

# TRA3105

# Computer and Business Translation 1

---

WEEK 6B

- (1) GENERAL CORPORA, CONCORDANCERS, AND  
SENTENCE DATABASES (PART 2)
- (2) LAB SESSION 3

# Simple Tools

---

1. Ludwig (<https://ludwig.guru/>)
2. Dante (<http://www.webdante.com/>): Sentences are categorised (Demonstration: consider the word “knowledge” and identify its usage / collocation patterns)

# Professional Corpora and Concordancers

Corpus of Contemporary American English  
(<http://corpus.byu.edu/coca/>)

The screenshot shows the web interface of the Corpus of Contemporary American English (COCA) at <http://corpus.byu.edu/coca/>. The browser window has a single tab titled "Corpus of Contemporary American English" and the address bar shows the URL. The website has a blue header with the title "Corpus of Contemporary American English" and a "new" badge. Below the header is a navigation bar with four tabs: "SEARCH", "FREQUENCY", "CONTEXT", and "HELP". The "SEARCH" tab is active. The main content area is divided into two columns. The left column contains a search form with a text input field, a "[POS]" dropdown menu, and buttons for "Find matching strings" and "Reset". Below the search form are links for "Sections", "Texts/Virtual", "Sort/Limit", and "Options". The right column contains a "NOT LOGGED IN" status bar and a "Hide help" checkbox. Below this is a large text box with information about the corpus, including a note about the "NOW Corpus" and a description of the COCA corpus itself. The text box also includes a link to "More help files".

Corpus of Contemporary American English

SEARCH FREQUENCY CONTEXT HELP

List Chart Collocates Compare KWIC

[POS]

Find matching strings Reset

Sections Texts/Virtual Sort/Limit Options

(Hide help) NOT LOGGED IN

In addition to this online interface, you can also download extensive data for offline use -- full-text, word frequency, n-grams, and collocates data. You can also access the data via WordAndPhrase (including the ability to analyze entire texts that you input).

For more recent data, try the NOW Corpus. Every day 4-5 million words of data (about 10,000 new texts) are added to the corpus. This means that it has 164 million words of data from just the past month and 1.4 billion words from the past year (and 3.3 billion words total).

The Corpus of Contemporary American English (COCA) is the largest freely-available corpus of English, and the only large and balanced corpus of American English. COCA is probably the most widely-used corpus of English, and it is related to many other corpora of English that we have created, which offer unparalleled insight into variation in English.

The corpus contains more than 520 million words of text (20 million words each year 1990-2015) and it is equally divided among spoken, fiction, popular magazines, newspapers, and academic texts.

Click on any of the links in the search form to the left for context-sensitive help, and to see the range of queries that the corpus offers. You might pay special attention to the comparisons between genres and years and the (new) virtual corpora, which allow you to create personalized collections of texts related to a particular area of interest.

More help files

# Getting Monolingual Examples

1. Click "List"

2. Your word

4. Click this.

3. Click "Sections":  
Text type, year,  
spoken...

Corpus of Contemporary American English

SEARCH | CONTEXT | OVERVIEW

**List** | Chart | Collocates | Compare | KWIC

fiscal cliff [POS]

Find matching strings | Reset

☐ Sections | Texts/Virtual | Sort/Limit | Options

1 IGNORE 2 IGNORE

SPOKEN  
FICTION  
MAGAZINE  
NEWSPAPER  
ACADEMIC

SHOW Determines whether the frequency is shown for each "section" of the corpus (in the case of COCA, the genre or year). For example, the synonyms of *beautiful* in [each section](#) and [overall](#).

Select a section: (sub-)genre or (set of) year(s). [Click here](#) for more examples of change over time.

<a href="#">*ize verbs in ACADEMIC</a>	<a href="#">Past tense verb + over in SPOKEN</a>
<a href="#">*ment in ACADEMIC</a>	<a href="#">Nouns near green in 2000-2009</a>
<a href="#">ADJ + track in NEWSPAPERS</a>	<a href="#">Noun near chair in FIC</a>
<a href="#">Hard + NOUN in MAGAZINES</a>	<a href="#">Synonyms of smart in FICTION</a>
<a href="#">Verbs in MAGAZINES-Sports</a>	<a href="#">Nouns in NEWSPAPERS-Money</a>
<a href="#">Adjectives in ACADEMIC-Medicine</a>	<a href="#">Adverbs in FICTION-Movies</a>

(Optional) Select a second (set of) section(s) against which to compare the sections chosen above

<a href="#">*ize verbs in ACAD vs MAG</a>	<a href="#">Past tense verb + over in FIC vs ACAD</a>
<a href="#">*ment in ACAD vs FIC</a>	<a href="#">Nouns near green in 2005-2015 vs 1990-1999</a>
<a href="#">ADJ + track in NEWS vs SPOK</a>	<a href="#">Nouns near chair in ACAD vs FIC</a>
<a href="#">Hard + NOUN in MAG vs ACAD</a>	<a href="#">Synonyms of smart in FIC vs NEWS</a>
<a href="#">ADJ in NEWS-Money vs NEWS</a>	<a href="#">Nouns in MAG-Sports vs MAG</a>
<a href="#">ADJ in ACAD-Medicine vs ACAD</a>	<a href="#">Phrasal verbs with up 2005-2015 vs 1990-1994</a>



SEARCH

FREQUENCY

CONTEXT

OVERVIEW

SEE CONTEXT: CLICK ON WORD OR SELECT WORDS + [CONTEXT] [\[HELP...\]](#)[COMPARE](#)

		CONTEXT	FREQ	
1		FISCAL CLIFF	472	<div></div>
				0.703 seconds



Click this.



SEARCH

FREQUENCY

CONTEXT

OVERVIEW

FIND SAMPLE: [100](#) [200](#)

PAGE: &lt;&lt; &lt; 1 / 5 &gt; &gt;&gt;

