Scientific Writing in English: Techniques and Tools

Using Corpus of Contemporary American English (COCA)

Ethel Schuster, Ph.D.

Northern Essex Community College, USA

eschuster@necc.mass.edu

Ethel Schuster, Ph.D.

1

Questions

- What makes a phrase or text sound natural?
- How to recognize a natural vs. an unnatural phrase/text?
- How to write natural/fluent text?
- What computational tools does Corpus Linguistics offers us?
  - Write better?

Ethel Schuster, Ph.D.

2

Language

- Made up of prefabricated "chunks"
  - Fixed or semi-fixed
    - combinations of words
  - Master them
  - language sounds
    - natural
    - fluent

Ethel Schuster, Ph.D.

3

Types of "Chunks"

- **Colligations**
- **Collocations**

Ethel Schuster, Ph.D.

Colligations

- **Combinations of words**
  - **in grammatical categories**
  - **e. g., verb, noun, preposition**

Ethel Schuster, Ph.D.

5

Examples

1. LOOK + preposition
   - He was looking at component
   - He was looking for the answer
2. BELIEVE + that clause
   - He believed that he would survive
3. TRY + to infinitive
   - He tried to run the marathon

Ethel Schuster, Ph.D.

Collocations

- Words "strung" together as units
  - Will sounds "right" to native speaker
  - E. g., "dream + about/of" (instead of "dream + with")

Ethel Schuster, Ph.D.

Types of Collocations

- **Verbal**
- **Nominal**
- **Adjectival**
- **Adverbial**

Ethel Schuster, Ph.D.

8

Example: Verbal

Nouns can be subject or object of verb

1. N + V

"The team identified the problem..."

2. V + N

"Our group has done the research.."

Ethel Schuster, Ph.D.

Example: Nominal

Nouns as compounds

1. N + N

"There is little research data related to obesity and THR"

2. N + of + N

"From the patient's point of view, they are likely to have little expectation that you can solve their problem"

Ethel Schuster, Ph.D.

Example: Adjectival

Adjectives combine with certain nouns

1. Adj + N

"They conducted a longitudinal study of 54 children from first through fourth grades"

"The present/current study employed both elicited and spontaneous language measures"

Ethel Schuster, Ph.D.

Example: Adverbial

- Verbs and adjectives combine with certain adverbs
- Order is usually fixed

"He had a fairly extensive knowledge of the colonies and their operation"

"Although widely acknowledged among traditional cultures, sleep paralysis is one of the most …"

Ethel Schuster, Ph.D.

12

Exception: Order NOT Fixed

- Areas with overlapping high values could be carefully studied through risk assessment.
- The team carefully studied every part of the virus.
- I'm carefully studying my setup and becoming very familiar with it.
- This material should be approached with an open mind, studied carefully and critically considered.
- The question would have to be studied carefully first.

Ethel Schuster, Ph.D.

13

Binomials

- Two words belonging to same grammatical category
- For example, verbs, nouns, adjectives, linked by conjunction or preposition

Ethel Schuster, Ph.D.

Order?

- Reversible = order can be inverted
    social or political or
    political or social
- Irreversible = order is fixed
    come and go or
    better or worse

Ethel Schuster, Ph.D.

Corpus Linguistics

- Based on assumption that language is probabilistic
  - certain combinations of words have higher probability of occurring than others, although these other ones may be perfectly grammatical

Ethel Schuster, Ph.D.

Examples

- Why should *increase or decrease* be more recurrent that *decrease or increase*?

- Why is *results suggest* much more frequent than *results suggested* in academic language?

- Why is *fill the gap* preferred to *bridge the gap* in academic discourse although latter is recurrent in everyday language?

Ethel Schuster, Ph.D.

Corpus of Contemporary American English (COCA)

- Corpus with over 460 million words, with texts ranging from 1990 to the current year (they are updated annually)

- Covered genres: from spoken language and fiction to journalistic and academic

- www.americancorpus.org

Ethel Schuster, Ph.D.

