists intelligences and your combined multiple-intelligence quotient (MIQ). Some Celebrity Squares <p> We gave the Multiple Intelligence Test to five points out caveats: <f> While Multiple Intelligence Theory is consistent seven ways of knowing, or the Multiple Intelligence Teaching Application ( deliberate decision to write about 'multiple intelligences": multiple" to the branches, you will believe in multiple intelligences. And even though the

## **Appendix 3: Gardner's text as presented in the textbook**

### **Intelligence in Seven Steps**

<sup>1</sup> The concept of intelligence, a very old one, has been employed in the most varied ways over the centuries. During the past century, there has been considerable movement on the "intelligence front," and this trend shows no sign of stopping. In this essay I briefly describe seven historical steps, or phases, in the development of thinking about intelligence, focusing in particular on work inspired by the theory of multiple intelligence.

#### **Lay Conceptions**

<sup>2</sup> Until this century, the word "intelligence" has been used primarily by ordinary individuals in an effort to describe their own mental powers as well as those of other persons. Consistent with ordinary language usage, intelligence has been used in anything but a precise manner. Individuals living in the West were called intelligent if they were quick or eloquent or scientifically clever or wise. In other cultures, the individual who was obedient, or well behaved, or quiet, or equipped with magical powers, may well have been referred to by terms which have been translated as intelligent.

<sup>3</sup> For the most part, the word intelligent was used in a positive way; however, its imprecision can be readily displayed by a recognition that it has been applied to nearly all of the American presidents in this century, even though it is doubtful that any two of the presidents exhibited similar kinds of minds. Perhaps ironically, Herbert Hoover and Jimmy Carter, two of Americas least successful presidents, both of whom were engineers, probably came closest to the lay idea of intelligence. It may be worth noting that they became distinguished by their behavior as *ex*-presidents.

#### **The Scientific Turn**

<sup>4</sup> In a sequence of events that is by now familiar, Alfred Binet responded to requests from Parisian ministers at the turn of the century by creating the first intelligence test. It then became possible to estimate an individual's intelligence by noting his or her performance on a deliberately heterogeneous set of items. Used first clinically for "at risk" Parisian elementary schoolchildren, the intelligence test became "normed" on Californian middle-class children and was administered quite widely, thanks in large part to the efforts of Lewis Terman at Stanford University. By the 1920's and 1930's, intelligence tests (and their product, an individual's IQ) had become deeply ingrained not only in American society but also in many other parts of the world.

#### **Pluralization of Intelligence**

<sup>5</sup> While intelligence was initially perceived as a unitary concept which could be captured by a single number, a debate soon arose about whether the concept could legitimately be broken into components. Such researchers as L.L. Thurstone and J.P. Guilford argued that intelligence was better conceived of as a set of possibly independent factors. In recent years, supported by findings from fields such as artificial intelligence, developmental psychology, and neurology, a number of investigators have put forth the view that the mind consists of several independent modules or intelligences.

<sup>6</sup> In my own theory of multiple intelligences,<sup>6</sup> I argue that human beings have evolved to be able to carry out at least eight separate forms of analysis:

<b>Linguistic Intelligence</b>	As in a poet. Involves reading, writing, speaking, and conversing in one's own or foreign languages. It may be exercised through reading interesting books, playing word board or card games, listening to recordings, using various kinds of computer technology, and participating in conversation and discussions.
<b>Logical-Mathematical Intelligence</b>	As in a scientist. Involves number and computing skills, recognizing patterns and relationships, timeliness and order, and the ability to solve different kinds of problems through logic. It may be exercised through classifying and sequencing activities, playing number and logic games, and solving various kinds of puzzles.
<b>Musical Intelligence</b>	As in a composer. Involves understanding and expressing oneself through music and rhythmic movements or dance, or composing, playing, or conducting music. It may be exercised by listening to a variety of recordings, engaging in rhythmic games and activities, and singing, dancing, or playing various instruments.
<b>Spatial Intelligence</b>	As in a sculptor or airplane pilot. Involves visual perception of the environment, the ability to create and manipulate mental images, and the orientation of the body in space. It may be developed through experiences in the graphic and plastic arts, sharpening observation skills, solving mazes and other spatial tasks, and exercises in imagery and active imagination.
<b>Bodily Kinesthetic Intelligence</b>	As in an athlete or dancer. Involves physical coordination and dexterity, using fine and gross motor skills, and expressing oneself or learning through physical activities. It may be exercised by playing with blocks and other construction materials, dancing, playing various active sports and games, participating in plays or make-believe, and using various kinds of manipulatives to solve problems or to learn.
<b>Interpersonal Intelligence</b>	As in a salesman or teacher. Involves understanding how to communicate with and understand other people and how to work collaboratively. It may be exercised through cooperative games, group projects and discussions, multicultural books and materials, and dramatic activities or role-playing.
<b>Intrapersonal Intelligence</b>	Exhibited by individuals with accurate views of themselves. Involves understanding one's inner world of emotions and thoughts, and growing in the ability to control them and work with them consciously. It may be exercised through participating in independent projects, reading illuminating books, journal-writing, imaginative activities and games, and finding quiet places for reflection.
<b>Naturalist Intelligence</b>	As in geologists, zoologists, anthropologists, and others who understand and interact with the natural environment. Involves understanding the natural world of plants and animals, noticing their characteristics, and categorizing them; it generally involves keen observation and the ability to classify other things as well. It may be exercised by exploring nature, making collections of objects, studying them, and grouping them.

