Data-driven learning: Exploring the affordances of corpora and DDL for SPOCS

Dr. Peter Crosthwaite (University of Queensland)

- Data-driven learning (DDL; Johns, 1991) involves the investigation of language corpora through printed concordance materials or students' direct, hands-on use of corpus query tools

- DDL has been used for a range of purposes including language acquisition, genre awareness, and understanding discipline specificity

- DDL increasingly used in tertiary contexts to enhance the teaching of disciplinary postgraduate thesis writing

Introduction – Data-Driven learning?

Data-driven learning?

USING CUTTING-EDGE
SOFTWARE

TO REVEAL THE
STATISTICAL PATTERNS
OF LANGUAGE USE

Does it work?

Creates plentiful opportunities for 'learner-centered' focus on form (Bernadini, 2004)

Statistical (and increasingly visual) nature of corpus output facilitates constructivist/connectionist approaches to language learning

Self-guided nature of students' corpus engagement for DDL results in improved learner autonomy.

Increasingly relevant pedagogy for modern digitally oriented learners looking for alternatives to dictionaries or translation websites.

Medium to large effect sizes found across a wealth of studies and thousands of research participants in recent meta-analyses (e.g. Boulton & Cobb, 2017).

Disciplinary DDL?

- Substantial cross-disciplinary variation in the language features employed across the hard sciences, social sciences, and arts and humanities (e.g. Hyland, 2000)

- Language reference resources employed also different (e.g. Steel, 2012).

- Significant variation in the uptake and usage of corpora for DDL.

- Large-scale DDL studies focusing on corpus use across multiple disciplines are still rare

- Little is currently known about postgraduate students' disciplinary corpus use or query habits.

- "what students report to be doing or what we assume they are doing when we observe them might be quite distant from what they are actually doing" (Pérez-Paredes et al, 2011: 235).

Project 1 – Improving disciplinary writing through corpora
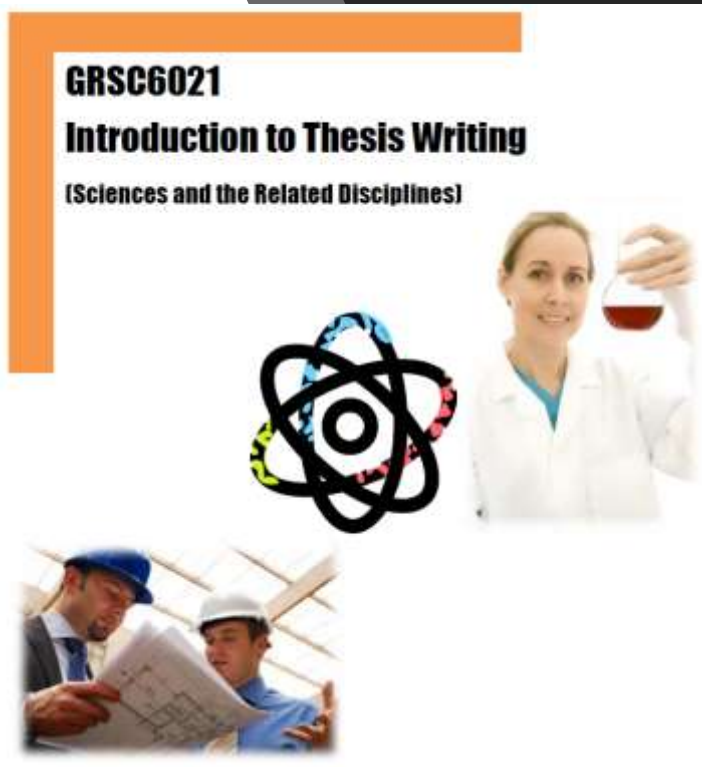
- HKU - Teacher-Development-Grant funded project headed by Dr. Lillian Wong of CAES and Dr. Peter Crosthwaite (now UQ).

- Aimed to build multidisciplinary corpus of successful Ph.D. and M.Phil theses from all faculties/schools at HKU

- Built new corpus platform specifically for DDL that captures users' corpus usage habits and query patterns

- Embedded into course materials for existing Arts & Humanities, Sciences-focused graduate thesis writing courses at HKU.

Participants

- 327 postgraduate PhD and MPhil students enrolled on the thesis writing courses during the data collection period
  - 89 = humanities and related disciplines classes
  - 238 = sciences and related disciplines)

- Medicine and Engineering students constitute the two largest cohorts in sciences, Education largest group within humanities.

- Post-course questionnaire on referencing tools (n=93) - Students "often" used dedicated language learning apps on smartphones (53%), social media (41.7%), dictionaries (47.1%), and spellcheckers or grammar checkers (33%).

- 3.1% "often" used corpora, 57.4% have never heard of a corpus or had never used one for writing.

Corpus size

Table 1. Word counts per faculty in the HKGC

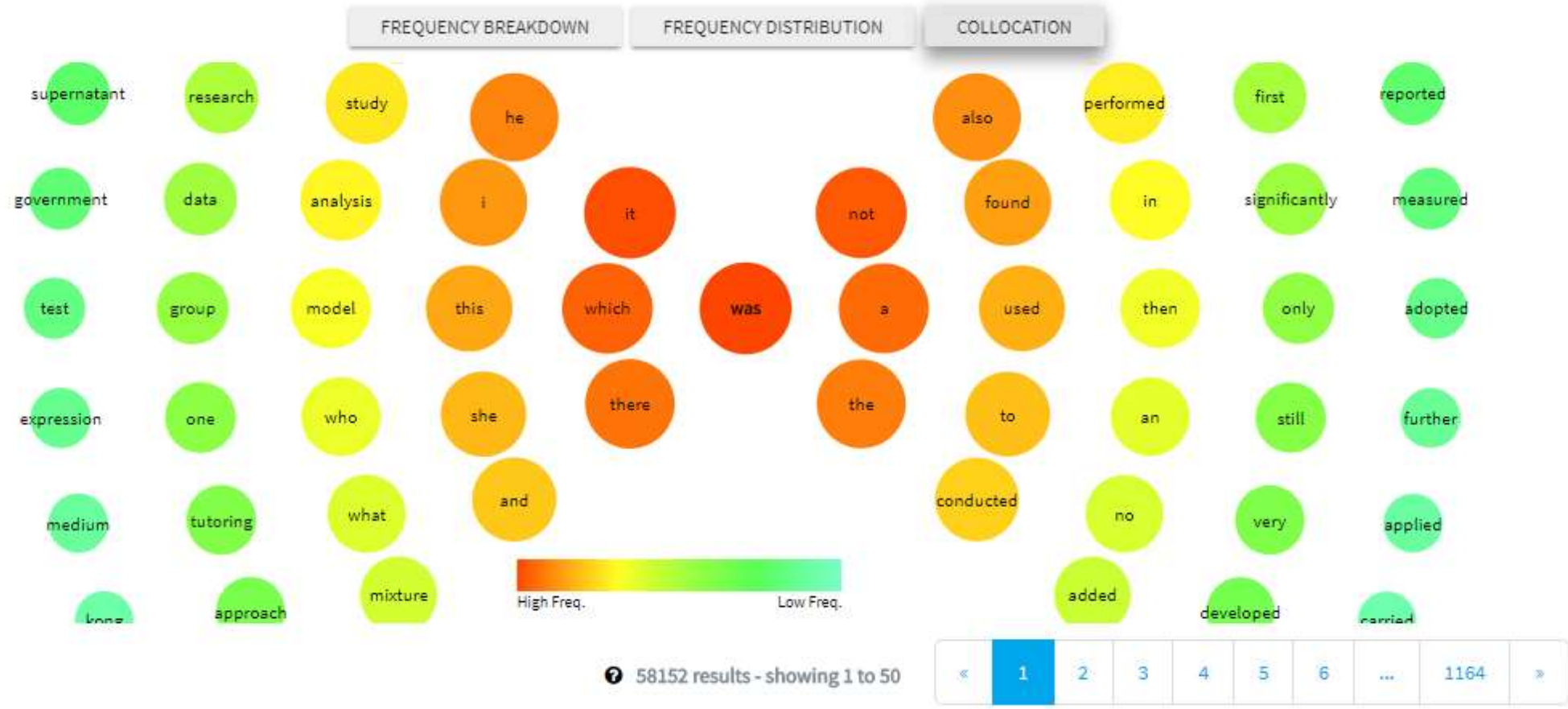| Faculty	Number of words	Proportion (%)
Arts	2,260,154	20.8
Education	2,108,264	19.4
Medicine	1,877,060	17.3
Social sciences	1,217,307	11.2
Law	912,170	8.4
Engineering	811,025	7.5
Architecture	788,474	7.2
Science	438,082	4
Business and economics	401,057	3.7
Dentistry	55,793	0.5

