The EMI Corpus of Student Academic Writing: A new resource in EAP materials design

Dana Gablasova, Raffaella Bottini, Vaclav Brezina and Luke Harding
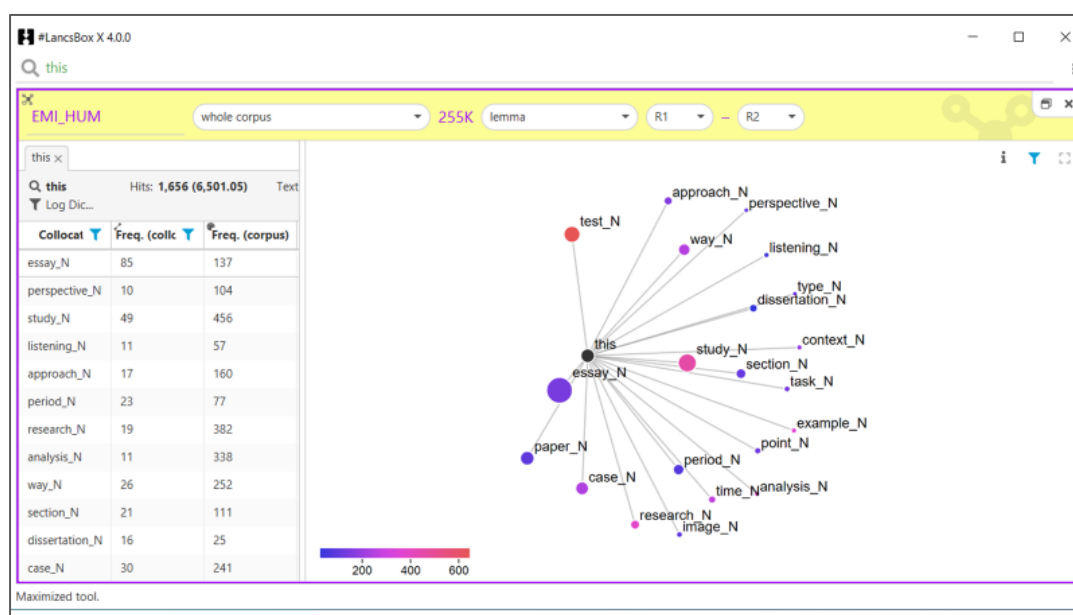*Lancaster University*

Activity 1 Using corpus evidence to understand language use in student writing

In this activity, we are going to focus on 'this + noun' structure in student academic writing. Look at the nouns that collocate with 'this' in the three EMI sub-corpora: HUM, ENG, and BUS. The collocation graphs contain nouns collocates of 'this', with 10 or more occurrences in each of the sub-corpora, using the span R1-R2.

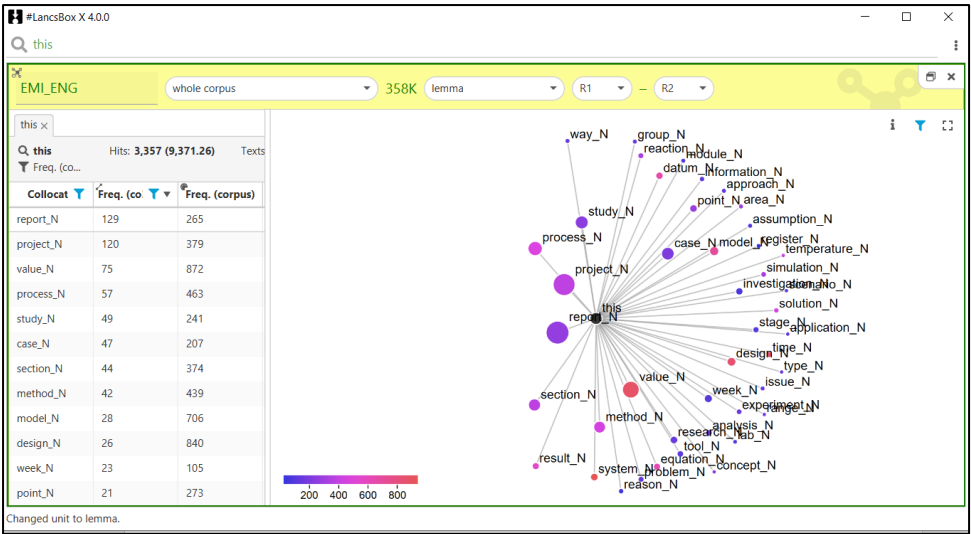Working in a pair or a small group, use the data to consider the following questions:

Q1: What type of nouns (e.g. in terms of semantic properties) do we see in the data?

