

Parallel EAP Corpora Search Engine User Guide – Table of Contents

1. Interface	2
2. Basic Mode / Advanced Mode.....	2
2.1 POS and Semantic Simple Search	2
2.2 POS Advanced Search	4
2.3 Semantic Advanced Search.....	6
3. Search Parameters in POS and Semantic Search.....	6
3.1 POS Tag and Semantic Tag	6
3.2 Match Mode	8
3.3 Concordance Length	9
3.4 Scope.....	10
4. Interpreting Search Results in POS and Semantic Search	11
5. POS Word List.....	13
5.1 Collocation in POS Search.....	13
5.2 Collocation Context in POS Search	14
6. Semantic Word List	14
6.1 Collocation in Semantic Search	15
6.2 Collocation Context in Semantic Search	16
7. Comparison of Results in POS Search.....	16
8. Comparison of Results between POS and Semantic Search.....	18

Parallel EAP Corpora Search Engine User Guide

1. Interface

The Parallel EAP Corpora search engine provides a set of features which are grouped into different tabs as shown below.

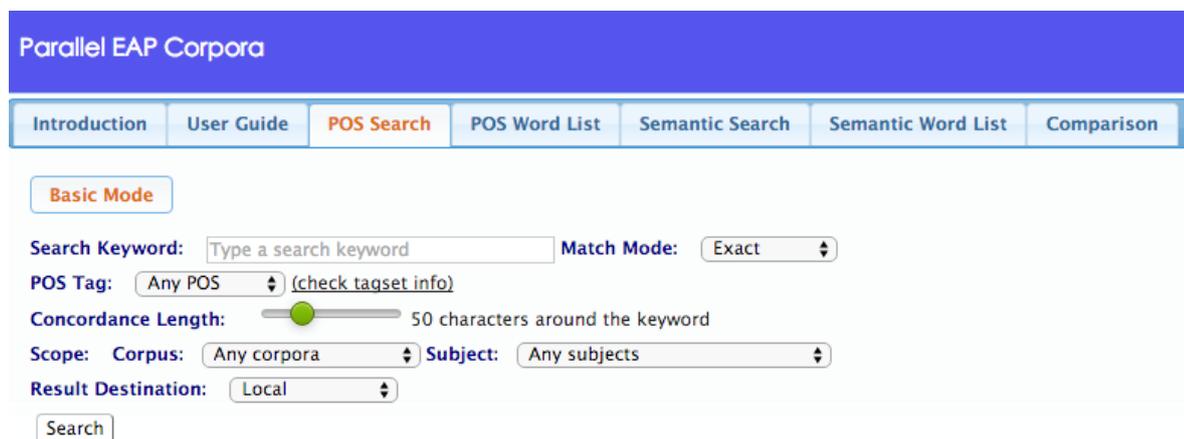


Figure 1: Interface of the Parallel EAP Corpora

1. Introduction: describe the nature of the parallel EAP Corpora
2. User Guide: contain the user manual on how to use the parallel EAP Corpora in textual description accompanied by a video
3. Part-of-speech (POS) Search: search the corpora based on POS tags
4. POS Word List: list the type of words (in different word classes) resulted from POS Search
5. Semantic Search: search the corpora based on semantic tags
6. Semantic Word List: list the type of words (in different semantic categories) resulted from Semantic Search
7. Comparison: juxtapose two different search results (i.e. POS vs. Semantic) for comparison

2. Basic Mode / Advanced Mode

2.1 POS and Semantic Simple Search

The POS Search tab and Semantic Search tab are shown below. This allows searches based on a number of search parameters. The simplest search is to enter a word in the Search Keyword input box and press ENTER, without modifying any other search parameters. For example, the word “study” is entered in the search box.

Search Keyword:

The **POS search** results are shown below.

Basic Mode

Search Keyword: study Match Mode: Exact

POS Tag: Any POS (check tagset info)

Concordance Length: 50 characters around the keyword

Scope: Corpus: Any corpora Subject: Any subjects

Result Destination: Local

Search Search for "study" with any tags, using "Exact" match mode on any subjects in any corpora

Tag Display Off

Show 10 entries Word Filter:

Word	Contents	Corpus	Subject	Genre
study	of data are chosen for the study: works from secondary school	Learner	Comparative Language Studies	Essays
study	meaning very effectively . The study could have been beneficial to	Learner	Comparative Language Studies	Essays
study	on how comparative languages study could inform teaching/ leaning	Learner	Comparative Language Studies	Essays
study	. This portfolio covers the study of the comparisons of the	Learner	Comparative Language Studies	Essays
study	Background of study	Learner	Comparative Language Studies	Essays
study	The aim of my study is to investigate how one	Learner	Comparative Language Studies	Essays
study	Linguistics is the scientific study of language . It contains a	Learner	Comparative Language Studies	Essays
study	implementing comparative language study in Hong Kong secondary	Learner	Comparative Language Studies	Essays
study	Comparative Language Study in Secondary Classroom	Learner	Comparative Language Studies	Essays
studv	teachina that comparative lanuaae study can contribute in enhancing the	Learner	Comparative Lanuaae Studies	Essavs

Figure 2: POS search results

The semantic search results are shown below.

Basic Mode

Search Keyword: study Match Mode: Exact

Semantic Tag: Any Main Tags (check tagset info) (check intro of tagset)

Concordance Length: 50 characters around the keyword

Scope: Corpus: Any corpora Subject: Any subjects

Result Destination: Local

Search Search for "study" with any tag, using "Exact" match mode on any subjects in any corpora

Tag Display Off

Show 10 entries Word Filter:

Word	Contents	Corpus	Subject	Genre
study	data are chosen for the study : works from secondary	Learner	Comparative Language Studies	Essays
study	very effectively . The study could have been beneficial to students	Learner	Comparative Language Studies	Essays
study	Background of study	Learner	Comparative Language Studies	Essays
study	@@The aim of my study is to investigate how one domain	Learner	Comparative Language Studies	Essays
study	Linguistics is the scientific study of language . It contains a	Learner	Comparative Language Studies	Essays
Study	Case Study	Learner	Comparative Language Studies	Essays
Study	maintained . The case study in the preceding part reports	Learner	Comparative Language Studies	Essays
Study	Case Study	Learner	Comparative Language Studies	Essays
study	In this case study , I will examine the reading	Learner	Comparative Language Studies	Essays

Figure 3: Semantic search results

There is a toggle button to switch between two modes, namely “Basic Mode” and “Advanced Mode”. Its label will also change according to the current status.

In POS Search Basic Mode, fewer search parameters are available and users are not required to specify these features in detail. Users will only need to select the parts of speech for the keyword in the search box via the POS Tag attribute.

In a similar vein, fewer search parameters are available in the Semantic Search Basic Mode and

users are not required to specify them in detail. Users will only need to select the main semantic labels for the keyword they would like to search in the search box via the Semantic Tag attribute (see below).