CLICK FOR MORE CONTEXT



SAVE LIST

CHOOSE LIST

CREATE NEW LIST



1	2015	SPOK	CNN	A	B	C	and through the reforms that we did with him several years ago on our <b>fiscal cliff</b> . And now he's come back again. Let me just also say this
2	2015	SPOK	Fox	A	B	C	n't do comprehensive tax reform, we've raised rates on individuals to the <b>fiscal cliff</b> deal, and there's no way we can bring them back down, so
3	2015	SPOK	PBS	A	B	C	but they insist his focus is to govern. Example? Deals over the <b>fiscal cliff</b> that McConnell made with Vice President Biden. TRENT-LOTT-(R), -Fo# If you go back and
4	2015	SPOK	PBS	A	B	C	And they have been dealing with more regulatory environment, higher taxes through the <b>fiscal cliff</b> deal, high commodity costs, and Tsedeye brought up a good point
5	2015	SPOK	ABC	A	B	C	(ILLINOIS) We tried this tax hike on the wealthy thing with the <b>fiscal cliff</b> , which I voted for, by the way, because, to me,
6	2015	NEWS	USAToday	A	B	C	2134 Forget the <b>fiscal cliff</b> . Though consumer confidence may be faltering amid Washington's protracted budget and tax battles
7	2014	SPOK	NPR	A	B	C	that changed this country. I remember having a discussion with colleagues during the <b>fiscal cliff</b> fight or the government shutdown fight and saying, wow, we used to
8	2014	NEWS	WashPost	A	B	C	141 Forget the " <b>fiscal cliff</b> ": When it comes to the nation's most pressing concerns, other matters
9	2013	SPOK	ABC	A	B	C	. It's disgusting. From the White House on down, this whole <b>fiscal cliff</b> resolution that took 18 months to get to where we are, even Simpson-Bowles said
10	2013	SPOK	ABC	A	B	C	neglect electric scooters, because this really tells you how Washington works. The <b>fiscal cliff</b> deal - scooters are made in Oregon. Senator Wyden of Oregon said it is
11	2013	SPOK	ABC	A	B	C	(FOX NEWS ANCHOR) But so - but the point is that this <b>fiscal cliff</b> is a big, fat lie. It did have some stimulus for its special
12	2013	SPOK	ABC	A	B	C	is what are they going to do now, faced with not only another <b>fiscal cliff</b> , but also this debt ceiling? I mean, that the showdown over the
13	2013	SPOK	ABC	A	B	C	we go right up to that deadline, because it's not like the <b>fiscal cliff</b> . We're talking about the full faith and credit of the United States.
14	2013	SPOK	ABC	A	B	C	, partly because of the distraction of this debt - I mean, this <b>fiscal cliff</b> debate. But in the case of Senator Hagel, with everyone coming from different
15	2013	SPOK	ABC	A	B	C	's going to be so much blood on the floor with regard to the <b>fiscal cliff</b> and so many other issues that are coming up, the debt ceiling. GEORGE-STEPHANOPOU#
16	2013	SPOK	ABC	A	B	C	, George Will, that this must follow all of the debates, the <b>fiscal cliff</b> debates we're going to be having over the next three months. GEORGE-WILL# (Off-camera)
17	2013	SPOK	ABC	A	B	C	going to think about anything along those lines until we're done with the <b>fiscal cliff</b> mess, the debt ceiling, the funding of the government, where at least
18	2013	SPOK	ABC	A	B	C	we've had in the past few years. We just got through the <b>fiscal cliff</b> thing. It is very strange to live in a great sophisticated wealthy modern democracy
19	2013	SPOK	ABC	A	B	C	to the well. We let the President get what he wanted on the <b>fiscal cliff</b> . But when it comes to the debt ceiling, he's going to have
20	2013	SPOK	ABC	A	B	C	YORK TIMES) It's incredibly scary. This is much scarier than the <b>fiscal cliff</b> , much scarier than any of the other things out there because we don't
21	2013	SPOK	ABC	A	B	C	PEGGY-NOONAN# My goodness. That appeared to be the White House attitude on the <b>fiscal cliff</b> just a month or two ago. PAUL KRUGMAN (COLUMNIST FOR THE NEW
22	2013	SPOK	ABC	A	B	C	I'm confident we don't have to 2017. GEORGE-STEPHANOPOU# (Off-camera) Coming off the <b>fiscal cliff</b> negotiations and the whole experience with that, is that realistic

Examples

# Finding Collocations

The screenshot shows the COCA website interface. At the top, there's a navigation bar with 'SEARCH' and 'FREQUENCY' tabs. Below this, there are links for 'List', 'Chart', 'Collocates', and 'Compare KWIC'. The 'Collocates' tab is selected. In the search area, the 'Word/phrase' field contains 'skill\_nn\*' and the 'Collocates' field contains '\_j\*'. There are buttons for 'Find collocates' and 'Reset'. A sidebar on the right contains a '(HIDE HELP)' button, a 'NOT LOGGED IN' status, and a 'COLLOCATES display' section with explanatory text and links for more information.

1. Click "Collocations"
2. Type your expression (e.g. skill).
3. Select the POS of your expression (optional): a POS tag (e.g., \_nn\*) will be added automatically.
4. Select the POS of collocates (optional): a POS tag (e.g., \_v\*) will be added automatically.
5. Click "Find collocates".

1. Click "Collocations"
2. Type your expression (e.g. skill).
3. Select the POS of your expression (optional): a POS tag (e.g., \_nn\*) will be added automatically.
4. Select the POS of collocates (optional): a POS tag (e.g., \_v\*) will be added automatically.
5. Click "Find collocates".

Corpus of Contemporary

corpus.byu.edu/coca/

new

Corpus of Contemporary American English

?

SEARCH

FREQUENCY

CONTEXT

HELP

SEE CONTEXT: CLICK ON WORD OR SELECT WORDS + [CONTEXT] [HELP...]

COMPARE

		CONTEXT	FREQ	
1	<input type="checkbox"/>	NEW	287	
2	<input type="checkbox"/>	SOCIAL	245	
3	<input type="checkbox"/>	IMPORTANT	235	
4	<input type="checkbox"/>	GREAT	225	
5	<input type="checkbox"/>	ACADEMIC	161	
6	<input type="checkbox"/>	TECHNICAL	158	
7	<input type="checkbox"/>	HIGH	155	
8	<input type="checkbox"/>	SPECIFIC	153	
9	<input type="checkbox"/>	DIFFERENT	127	
10	<input type="checkbox"/>	PHYSICAL	120	
11	<input type="checkbox"/>	BASIC	117	
12	<input type="checkbox"/>	NECESSARY	108	
13	<input type="checkbox"/>	PARTICULAR	104	
14	<input type="checkbox"/>	CRITICAL	103	
15	<input type="checkbox"/>	HIGHER	102	
16	<input type="checkbox"/>	OTHER	94	
17	<input type="checkbox"/>	COGNITIVE	90	
18	<input type="checkbox"/>	ESSENTIAL	87	
19	<input type="checkbox"/>	LOW	87	
20	<input type="checkbox"/>	GOOD	86	
21	<input type="checkbox"/>	KEY	83	
22	<input type="checkbox"/>	POLITICAL	81	
23	<input type="checkbox"/>	MUSICAL	79	
24	<input type="checkbox"/>	PSYCHOLOGICAL	74	
25	<input type="checkbox"/>	INDIVIDUAL	73	
26	<input type="checkbox"/>	GREATER	71	
27	<input type="checkbox"/>	CONSIDERABLE	68	
28	<input type="checkbox"/>	SIGNIFICANT	64	
29	<input type="checkbox"/>	REAL	63	
30	<input type="checkbox"/>	SPECIAL	63	

Click this.



Corpus of Contemporary American English

SEARCH FREQUENCY CONTEXT HELP

FIND SAMPLE: [100](#) [200](#)  
PAGE: << < 1 / 3 > >>

CLICK FOR MORE CONTEXT [?]