Q2: What is the pattern for each of the disciplines?

Q3: In what way are the patterns in the three disciplines similar to/different from each other?

Q4: What are possible pedagogical implications of these findings?
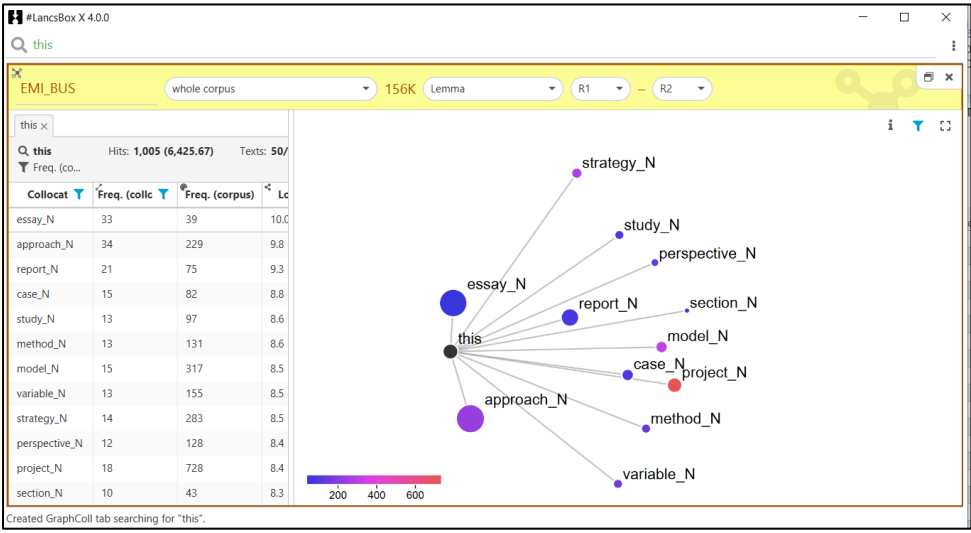
HUMANITIES & SOCIAL SCIENCES

SPACE FOR NOTES:

_____

_____

_____

ENGINEERING

BUSINESS & MANAGEMENT

SPACE FOR NOTES

_____

_____

_____

_____

_____

_____

_____

_____

_____

Activity 2: Using corpus evidence to understanding how less and more successful student writers differ from each other

In this activity, we are going to consider what corpus data can reveal about differences between less and more successful student writing. The students writing from Business & Management was divided into two categories: i) the writing that received higher marks (merit and distinction) and ii) the writing with lower marks (pass). The two sub-corpora were searched for the occurrence of words related to different dimensions of academic writing.

Frequency of target words in lower- and higher-scored Business essays

Linguistic category	Example words	Relative frequency per 10,000 words	
		Lower-scored essays	Higher-scored essays
Adjectival expressions of **certainty**	likely	156.7	196.9
	possible	940.2	492.2
	typical	52.2	69.2
Indefinite pronouns for **generalizability**	anything	104.5	32.8
	something	156.7	65.5
	everyone	173.4	43.6
Words for **hedging**	suggest	0	76.1
	might	783.5	63.6
Words for **discourse organisation**	additionally	52.8	98.4
	thus	64.2	393.8

Working in a pair or a small group, use the data to consider the following questions:

Q1: In what way is the use of the target linguistic features in more and less successful writing of Business students similar or different?

Q2: What are the possible reasons for the differences in the two sub-corpora?

Q3: How could we explore these findings further?