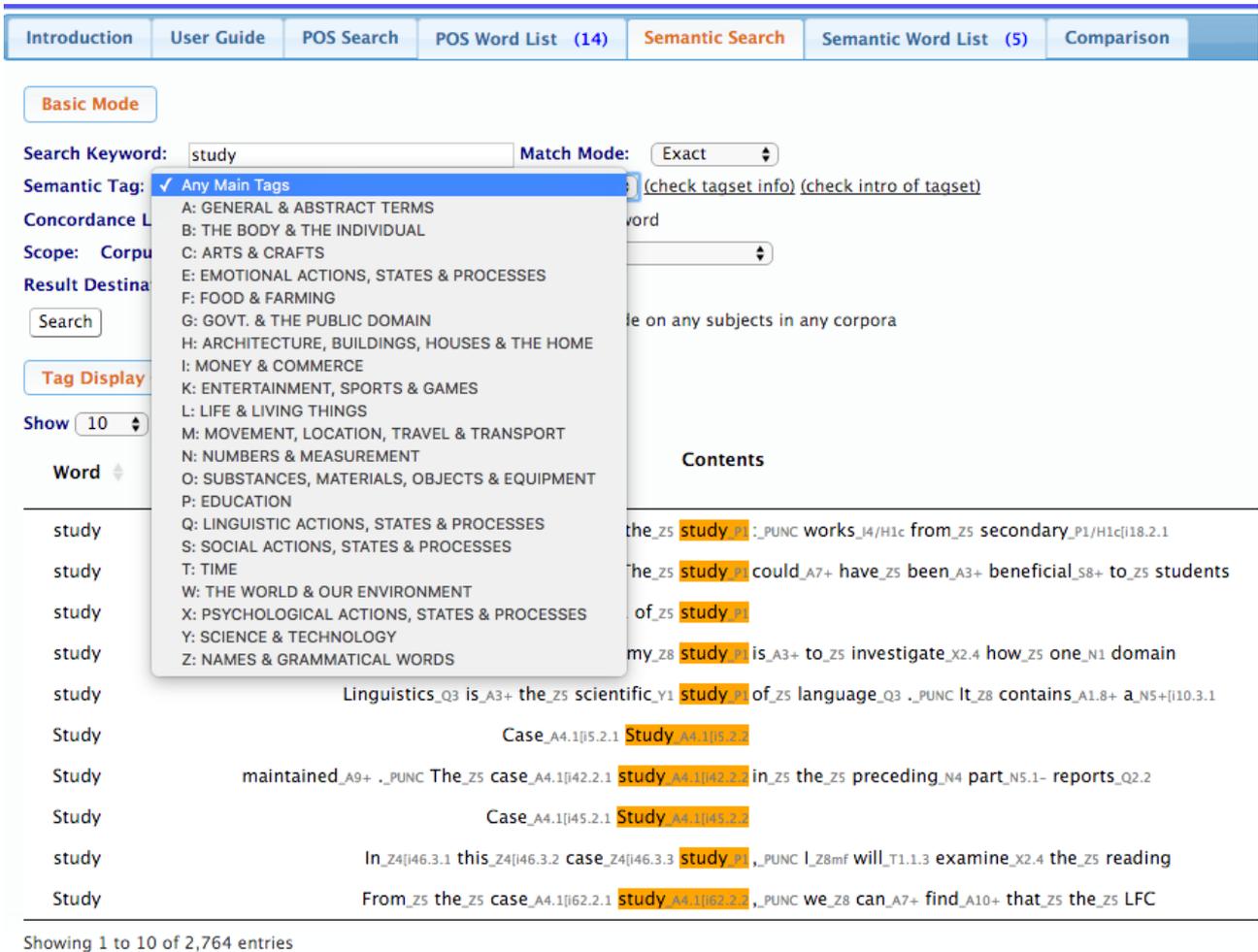


Figure 4: Semantic Tags in Basic Mode

2.2 POS Advanced Search

In Advanced Mode, in addition to the primary level of POS tags, more detailed categories are available, as shown in the figure below. Users may narrow down the search by selecting a secondary level of POS tags to search.

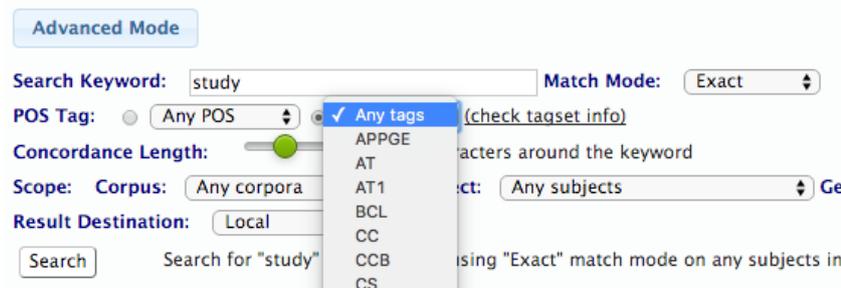
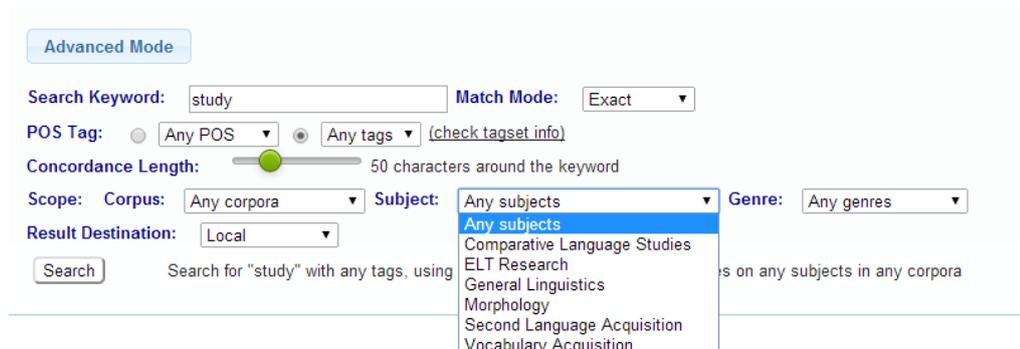


Figure 5: Secondary level tags for POS Search Advanced Mode

For detailed information on the POS tag sets and their meanings you may refer to this website <http://ucrel.lancs.ac.uk/claws7tags.html>.

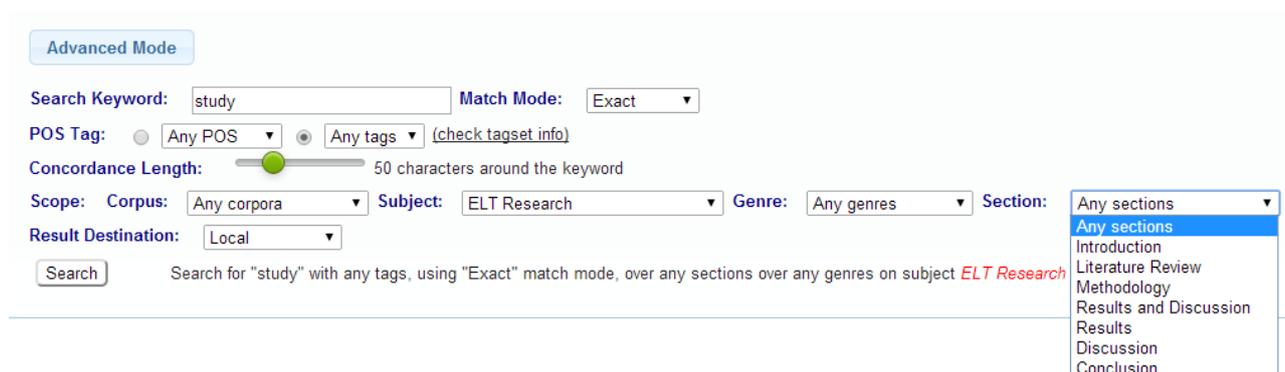
In Advanced Mode, users can choose to limit the scope of search by the “Subject” attribute (see below).



The screenshot shows the 'Advanced Mode' search interface. The search keyword is 'study' and the match mode is 'Exact'. The POS tag is set to 'Any tags'. The concordance length is 50 characters. The scope is set to 'Any corpora'. The 'Subject' dropdown menu is open, showing options: 'Any subjects', 'Comparative Language Studies', 'ELT Research', 'General Linguistics', 'Morphology', 'Second Language Acquisition', and 'Vocabulary Acquisition'. The 'Genre' is set to 'Any genres' and the result destination is 'Local'. A search button is visible.