Corpus platform – The HKGC

- Presenting students with multiple sources of information aids learning under a constructivist approach (Flowerdew, 2015)

- "the more possible starting points a corpus offers for exploitation, the more likely it is there exists an appropriate starting point for a specific learner" (Widmann, Kohn & Ziai, 2011, p.168)

Corpus platform – Collocation output

Frequency distribution across sections and disciplines

Frequency breakdown
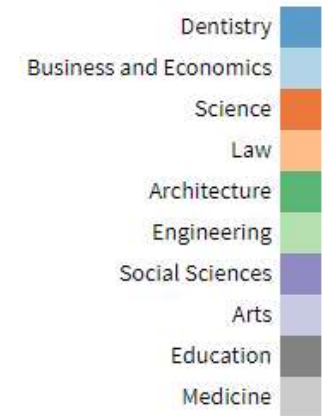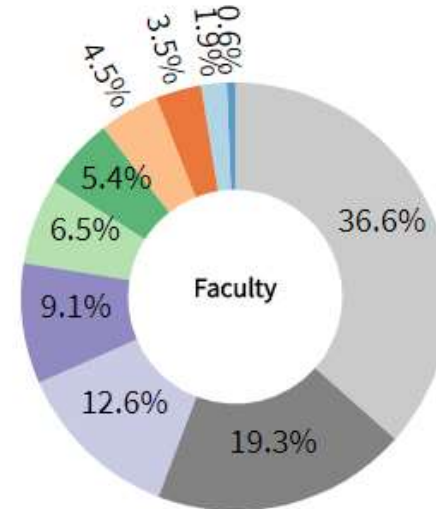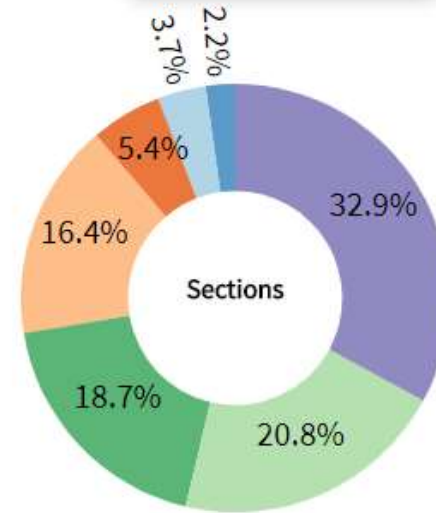
Frequency Breakdown (instances / per million words)
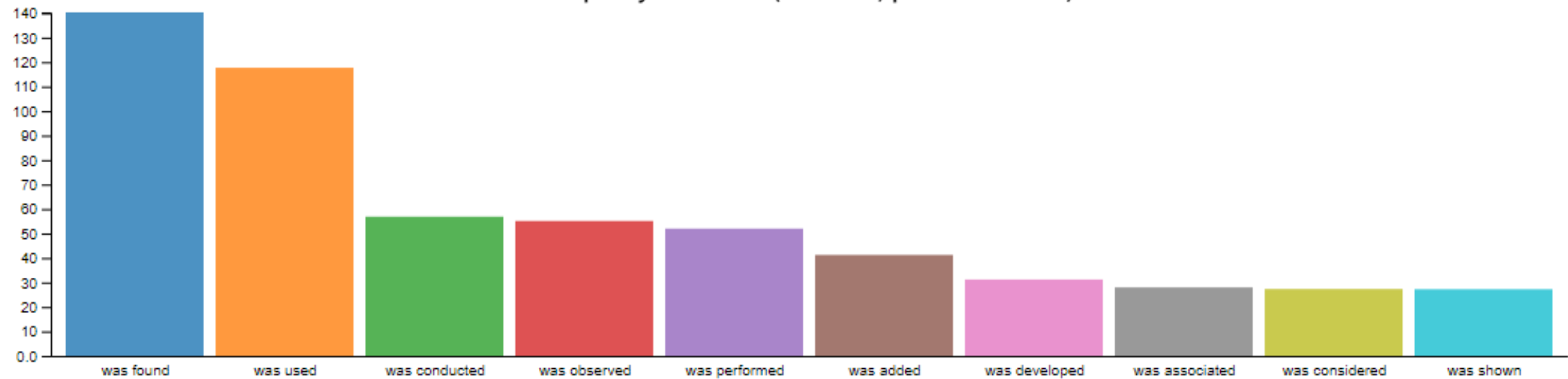
Learner tracking parameters

The corpus platform tracked learners' corpus use according to the following parameters:

- (1) User ID
- (2) time, date and duration of user log-in to the platform
- (3) individual corpus query syntax
- (4) any filters applied to corpus query results (i.e. searches by thesis section/faculty and sub-discipline)
- (5) corpus function used (i.e. concordance, frequency breakdown, etc.)

DDL materials

Course curriculum covers variation in thesis structure; and language and discourse features used in reviewing research literature, identifying the research gap, explaining methodology, reporting and discussing results, etc.

DDL activities were used to make the presentation of these structures more interactive and to promote both 'top-down' and 'bottom-up' learning where students combine analyses of longer sequences of texts with corpus-based investigations of grammar and lexis (Charles, 2014).

DDL materials

Step 3: **Point out the importance of knowing more about that topic AND/OR a gap in current knowledge**

- The importance of one's research can be indicated by a reference to a particular problem or the limitation(s) of existing research.

- To do so, gap or problem statements, which are very common in RA/theses across topics and fields, are often used. They often include negative words and expressions such as *little research, few studies, no work*, or words beginning with in- (e.g., *incomplete, insufficient*).

- A second important feature that frequently occurs in gap statements is the contrastive signal word - e.g., *However, While, Although.* These words introduce contrasts or problems in relation to the part of the thesis chapter which has preceded them.

- A third common feature of gap statements is that they often occur at the beginnings of new paragraphs

---

Enter the term 'little research', while checking 'introduction' in the search field.
1) Write 5 sentences containing the words 'little research' in the middle
   a) [In Hong Kong]_____ little research _____[has been conducted in relation to gay culture.]
   b) [there is]_____ little research _____[addressing this issue.]
   c) [there has been]_____little research _____[related to forgiveness or reconciliation.]
   d)[there has been]____ little research _____[on the pooling effect.]
   e) [very]_____ little research _____[has examined these variables.]

