2.3.4 Employing corpora for language learning

Activities

Activity 1: What are corpora?

- Do you know what language corpora are? If not, can you guess what they are and what their function might be in language teaching/learning?

Write your views in the text box below.

When you finish, watch the video "Corpus Linguistics: The Basics" to check your guesses. You can find the link in the "Supplementary Materials" part of this Section.

Activity 2: What can corpora reveal about English?

Question 2.1

- What do you expect corpora might reveal about English?

Write your views in the text box below.

Question 2.2

- Can you guess the most frequent words in English – and in your own language, as well?

Write your views in the text box below.

Question 2.3

Have a look at the following table with 2 columns with the 50 most frequent words in written (W) and in spoken (S) English.

- Think of what you answered in Question 2.2. Is there anything that matches your guesses?
- What do you notice that you did not expect to find?
- Would you use this table and this task with your students?

W	S	W	S	W	S	W	S	W	S
1. THE	1. THE	11. IT	11. IN	21. BE	21. THEY	31. ARE	31. FOR	41. THEIR	41. DON'T
2. TO	2. I	12. ON	12. WAS	22. MY	22. WELL	32. AN	32. THIS	42. SHE	42. SHE
3. OF	3. YOU	13. HE	13. IS	23. HAVE	23. WHAT	33. THIS	33. JUST	43. WHO	43. THINK
4. A	4. AND	14. IS	14. IT'S	24. FROM	24. YES	34. HAS	34. ALL	44. IF	44. IF
5. AND	5. TO	15. WITH	15. KNOW	25. HAD	25. HAVE	35. BEEN	35. THERE	45. HIM	45. WITH
6. IN	6. IT	16. YOU	16. NO	26. BY	26. WE	36. UP	36. LIKE	46. WE	46. THEN
7. I	7. A	17. BUT	17. OH	27. ME	27. HE	37. WERE	37. ONE	47. ABOUT	47. AT
8. WAS	8. YEAH	18. AT	18. SO	28. HER	28. DO	38. OUT	38. BE	48. WILL	48. ABOUT
9. FOR	9. THAT	19. HIS	19. BUT	29. THEY	29. GOT	39. WHEN	39. RIGHT	49. ALL	49. ARE
10. THAT	10. OF	20. AS	20. ON	30. NOT	30. THAT'S	40. ONE	40. NOT	50. WOULD	50. AS

Write your views in the text box below.

Activity 3: Extracts from English spoken corpora

Have a look at the following extracts from spoken English corpora. They can be used by learners to "observe" spoken language.

- What do you notice? What can you ask your learners to notice? (e.g., use of fillers, like 'erm'; short forms, as 'cos'; repetitions; slang forms; discourse markers as *Well*, *You know*, *I mean*…)

1. … only five of us. But eh. Right, well, we'll have to, but you know …
2. That's right. Yes, I mean that would be very good …
3. … I can come on the Friday. Sixth is the Friday. Yeah, well I mean I think we need to …
4. What car they going in? Yes, well eh. Ten till two, usual? Ten till two, yes …
5. … we, we produced erm, eh I mean we talked to young people and …
6. … we've already been told that haven't we? Well I. Erm, if I. Do you, do you want me …
7. A place in Harlow? Well, I, I, I have asked for that …
8. Right, so I mean this is … Listen, this is …
9. Yes, I've got plenty. Okay, well if you've got it, that's alright, I'll see you afterwards …
10. … the twenty fourth of September isn't it? Yes, twenty fourth. Well, it'll, it'll be the er, the twenty fourth …
11. Don't leave it too long Norman. Well yeah, I mean, your in the second week aren't ya?
12. … we can come up with ideas, like the idea of that? Well I think the ideas should come from …
13. Can we go ahead, or not? Can we go ahead. Well it's up to you. Right. What'll you suggest?
14. Is this on yet? Yeah. Oh. Okay well, good morning …

Write your views in the text box below.

Activity 4: Corpora for discovering authentic use of English

Corpora for language learning can be a very useful tool for exploring authentic uses of English.

- How can language corpora help EL teachers to better scaffold their students' learning?
- How can language corpora enhance an ELF-aware approach in ELT?
- What might be the pedagogical advantages of using Data Driven Learning?
- Make a list of what you foresee as pros and cons of using corpora in the ELT class.

Write your views in the text box below.

Additional Activity: Example from VOICE

Look at the extract below and notice how 'help' is being used by non-native speakers of English in the VOICE corpus.

- Is there any use that is not traditional? Which one?
- Is communication in any way affected? How?
- Do speakers use strategies to make themselves understood? Which ones?

Write your views in the text box below.