Figure 6: The “Subject” attribute in POS Search Advanced Mode

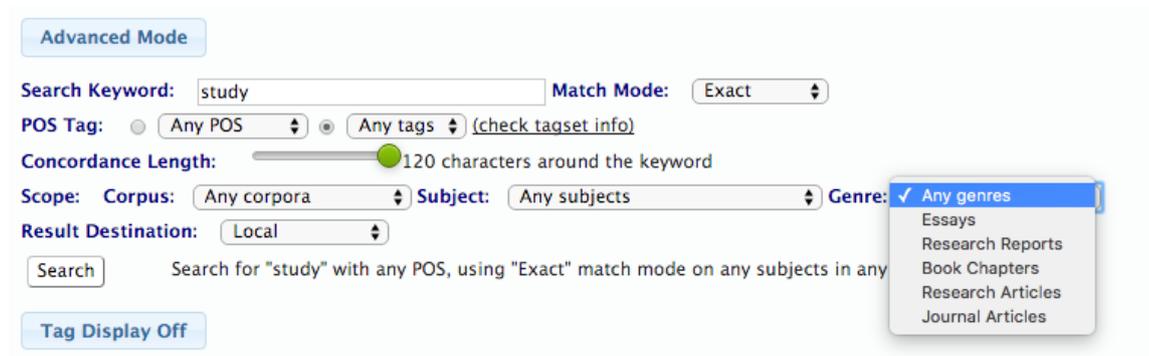
In Advanced Mode, if user selects “ELT Research” in the “Subject” attribute, the “Section” attribute will appear (see below).



The screenshot shows the 'Advanced Mode' search interface. The search keyword is 'study' and the match mode is 'Exact'. The POS tag is set to 'Any tags'. The concordance length is 50 characters. The scope is set to 'Any corpora'. The 'Subject' is set to 'ELT Research'. The 'Section' dropdown menu is open, showing options: 'Any sections', 'Introduction', 'Literature Review', 'Methodology', 'Results and Discussion', 'Results', 'Discussion', and 'Conclusion'. The 'Genre' is set to 'Any genres' and the result destination is 'Local'. A search button is visible.

Figure 7: The “Section” attribute in POS Search Advanced Mode

In Advanced Mode, users can choose to limit the scope of search by the “Genre” attribute (see below).



The screenshot shows the 'Advanced Mode' search interface. The search keyword is 'study' and the match mode is 'Exact'. The POS tag is set to 'Any tags'. The concordance length is 120 characters. The scope is set to 'Any corpora'. The 'Subject' is set to 'Any subjects'. The 'Genre' dropdown menu is open, showing options: 'Any genres', 'Essays', 'Research Reports', 'Book Chapters', 'Research Articles', and 'Journal Articles'. The result destination is 'Local'. A search button and a 'Tag Display Off' button are visible.

Figure 8: The “Genre” attribute in POS Search Advanced Mode

2.3 Semantic Advanced Search

Like the case of POS Search, only the primary level tags are available in Basic Mode, whereas in Advanced Mode, the secondary level tags are available for selection when any primary level tag is chosen (see below).

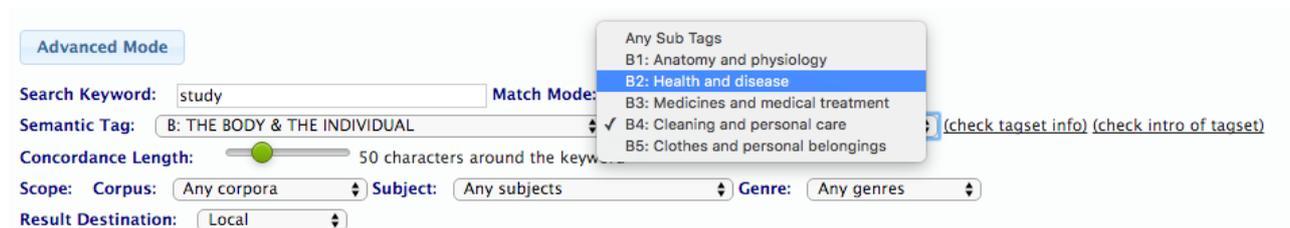


Figure 9: Secondary level tags for Semantic Search Advanced Mode

Apart from the Semantic Tag parameter, the search parameters of Semantic Search are the same as those of POS Search; the results are also displayed in a similar way (see below).

Word	Contents	Corpus	Subject	Genre
study	data_x2.2/x2.4 are_25 chosen_x7+ for_25 the_25 study :_PUNC works_j4/H1c from_25 secondary_p1/H1c 18.2.1	Learner	Comparative Language Studies	Essays
study	very_A13.3 effectively_A5.4+ _PUNC The_25 study could_A7+ have_25 been_A3+ beneficial_S8+ to_25 students	Learner	Comparative Language Studies	Essays
study	Background_T1.1.1 of_25 study	Learner	Comparative Language Studies	Essays
study	@@The_z99 aim_x7+ of_25 my_28 study is_A3+ to_25 investigate_x2.4 how_25 one_N1 domain	Learner	Comparative Language Studies	Essays
study	Linguistics_Q3 is_A3+ the_25 scientific_Y1 study of_25 language_Q3 _PUNC It_28 contains_A1.8+ a_N5+ 10.3.1	Learner	Comparative Language Studies	Essays
Study	Case_A4.1 5.2.1 Study_A4.1 5.2.1	Learner	Comparative Language Studies	Essays
Study	maintained_A9+ _PUNC The_25 case_A4.1 42.2.1 study_A4.1 42.2.1 in_25 the_25 preceding_N4 part_N5.1- reports_Q2.2	Learner	Comparative Language Studies	Essays
Study	Case_A4.1 45.2.1 Study_A4.1 45.2.1	Learner	Comparative Language Studies	Essays
study	In_Z4 46.3.1 this_Z4 46.3.2 case_Z4 46.3.3 study _PUNC I_Z8mf will_T1.1.3 examine_x2.4 the_25 reading	Learner	Comparative Language Studies	Essays
Study	From_25 the_25 case_A4.1 62.2.1 study_A4.1 62.2.1 _PUNC we_28 can_A7+ find_A10+ that_25 the_25 LFC	Learner	Comparative Language Studies	Essays

Figure 10: Search results in Semantic Search

For detailed information on the semantic tag sets and their meanings you may refer to this website <http://ucrel.lancs.ac.uk/usas/>.

3. Search Parameters in POS and Semantic Search

3.1 POS Tag and Semantic Tag

The POS Tag can be specified to narrow down the search results. The list of possible values are shown in the following figure.

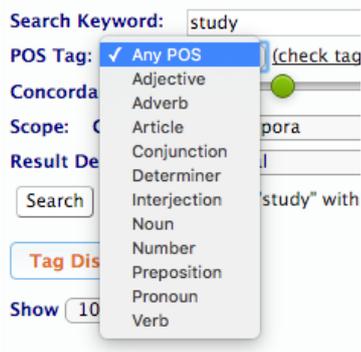


Figure 11: POS Tag

For example, only “Noun” will be included in the search results with this setting (see below):



Figure 12: Selecting “Noun” POS Tag to search the word “study”

Results: The tag display is turned on to show that the matched keywords have the POS Tag NN1.

first_MD language_NN1 . In_II this_DD1 study_NN1 the_AT Hong_NP1 Kong_NP1 EFL_NP1 learner_NN1
 This_DD1 study_NN1 suggests_VVZ that_CST the_AT nuclear_JJ
 1989_MC an_AT1 official_JJ large-scale_JJ study_NN1 of_JO the_AT long_JJ term_NN1 nature_NN1 of_JO
 . These_DD2 activities_NN2 are_VBR used_VVN to_TO study_VV whether_CSW31 or_CSW32 not_CSW33 NETs_NN2
 The_AT findings_NN2 of_JO the_AT present_JJ study_NN1 indicate_VV0 that_CST primary_JJ students_NN2 '
 the_AT outcome_NN1 of_JO the_AT present_JJ study_NN1 proves_VVZ that_CST the_AT NET_NN1 's_GE
 a_RR21 lot_RR22 with_IW her_APPGE liberal_JJ study_NN1 . She_PPHS1 pointed_VVD out_RP that_CST the_AT
 ?_? The_AT purpose_NN1 of_JO the_AT present_JJ study_NN1 is_VB2 to_TO investigate_VVI the_AT above_JJ
 abilities_NN2 participated_VVD in_II the_AT study_NN1 . The_AT participants_NN2 's_GE ages_NN2
 . Considering_CS21 that_CS22 the_AT current_JJ study_NN1 " effects_NN2 of_JO Big_JJ Book_NN1

Figure 13: POS Tag Search Results

The Semantic Tag can be specified to narrow down the search results. The list of possible values is shown in the following figure.