These ideas have attracted some attention on the part of educators seeking a more comprehensive and individualized educational system. Recently my colleagues and I have been exploring certain educational implications of the theory in our own research.

### **Contextualization**

<sup>7</sup> As initially put forth, most theories of intelligence - whether singular or multiple-have assumed that intelligences are simply biological beings or potentials, which exist "in the head" (and "in the brain") and can be measured reliably, independent of context. While the theory of multiple intelligences was deliberately formulated to take into account the unfolding of intelligence in different cultures, it nevertheless suffered in its early formulations from an "individual-centered" bias. Most students of intelligence however, are now coming to the realization that intelligence cannot be conceptualized, or measured with accuracy, independent of the particular contexts in which an individual happens to live, work, and play, and of the opportunities and values provided by that setting. Bobby Fischer might inherently have had the potential to be a great chess player, but if he had lived in a culture without chess, that potential would never have been manifested, let alone actualized. Intelligence is always an interaction between biological tendencies and opportunities for learning in a particular cultural context.

<sup>8</sup> Project Spectrum, a curriculum-and-assessment project for young children, is one reflection of this view. We initially designed the project to determine whether young children exhibit distinctive profiles of intelligences, but we soon came to realize that intelligences could not be measured out of

context; instead we had to create new environments, contexts more like children's museums than like traditional schoolrooms, in which children's intellectual tendencies had an opportunity to be elicited and practiced. Only then could some kind of meaningful assessment become possible.

### **Intelligence as Distributed**

<sup>9</sup> Closely related to the trend toward the contextualization of intelligence is the realization that significant parts of intelligence are *distributed*. The essential understanding here is, again, that not all intelligence is in the head. But rather than existing simply in the general context wherein a person lives, much of everyday intelligence can be located in the human and non-human resources with which individuals work, and on which they come to depend in their productive work. Typically these are thought of as non-human resources, such as books, notebooks, computer files, and the like. And it is true that in a literate world, much on which the productive individual depends is contained in these materials.

<sup>10</sup> It is also appropriate, however, to think of other individuals as part of one's "distributed intelligence." Most workers do not depend exclusively on their own skills and understanding; rather, they assume the presence of others in their work environments with whom they can regularly interact. This view is brought home vividly when one considers an office that is being computerized. Rarely does all relevant knowledge reside with a single individual; much more commonly, different office members exceed the novice level in different areas of hardware or software expertise. In our terms, intelligence about computers is "widely distributed" across individuals under such circumstances.

<sup>11</sup> Our own efforts to examine the "distributed nature" of intelligence can be seen in two of our projects. In Arts Propel, a cooperative project in arts and humanities assessment, we ask students to keep detailed "processfolios" - complete records of their involvement in a project, from initial conception through mid-point sketches and drafts, ultimately to new plans that grow out of the final completed project. We believe that students' learning is significantly improved when they can have an on-going dialogue with the record of their previous efforts, as captured in these constantly evolving processfolios.

<sup>12</sup> In the Key School, an Indianapolis elementary school, children are exposed each day to contexts that develop each of the intelligences. As part of their regular work at this experimental public school, students carry out each year three theme-related projects. Our research interest is in developing methods whereby these projects can be evaluated in a fair and comprehensive way. Part of that evaluation centers on the ways in which participation in a project has been cooperative: the human and non-human resources involved in preparation of the project, the help given by others in the presentation of the project, and the reactions of other individuals - peers as well as experts - to the final project. By deliberately including these "extra-individual" elements in our evaluations, we hope to bring home to the community the importance of "distributed aspects of intelligence." At the same time we want to get rid of the common notion that all skill and learning must exist within a single brain, be that brain at home, at school, or at the work place.

<sup>13</sup> Thus far, I have spoken of historical steps that have already been carried out, or which at least are being taken at the present time. My last two steps represent hopes for future work on intelligence in our own laboratory and in others around the world.