2) Enter the term '* little research' then click 'frequency breakdown'. What *collocates* of 'little research' are common? 'Very'.

3) Enter 'little research' and click 'frequency distribution'. Is the phrase 'little research' commonly used in your discipline? Note: it mostly occurs in the Division of Learning, Development and Diversity and the Division of Policy, Administration and Social Sciences Education.

The following methodology section is abridged from the one in a paper entitled "Relationship between impulsive sensation seeking traits, smoking, alcohol and caffeine intake, and Parkinson's diseases". Complete the passage with the appropriate verb form.

METHODS

Patients

Consecutive outpatients of Caucasian descent fulfilling Queen Square Brain Bank criteria for PD[18] _____ **(undergo)** a Mini-Mental State Examination[19] (MMSE) administered by the examining physician and _____ **(invite)** to participate if the MMSE score was >26. We _____ **(exclude)** patients with significant cognitive decline because of the requirement to complete the behavioural and depression rating scales. The Unified Parkinson's Disease Rating Scale (UPDRS)[20] part II _____ **(rate)** for the "on" state and patients provided a list of all current medications and their dosages. Demographic data including age, sex, and age at onset of symptoms of PD _____ **(also collect)**.

p.74 - Tense and voice in the methodology chapter

Using corpora to resolve morphosyntax

Use the wildcard * after a word's base form as well as the frequency breakdown and collocation functions to get an idea about which tense to use for each verb in the extract. Note that some of these verbs will be in the passive voice, requiring the addition of the correct 'be' verb. Also use the wildcard * before the search item to represent the verb-to-be if necessary, and check the 'methodology' section. The first has been done for you:

1) Under*, click 'frequency breakdown'. Two phrases are possible from the available options. Click on either, and use the concordance lines to make a decision.
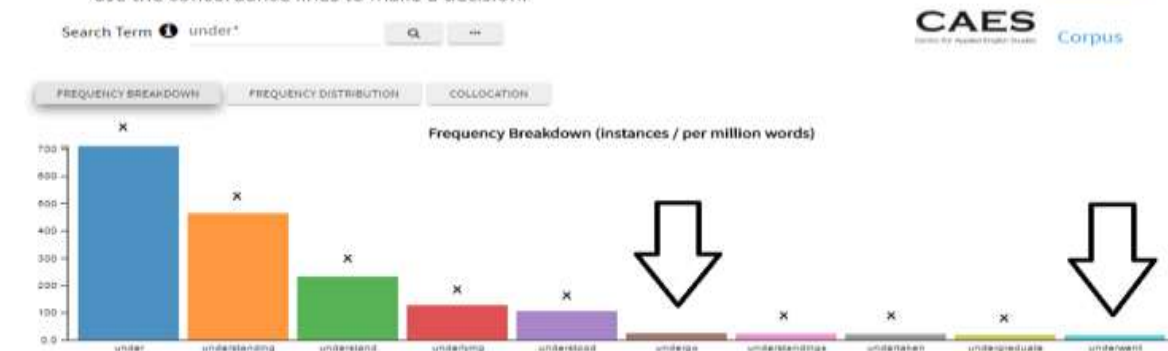
2) [were] Invite [invited]
3) Exclude [excluded]
4) Also collect [Look for words before 'Also' as well as the right form of 'collect'] [also collected]
5) Provide [provided]
6) [Were] give [given]
7) Include [included]
8) Collect [collected]
9) Assess [assessed]
10) Range [ranged]
11) [Was] estimate [was estimated]
12) [Was] convert [was converted]

DDL materials – replacing old gap-fill activities

Example DDL task – Referring to yourself in the introduction (Taken from Wong, 2019)

Corpus task: Referring to yourself in the introduction

Using the HKG corpus, search for 'My', remembering to select the 'introduction' section in the search menu.

1) **Is 'My' commonly used in thesis writing?**
   ['My' has a frequency of 55.54 encounters per million words, which is above the Biber's (2006) cut-off point 40 instances per million words. In short, 'my' is quite commonly used in thesis writing.]

2) **What words are usually found after 'My' (click 'collocates', and look for words that occur _after_ 'My')**
   ['Study', 'research' and 'thesis'.]

3) **In what discipline is 'My' typically used? (click 'frequency breakdown') Is it common in your discipline? Why do you think this is the case?**
   ['My' is typically used in the Policy, Administration and Social Sciences Education, Humanities, Sociology, Social Work and Social Administration and English Language Education etc.]

4) **Repeat 1-3 for 'Our' - what do you notice?**
   ['Our' has a frequency of 63.79 encounters per million words, which is above the Biber's (2006) cut-off point 40 instances per million words. In short, 'our' is quite commonly used in thesis writing.
   'Understanding', 'knowledge', 'study' and 'paper'.
   'Our' is typically used in the Computer Science, Economics and Finance and Humanities etc.]

5) **Repeat 1-3 for 'We' - what do you notice?**
   ['We' has a frequency of 151.17 encounters per million words, which is way above the Biber's (2006) cut-off point 40 instances per million words. In short, 'we' is commonly used in thesis writing.
   'Can', 'have', 'will', 'are', 'also' and 'propose'.
   'We' is typically used in the Computer Science, Economics and Finance and Humanities as well.]

6) **Based on 1-5, under what circumstances can we refer to ourselves directly in the introductions of our theses?**
   [E.g. We can refer to ourselves directly when we give ourselves credits on the thesis.]

Results – Overall platform use (over 1 semester)

- 258 unique users
- 11,436 accumulated searches
- 449 accumulated site visits
- 2,498 searched (unique) queries
- Frequency of unique queries is far higher than those featuring as exemplars in the course materials - indicative of substantial variation and innovation among users in the queries made.
- Indications of continued corpus use beyond the final taught class and the data collection cut-off date.

Most frequent corpus query terms

- Although the most frequent query terms were exemplars from course materials, many query terms were of the students' own making.
- Students are choosing to query the corpus when encountering terms from the reading or disciplinary course content
- They are doing this without explicit prompting in an autonomous fashion – a key tenet of the affordances of DDL for language learning.
- Flexibility in the use of wildcard and POS queries for unique queries outside those in the DDL materials

Table 4. Most frequent corpus query terms

Query syntax	No. of queries	Query syntax	No. of queries
this research	316	indicate	30
Studies have * that	231	will	29
my	221	It is hoped that	28
little research	175	null hypothesis	28
This chapter	172	show	26
suggest	106	may	24
describe	99	*_adj risk	24
our	84	found_v	23
research shows	81	Show*	22
argue	77	*variable	22
few research	60	strongly	22
find_v	58	substantially	22
Possibly	55	show_v	22
no studies	55	studies have*that	21
research question	49	no research	21
research	47	hopefully	19
as * as	44	studies have*	19
*little research	43	research has*	18
studies have shown that	42	claim	18
research has*	39	studies have	18
we	37	deriv*	17
*have shown that	35	the questionnaire *_v	17
limited studies	33	*_adj studies	16
*_adv understood	32	few studies	16

Top ten users

- Even among top 10 users, there is significant variation in their corpus usage
- This variation suggests certain users were looking for different information from the corpus, or better understood / were more comfortable with the format of the output they were receiving from the given function