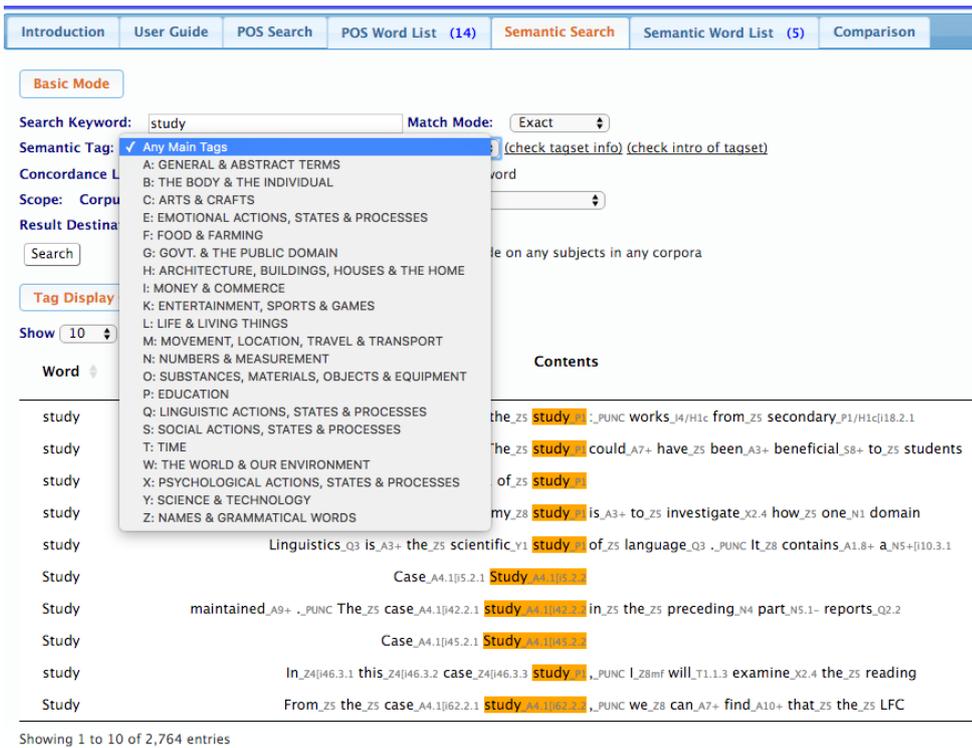


Figure 14: Semantic Tag

3.2 Match Mode

When the default match mode “Exact” is selected, the exactly matched word (“study” in this case) will be searched and included in the results (see below).

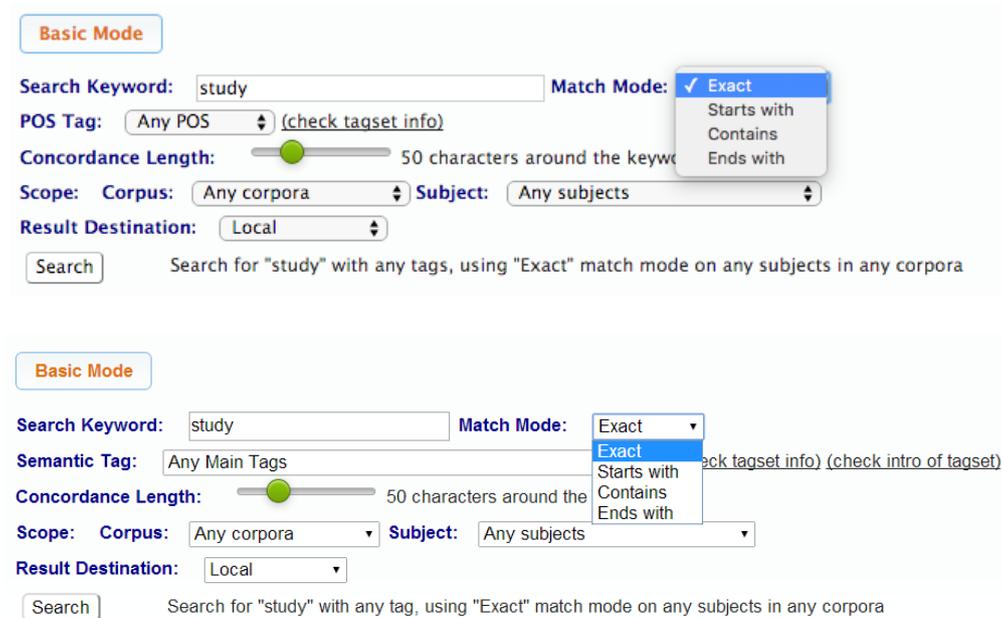


Figure 15 and 16: Respectively show the match mode function in both POS and Semantic Search

The match mode parameter allows you to adjust the mode of matching. There are four matching

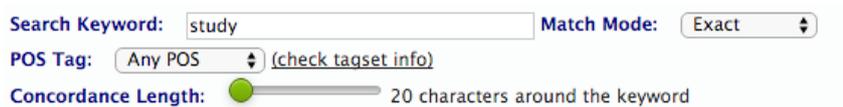
modes, namely “Exact”, “Starts with”, “Contains” and “Ends with”:

1. “Exact” is the default mode of search. Search results using this matching mode will include lexical items exactly the same as the keyword, i.e. “study”.
2. For “Starts with”, any lexical items starting with the keyword entered will be identified, like “studying”, “studied”, etc.
3. “Contains” mode will search for any lexical items that contain the keyword entered, like “case-study”, “self-studying”, etc.
4. “Ends with”: any lexical items ending with the keyword entered will be identified, like “case-study”, “meta-study”, but not “studying” nor “study-leave”.

Naturally the “Contains” match mode should provide the most number of search results among all four modes.

3.3 Concordance Length

This parameter is used to control the length of contents to be displayed in the results. Users can slide the circle to change the desired value. The minimum length is 20 characters (see below).



The screenshot shows a search interface with the following elements: a search box containing the text 'study', a 'Match Mode' dropdown menu set to 'Exact', a 'POS Tag' dropdown menu set to 'Any POS' with a '(check tagset info)' link, and a 'Concordance Length' slider. The slider is positioned at the far left, with a green circle indicating the current value of 20 characters around the keyword.

Figure 17: Concordance Length (Minimum Length)

The length of the contents displayed will be shorter (see below).

for the study: works
. The study could have
languages study could inform
covers the study of the

Figure 18: The results displayed shows 20 characters of the sentence containing the keyword

The maximum allowed is 120 characters (see below).



The screenshot shows the same search interface as Figure 17, but the 'Concordance Length' slider is now positioned further to the right, with a green circle indicating the current value of 120 characters around the keyword.

Figure 19: Concordance Length (Maximum Length)

The larger contexts of the results can also be examined (see below).