1	2015	ACAD	GeneralMusicToday	A	B	C	and Forness (1996) found that 75% of students with learning disabilities manifest <b>social skill</b> deficits that distinguish them from students without learning di
2	2015	ACAD	GeneralMusicToday	A	B	C	self) and across most dimensions of <b>social</b> competence. The researchers concluded that <b>social skill</b> deficits appear to be an integral part of the learning-disa
3	2015	ACAD	GeneralMusicToday	A	B	C	learn to play an instrument. # Assertiveness # Being assertive is an important <b>social skill</b> . It involves being able to express one's self effectively and directly, i
4	2015	ACAD	ResPracSevereDisab	A	B	C	to close relationships, brothers and sisters with severe disabilities need full access to <b>social skill</b> interventions and opportunities, various methods and mod
5	2015	ACAD	ResPracSevereDisab	A	B	C	siblings should include information about the full range of interventions and approaches related to <b>social skill</b> , communication, and positive behavior supp
6	2015	ACAD	AmericanSecondary	A	B	C	) is a viable option for secondary students regardless of their academic ability, <b>social skill</b> competence, or intentions for post-secondary education. However
7	2015	ACAD	ExceptionalChildren	A	B	C	Education, 4, 4758. **37;41411;TOOLONG. S. (2001). Empirically based <b>social skill</b> intervention: A case for quality-of-life improvement. Behavioral Disorders,
8	2015	ACAD	TeachingExceptional	A	B	C	one possible skill illustrated in the text is conflict resolution, which is a <b>social skill</b> . In their lesson, Ms. Perez and Ms. Hughes use an interactive journal activi
9	2014	ACAD	ExceptionalChildren	A	B	C	much of the data were teacher ratings (e.g., behavior problem rating and <b>social skill</b> ratings), and would be expected to be highly correlated. However, as
10	2014	ACAD	ExceptionalChildren	A	B	C	between behavior problem rating (i.e., externalizing behavior and internalizing behavior) and <b>social skill</b> rating were not highly correlated. Third, the PEELS c
11	2014	ACAD	ExceptionalChildren	A	B	C	replication, and evidence of generalization. # Reichow and Volkmar (2010) reviewed <b>social skill</b> intervention research published between 2001 and 2008 con
12	2014	ACAD	ExceptionalChildren	A	B	C	is central to the conclusion that strong evidence is available to support adoption of <b>social skill</b> interventions for preschoolers with ASD. This report presents
13	2014	ACAD	InstrPsych	A	B	C	book " <b>Social Skills</b> in the Classroom ", Stephens (1992) divided <b>social skill</b> into four main domains: environmental, interpersonal, self-related, and task-relate
14	2014	ACAD	EducTreatmenChildren	A	B	C	teacher determines, based on his/her judgment, that a student has mastered a <b>social skill</b> , the teacher gives the student a certificate of mastery. # Module 1
15	2014	ACAD	CollegeStud	A	B	C	p. 95) as being comprised of the attributes of self-regulation, self-awareness, <b>social skill</b> , empathy and motivation. While the components of Emotional Intell
16	2014	ACAD	PracticeNurse	A	B	C	, so these two go together. Later, language development is also a <b>social skill</b> and so is related to conditions that affect <b>social</b> skills, such as autism. #
17	2014	ACAD	TeachingExceptional	A	B	C	instruction all can further strengthen parent and teacher engagement. # Consider students with <b>social skill</b> deficits who are in online learning. Online instru
18	2014	ACAD	TeachingExceptional	A	B	C	interventions, including small group video-chats where the teacher and select students work on <b>social skill</b> development. A lesson might feature explicit inst
19	2014	ACAD	TeachingExceptional	A	B	C	story. Videos could also be made allowing the student to contextualize the particular <b>social skill</b> being taught. Table 3 lists some tec
20	2014	ACAD	TeachingExceptional	A	B	C	<b>skill</b> being taught. Table 3 lists some technology-based solutions to further enrich the <b>social skill</b> learning process. Here, parents are
21	2014	ACAD	TeachingExceptional	A	B	C	Disabilities in Fully Online Schools During Lessons # Table 3. Technology-Based Solutions for <b>Social Skill</b> Development # Figure 1. Lu
22	2014	ACAD	TeachingExceptional	A	B	C	is the rich opportunity for peer interactions. The context of peer interactions facilitates <b>social skill</b> learning, knowledge acquisition, and a <b>development of soc</b>
23	2014	ACAD	SchoolPsych	A	B	C	any beneficial response the First Step students initially showed. Generalization and maintenance of <b>social skill</b> interventions requires embedding the target
24	2014	ACAD	TeachingExceptional	A	B	C	e.g., Temple Grandin, an internationally renowned author) they often experience significant <b>social skill</b> deficits. Childhood Disintegrative Disorder (CDD) CDD
25	2014	ACAD	TeachingExceptional	A	B	C	given the increasing prevalence rates and wide range of educational, verbal, and <b>social skill</b> deficits associated with this disability. Note: This manuscript was
26	2014	ACAD	LearningDisability	A	B	C	SSRS academic competence added significantly to the model. By contrast, the SSRS <b>social skill</b> and problem behavior subtests and the SWAN hyperactivity d
27	2014	ACAD	LangSpeechHearing	A	B	C	of Oregon. Across children who presented with varied communication, cognitive, and <b>social skill</b> profiles at baseline, they found overall increases in languag
28	2013	ACAD	SchoolPsych	A	B	C	also included <b>social</b> skills training whereby each day during the morning meeting a particular <b>social skill</b> was introduced to the class in a lesson lasting appr

Examples

# Analysing texts with a corpus

Word and Phrase (<http://www.wordandphrase.info>)

The screenshot shows the website 'WORD AND PHRASE .INFO' with a navigation bar and a main content area. The main content area is divided into two sections: 'Frequency list' and 'Input/analyze texts'.

**Frequency list**

RANK #	POS	WORD	DF	FC	PC	NEWS	ACAD	
1	#000	THE	7474	287	1597	615	611	244
2	#001	AND	5994	121	1211	5015	1576	103
3	#002	OF	4767	112	111	5115	915	128
4	#003	TO	4574	104	102	501	111	212
5	#004	IN	4374	101	101	485	111	176
6	#005	FOR	3974	101	101	100	100	100
7	#006	ON	3574	101	101	100	100	100
8	#007	AT	3174	101	101	100	100	100
9	#008	BY	2774	101	101	100	100	100
10	#009	WITH	2374	101	101	100	100	100

**Input/analyze texts**

Enter text to analyze (max 1000 characters) and click on the button 'ANALYZE' to get the results. The results will be displayed in the table below.