Table 5. Ten most frequent corpus users' usage history

Rank by query frequency	Query frequency	Faculty filter?	Concordance	Collocation	Frequency breakdown	Frequency distribution	Sort left?	Remarks
1	296	Business, Medicine, Education	212	12	52	21	0	Used wildcards and POS _v. Queries include *margin, correlation, negate, theory, accumulation.*
2	239	Science, Medicine	170	48	14	7	18	Only used wildcards to derive suffixes. Queries include *large amount, residue, sentence, quotation, quantities.*
3	198	Medicine	104	37	47	10	0	Combined use of wildcards and a range of POS tags. Queries include *questionnaire, consumption, proportion.*
4	197	Education, Social sciences	130	39	14	14	44	Used wildcards and POS _v _n. Queries include *integration, creation, learn, entangled, provoke.*
5	181	All	122	21	25	13	0	Used wildcards and POS _v _n. Queries include *questionnaire, survey, derive, future work, knowledge gap.*
6	168	Default	108	55	3	2	0	Only used wildcards to derive suffixes. Queries include *judicial, classical literature, monograph, weakness.*
7	152	Education	121	15	9	7	25	No use of wildcards/POS tags. Queries include *little, thorough, word choice.*
8	151	All	146	1	2	2	1	Only used wildcards to derive suffixes. Queries include *STEM, self-directed, schools.*
9	150	Science, Medicine	115	8	18	9	0	Used wildcards and POS _adj _v. Queries include *majority, consumption, risk.*
10	143	Education, Engineering	99	6	31	7	0	Combined use of wildcards and a range of POS tags. Queries include *questionnaire, motivation, excretion.*

Discipline-specific use

- Greater use of the corpus in the sciences
- Proportion of queries for the arts and humanities/social sciences is higher than their 27% enrolment ratio, accounting for around 34% of queries in all.
- Queries by subdiscipline heavily skewed towards the sciences.

Table 9. Queries filtered by faculty[1]

Faculty/Faculties	Queries	Proportion by faculty (in this list only)	Proportion of all HKGC queries
Medicine	1,098	24.6%	9.6%
Engineering	1,032	23.1%	9.0%
Education	819	18.4%	7.2%
All*	517	11.6%	4.5%
Arts	208	4.7%	1.8%
Dentistry	173	3.9%	1.5%
Architecture	144	3.2%	1.3%
Law	139	3.1%	1.2%
Business and economics	125	2.8%	1.1%
Education + Social science	111	2.5%	1.0%
Medicine + Science	92	2.1%	0.8%

[1]Not all selected options are shown.

Figure 7. Proportional use of query functions employed for queries filtered by faculty

Query functions filtered by faculty
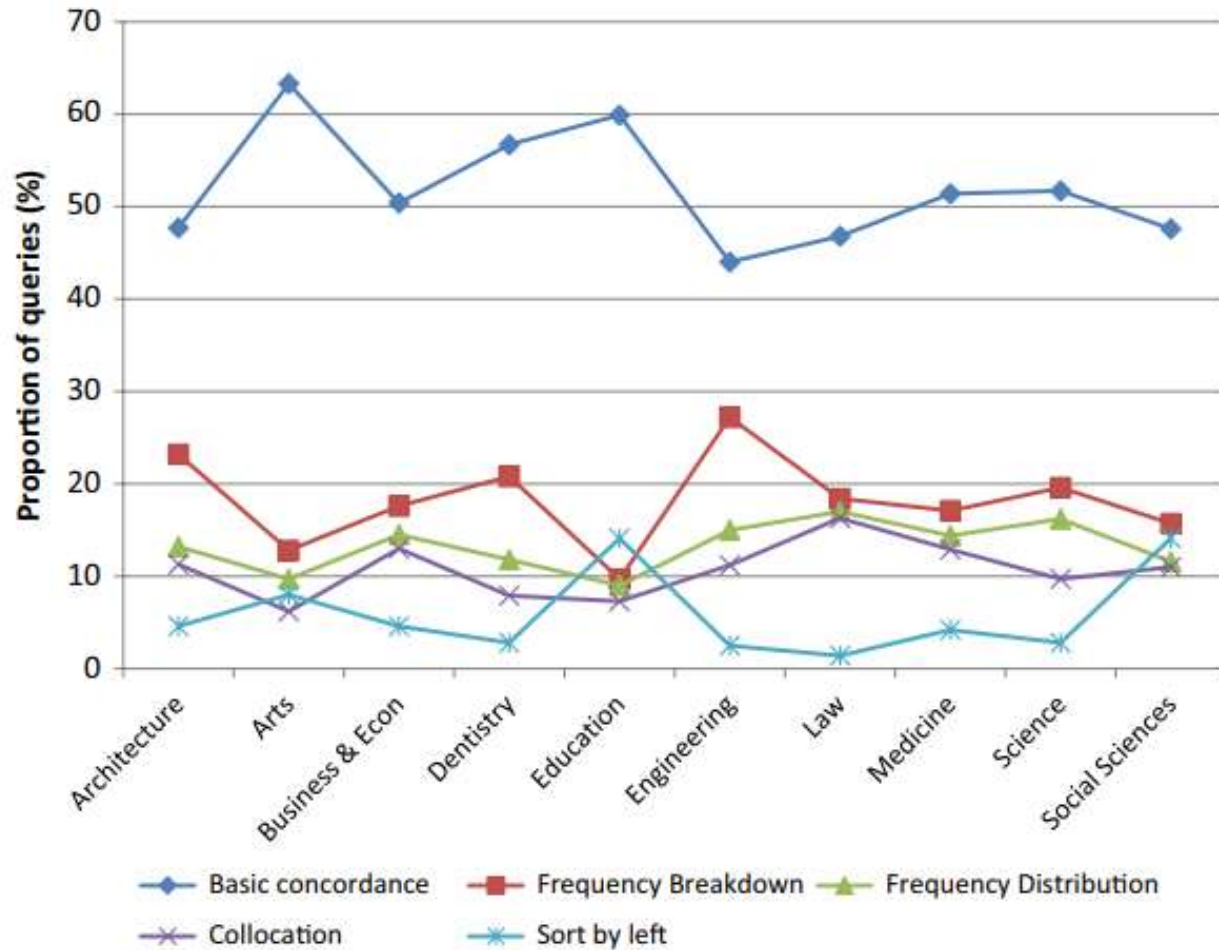
- Those in the physical/life sciences prefer to query the corpus for statistical information

- Those in arts and humanities and social sciences disciplines prefer to query the corpus for textual information

Unique queries by facuilty

Table 10. Top five unique queries by faculty (cut-off frequency = 5)

Faculty	Query syntax	Freq.	Faculty	Query syntax	Freq.
Architecture	cultural landscape	16	Engineering	describe	47
	paradigm	12		possibility	28
	landscape	7		machine learning	28
	show_v	7		novel	18
	"garden City"	6		simulations	7
Arts	show*	12	Law	show *_v	13
	clean	6		include*	9
				data have shown	7
				thesis	6
Business	marketing	16	Medicine	no studies	24
	*margin	8		it is hoped that	23
	blockchain	8		the questionnaire *_v	22
				consum*	18
				limited studies	16
Dentistry	describe	20	Science	describe	34
	hypothesis	6		possibly due	25
				no studies	14
				hopefully	10
				The *of the questionnaire	7
Education	understanding	24	Social	tackle	13
	learn *_n	16	sciences	gender	9
	emotion	11		tourism	9
	show*	10		entangled	6
	evaluation_n	9		gap	6

Discussion – Take-home successes

Most comprehensive analysis yet of the corpus usage characteristics of students engaged in DDL for disciplinary thesis writing.