Q4: What are possible pedagogical implications of these findings? How could we use these findings in teaching and materials design?

SPACE FOR NOTES

_____

_____

_____

_____

_____

_____

Activity 3 Using corpus evidence to understand how student and expert writers differ from each other

In this activity, we are going to focus on 'this + verb' structure in student and expert academic writing. Look at the verbs that collocate with 'this' in two corpora: i) Academic writing in the British National Corpus 2014 and ii) the EMI_HUM. The collocation graphs contain verb collocations of 'this', with 10 or more occurrences in each of the corpora, using the span R1-R2.
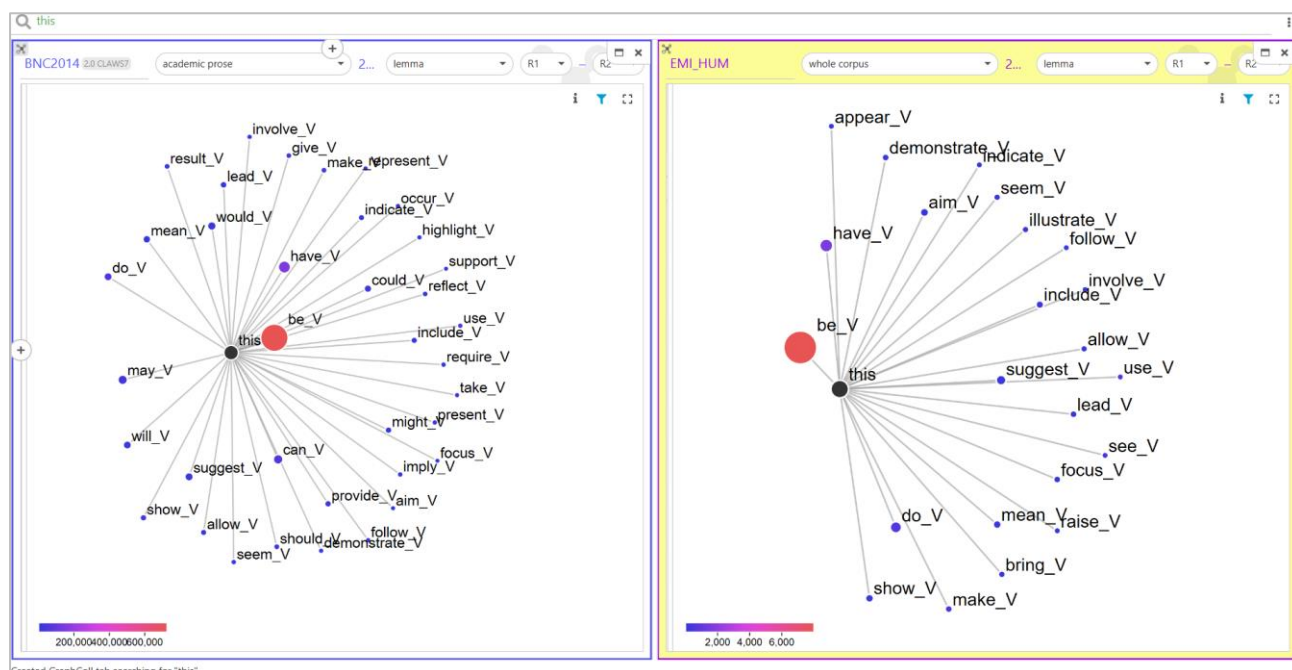
Working in a pair or a small group, use the data to consider the following questions:

Q1: In what way are the patterns in EMI_HUM and BNC2014 (academic writing) similar or different? (e.g. Is there an overlap between the two groups in their choice of verbs? Is the range of verbs similar or different?)
Q2: What are the possible reasons for the differences in the two sub-corpora?
Q3: What are possible pedagogical implications of these findings? Can we use these findings in designing EAP materials?

'THIS + VERB' COLLOCATIONS in BNC2014 (academic writing) and EMI_HUM

SPACE FOR NOTES

_____

_____

_____

_____

_____

_____