Example text: `the word and phrase .info website is a free online resource for word and phrase analysis. it provides a frequency list of words and phrases in the english language. the website is updated regularly with new data.`

RESULTS: WORD AND PHRASE .INFO (1000 CHARACTERS) (1000 CHARACTERS)

Words and phrases from: K

www.wordandphrase.info/analyzeText.asp

WORD AND PHRASE ANALYZER

FREQUENCY LISTS ANALYZE TEXTS

ENTER TEXT BELOW:  MY TEXTS

What is fiscal cliff?

SEARCH CLEAR HELP

WORD PHRASE

Create phrases from the text and search for related phrases

1. Click this.

2. Select words to form a phrase.

3. Click this.

4. View examples and collocations.

SEE LISTS

FREQ RANGE

1-500	501-5000	> 5000	ACAD
4 WORDS	50 %	50 %	0 %

HELP

1. SELECT WORD:  fiscal

2. CHANGE TO: EXACT WILDCARD ALL FORMS SYNONYMS PART OF SPEECH ANY WORD REMOVE

SEARCH CLEAR HELP

SPK	PC	GENRE	TEXT	WORD	FORMS	SYNONYMS	PART OF SPEECH	ANY WORD	REMOVE
1	SPK		And he said we're going to look at this	look	at				
2	SPK		what I thought something had to do with	do	with				
3	SPK		for some capacity to deal with this	deal	with				
4	SPK		the way, we've been reporting on the	report	on				
5	SPK		return to cut, that's why we have the	have	the				
6	SPK		and how do their plans for dealing with the	deal	with				
7	NEWS		improved. As we progress through this year, the	year					
8	NEWS		government spending, which is already headed off a	head	off				
9	NEWS		all over the world, it's training off the	train	off				
10	NEWS		that what Federal Reserve chairman Ben Bernanke calls a "	call	a				
11	NEWS		headlong off what Bernanke has called the "	call	the				
12	SPK		is exactly right, our country is headed for a	head	for				
13	SPK		into working groups to address various issues in the	address	in				
14	SPK		about it, said that that's what	say					
15	SPK		morning went discussions are going on to deal with that	deal	with				
16	SPK		meetings on Capitol Hill aimed at avoiding the nation's	avoid	the				
17	SPK		accumulation of debt that is sending us right over the	send	over				
18	SPK		to do is to find some way to avoid the	avoid	the				

4

Frequency: You may want to focus on the more difficult words

Your text

# Synonyms

## Definition

Genre: Make sure that your words are appropriate

## Collocations

## Examples

# Demonstration: Consider the following sentences

---

- That is gorgeous(?). (Should the word “gorgeous” be used in an academic paper?)
- I look forward to see(?) you. (Should the word “see” be used here?)

# Corpus of Global Web-based English

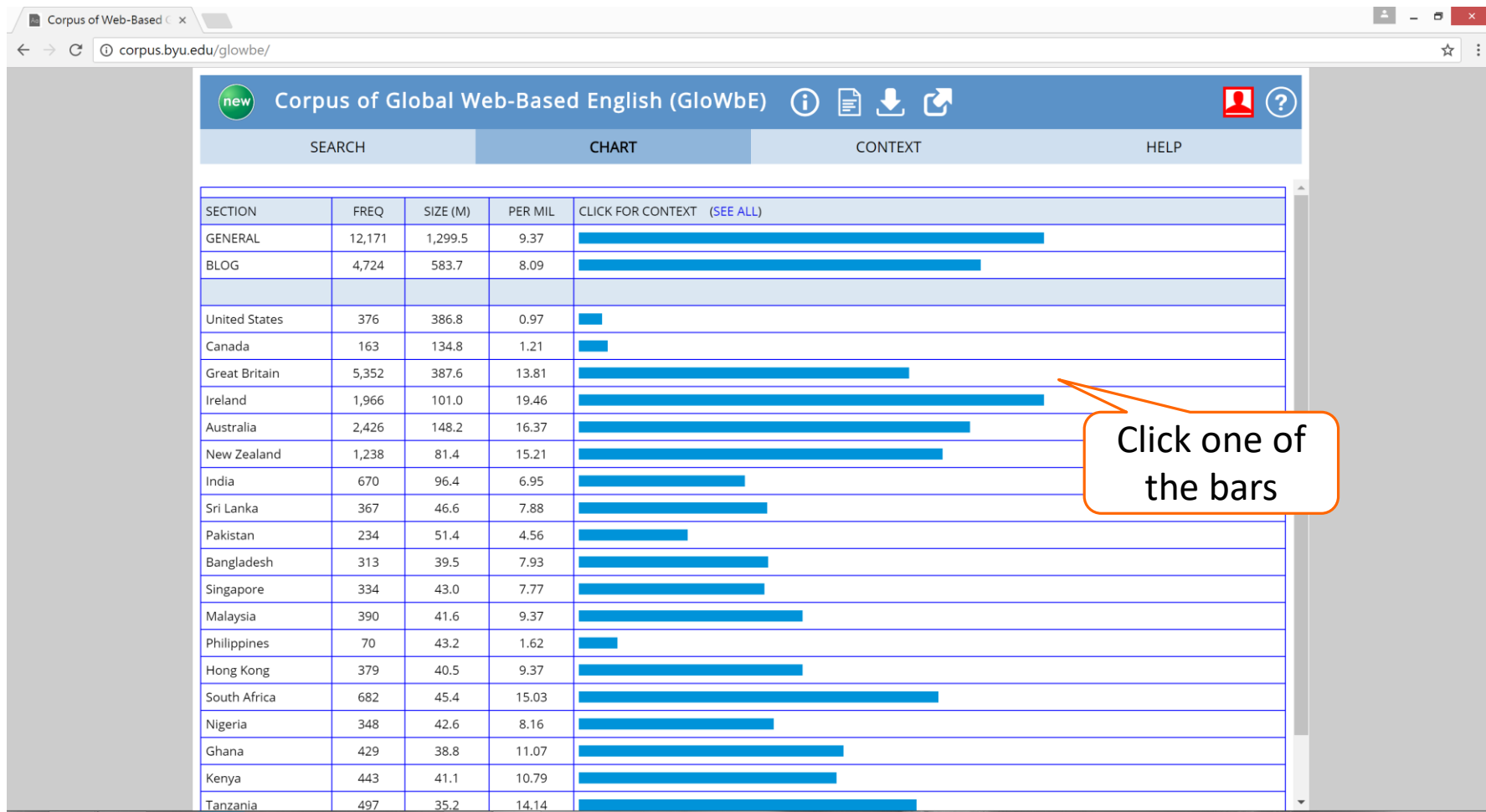
(<http://corpus.byu.edu/glowbe/>)

The screenshot shows the homepage of the Corpus of Global Web-Based English (GloWbE) website. The browser address bar shows [corpus.byu.edu/glowbe/](http://corpus.byu.edu/glowbe/). The website has a blue header with the title "Corpus of Global Web-Based English (GloWbE)" and navigation tabs: SEARCH, FREQUENCY, CONTEXT, and HELP. Below the header, there are links for "List", "Chart", "Collocates", and "Compare KWIC". A search box contains the word "organise" and a "[POS]" button. Below the search box are buttons for "See frequency by section" and "Reset". To the right, there is a "CHART display" section with a "(Hide help)" checkbox and a "NOT LOGGED IN" status. The "CHART display" section contains text explaining the CHART option and providing examples of search results.