Short length of time required for corpus uptake via our unguided approach to the DDL materials

Students often went **beyond the provisions of their assigned tasks** to freely experiment under their own autonomy "unconstrained by the imposition of specific data-driven learning tasks set by the teacher" (Hafner & Candlin, p.306).

Sustained and autonomous corpus use both during and *after* the writing course

Significant interdisciplinary variation found in the usage of particular corpus functions and query syntax

Future opportunities

More activities needed in line with some of the disciplinary preferences we have outlined in this research.

Need corpora that facilitate analysis of "higher-level" phenomena at the discourse or genre level (Boulton, Carter-Thomas & RowleyJolivet, 2012: 3) e.g. genre-annotated corpora for DDL (e.g. Cotos, Link & Huffman, 2017)

A logical extension is to create discipline-annotated corpora and employ these for DDL.

Of the 327 students enrolled, 69 (21%) had not attempted to query the corpus – need to know *why*.

Project 2: Taking DDL online?

- "It generally takes a corpus linguist to teach DDL writing classes" (Ädel, 2010)

- Few corpus linguists working on direct pedagogical applications of corpora in Australia

- Current shift towards blended / flexible learning in Australian universities

- Taking DDL training online can help 'spread the word' (Römer, 2009) about DDL both within and outside of Australia's borders

DDL in Australia…

- Corpus 'missionaries' have so far failed to 'spread the word' (Römer, 2009) to this remote part of the world.

- Despite the fact that "almost all Australian universities promote [...] student centred teaching practices which foster students' independent thinking and learning" (Conroy, 2010, p.863), there is little evidence of Australian universities' engagement with corpora or DDL.

The challenge

- No study exploring the provision of DDL training in an entirely online format, taking the teacher (almost) out of the equation.

- Need to develop resources that raise students' awareness of the affordances of DDL and provide plentiful opportunities for learners to gain skills in corpus consultation in the absence of a physical 'teacher'.

- No real 'connection' present between students and teachers outside of the materials themselves

- Move to an online only format could potentially impact learner engagement and opportunities for constructivist learning (vital for DDL) to occur (Kop, 2011).

Approach to materials

- Scaffolding embedded in activity design.
- Activities range from 'guided' tasks to 'unguided' tasks (Perez-Parades et. al, 2011)
- Activities follow Carter and McCarthy's (1995) "3 Is", namely:
  - (1) Illustration (looking at data)
  - (2) Interaction (discussion and sharing observations and opinions)
  - (3) Induction (making one's own rule for a particular feature).
- Interaction includes multiple-choice / dropdown questions, text/number entry activities and (occasionally) discussion forum response.
- Completion of activities leads students to Induction.

UQx SLATx
Improving writing through corpora: Data-driven learning?

Improving writing through corpora: Data-driven learning?
UQx

THE UNIVERSITY OF QUEENSLAND
AUSTRALIA

Enroll in SLATx

About This Course

Welcome to UQ SLATx 'Improving writing through corpora'!

This 5-hour course aims to provide you with the tools, knowledge and skills to improve the writing you do for your degree programme. As a novice writer for your discipline, you may make mistakes with your writing that you don't know how to solve. If you are a second-language user of English, this can be even more difficult. But what if there was a data-driven solution to your language problems? It just so happens that there is such a solution... The use of corpora, or large, principled collections of electronically searchable text – has been proven to really help writers understand how language is actually used across a wide range of contexts. Advances in technology have now made it much easier to get online access to this kind of data, numbering in the millions (or even billions) of words. Better still, we can search for examples of language in use across different disciplines, like engineering, the arts, physics, and so on, ensuring we can use the right terminology for the subjects we are writing.

Course Number		SLATx
Classes Start		Jun 11, 2018
Classes End		Jun 12, 2030
Estimated Effort		05:00

Question 1

2 points possible (ungraded)

(All of these questions use BAWE)

What is the 'raw' frequency of the word 'science' in the BAWE corpus?

What is the 'normalised' frequency of the word 'science' in the BAWE corpus? (per 1,000,000 words)

Submit

Q&A, general discussion

Hide Discussion

Topic: Section 2 - Q&A, general discussion / Topic-Level Student-Visible Label

Add a Post

Show all posts ▼   by recent activity ▼

? Question 2 "express" frequency as a verb   2
In question 2, what is the frequency of "express" as a verb. I typed all possible answers (216, 265, 143, 113, 92, 78 or the sum of all or partial sum) and i could not find the corr...

💬 Incorrect responses?   3
In some of the previous tests and general quizzes I am getting answers incorrect when there is no other options, for instance raw frequencies when it is clearly the answer. I d...

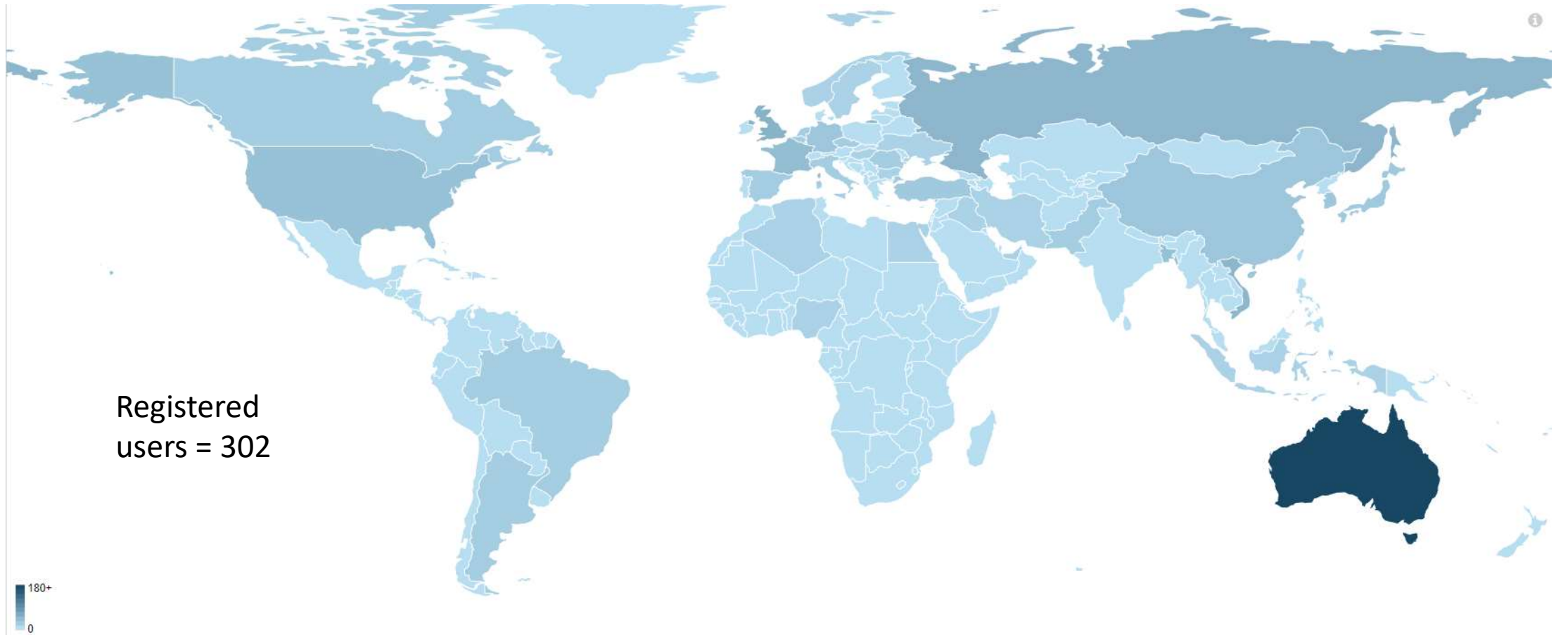## Different forms of the word 'research'

1/1 point (ungraded)

What information is there about the different forms of the word 'research' (e.g. 'researched', 'researching', etc.) on dictionary.com?