Chinese Learner English Corpus . 3 sources of data are chosen for the **study**: works from secondary school students , university English majors in 1st or 2nd year and English are capable of conveying the meaning very effectively . The **study** could have been beneficial to students learning both languages since they will follow and at last a reflection on how comparative languages **study** could inform teaching/ leaning in secondary classrooms will be given .

using the techniques of parallel concordance . This portfolio covers the **study** of the comparisons of the morphology , syntax , semantics and discourse in

Figure 20: The results displayed shows 120 characters of the sentence containing the keyword

3.4 Scope

By default, all texts in the corpus will be included in the POS and Semantic Search. Users, however, can choose to limit the scope of search by the corpus type (“learner” or “professional”), like below:

Search Keyword: study Match Mode: Exact

POS Tag: Any POS (check tagset info)

Concordance Length: 120 characters around the keyword

Scope: Corpus: Any corpora (dropdown menu open showing Learner Corpus, Professional Corpus) Subject: Any subjects

Result Destination: Local

Search for "study" with any POS, using "Exact" match mode on any subjects in any corpora

Search Keyword: study Match Mode: Exact

Semantic Tag: Any Main Tags (check tagset info) (check intro of tagset)

Concordance Length: 50 characters around the keyword

Scope: Corpus: Any corpora (dropdown menu open showing Any corpora, Learner Corpus, Professional Corpus) Subject: Any subjects

Result Destination: Local

Search for "study" with any tag, using "Exact" match mode on any subjects in any corpora

Figure 21 and 22: Respectively showing the “Corpus” type in POS and Semantic Search

Users can also limit the scope by subject in POS and Semantic Search (see below):

Search Keyword: study Match Mode: Exact

POS Tag: Any POS (check tagset info)

Concordance Length: 120 characters around the keyword

Scope: Corpus: Any corpora Subject: Any subjects (dropdown menu open showing Comparative Language Studies, ELT Research, General Linguistics, Morphology, Second Language Acquisition, Vocabulary Acquisition)

Result Destination: Local

Search for "study" with any POS, using "Exact" match mode on any subjects in any corpora

Tag Display Off

Search Keyword: study Match Mode: Exact

Semantic Tag: Any Main Tags (check tagset info) (check intro of tagset)

Concordance Length: 50 characters around the keyword

Scope: Corpus: Any corpora Subject: Any subjects (dropdown menu open showing Comparative Language Studies, ELT Research, General Linguistics, Morphology, Second Language Acquisition, Vocabulary Acquisition)

Result Destination: Local

Search for "study" with any tag, using "Exact" match mode on any subjects in any corpora

Tag Display Off

Figure 23 and 24: Respectively showing the “Subject” attribute in POS and Semantic Search

If user selects “ELT Research” in the “Subject” attribute, the “Section” attribute will appear in Advanced Mode (see below):

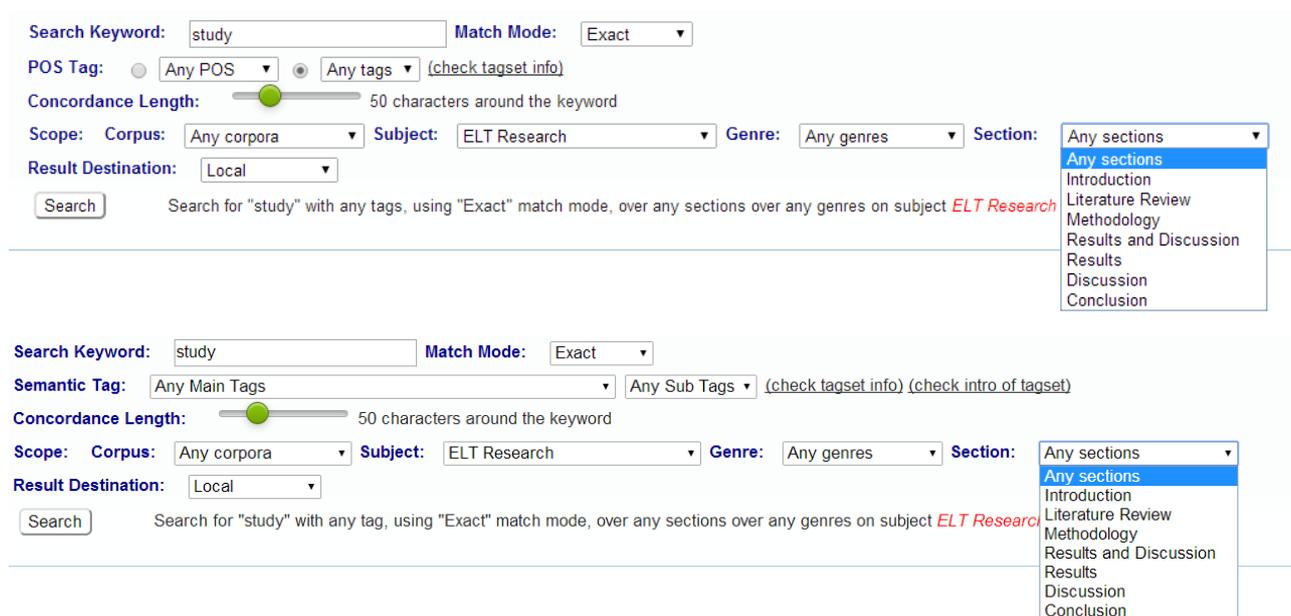


Figure 25 and 26: Respectively the “Section” attribute is displayed in POS and Semantic Search

4. Interpreting Search Results in POS and Semantic Search

The search results are tabulated for easy navigation with a number of columns indicating their different attributes. The results are based on sentence units, which means if a sentence has more than one match, the sentence will only appear once in the result section (see below).

Word	Contents	Corpus	Subject	Word Filter
study	of data are chosen for the study : works from secondary school	Learner	Comparative Language Studies	'study
study	meaning very effectively . The study could have been beneficial to	Learner	Comparative Language Studies	meta-study
study	on how comparative languages study could inform teaching/ leaning	Learner	Comparative Language Studies	self-study
study	. This portfolio covers the study of the comparisons of the	Learner	Comparative Language Studies	study
study	Background of study	Learner	Comparative Language Studies	three-study
study	The aim of my study is to investigate how one	Learner	Comparative Language Studies	Essays
study	Linguistics is the scientific study of language . It contains a	Learner	Comparative Language Studies	Essays
study	implementing comparative language study in Hong Kong secondary	Learner	Comparative Language Studies	Essays
study	Comparative Language Study in Secondary Classroom	Learner	Comparative Language Studies	Essays
study	teaching that comparative language study can contribute in enhancing the	Learner	Comparative Language Studies	Essays

Showing 1 to 10 of 2,775 entries

1. Word column: indicates the matched word
2. Contents column: shows the sentence containing the matched word. The length of the sentence fragment shown depends on the “Concordance Length” parameter in the search

- form
3. Corpus column: shows from which corpus this result comes from: “Learner” or “Professional”
 4. Subject column: shows the subject to which this result belongs
 5. Genre column: shows the genre of the searched text
 6. Word Filter (a selection list at top right corner): displays only the results of the selected word among the words in the Word Filter list
 7. Show n entries (at top left corner): changes the number of entries on one page. There are choices of 10, 25, 50, 100 entries on one page
 8. Tag Display switch (a button at top left corner): it enables users to switch between two modes whether to show the tags of each word or not. The default is off. When this is switched on, the button will become “Tag Display On”, and the results will be changed accordingly (see below):

