



THE UNIVERSITY OF UTAH

L2TReC Webinar Series

Creating Data-Driven Pedagogical Materials Using Learner Corpora: A Guide for Language Teachers

Speaker: Dr. Elnaz Kia

Thursday, May 6, 2021, 3:00-4:00 PM MDT



L2TReC

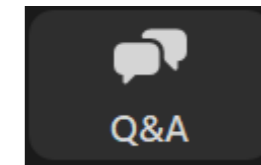
**SECOND LANGUAGE
TEACHING & RESEARCH CENTER**

THE UNIVERSITY OF UTAH

Webinar Layout & Zoom Features

- Three sections

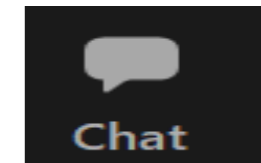
- Q&A



- Discussion breaks



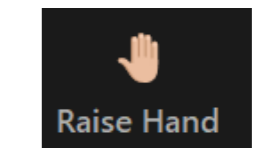
- Chat



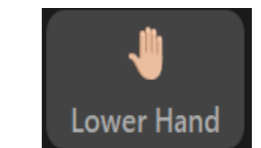
- Reflection Stops



- Raise/lower hand

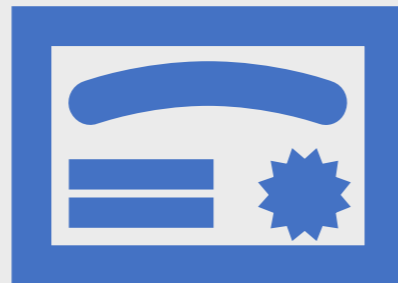


- Polls



- Q&A





Where in the world are you located?



Salt Lake City, UT



Flagstaff, AZ



Purpose

- Use simple corpus-linguistic techniques to explore learner language
- Use learner corpora to create customized and authentic materials



Outline

1. Introduction to Corpus Linguistics
2. Pedagogical Applications of Learner Corpora
3. Developing Data-Driven Activities
 - a. Example data-driven activities and studies
 - b. Interpret corpus search results and create data-driven activities



THE UNIVERSITY OF UTAH

Section 1

Introduction to Corpus Linguistics

Corpus → Corpora (plural)

- A **large, principled** collection of **natural** texts
- Analyzed using both **automatic** and **interactive** computer techniques

(Biber, Conrad & Reppen, 1998)



And the terrorists bombed in the clash of cultures

What's on your mind today?

*For a fair skin use our products
Hi, it's been a while, how do you do?*

Dr. Mark Davies



Corpus of Contemporary American English (COCA)

<http://ola-onyx.blogspot.com/>

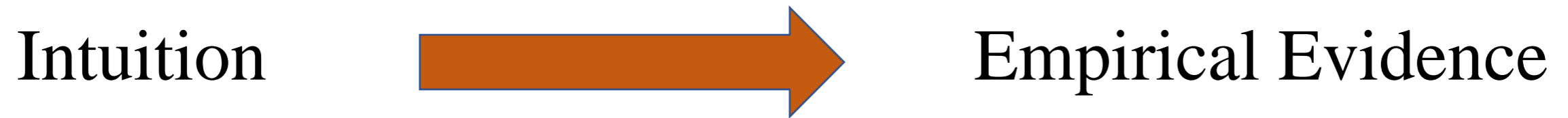
Corpus → Corpora (plural)

- **Large:**
 - Depending on the type and purpose of the corpus
- **Text:** any language produced – spoken or written
- **Principled:** organized around your goals
- **Natural Language**
- **Interactive**