Using COCA

- Main tool available on both sites is a concordancer
  - Brings you ALL occurrences in corpus of the word you are studying
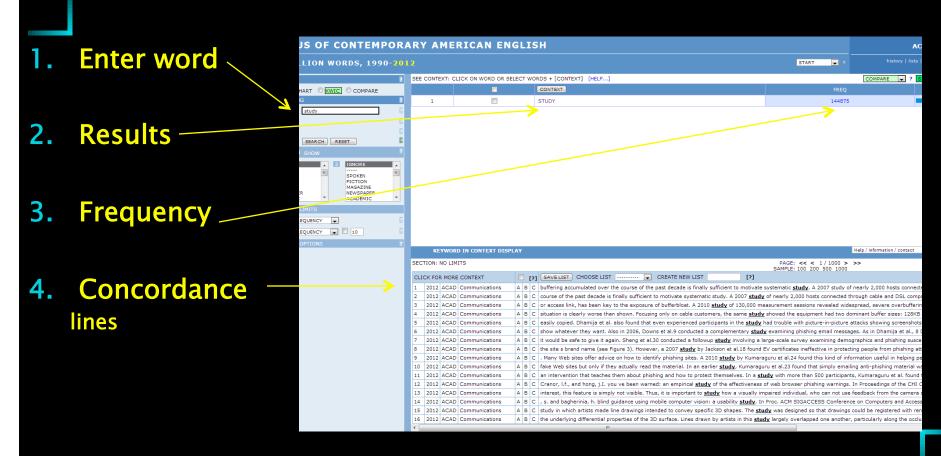 - Called KWIC (Keyword In Context) concordance, because word is shown in context

Ethel Schuster, Ph.D.

Let's use it!

- **Sample words**
- **Suggestions**
- **See tutorials**

Ethel Schuster, Ph.D.

20

COCA

- Main tool available is concordancer
  - Shows ALL occurrences of word(s)
  - Called a KWIC (Keyword In Context) concordance, i.e., word shown in context
  - Larger portion of context obtained by clicking on concordance line
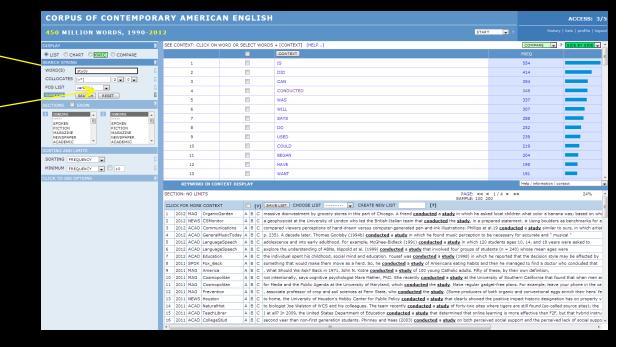
Ethel Schuster, Ph.D.

Search: study

1. **Enter word**

2. **Results**

3. **Frequency**

4. **Concordance** lines

Search: study "co-occurs" with ?

1. **Enter word**

2. **Enter category, e.g., any verb**

Can be done with COLLOCATES

Ethel Schuster, Ph.D.

Results?

- Show list of verbs that occur in specified position(s)
- Can identify *conducted* as frequent collocate, which would give us verbal collocation:

  conduct a study.

Ethel Schuster, Ph.D.

Results?

- **Go down in the list, find**
  - *Conduct* **(position 20) with 117 occurrences**
  - *conducted* **shows 348 occurrences**

Ethel Schuster, Ph.D.

25

What we learn?

- This also tells us that
  conducted a study is <u>twice</u> as common as *conduct a study*

  In other words, the collocation occurs more frequently in its past form

Ethel Schuster, Ph.D.

Verbs + "study"

- See first 20 verbs that occur 1 or 2 positions to the right of *study*

  - *conducted* is among them

Ethel Schuster, Ph.D.

27

	CONTEXT	FREQ
1	WAS	7175
2	IS	3464
3	WERE	2530
4	FOUND	2398
5	CONDUCTED	1487
6	HAS	1162
7	ARE	1109
8	SHOWED	927
9	PUBLISHED	882
10	BE	812
11	EXAMINED	737
12	REPORTED	714
13	HAD	674
14	DID	663
15	SHOWS	559
16	HAVE	549
17	WILL	538
18	SUGGESTS	530
19	SUGGEST	488
20	INCLUDED	466

Other verbs?