○ The different variants of 'research' are clearly shown

○ There are lots of examples of each type

◉ That information is not clearly shown here ✔

Registered users = 302

Geography Metrics

36	Australia	United Kin...	Russia
Total Countries or Regions Represented	Top Country or Region by Enrollment	Second Country or Region by Enrollment	Third Country or Region by Enrollment
	59% of learners	4% of learners	3% of learners

Course outline

Module 1: Swapping your dictionary or translation website with corpora

Module 2: Basic corpus functions and understanding concordance output

Module 3: Using corpora to resolve lexical errors

Module 4: Using corpora to resolve grammatical errors

Module 5: Using corpora to understand disciplinary / register variation

'The data suggest that'?

1/1 point (ungraded)

There might also be confusion about which form of a word to use in a multi-word expression. For example, what happens if I want to know whether the correct expression is 'the data suggest that...' or 'the data suggestS that...'?

Enter 'the data suggest that' or 'the data suggests that' into the dictionary. What information do you receive?

○ Lots of examples of both, with added frequency information to tell you form which is the most common

◉ No information at all, or at least not very much ✔

○ Lots of examples of both that you can use to formulate a rule for the correct answer

Module 1: Google Translate fail

Module 2: Introduction to BAWE (Language detectives-in-training)

- Google Scholar-as-corpus
- SketchEngine for Language Learning (SKELL)
- BAWE in SketchEngine Open
  - Simple, Lemma, POS search
  - Wildcards
  - Frequency > node forms/tags
  - Collocation

What happens if we **don't know exactly which word we want to search for** in a multi-word expression, or if we just want to look at some available options for what *might* appear in a multi-word expression?

Here, we can tell Sketch Engine that we want to search for a missing word by using the **wildcard symbol '*'** within, or at the beginning or end of a search. In this example, we will reconsider the expression **'the data suggest that'**, but *replace 'suggest' with a wildcard*. Please follow the search term shown in the image below:

Query the, data, .*, that 18 (2.16 per million) ⓘ

text#207	[2]. The linear relationship exhibited by	the data demonstrates that	the MCA was correctly calibrated and hence
text#331	have bought a Book of the Dead of their own and if	the data shows that	they could not have afforded it, then we really
text#409	an average 6.25% in the 40 countries analysed.	The data suggests that	all countries (except for Kazakhstan) have
text#871	general factor and then many specific factors,	the data suggests that	there are also group factors, linking in for
text#1167 ...	interactions in large numbers of crystals.	The data implies that	strong hydrogen bonds, such as O·H···O, favour
text#1413 ...	the fact that it explains much less variation in	the data implies that	the original data matrix might be quite far from
text#1445 ...	Normal distribution, the linear dependency in	the data suggests that	a Gaussian random walk model for the log prices
text#1447 ...	quadratic effects. One concern when examining	the data is that	gestational age is measured in weeks. A more
text#1819 ...	in Table 5.2, and in Appendix 2. PV diagrams of	the data show that	higher peak pressures are attained with WOT at
text#2040 ...	you say about /r/ in syllable final position? <p>	The data indicates that	the allophones of /r/: </p> 1. appear in initial and
text#2040 ...	. For example, train [t ern] and pray [p er]. <p>	The data indicates that	the allophones of /l/: </p> 1. occur in initial,
text#2040 ...	. For example bell [beɫ]☐and☐fold [fouɫd].	The data indicates that	/r/ in syllable final position is not
text#2090 ...	<p> The fact that heteroskedasticity is present in	the data means that	the least squares estimator is still a linear
text#2410 ...	architecture document, is a wise one.	The Data Model that	is used by the company is implemented through
text#2557 ...	, 802.11 is well known to not perform at	the data rates that	should be possible. Aside the reasons already
text#2637 ...	firewall at one time it is possible to encrypt	the data so that	no third party can read them. Encrypting data
text#2651 ...	large ridge (Appendix 1). Closer inspection of	the data reveals that	the lowest pressure and temperature readings
text#2653 ...	of mate is not limited to a particular age group.	The data show that	the vocative is distributed in all age groups,

Wildcard search: 'in this paper...'

Let's try searching using wildcards again

Enter 'in this paper we *'. What are the three words that fit in the wildcard slot?

[ ]

[ ]

[ ]

Submit

Frequency of POS for 'influence'

1/1 point (ungraded)

You can see from the results above that 'influence' in the form VVN is more frequent than VVI (infinitive form e.g. '...to influence x') or VVO (the base form of the lexical verb, e.g. ...this may influence x)

From your intepretation of these frequency results, what can we say about 'influence' as a verb in academic writing?

○ The word 'influence' as a verb is most likely to be used in passive constructions than in other forms

○ The word 'influence' as a verb is more likely to be used in infinitive constructions ('...to influence x') than use in the base form (...this may influence x)

○ The word 'influence' as a verb is less likely to be used in infinitive constructions than in passive constructions

◉ All of the above ✔

Module 3: Resolving lexical errors

word form

"Using the corpus is a useful <u>active</u> [activity] for learning language"

word choice

"We <u>insure</u> [ensure] the very best deal on your new car"

collocation

particles ("the demand <u>of</u> [for] potatoes")

phrasal verbs ("I need more time to work <u>over</u> [out] the solution"

common phrases

"On <u>an another</u> [the other] hand"

Example: Demand 'of' potatoes?

Query the, demand, of 91 (10.92 per million) ℹ️

Page 1 of 5 Go Next | Last

text#26	wives, subordinating their civilian lives to **the demands of** the state, XC urges civilians both to accept
text#67	coping processes that people use to manage **the demands of** particular stressful events. The cognitive
text#67	healthy' drinkers in their ability to cope with **the demands of** everyday life, and that deficiencies in more
text#79	the supply of indulgences was in response to **the demands of** the laity, and was perhaps a manifestation of
text#107	, the staging of his work failed to meet **the demands of** his emotional capacity. </p> '... By placing his
text#170	children. Literally, women just had to endure **the demands of** her husband. There was no practical route of
text#179	which would guide his research. In short, **the demand of** the structure and the supply of facts result in a
text#259	of poor legal service [62], yet it responds to **the demands of** the critical consumer, whether it will go far
text#298	end in failure and government pandering to **the demands of** interest groups, whether pensioners or tenants
text#365	interest rate between t and t+1. This means that **the demand of** money depends positively on price level and
text#406	reached nearly £ 2 m in year 4. Speculating **the demand of** our products correctly, all of our 54,000 units
text#422	between traditional religious practises and **the demands of** political authority. Significantly, the play
text#484	contemporary British play. How fully did it met **the demands of** the script, and did it constitute an
text#507	, tailors and grocery markets which catered for **the demand of** those involved in the industry. (XL about
text#513	yearning for a professional position. Despite **the demands of** this existence however, disdain for suburbia
text#544	and transforming them in accordance with **the demands of** a modern liberal state Nye, 'Just don't
text#586	tensions, manorial tithes and taxation. **The demands of** the revolt were formulated within twelve
text#618	lowest and make most cash flow. We forecasted **the demand of** standard and XL products by last year's demand
text#618	shift for peak time production. Because of **the demand of** accessories grew to double amount' we order
text#618	almost meet the condition of EOQ' for example **the demand of** 4 weeks is constant and can be forecasted. All

Page 1 of 5 Go Next | Last

What's the connection?