1. Click "Chart".

2. Type your expression.

3. Click this.



Corpus of Web-Based English (GloWbE)

SEARCH CHART CONTEXT HELP

SECTION: Great Britain (5,352)  
 FIND SAMPLE: 100 200 500 1000  
 PAGE: << < 1 / 54 > >>

CLICK FOR MORE CONTEXT ☐ [?] SAVE LIST CHOOSE LIST CREATE NEW LIST  [?]

1	GB G	telegraph.co.uk	A	B	C	In seeing a house that is ready to move into. " # 8 Just <b>organise</b> yourself # A third of sales fall through, so don't give your buyer
2	GB G	...oneysavingexpert.com	A	B	C	Barclaycard), Halifax, HSBC, Lloyds TSB and Santander don't, so <b>organise</b> a payment by contacting your tax office, or call the income tax helpline on
3	GB G	i-scribbles.com	A	B	C	outline Creating an outline for your story can be useful because it helps you to <b>organise</b> your ideas and think about how you want your story to start and finish a
4	GB G	guardian.co.uk	A	B	C	as well as truly independent trade unions. The hope is that the left will <b>organise</b> further on these things, bring people round to recognise the need to struggle for
5	GB G	independent.co.uk	A	B	C	justify. So you have got people in barracks playing cards because they can't <b>organise</b> exercises since they don't have the money or the kit. That is not
6	GB G	...lghereducation.co.uk	A	B	C	becoming multiskilled workers. We are having to manage lots of projects and priorities and <b>organise</b> our own world of work and develop new skills all the time, "
7	GB G	greenpeace.org.uk	A	B	C	individual, or with a couple of friends. Download and display some posters, <b>organise</b> a film night, write a letter to the local paper petition your friends,
8	GB G	greenpeace.org.uk	A	B	C	a film night, write a letter to the local paper petition your friends, <b>organise</b> a campaign themed gig or... # Ever had a burning desire to try something
9	GB G	scotland.org	A	B	C	of the younger generation go to dance clubs and party all night. Some cities <b>organise</b> street parties or concerts in the park, with local musicians. My parents usual
10	GB G	...wedding-insider.com	A	B	C	and civil weddings but the principles are the same. The key thing is to <b>organise</b> it and make sure everyone know the situation. # If you're getting married
11	GB G	lemansrace.com	A	B	C	sites offer better facilities and transport to and from the circuit. You can even <b>organise</b> a pre-erected tent in the grounds of a Chateau. # If I camp on-circuit
12	GB G	...lcomeimageawards.org	A	B	C	were initially mixed randomly on a petri dish. As the bacteria grow, they <b>organise</b> themselves into reproducible patterns and shapes that can be predicted with m
13	GB G	...c.clearerchannel.org	A	B	C	, resource centre or community cafe. Containing concrete advice on how to open, <b>organise</b> and maintain these spaces and vital information for those who aim to
14	GB G	guardian.co.uk	A	B	C	is simply accepted that the highest rate of tax is without purpose because so many <b>organise</b> their affairs not to pay it -- even famous knights such as Philip
16	GB G	ipswichstar.co.uk (1)	A	B	C	they can prepare? If I get the team from Burnley now I can <b>organise</b> myself far better than if I suddenly see something an hour before kick-off. I
17	GB G	guardian.co.uk	A	B	C	the lives of other people, but the discussion amongst all mankind about how to <b>organise</b> and run our world; how to improve and shape it. Such people are
18	GB G	telegraph.co.uk	A	B	C	most recognisable leaders of the Tea Party movement. Armeys well-financed group FreedomWorks helped <b>organise</b> the first Tea Party tax day protests in
19	GB G	telegraph.co.uk	A	B	C	FreedomWorks helped organise the first Tea Party tax day protests in 2009 and has helped <b>organise</b> many of the Tea Party rallies and protests against health car
20	GB G	telegraph.co.uk	A	B	C	of February 2009. # A founding member of Tea Party Patriots, she helped <b>organise</b> the waves of protests against healthcare reform the following summer at tow
21	GB G	...ecutivesontheweb.com	A	B	C	A leading banking client now requires a Communication Manager, Corporate Banking to plan, <b>organise</b> , control and deliver complex communications and to sup
22	GB G	...nshistorynetwork.org	A	B	C	the simple expedient of not listening when I returned. Two attempts were made to <b>organise</b> meetings on my behalf to discuss the conference, one at the Womer
23	GB G	britishcouncil.org	A	B	C	test is held in our offices in Bern three times a month, we also <b>organise</b> tests in Zurich and Geneva eight times a year. For the exact dates,