Tag Display On

Show 10 entries Word Filter: self-study

Word	Contents	Corpus	Subject	Genre
self-study	are_VBR claimed_VVN suitable_JJ both_RR_for_IF self-study_NN1 and_CC instructed_VVN study_NN1 , however_RR ,	Learner	Vocabulary Acquisition	Essays
self-study	words_NN2 and_CC continuous_JJ interest_NN1 in_I self-study_NN1 .	Learner	Vocabulary Acquisition	Essays
self-study	A_AT1 taskforce_NN1 was_VBD2 set_VVN up_RP_to TO study foreign_JJ teaching_NN1 methods_NN2 , and_CC	Professional	General Linguistics	Book Chapters
self-study	, Murphy_NP1 's_GE (1985_MC) self-study book_NN1 , English_JJ Grammar_NN1 in_I Use_NN1	Professional	General Linguistics	Book Chapters
self-study	activities_NN2 , the_AT record_NN1 of_IO self-study_NN1 on_I the_AT web_NN1 , interviews_NN2 or_CC	Professional	Vocabulary Acquisition	Journal Articles
self-study	, on_I the_AT bus_NN1 , in_I the_AT self-study room_NN1 etc_RA ,) , " : When_RRQ do_VDO	Professional	Second Language Acquisition	Journal Articles
self-study	knowledge_NN1 and_CC the_AT ability_NN1 of_IO self-study_NN1 in_I the_AT process_NN1 of_IO doing_VDG	Learner	ELT Research	Research Reports
self-study	'_GE " : self-awareness_NN1 " : and_CC " : self-study_NN1 " : (Jun_NN1 , n.d._NNU) as_CSA	Learner	ELT Research	Research Reports

Showing 1 to 8 of 8 entries (filtered from 2,775 total entries) Previous 1 Next

Figure 27: Results generated when Tag Display is On

When a concordance line is clicked, a window will appear showing the whole essay, with the currently matched sentence highlighted (see below):

Basic Mode

Search Keyword: study Match Mode: Ends with

POS Tag: ILS 04t.txt

Concordance

Scope: Cor

Result Desti

Search

Tag Display

Show 10

Word

self-study

self-study

self-study

self-study

self-study

self-study

self-study

self-study

self-study

Showing 1 to

Show/hide Tags Close

1 Next

The Sino-Soviet schism in the early 1960s resulted in China being more open to western ideas and collaboration , as the nation pursued economic rather than political goals . The Chinese authorities decided that economic modernization would require oral competence in the workforce as well as strong reading skills . A taskforce was set up to study foreign teaching methods , and this investigation led to experiments with the audiolingual method , which had been introduced to China by teachers from overseas . At that time , the audiolingual method had established a significant position in the international market , as it was easily packaged into commercial courses for self-study or classes , and its use of modern technology was appealing to teachers and learners who wished to feel that they were at the cutting edge . The method had grown in the USA from the need in World War II to train personnel rapidly in order to cope with military situations that would require foreign languages , such as questioning captured soldiers or administering displaced persons . The language drills developed for this purpose were then allied to Skinner 's work on Behaviorism with animals . Once again , the choice of this method was seen as appropriate for China as learning through repetition was similar to common approaches to learning Chinese characters .

Figure 28: The highlighted text of the entire paragraph containing the keyword

The “Show / hide Tags” button at the bottom functions similarly and will toggle the display of tags for the essay. Clicking the “Close” button will dismiss the essay display.

5. POS Word List

After every search, a word list will be generated consisting of all matched words according to the search parameters. For example, a search of “study” with “Contains” match mode will result in the following word list:

Please click any entry to check its collocation details

Show entries

Text	POS Tag	Frequency	Freq Per Million
study	NN1 [?]	2155	1,812.257
studying	VVG [?]	172	144.644
study	VVI [?]	71	59.708
study	VVO [?]	29	24.388
self-study	NN1 [?]	3	2.523
case-study	JJ [?]	2	1.682
self-study	JJ [?]	2	1.682
study-in	NN1 [?]	1	0.841
study-the	AT [?]	1	0.841
three-study	JJ [?]	1	0.841
study'	NN1 [?]	1	0.841
study-wise	JJ [?]	1	0.841
meta-study	NN1 [?]	1	0.841
'study	NN1 [?]	1	0.841

Showing 1 to 14 of 14 entries Previous **1** Next

Figure 29: POS Word List generated for the word “study” with match mode “Contains”

The list shows the matched word, its POS tag, its frequency in the corpus, and the adjusted frequency per million for cross-corpus comparison. The description of each POS Tag can be examined by clicking the “[?]” hyperlink beside each POS Tag. For example, clicking the link beside the NN1 tag will show the following popup message:

The page at localhost says:
singular common noun (e.g. book, girl)

Prevent this page from creating additional dialogs.

OK

Text	POS Tag	Frequency	Freq Per Million
study	NN1 [?]		
studying	VVG [?]		
study	VVI [?]	71	59.708
study	VVO [?]	29	24.388
self-study	NN1 [?]	3	2.523
case-study	JJ [?]	2	1.682
self-study	JJ [?]	2	1.682
study-in	NN1 [?]	1	0.841
study-the	AT [?]	1	0.841
three-study	JJ [?]	1	0.841

Figure 30: The explanation of the POS Tag

5.1 Collocation in POS Search

From the word list generated above, clicking any word will generate another list showing collocation details of the word. For example, clicking “studying” with POS VVG will result in the following collocation details, shown below in the word list table.

Collocation of studying_VVG

Please click any collocated word to display the associated sentence

Tag Display Off

Show 10 entries

-5	-4	-3	-2	-1	Word	+1	+2	+3	+4	+5
of (11)	What (10)	they (8)	did (6)	students (19)	studying (172)	English (26)	and (11)	the (13)	of (9)	the (11)
about (5)	of (8)	of (8)	to (5)	in (16)		the (21)	language (11)	language (9)	the (4)	language (6)
in (4)	the (7)	I (6)	of (5)	for (13)		in (13)	in (9)	a (7)	foreign (4)	a (5)
sure (4)	and (6)	and (5)	do (4)	and (8)		a (12)	the (8)	of (5)	they (3)	and (5)
a (3)	students (4)	the (5)	they (4)	were (7)		French (5)	for (5)	is (4)	to (3)	at (3)
for (3)	a (3)	a (3)	strategies (4)	learners (6)		and (3)	at (5)	for (3)	in (3)	are (3)
was (3)	their (3)	students (3)	and (4)	when (6)		abroad (3)	an (4)	English-speaking (3)	it (2)	it (3)
and (2)	me (2)	reasons (2)	for (3)	are (5)		English (3)	is (4)	college (3)	all (2)	of (3)
have (2)	to (2)	group (2)	than (3)	been (4)		Japanese (3)	foreign (4)	that (2)	and (2)	words (2)
styles (2)	do (2)	who (2)	university (3)	prefer (4)		with (3)	others (3)	to (2)	was (2)	classes (2)

Showing 1 to 10 of 114 entries

Previous 1 2 3 4 5 ... 12 Next

Figure 31: Collocation results generated for the word “studying” with POS VVG

The collocation results show the preceding and following 5 words of the chosen word, sorted by descending frequency (indicated in the parentheses) at each position.