1/1 point (ungraded)

Look at the results gained from the concordances above. What connects these things in terms of what they are?

Choose an answer from the drop-down menu

They are all people or things that are making demands ▼	✔

Select an option
They are all things being demanded
They are all people or things that are making demands
They are all types of demands

SUBMISSION HISTORY

'the state', 'everyday life', 'her husband', 'the critical consumer', 'interest groups', 'political authority', 'those involved in the industry, and *'a modern liberal state'.*

Frequency list

Frequency limit: [0] [Set limit]

Page [1] [Go] Next >

word	Frequency	Items: 67 || Total frequency: 361
P | N the demand for	121	
P | N the demands of	62	
P | N the demand of	26	
P | N The demand for	20	
P | N the demand curve	17	
P | N the demand is	10	
P | N the demand and	6	
P | N the demand .	6	
P | N the demand ,	6	
P | N the demands on	4	
P | N the demands for	4	
P | N the demand was	4	
P | N the demand side	3	
P | N the demand in	3	
P | N The demands of	3	

From the answer above, I now have a sense of what 'the demand for' means, compared with that of 'the demand of':

'The demand **of** X' - X refers to *who is demanding something*

'The demand **for** X' - X now refers to *something being demanded*

As our original sentence mentioned **'potatoes'**, this sounds like something being demanded (as potatoes cannot 'demand' anything!).

Therefore, *I can guess that 'the demand for' is the correct answer, and not 'the demand of'!*

STAFF DEBUG INFO

Query **the, demand, .*** 361 > Positive filter **the, demand, for** 121 (14.51 per million) ⓘ

Page [1] of 7 [Go] Next | Last

text#32	about. For firms like Ferrari and Porsche, **the demand for** their products are inelastic because they
text#81	'foreignness', and the mercantilist fear that **the demand for** luxury imports will drain the metropolis of its
text#81	advocates of luxury (notably Mandeville) that **the demand for** luxury would "lead to the expansion of commerce
text#81	, and justified the state's act of harnessing **the demand for** luxury as an engine for economic growth.
text#129	people merely needed a leader. He stresses that **the demand for** reform had become violent by the sixteenth
text#129	his allies. However, this was primarily due to **the demand for** Luther's teachings - many made great fortunes
text#130	for the movement which grew so quickly that **the demand for** preachers became too high for Geneva to cope, so
text#173	zeal did not lie in one particular issue, but in **the demand for** a more equal and just society, and they were
text#217	rates, cause increased hiring costs reduces **the demand for** workers. Countries with strong EPL have
text#220	increase, so does spending and therefore **the demand for** imports both by consumers (with more
text#234	Y 1 through the multiplier effect so increasing **the demand for** money. Given fixed money supply the money
text#285	increasing precarity of the labour market and **the demand for** flexibility of employment signifies that 'the
text#285	developments, the service sector grows and **the demand for** knowledge work increases (Giddens 2000;
text#285	to the contrary is very persuasive. Regarding **the demand for** high-skilled labour, Rifkin (1995) argues
text#285	' (Warhurst and Thompson 1998:5). Thus, **the demand for** knowledge workers is presumably limited to
text#300	and provides vital data concerning **the demand for** certain objects during the Bronze and Iron Ages
text#367	at point B is equal to the slope at point C. Since **the demand for** clothes increase, and clothes being labour
text#367	increase, and clothes being labour intensive, **the demand for** labour increases proportionately more than
text#367	labour increases proportionately more than **the demand for** capital. Thus it follows that the wage-rental
text#376	interest rate (i.e. R) and loans (i.e. L) locus, **the demand for** loanable funds could be summarised as follows. </p>

Page [1] of 7 [Go] Next | Last

Module 5: Disciplinary / Register variation with 'Word Sketch Difference'

"Believe"

and/or	0	7	0	0.01
understand	0	1	--	7.8
accept	0	1	--	9.0
see	0	1	--	9.0
threaten	0	1	--	10.5
educate	0	1	--	10.8
believe	0	2	--	11.8

object	57	204	0.23	0.16
neutrino	2	0	10.1	--
matter	7	0	9.8	--
verne-	1	0	9.1	--
thinking/case	1	0	9.1	--
centralization	1	0	9.1	--
chromate	1	0	9.1	--
studio	1	0	9.0	--
motorist	1	0	9.0	--
emo-tion	1	0	9.0	--
rsa-	1	0	8.9	--
detection	1	0	8.7	--
stiffness	1	0	8.6	--
universe	2	1	9.2	7.0
man	1	5	5.7	7.7
polixene-	0	1	--	7.3
gender	0	2	--	7.7
america-	0	2	--	7.9
scientist	0	2	--	8.0
waterland-	0	2	--	8.2
neaira-	0	2	--	8.3
thunder	0	2	--	8.3
structuralism	0	2	--	8.3
phaedra-	0	2	--	8.3
girl	0	3	--	8.4
proposition	0	3	--	8.5

subject	48	418	0.19	0.33
astronomer	2	0	10.2	--
huygens-	1	0	9.4	--
pacey-smith-	1	0	9.4	--
cosmologist	1	0	9.4	--
trotman-dickenson-	1	0	9.4	--
cosmology	1	0	9.3	--
llp-	1	0	9.3	--
administrator	1	0	9.2	--
cauchy	1	0	9.2	--
bacon-	1	0	9.1	--
auditor	1	0	9.1	--
nasa-	1	0	9.0	--
scientist	8	1	10.5	5.9
agent	0	4	--	8.0
smith-	0	4	--	8.0
jackson-	0	4	--	8.1
ranke-	0	4	--	8.2
berkeley-	0	4	--	8.2
burke-	0	4	--	8.2
swinburne-	0	4	--	8.2
greek	0	5	--	8.4
schopenhauer-	0	5	--	8.5
plato-	0	6	--	8.6
nietzsche-	0	8	--	9.0
historian	0	12	--	9.3

modifier	15	116	0.06	0.09
foolishly	1	0	10.8	--
extremely	1	0	5.6	--
previously	1	0	5.6	--
however	2	1	7.2	5.9
now	2	2	5.7	5.6
strongly	1	2	6.5	7.2
therefore	1	3	4.0	5.5
only	1	3	3.0	4.6
widely	1	4	6.0	7.7
generally	1	4	5.4	7.2
also	3	13	3.3	5.4
either	0	1	--	5.4
soon	0	1	--	6.7
passionately	0	1	--	8.0
innocently	0	1	--	8.1
wholeheartedly	0	1	--	8.1
erroneously	0	1	--	8.1
unquestioningly	0	1	--	8.1
fervently	0	1	--	8.1
arrogantly	0	1	--	8.1
instead	0	3	--	8.1
certainly	0	4	--	8.3
sincerely	0	2	--	9.1
truly	0	7	--	9.3
firmly	0	6	--	9.8

s4529292
2 months ago - endorsed 2 months ago by **uqpcros1** (Staff)

a) The most important thing I learned from this module is the use of the corpus software as it is a very helpful tool that can be used for my time in university.

b) This module helped me become aware of words or phrases that need to be avoided in order to have an academic tone. In addition, I will be able to use different types of metadiscourse taught in this module in my essays.

c) As a writer, I will make sure to know which tone to use when writing and make use of the corpus software.