Examples

Websites



# Demonstration

---

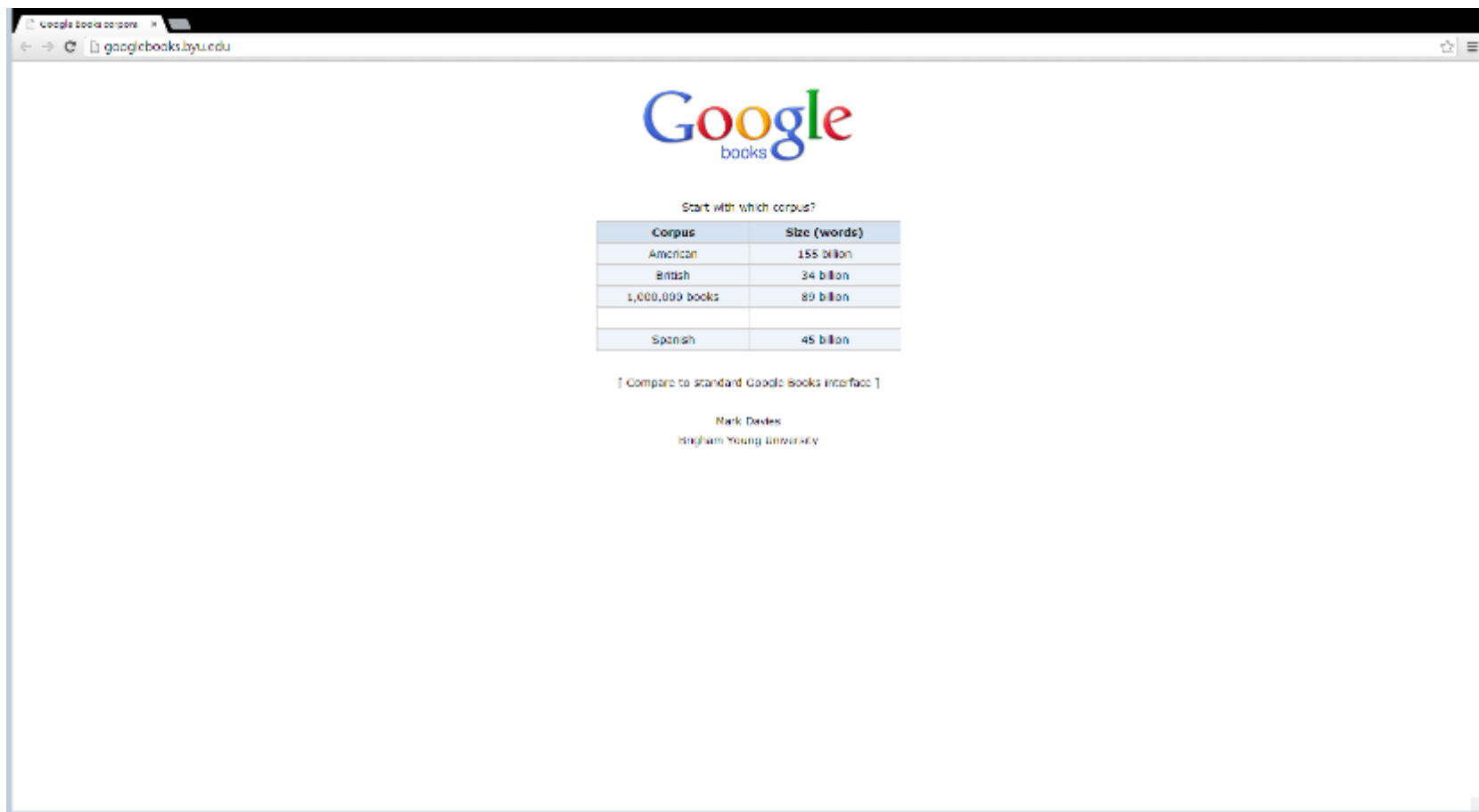
- They organized / organised a campaign.  
(Should we use “organize” or “organise” if the TL is British English?)
- Should we translate 參選 into “stand for election” or “run for office” if the TL is American English?

# Now Corpus (<http://corpus.byu.edu/now/>)

The screenshot shows the NOW Corpus website in a web browser. The browser's address bar displays [corpus.byu.edu/now/](http://corpus.byu.edu/now/). The website has a blue header with the title "NOW Corpus (News on the Web)" and several icons. Below the header is a navigation bar with four tabs: "SEARCH", "FREQUENCY", "CONTEXT", and "HELP". The "SEARCH" tab is currently selected.

The main content area is divided into two columns. The left column contains a search interface with a text input field, a "[POS]" label, and two buttons: "Find matching strings" and "Reset". Below the input field, it says "Search by date" and "Sections Texts/Virtual Sort/Limit Options". The right column contains a "NOT LOGGED IN" status and a "See the words of 2016" section. This section includes a green box with text about downloading the corpus for offline use, followed by a paragraph describing the corpus's size (3.9 billion words of data) and growth (4-5 million words of data each day). It also mentions that users can see what is happening with the language this week, not just 10 or 20 years ago, and provides examples of queries like "frequency of words" and "new words and phrases". A "Five minute tour" link is at the bottom of the right column.

# Google Books corpus (<http://googlebooks.byu.edu/>)

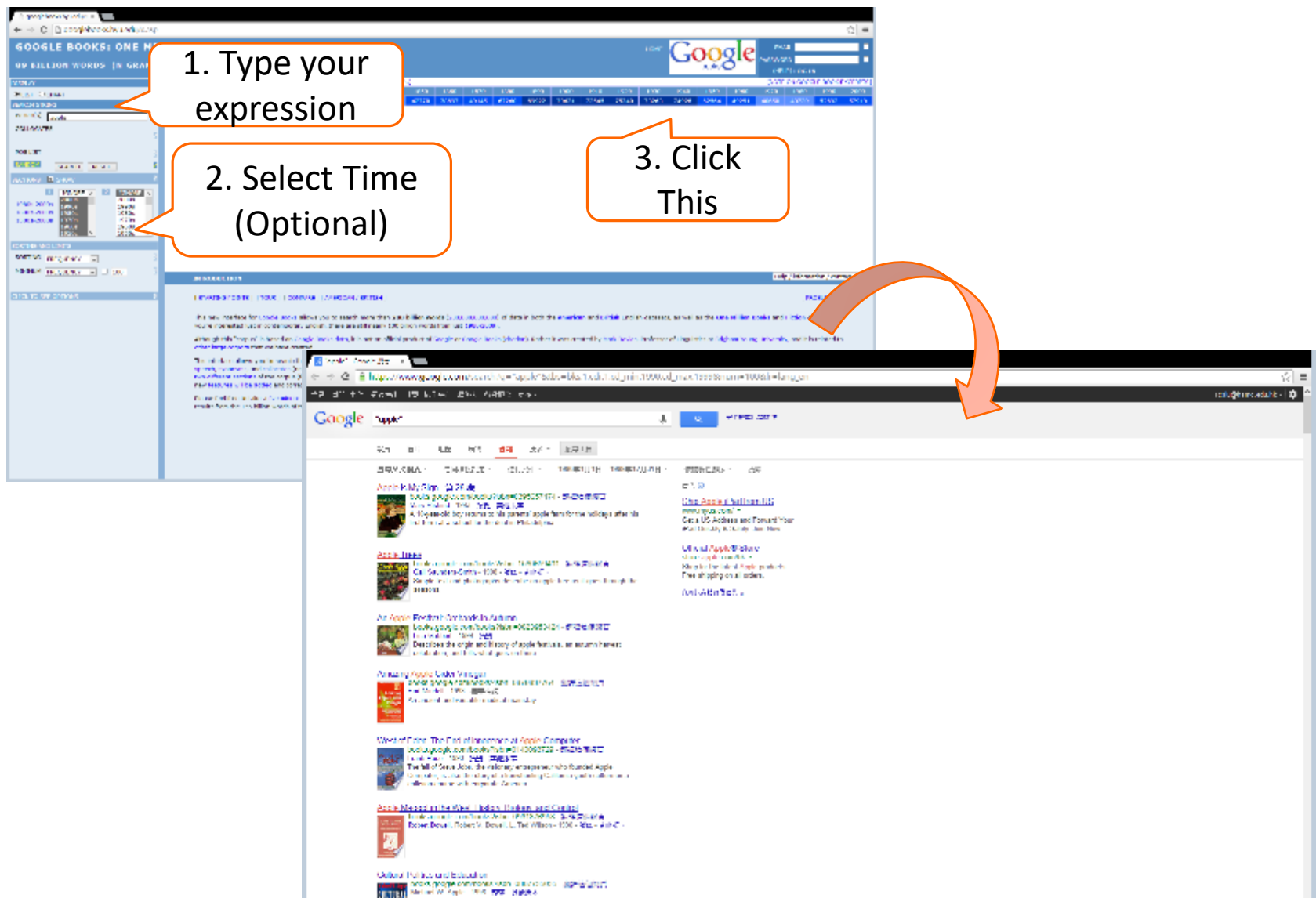


The screenshot shows a web browser window with the address bar displaying "googlebooks.byu.edu". The page features the Google Books logo at the top center. Below the logo, the text "Start with which corpus?" is displayed. A table lists four corpora: American (155 billion words), British (34 billion words), 1,000,000 books (89 billion words), and Spanish (45 billion words). Below the table, there is a link "[ Compare to standard Google Books interface ]" and the name "Mark Davies" followed by "Brigham Young University".