- Related to what the *study* investigated
  - what the *study* ...
    showed/shows/ suggests

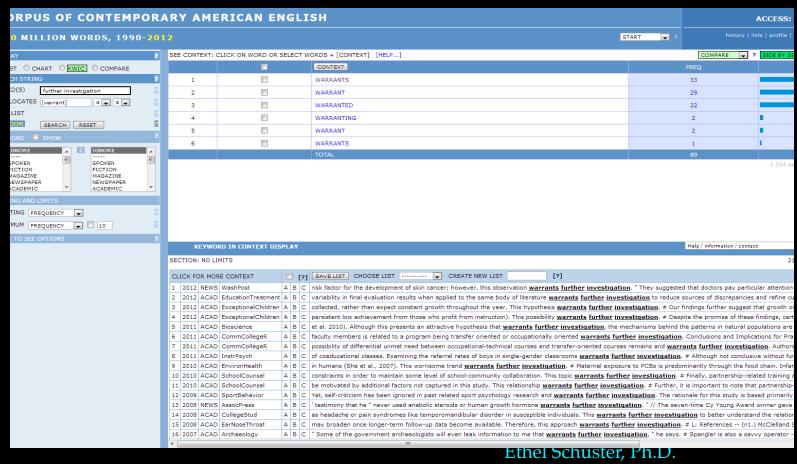>> Examine concordance lines for each verb to establish exact context in which collocation is customarily used

Ethel Schuster, Ph.D.

NOTE

- Unexpected results occur
- E.g., adjectives + "investigation"?
  Yields "further investigation"

Ethel Schuster, Ph.D.

30

Look at Concordances

- **This collocation occurs frequently with the verb WARRANT, as in**
  - ✓ These results are consistent with previous findings and warrant further investigation…
  - ✓ … this observation warrants further investigation.
  - ✓ Because further investigation of these treatments is warranted, the committee is…
  - ✓ … we believe that further investigation on this topic is warranted.
  - ✓ This hypothesis warrants further investigation.
  - ✓ This possibility warrants further investigation.

Ethel Schuster, Ph.D.

Longer conventional expression? Search…

Ethel Schuster, Ph.D.

32

Findings?

- **Forms in which WARRANT occurs:**
  - **89 occurrences.**
- **Simple <u>present</u> tense form (62) is more frequent than the <u>past</u> tense (22)**

Ethel Schuster, Ph.D.

33

Examine concordance lines...

- Most recurrent structure is
- *Det + WARRANT + further investigation*

- See...

Findings…

Determiner	Noun Phrase	WARRANT	
This	hypothesis	warrants	further investigation.
This	worrisome trend	warrants	further investigation.
This	x.x	warrants	further investigation.
These	findings	warrant	further investigation.
That	x.x	warranted	further investigation.

Concordance for *warranted*

- **Verb in passive construction, e.g.,**

  ✓ … saying no <u>further</u> <u>investigation</u> is <u>warranted</u>.

  ✓ … reasonable grounds to believe that <u>further</u> <u>investigation</u> is <u>warranted</u>.

  ✓ <u>Further</u> <u>investigation</u> is <u>warranted</u> to test the thesis that …

  ✓ … and found that no <u>further</u> <u>investigation</u> was <u>warranted</u>.

Ethel Schuster, Ph.D.

36

Longer conventional expression? Search...

- We find

    "reasonable grounds to believe that further investigation is warranted"

- In 5 of the 22 concordance lines

Ethel Schuster, Ph.D.

Examples found

- ✓ ... that the Department of Justice find that no reasonable grounds to believe that further investigation is warranted.
- ✓ ... saying there are no reasonable grounds to believe that further investigation is warranted.
- ✓ ... if she determines that there are reasonable grounds to believe that further investigation is warranted.
- ✓ ... the Attorney General concludes that "there are no reasonable grounds to believe that further investigation is warranted..."
- ✓ the Attorney General determines that reasonable grounds do exist to believe that further investigation may be warranted.

Ethel Schuster, Ph.D.

38

Closing Remarks

- Sometimes field of work is too specific
- Corpora do not have enough material in domain to answer all your questions

>> Should build customized corpus

Ethel Schuster, Ph.D.