### **Nurturance of Intelligence**

<sup>14</sup> Even though our efforts to understand intelligence have been advancing, we still know very little about how to nurture intelligence, be it conceptualized in unitary or pluralistic fashion, in individual-centered, contextualized, or distributed form. Yet surely our efforts to understand intelligence as scientists can best be realized by a demonstration that intelligence can be nurtured in

particular educational settings, using strategic pedagogical or facilitating techniques. Here lies one important challenge for the future.

**Humanizing Intelligence**

<sup>15</sup> In the Key School, understanding the nature of the human mind in all of its complexity is no easy task, and a complete understanding may well exceed human investigative capacities. But understanding intelligence-and even knowing how better to develop it-does not suffice in itself. Any human capacity can be used for ill as well as for good; and it is part of our responsibility as human beings living on a single troubled planet to try to use our competences, our intelligences, in morally responsible ways. This assignment cannot fall exclusively on the shoulders of researchers; nor can we simply afford to pass this responsibility on to others.

<sup>16</sup> The human being is also more than his or her intellectual powers. Perhaps more crucial than intelligence are motivation, personality, emotions, and will. If we are ever to obtain a comprehensive and fully integrated picture of human beings, we need to combine our insights about cognition with comparable insights in respect to these other aspects of the human being. Perhaps, indeed, a different view of human nature will result from this activity of synthesis.

<sup>17</sup> Obviously so grand an undertaking requires the highest degree of "distributed collaboration" among researchers, educators, and the general public. Although the task is great, the advances made in understanding over the past decade give one some reason for optimism.

**Appendix 4: composition of the usacad corpus (Oakey, personal communication)**

**Usacad corpus**

<b>Text type</b>	<b>number of tokens (08/07/2009)</b>	<b>number of texts</b>	<b>Dates</b>
American Academic textbooks*	6,341,888	31	1993-96

**\*Disciplines represented by textbooks:**

6 Politics & Government, 5 Psychology, 5 Sociology/Anthropology, 5 History, 4 Gender studies, 3 Economics, 1 Health, 1 Biology, 1 Oceanography

## References

- Barnbrook, G. (1996). **Language and computers: A practical introduction to the computer analysis of language**. Edinburgh: Edinburgh University Press.
- Becker, J. (1975). **The phrasal lexicon**. In Proceedings of the 1975 workshop on Theoretical issues in natural language processing. pp 60-63.
- Cheng, W., Greaves, C., Sinclair, J., & Warren, M. (2008). Uncovering the extent of the phraseological tendency: Towards a systematic analysis of concgrams. **Applied Linguistics**, 30 (2): 236-252.
- Ellis, N.C. (2008). 'Phraseology: The periphery and the heart of language'. In Meunier, F. & Granger, S. & Meunier, F. (eds.) **Phraseology in foreign language learning and teaching**. Amsterdam/Philadelphia: John Benjamins. pp 1-13.
- Erman, B. & Warren, B. (2000). The idiom principle and the open choice principle. **Text**, 20 (1): 29-62.
- Flowerdew, L. (2004). 'The argument for using English specialized corpora to understand academic and professional language'. In Connor, U. & Upton, T. (eds.). **Discourse in the professions**. Amsterdam/Philadelphia: John Benjamins. pp 11-33.
- Francis, G. (1993). 'A corpus-driven approach to grammar: Principles, methods and examples'. In Baker, M., Francis, G., Tognini-Bonelli, E. (eds.). **Text and technology: In honour of John Sinclair**. Amsterdam/Philadelphia: John Benjamins. pp 137-156.
- Francis, G., Hunston, S., Manning, E. (1996). **Collins cobuild grammar patterns 1: Verbs**. London: Harper Collins
- Francis, G., Hunston, S., Manning, E. (1998). **Collins cobuild grammar patterns 1: Nouns and adjectives**. London: Harper Collins
- Gardner, H. (1991). Intelligence in seven steps [online]. **New Horizons for Education**. Available at [http://www.newhorizons.org/future/Creating\\_the\\_Future/crfut\\_gardner.html](http://www.newhorizons.org/future/Creating_the_Future/crfut_gardner.html) [Accessed 15/06/2009]
- Gries, S.Th. (2008). 'Phraseology and linguistic theory: A brief survey'. In Granger, S. & Meunier, F. (eds.). **Phraseology: An interdisciplinary perspective**. Amsterdam/Philadelphia: John Benjamins. pp 3-26.
- Groom, N. (2007). **Phraseology and epistemology in humanities writing : a corpus-driven study**. PhD Thesis. Birmingham: University of Birmingham.
- Hoey, M. (2000). 'A world beyond collocation: new perspectives on vocabulary teaching'. In Lewis, M. (Ed.) **Teaching Collocation: Further developments in the lexical approach**. Hove: Language Teaching Publications. pp 224-243.
- Hoey, M. (2005). **Lexical Priming**. Oxon: Routledge.