Corpus	Size (words)
American	155 billion
British	34 billion
1,000,000 books	89 billion
Spanish	45 billion

[ Compare to standard Google Books interface ]

Mark Davies  
Brigham Young University



# Monolingual Corpora (Chinese)

---

- 中央研究院現代漢語平衡語料庫  
(<http://asbc.iis.sinica.edu.tw/>)
- 北京大學中國語言學研究中心現代漢語語料庫  
([http://ccl.pku.edu.cn:8080/ccl\\_corpus/](http://ccl.pku.edu.cn:8080/ccl_corpus/))

# Demonstration

---

- 感同身受 vs 身同感受

# Bilingual / Multilingual Corpora / Sentence Databases

---

1. Tatoeba (<http://tatoeba.org/eng>)
2. 必應 (<http://cn.bing.com/dict/?mkt=zh-cn>)
3. iCiBa句庫 (<http://dj.iciba.com/>)
4. 句酷 (<http://www.jukuu.com/>)
5. Line Dict (<http://ce.linedict.com/dict.html#/cnen/home>)
6. Linguee (<http://www.linguee.com/english-chinese/page/about.php?source=auto>)

# Bilingual / Multilingual Corpora / Sentence Databases

---

1. Glosbe (<https://glosbe.com/>)
2. MyMemory (<http://mymemory.translated.net>)



# Demonstration

---

Translate the following expressions with multilingual corpora / sentence databases:

- 改革開放
- 科學發展觀
- 十九大
- 上調人民幣貸款基準利率

# Lab Session 3

(See Handout 11)

Lab Session 3  
(See Handout 11)  
Discussion

# Lab Session 3 (Discussion: Question 1)

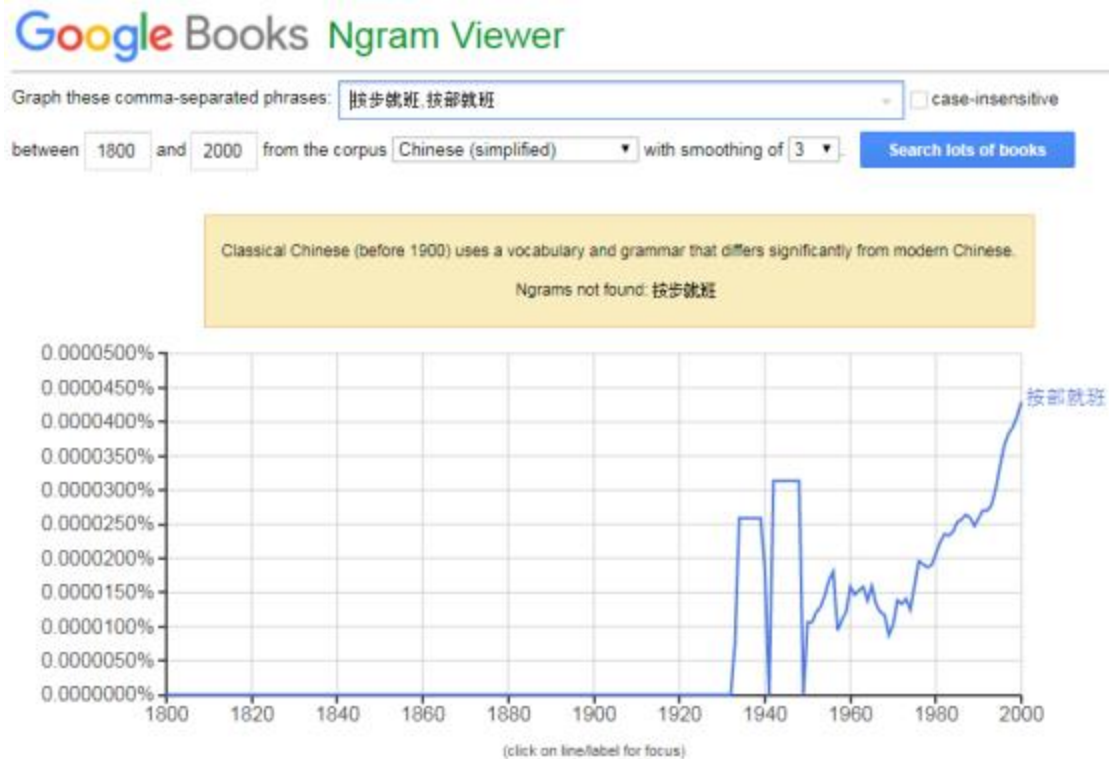
---

## Additional Tools

- Google Ngram Viewer (<https://books.google.com/ngrams>)
- Netspeak (<http://www.netspeak.org/>)

# Lab Session 3 (Discussion: Question 2)

## Additional Tool: Google Ngram Viewer



# Lab Session 3 (Discussion: Question 3)

---

## Tip 1

site:.info.gov.hk



The screenshot shows a Google search interface. The search bar contains the text "一刀切" site:.info.gov.hk. Below the search bar, the "All" tab is selected. The search results show "About 410 results (0.36 seconds)". The first result is titled "標準工時委員會展開工時政策方向諮詢 (附短片)" and includes the URL [www.info.gov.hk/gia/general/201604/25/P201604250434.htm](http://www.info.gov.hk/gia/general/201604/25/P201604250434.htm) and a "Translate this page" link. The snippet text reads: "Apr 25, 2016 - 僱員及僱主均認同工時政策需充分考慮及照顧不同行業或職業的運作情況及實際需要，劃一（「一刀切」）規管所有行業工時可能會對個別行業甚至整體社會和經濟帶來嚴重影響。」「參考上述工作所得的結果，以數據為依歸，委員會歸納以下原則及建議以進一步探討適合香港的工時政策方向：(i) 以立法方式規管僱員工時 ..."