5.2 Collocation Context in POS Search

Clicking any collocated word in the above table will further show the actual context where the chosen word and collocated words occur; see the example below:

Collocation Context of studying_VVG and English_NN1 at position 1

Tag Display Off

Show 10 entries

Corpus	Subject	Genre	Contents
professional	ELT Research	Research Articles	This is the background to an initiative by researchers at the University of Michigan, who aim to establish a dataset of frequent formulaic chunks that are of use to those studying English for academic purposes. One of the problems in constructing such a resource lies in defining what is or is not recognized as a formulaic chunk. Ellis, Simpson-Vlach, and Maynard (this issue) investigate three factors which cause a language user to regard a chunk as a linguistic unit. They are its length, the cumulative frequency of the components of the chunk, and mutual information (MI), the extent to which the components of the chunk co-occur across the corpus in question. The writers report that the most important factor for L1 users proved to be MI but that for L2 users tested, it was cumulative frequency. This result suggests that, even at quite an advanced level, L2 users continue to process the formulaic chunk as if it were a set of independent words.
professional	ELT Research	Research Articles	Natasha also displayed numerous transfer errors that stem from differences between the Russian and English morphosyntactic systems. Among these errors were (a) omission of subject pronouns, which is acceptable in Russian but not in English ; for example, in a comment about her parents she said, "Know no English " (interrogation transcript, p. 2); (b) erroneous tense assignment: Because Russian has a single present tense, she substituted present simple for present progressive and stated that her sister "learns English at school" (p. 3) instead of studying English ; (c) omission and incorrect assignment of articles because Russian has no articles, for example, "I never had the dog" (p. 14); "there was a police" (p. 31); and "I have a Adidas shoes" (p. 33); (d) omission and incorrect assignment of prepositions, driven by Russian preposition usage, for example, "pays more attention on chemistry" (p. 3); "I went on the taxi" (p. 10); "I was waiting him at his car" (p. 28); "and he just says me" (p. 29). She also exhibited difficulties with negation, seen in statements such as "I'm very not about the law part. I mean it 's not know" (p. 15).
professional	ELT Research	Research Articles	A cross-sectional study was designed, with three groups of participants: two groups of French learners of English and a control group of L1 English (L1Eng) speakers. French learners of English were chosen because French is an alphabetic language, like English : In both these languages readers store the immediate products of decoding in the phonological loop, rather than in the visuospatial sketchpad. The L1 French learners of English were the members of one class each from the middle and upper state secondary schools of a provincial French town; they were from monolingual French families and were accustomed to southern British English accents in their classrooms. The poor L2 comprehender (PoorC) group comprised 21 adolescents: 13 girls and 8 boys with a mean age of 14 years, 7 months, in their fourth year of studying English .

Figure 32: More contexts and information shown for the collocated words

6. Semantic Word List

After every search, a word list will be generated consisting of all matched words according to the search parameters. For example, a search of “study” with “Exact” match mode will result in the following word list:

Text	Semantic Tag	Frequency	Freq Per Million
study	P1 [?]	3690	1,712.173
study	Z2 [?]	58	26.912
Study	A4 [?]	46	21.344
study	H2 [?]	2	0.928
study	Q4 [?]	1	0.464

Showing 1 to 5 of 5 entries

Figure 33: Semantic Word List generated for the word “study” with match mode “Exact”

The list shows the matched word, its Semantic tag, its frequency in the corpus, and the adjusted frequency per million for cross-corpus comparison. The description of each Semantic Tag can be examined by clicking the “[?]” hyperlink beside each tag. For example, clicking the link beside the Z2 tag will show the following popup message:

Text	Semantic Tag	Frequency	Freq Per Million
study	P1 [?]	3690	1,712.173
study	Z2 [?]	58	26.912
Study	A4 [?]	46	21.344
study	H2 [?]	2	0.928
study	Q4 [?]	1	0.464

The page at corpus.ied.edu.hk says:

Geographical names

Figure 34: The explanation of the Semantic Tag

6.1 Collocation in Semantic Search

From the word list generated above, clicking any word will generate another list showing collocation details of the word. For example, clicking study with Semantic Tag Z2 will result in the following collocation details; see below in the word list table.

Collocation of study_Z2

Please click any collocated word to display the associated sentence

Tag Display On

Show 10 entries

-5	-4	-3	-2	-1	Word	+1	+2	+3	+4	+5
I_Z8 (5)	I_Z8 (4)	I_Z8 (8)	began_T2 (4)	to_Z5 (42)	study_Z2 (58)	English_Z2 (56)	in_Z5 (5)	the_Z5 (6)	I_Z8 (5)	a_Z5 (2)
K_Z5 (2)	motivation_A2 (2)	the_Z5 (3)	motivation_A2 (3)	I_Z8 (5)		CA_Z2 (1)	words_Q3 (4)	my_Z8 (2)	Instrumental_A11 (3)	remember_X2 (2)
of_Z5 (2)	students_P1 (2)	to_Z5 (2)	me_Z8 (3)	us_Z8 (2)		LLS_Z2 (1)	vocabulary_Q3 (3)	in_Z5 (2)	to_Z5 (3)	still_T2 (1)
school_P1 (2)	not_Z6 (2)	my_Z8 (2)	like_E2 (2)	me_Z8 (2)			for_Z5 (3)	because_A2 (2)	the_Z5 (2)	years_T1 (1)
school_H1 (2)	and_Z5 (2)	study_P1 (2)	K_Z5 (2)	not_Z6 (2)			is_A3 (2)	more_A13 (2)	kept_A9 (1)	not_Z6 (1)
my_Z8 (1)	students_S2 (2)	are_Z5 (2)	motivated_A2 (2)	this_M6 (1)			from_Z5 (2)	not_Z6 (2)	efficiently_X9 (1)	rewards_S1 (1)
that_Z8 (1)	the_Z5 (2)	everyone_N5 (1)	required_X7 (2)	begun_T2 (1)			was_A3 (2)	it_Z8 (2)	prefer_E2 (1)	middle_P1 (1)
some_N5 (1)	thus_Z5 (1)	why_A2 (1)	help_S8 (2)	can_A7 (1)			language_Q3 (2)	a_Z5 (2)	classmates_P1 (1)	According_Z5 (1)
form_A4 (1)	very_A13 (1)	necessary_S6 (1)	hard_O4 (1)	to_S6 (1)			As_Z5 (2)	rarely_N6 (1)	2_A4 (1)	please_E4 (1)
generally_A6 (1)	before_Z5 (1)	participants_S1 (1)	in_Z5 (1)	we_Z8 (1)			just_A14 (2)	men_S2 (1)	a_Z5 (1)	punished_G2 (1)

Showing 1 to 10 of 50 entries

Previous 1 2 3 4 5 Next

Figure 35: Collocation results generated for the word “study” with Semantic Tag Z2

The collocation results show the preceding and following 5 words of the chosen word, sorted by descending frequency (indicated in the parentheses) at each position.

6.2 Collocation Context in Semantic Search

Clicking any collocated word in the above table will further show the actual context where the chosen word and collocated word occur; see an example below:

Collocation Context of study_Z2 and vocabulary_Q3 at position 2

Tag Display Off

Show 10 entries

Corpus	Subject	Genre	Contents
learner	Vocabulary Acquisition	Essays	The study of lexical change and lexical variation is important in the development of the English language and the process of learning and teaching English vocabulary as well . Teaching the knowledge of lexical change and lexical variation helps students have a better understanding of the development of lexis and inspires them to study English vocabulary more consciously and effectively .
learner	ELT Research	Research Reports	This essay aims at study how English language learners learn English vocabulary . The author has interviewed some English language learners , which include English major and non-English major students . The content of the interview include how they learn new English words and the strategies they used to remember English vocabulary . From the interview the author got the result that there are many different ways learners used to learn and remember English vocabulary , among which use a context or a sentence is the most popular method . And then the author offer some suggestions of how to study English vocabulary more efficiently according to what has learned from the course-Lexis , Morphology and Semantics , that is we can use the word 's root to remember some words and also the synonyms and antonyms , etc . Since many English words have connection with others , using these rules may help us study English vocabulary much easier .

Showing 1 to 2 of 2 entries

Previous 1 Next

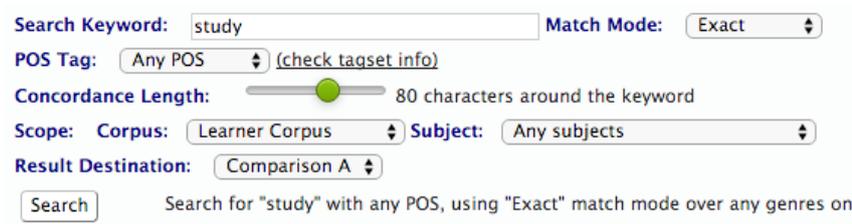
Figure 36: More contexts and information shown for the collocated words

7. Comparison of Results in POS Search

Sometimes users may need to juxtapose two different search results for easier comparison. For example, the exact match of “study” and “studying” results may be compared. One way to do this is to perform one search after another. This comparison feature makes it easy to do comparisons

among different search modes. The user can change the search parameters to perform the two searches one after one, and then choose the “Result Destination” to display the results on the Comparison tab.