Add a comment

sunnivaeinarse
2 months ago - endorsed 2 months ago by **uqpcros1** (Staff)

a) The most important thing I learned during this module was the importance of academic tone and how to write in an academic versus non-academic way.

b) This module has changed my knowledge of academic writing in the way that I learned new techniques in term of corpora.

c) After completing this course, I will be more careful in my 'go-together' words and academic writing.

Add a comment

Student feedback

Student feedback (from interviews)

The software is very nice because I can like specifically search for prepositions, that kind of thing? […] Here I can try to find words after, words before, for example. I can like very specifically do my search and change some parameters, so it's nice to have something that's more specific for what I want.

It [corpora] is very unique, I get it. Doing the course I actually see it is a veritable [sic] resource that can be used, right? You can't just plug in your entire thesis and it will come up with the answers, but it does give you that third party verification that sometimes when you get to a point in your writing when you are like 'which is how it is better written?' [sic], and having a corpus with infinite amounts of probabilities of how a word or phrase can be used, I think it has value.

Challenges for Version 2: Collaboration and teacher 'presence'
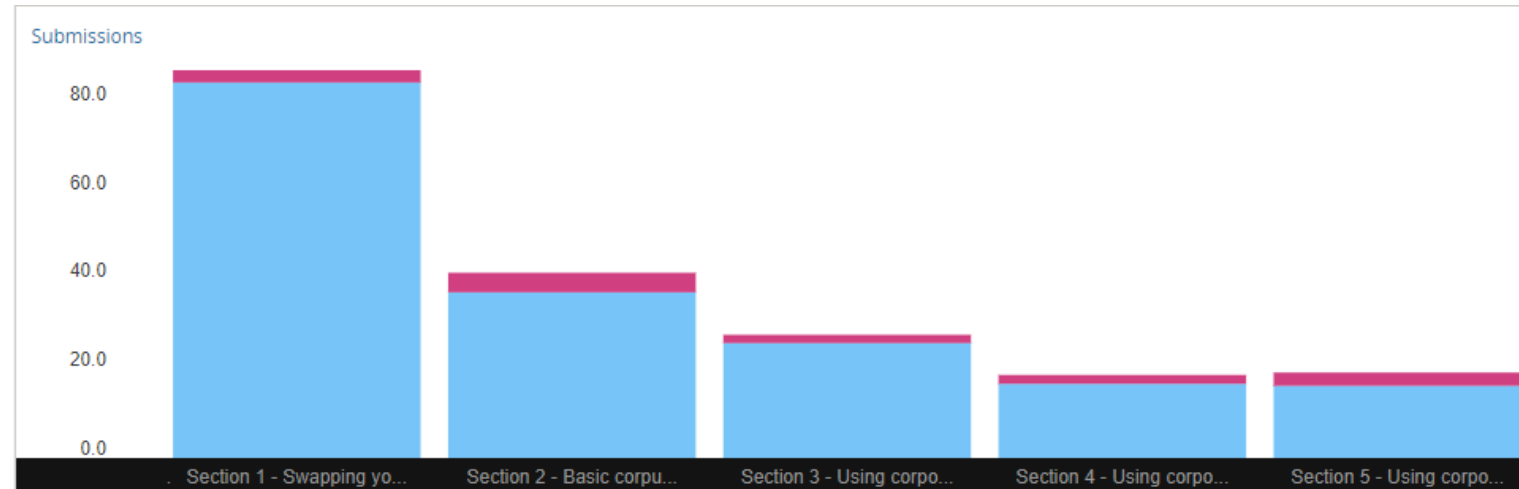
- Self-guided nature of the course may have impacted opportunities for constructivist learning to occur.

- Such learning is supposed to be one of DDL's main strengths, but can be lost in online-only instruction without a teacher or other students present (Siemens & Downes, 2009; Kop, 2011)

- Transition to online format placed responsibility for learning on the shoulders of the individual learner

- Answers to activities programmed in with a single 'correct' answer, negating opportunities for meaningful discussion and reflection on concordance output.

- Learners forced to self-manage their motivation to complete the course, manage their time appropriately, and (perhaps most importantly for DDL novices) demonstrate control over the technology itself (Bouchard, 2009).

Improving Writing Through Corpora – Version 2!

https://edge.edx.org/courses/course-v1:UQx+SLATx+2019/about
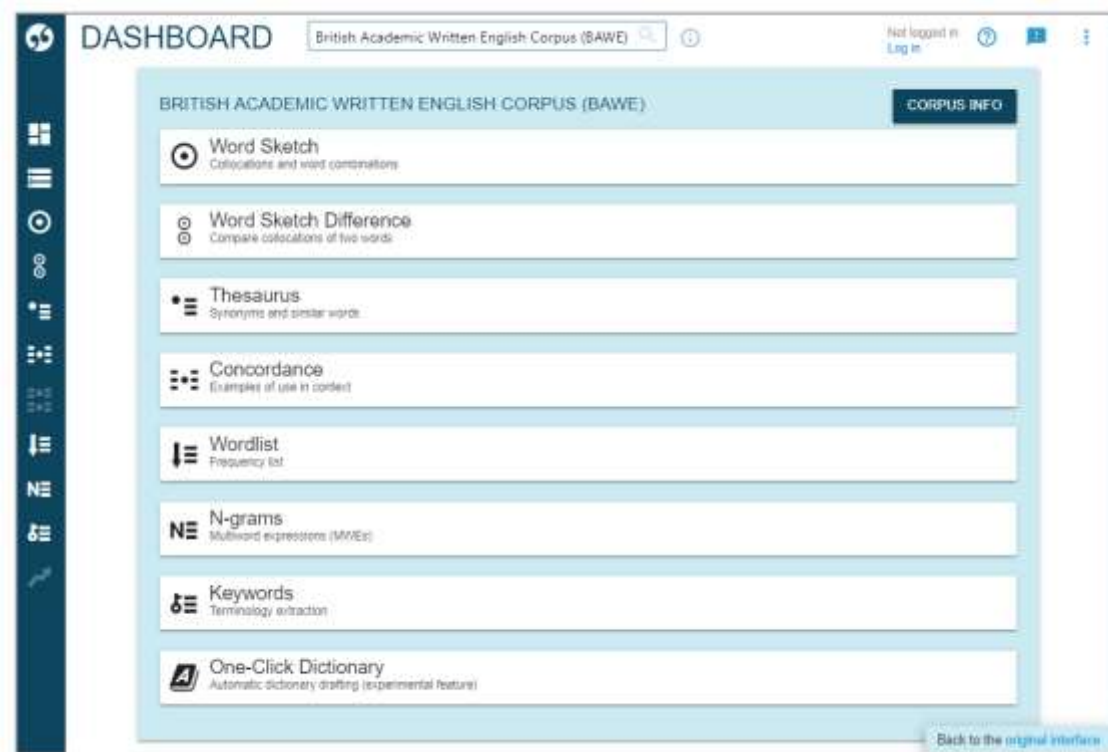
Improvements in Version 2 include:
A) All course images and functionality have been updated for the 'new' Sketch Engine interface.
B) New functions specific to the 'new' Sketch Engine interface are now included in the course (e.g. Good Dictionary EXamples (GDEX))
C) Course is now completely self-contained - no need for external assessments. Certificates of completion generated automatically upon completion of online activities.
D) Improved reflective component and opportunities for peer discussion.

New books in 2019!