- Hoey, M. (2006). 'Language as choice: What is chosen?' In Thompson, G. & Hunston, S. (eds.) **System and corpus: Exploring connections**. London: Equinox. pp 37-54.
- Hunston, S. (2002). **Corpora in applied linguistics**. Cambridge: Cambridge University Press.
- Hunston, S. (2007). Semantic prosody revisited. **International Journal of Corpus Linguistics**, 12 (2): 249-268
- Hunston, S. (2009). **What do we mean by 'pattern'?** University of Birmingham Open Distance Learning Summer School, July 8, 2009.
- Hunston, S. & Francis, G. (1998). Verbs observed: A corpus-driven pedagogic grammar. **Applied Linguistics**, 19 (1): 45-72.
- Hunston, S. & Francis, G. (2000). **Pattern Grammar**. Amsterdam/Philadelphia: John Benjamins.
- Johns, T.F. (1991). Should you be persuaded: Two samples of data-driven learning materials. **English Language Research Journal**, 4: 1-16.
- Lewis, M. (1993). **The lexical approach**. Hove: Language Teaching Publications.
- Mason, O. (2007). 'From lexis to syntax: the use of multi-word units in grammatical description'. In **The Proceedings of the 26th conference on Lexis and Grammar, Bonifacio, 2-6 October 2007**. Available from <http://phrasys.net/uob/om/publications> [Retrieved 15/06/2009]
- Mel'čuk, I. (1998). 'Collocations and lexical functions'. In Cowie, A.P. (ed.). **Phraseology: Theory, analysis, and applications**. Oxford: Oxford University Press. pp 23-54.
- Nattinger, J. & De Carrico, J. (1992). **Lexical phrases and language teaching**. Oxford/New York: Oxford University Press.
- Oakey, D. (2009). **Personal communication**
- Owen, C. (1996). Do corpora require to be consulted? **ELT Journal**, 50 (3): 219-224.
- Pawley, A. & Syder, F.H. (1983). 'Two puzzles for linguistic theory: Nativelike selection and nativelike competence.' In Richards, J.C. & Schmidt, R.W. (eds.). **Language and Communication**. London: Longman. pp 191-227.
- Philip, G. (2008). 'Reassessing the canon: 'Fixed' phrases in general reference corpora'. In Granger, S. & Meunier, F. (eds.). **Phraseology: An interdisciplinary perspective**. Amsterdam/Philadelphia: John Benjamins. pp 95-108.
- Sharp Laboratories of Europe (2008). **Just the Word** [online]. Available from <http://193.133.140.102/JustTheWord/>
- Sinclair, J.McH. (1991). **Corpus, concordance, collocation**. Oxford: Oxford University Press
- Sinclair, J. (1996). The search for units of meaning. **Textus** 9 (1): 75-106. (Reprinted in Sinclair 2004a: 24-48.)
- Sinclair, J. (1998). 'The lexical item'. In E. Weigand (ed.), **Contrastive Lexical Semantics**.

- Amsterdam: Benjamins. pp 1–24. (Reprinted in Sinclair 2004a: 131–48.)
- Sinclair, J. (2004a). *Trust the text*. London: Routledge
- Sinclair, J. (2004b). ‘Language and computing, past and present’. In Ahmad, K. & Rogers, M. (eds.). **New directions in LSP studies: Proceedings of the 14th European Symposium on Language for Special Purposes: Communication, culture, knowledge, 18 - 22 August 2003**. Guildford: University of Surrey. pp 1-12.
- Sinclair, J. (2008). ‘The phrase, the whole phrase and nothing but the phrase’. In Granger, S. & Meunier, F. (eds.). **Phraseology: An interdisciplinary perspective**. Amsterdam/Philadelphia: John Benjamins. pp 407-410
- Sinclair, J. & Renouf, A. (1991). ‘Collocational frameworks in English’. In Aijmer, K. & Altenberg, B. (eds.) **English Corpus Linguistics**. Harlow: Longman. pp 128-143.
- Stubbs, M. (2002) **Words and phrases: Corpus studies of lexical semantics**. Oxford: Blackwell.
- Stubbs, M. (2009). The search for units of meaning: John Sinclair on empirical semantics. **Applied Linguistics**, 30 (1): 115-137.
- Teubert, W. (2009). **Corpus and discourse**. University of Birmingham Open Distance Learning Summer School, July 10, 2009.
- Tognini-Bonelli, E. (2001). **Corpus linguistics at work**. Amsterdam/Philadelphia; John Benjamins.
- Walker, C. (2009). **Corpus linguistics workshop**. University of Birmingham Open Distance Learning Summer School, July 6 2009.
- Wray, A. (2002) **Formulaic language and the lexicon**. Cambridge: Cambridge University Press.
- Wray, A. (2008). **Formulaic language: Pushing the boundaries**. Oxford: Oxford University Press.