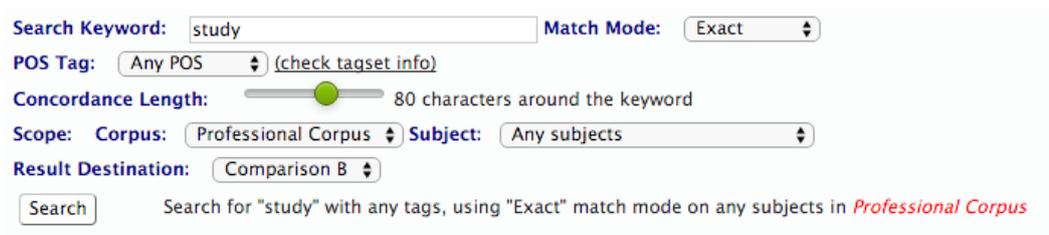
For example, by choosing the “study” with exact match from the “Learner Corpus” and placing the result at “Comparison A”, the result is shown below:



The screenshot shows a search interface with the following parameters: Search Keyword: study; Match Mode: Exact; POS Tag: Any POS; Concordance Length: 80 characters around the keyword; Scope: Corpus: Learner Corpus; Subject: Any subjects; Result Destination: Comparison A. A Search button is present, and the search description reads: Search for "study" with any POS, using "Exact" match mode over any genres on

Figure 37: Display search result “study” from Learner Corpus in Comparison A

Then choosing the “study” with exact match from “Professional Corpus” and placing the result at “Comparison B”, the result is shown below:



The screenshot shows a search interface with the following parameters: Search Keyword: study; Match Mode: Exact; POS Tag: Any POS; Concordance Length: 80 characters around the keyword; Scope: Corpus: Professional Corpus; Subject: Any subjects; Result Destination: Comparison B. A Search button is present, and the search description reads: Search for "study" with any tags, using "Exact" match mode on any subjects in Professional Corpus

Figure 38: Display search result “study” from Professional Corpus in Comparison B

On the Comparison tab, the first and second search results will be shown on the same page, as below.

Introduction User Guide POS Search POS Word List (3) Semantic Search Semantic Word List Comparison

Comparison A

POS Search: Search for "study" with any POS, using "Exact" match mode over any genres on any subjects in *Learner Corpus*

Tag Display Off

Show 10 entries

Word Filter:

Word	Contents	Corpus	Subject	Genre
study	Corpus . 3 sources of data are chosen for the study : works from secondary school students , university	Learner	Comparative Language Studies	Essays
study	of conveying the meaning very effectively . The study could have been beneficial to students learning	Learner	Comparative Language Studies	Essays
study	last a reflection on how comparative languages study could inform teaching/ leaning in secondary classrooms	Learner	Comparative Language Studies	Essays
study	of parallel concordance . This portfolio covers the study of the comparisons of the morphology , syntax ,	Learner	Comparative Language Studies	Essays
study	Background of study	Learner	Comparative Language Studies	Essays
study	The aim of my study is to investigate how one domain of syntax ,	Learner	Comparative Language Studies	Essays
study	Linguistics is the scientific study of language . It contains a number of	Learner	Comparative Language Studies	Essays
study	is a section about implementing comparative language study in Hong Kong secondary classroom .	Learner	Comparative Language Studies	Essays
study	Comparative Language Study in Secondary Classroom	Learner	Comparative Language Studies	Essays
study	a future English teaching that comparative language study can contribute in enhancing the learning of grammar	Learner	Comparative Language Studies	Essays

Showing 1 to 10 of 1,185 entries

Previous 1 2 3 4 5 ... 119 Next

Comparison B

POS Search: Search for "studying" with any POS, using "Exact" match mode over any genres on any subjects in any corpora

Tag Display Off

Show 10 entries

Word Filter:

Word	Contents	Corpus	Subject	Genre
studying	paper were written by some Form 1 students studying in a band 1 EMI (English as Medium of	Learner	Comparative Language Studies	Essays
studying	some complicated dative alternations by means of studying syntax . However , it does n't mean that	Learner	Comparative Language Studies	Essays
studying	to become a competent intra-cultural communicator , studying linguistic features is important but not all ,	Learner	Comparative Language Studies	Essays
studying	impoliteness in cross-cultural communication . Thus , studying politeness in the communication of a cultural group	Learner	Comparative Language Studies	Essays
studying	impoliteness in cross-cultural communication . Thus , studying politeness in the communication of a cultural group	Learner	Comparative Language Studies	Essays

© Copyright 2014. All rights reserved.

Figure 39: A side by side comparison on the word “study” from POS Search

8. Comparison of Results between POS and Semantic Search

The comparison tabs not only allow comparison between different searches from POS Search, but also enable comparison between POS Search and Semantic Search, or actually any combination. Choosing the “Result Destination” will display the results accordingly in the selected location in the Comparison tab.

The example below shows POS Search in “Comparison A” and Semantic Search in “Comparison B”.

Introduction User Guide POS Search POS Word List (3) Semantic Search Semantic Word List (5) Comparison

Comparison A

POS Search: Search for "study" with any tags, using "Exact" match mode on any subjects in any corpora

Show 10 entries Word Filter:

Word	Contents	Corpus	Subject	Genre
study	of_JO data_NN are_VBR chosen_VVN for_IF the_AT study_NN1 works_NN from_JJ secondary_JJ school_NN1	Learner	Comparative Language Studies	Essays
study	meaning_NN1 very_RG effectively_RR ... The_AT study_NN1 could_VM have_VHI been_VBN beneficial_JJ to_JJ	Learner	Comparative Language Studies	Essays
study	on_JJ how_RGQ comparative_JJ languages_NN2 study_NN1 could_VM inform_VVI teaching_JJ leaning_VVG	Learner	Comparative Language Studies	Essays
study	... This_OD1 portfolio_NN1 covers_VVZ the_AT study_NN1 of_JO the_AT comparisons_NN2 of_JO the_AT	Learner	Comparative Language Studies	Essays
study	Background_NN1 of_JO study_NN1	Learner	Comparative Language Studies	Essays
study	The_AT aim_NN1 of_JO my_APPGE study_NN1 is_VBZ to_TO investigate_VVI how_RRQ one_MC1	Learner	Comparative Language Studies	Essays
study	Linguistics_NN1 is_VBZ the_AT scientific_JJ study_NN1 of_JO language_NN1 ... It_PPH1 contains_VVZ a_AT1	Learner	Comparative Language Studies	Essays
study	implementing_VVG comparative_JJ language_NN1 study_NN1 in_JJ Hong_NP1 Kong_NP1 secondary_JJ	Learner	Comparative Language Studies	Essays
study	Comparative_JJ Language_NN1 Study_NN1 in_JJ Secondary_JJ Classroom_NN1	Learner	Comparative Language Studies	Essays

Comparison B

Semantic Search: Search for "study" with any tag, using "Exact" match mode on any subjects in any corpora

Tag Display On

Show 10 entries Word Filter:

Word	Contents	Corpus	Subject	Genre
study	data_X2.2)X2.4 are_ZS chosen_X7+ for_ZS the_ZS study_NN1 ...PUNC works_J4)H1c from_ZS secondary_P1)H1c)18.2.1	Learner	Comparative Language Studies	Essays

Figure 40: A side by side comparison on the word “study” between POS Search and Semantic Search