Google

"一刀切" site:.info.gov.hk

All Images News Maps Videos More Settings Tools

About 410 results (0.36 seconds)

標準工時委員會展開工時政策方向諮詢 (附短片)  
[www.info.gov.hk/gia/general/201604/25/P201604250434.htm](http://www.info.gov.hk/gia/general/201604/25/P201604250434.htm) ▼ Translate this page

Apr 25, 2016 - 僱員及僱主均認同工時政策需充分考慮及照顧不同行業或職業的運作情況及實際需要，劃一（「一刀切」）規管所有行業工時可能會對個別行業甚至整體社會和經濟帶來嚴重影響。」「參考上述工作所得的結果，以數據為依歸，委員會歸納以下原則及建議以進一步探討適合香港的工時政策方向：  
(i) 以立法方式規管僱員工時 ...

[简体版](#) | [English](#) | [寄給朋友](#) | [政府新聞網](#)

標準工時委員會展開工時政策方向諮詢（附短片）

\*\*\*\*\*

下稿代標準工時委員會發出：

標準工時委員會（委員會）今日（四月二十五日）展開為期三個月的工時政策方向諮詢，就探討中的工時政策方向收集意見，供委員會隨後在擬定向政府提交的報告時作參考。

委員會主席梁智鴻醫生說：「今次是一個全面、有前瞻性、有方向，及具備審慎內容的廣泛諮詢。由於工時政策極其複雜及具高度爭議性，並會對整體勞工市場、人力需求、僱傭關係、工作文化、營商環境、經濟發展和企業競爭力等不同層面帶來廣泛和深遠影響，我及委員會期望社會各界，特別是僱員及僱主，會積極參與這次諮詢。」

「過去三年，我和來自勞工界、工商界、學術界、社會各界及政府的委員抱着最大的誠意及持開放態度，在沒有前設立場的大前提下，本着以數據為依歸的原則進行討論，採取求同存異的方式，以釐定適合香港的工時政策方向及向政府提供意見。」

「委員會於二○一四年完成廣泛的第一階段諮詢及深入的工時統計調查，並檢視一系列可能影響工時政策的經濟及社會因素，收集了不少有用的數據及社會各界的意見。調查結果顯示『規定僱主及僱員須透過僱傭合約訂明工作時數、超時工作安排及超時工作的補償方法』普遍獲得僱員及僱主的贊同。他們亦普遍認同工作與生活平衡以及職業安全與健康的重要；而合適的工時政策或有助提高生產力及僱員工作表現。此外，勞工團體普遍建議標準工時每周44小時，超時工作補償為基本工資率的1.5倍；而僱主團體及僱主則普遍反對制定全港性劃一的工時法例，並認為透過清晰的僱傭合約，已能夠達致勞資雙方同意的工時安排。部分僱主亦傾向可接受工時政策應聚焦幫助社會上較低收入、議價能力較低及需無償超時工作的基層僱員。僱員及僱主均認同工時政策需充分考慮及照顧不同行業或職業的運作情況及實際需要，劃一（「一刀切」）規管所有行業工時可能會對個別行業甚至整體社會和經濟帶來嚴重影響。」

「參考上述工作所得的結果，以數據為依歸，委員會歸納以下原則及建議以進一步探討適合香港的工時政策方向：

- (i) 以立法方式規管僱員工時的政策；
- (ii) 由於不同行業及職業的情況不同，故工時政策應提供一定彈性，「一刀切」立法方式並不合適；

## 網上廣播

MP4 格式

MP4

標準工時委員會工時政策  
方向諮詢記者會

3

2

一刀切

## Press Releases

繁體版 | 簡體版 | Email this article | news.gov.hk

Standard Working Hours Committee launches Consultation on Working Hours Policy Directions (with video)

\*\*\*\*\*

The following is issued on behalf of the Standard Working Hours Committee:

The Standard Working Hours Committee (SWHC) today (April 25) launched the Consultation on Working Hours Policy Directions to collect views over a three-month period on working hours policy directions, which the SWHC will then use as a reference in preparing its report for the Government.

The Chairperson of the SWHC, Dr Leong Che-hung said, "This is a comprehensive, forward-looking, directional and extensive consultation with substantial content. In view of the fact that working hours policy is an extremely complex and highly controversial subject which carries widespread and significant implications for the overall labour market, manpower demand, employment relations, work culture, business environment, economic development and business competitiveness, the SWHC and I appeal for the active participation of the community, particularly employees and employers, in this consultation."

"Over the past three years, I and fellow members from the labour and business sectors, academia, the community and the Government have been acting sincerely with an open-mind and no pre-conceived stance on this subject. Adopting the spirit of forging common grounds and allowing differences, we will strive to formulate appropriate working hours policy directions for Hong Kong by using an evidence-based approach and offer advice to the Government."

"In 2014, the SWHC conducted an extensive first-stage consultation and an in-depth working hours survey, and examined a range of economic and social factors relating to the working hours policy. Through these we collected voluminous and useful data as well as views of the community. The findings of surveys revealed that employees and employers generally agreed with the approach of 'providing for stipulation of hours of work, overtime arrangements and overtime compensation in employment contracts'. They also in general recognised the importance of work-life balance and occupational safety and health. Productivity and work performance of employees might be enhanced by an appropriate working hours policy. In addition, labour organisations generally suggested a weekly standard of working hours at 44 hours, and overtime compensation at 1.5 times of the basic pay rate, while employers' associations and employers in general objected to introduction of uniform working hours legislation in Hong Kong and considered that employment contracts with clear stipulations would be sufficient for the purpose of achieving working hours arrangements as mutually agreed by employees and employers. Some employers also tended to accept a working hours policy targeting at helping grassroots employees with lower income, less bargaining power and who were required to work overtime without compensation. Employees and employers both recognised that a working hours policy had to take full account of the operational situations and practical

needs of different sectors or occupations. A uniform (across the board') working hours regulation for all sectors could result in serious impacts on individual sectors or even society and the economy as a whole."

### WEBCAST

MP4 format

MP4

Press conference on  
SWHC's Consultation on  
Working Hours Policy  
Directions



# Lab Session 3 (Discussion: Question 3)

---

## Tip 2

Step 1: `site:.legco.gov.hk`

Step 2: URL Modification

english	chinese
eng	chi
en	zh, tc, sc, ch
e	c

# Lab Session 3 (Discussion: Question 3)

---

## Example

[www.legco.gov.hk/yr07-08/english/bc/bc04/papers/bc040415cb1-1223-1-e.pdf](http://www.legco.gov.hk/yr07-08/english/bc/bc04/papers/bc040415cb1-1223-1-e.pdf) (English Version)

[www.legco.gov.hk/yr07-08/chinese/bc/bc04/papers/bc040415cb1-1223-1-c.pdf](http://www.legco.gov.hk/yr07-08/chinese/bc/bc04/papers/bc040415cb1-1223-1-c.pdf) (Chinese Version)



Statistics

and



Examples

are equally important.



# Lab Session 3

---

- Search Engines, Monolingual Corpora and Multilingual Corpora
- Concordancers
- Sentence Databases

Create your own  
multilingual corpora?  
We're going to talk about  
this around Week 10.