Corpus Linguistics

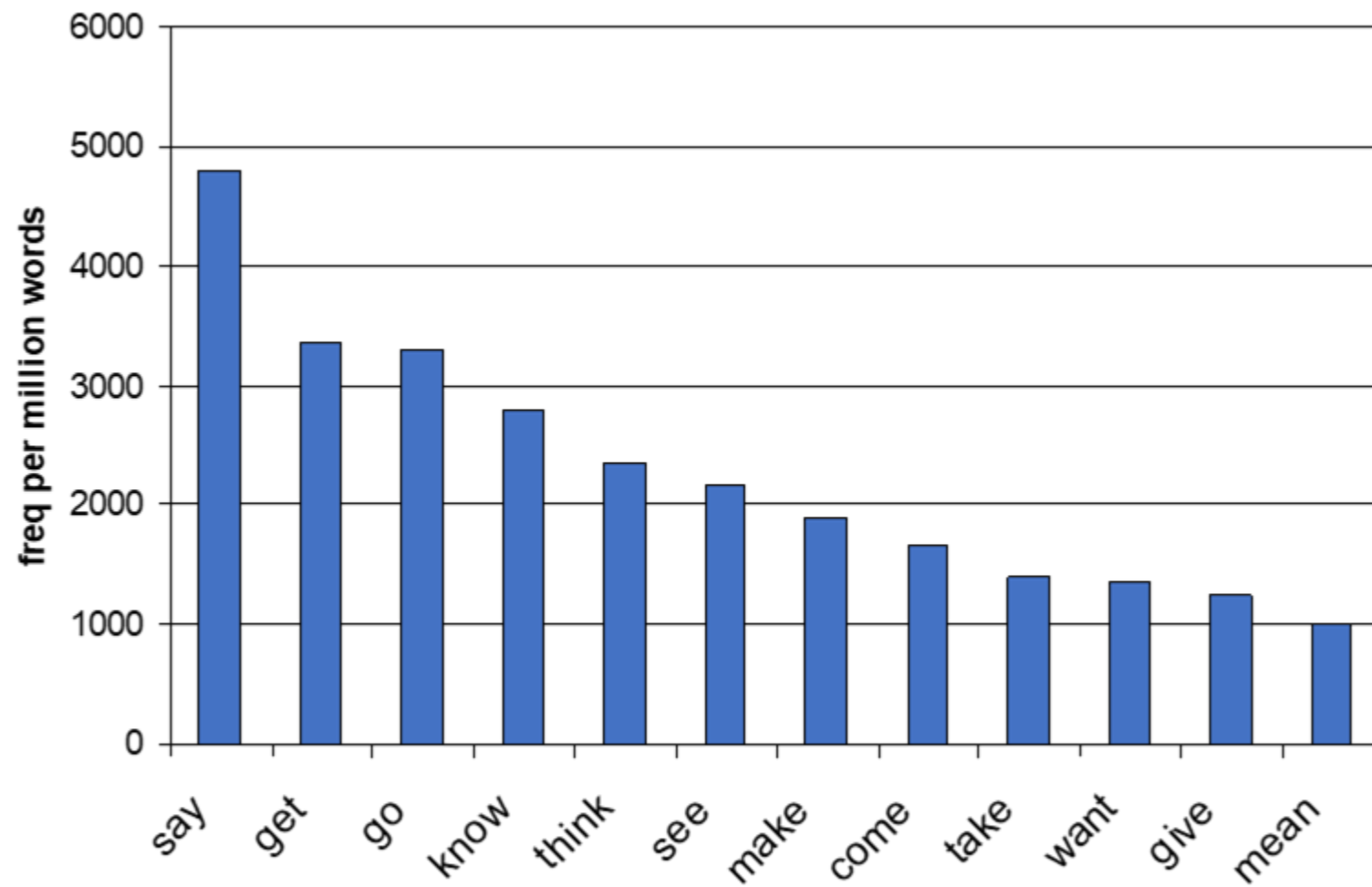
- How we use language
- How our language use varies
 - Audience
 - Production
 - Purpose
- Look at collections of **texts** and see how texts use different linguistic features to accomplish different purposes

Why Corpus Linguistics



Most Common Lexical Verbs in English

(From LGSWE; Biber et. al. 1999)





THE UNIVERSITY OF UTAH

Section 2

Pedagogical Applications of Learner Corpora

Learner Corpora

Collection of texts produced by the learners of an additional language

Why learner data?

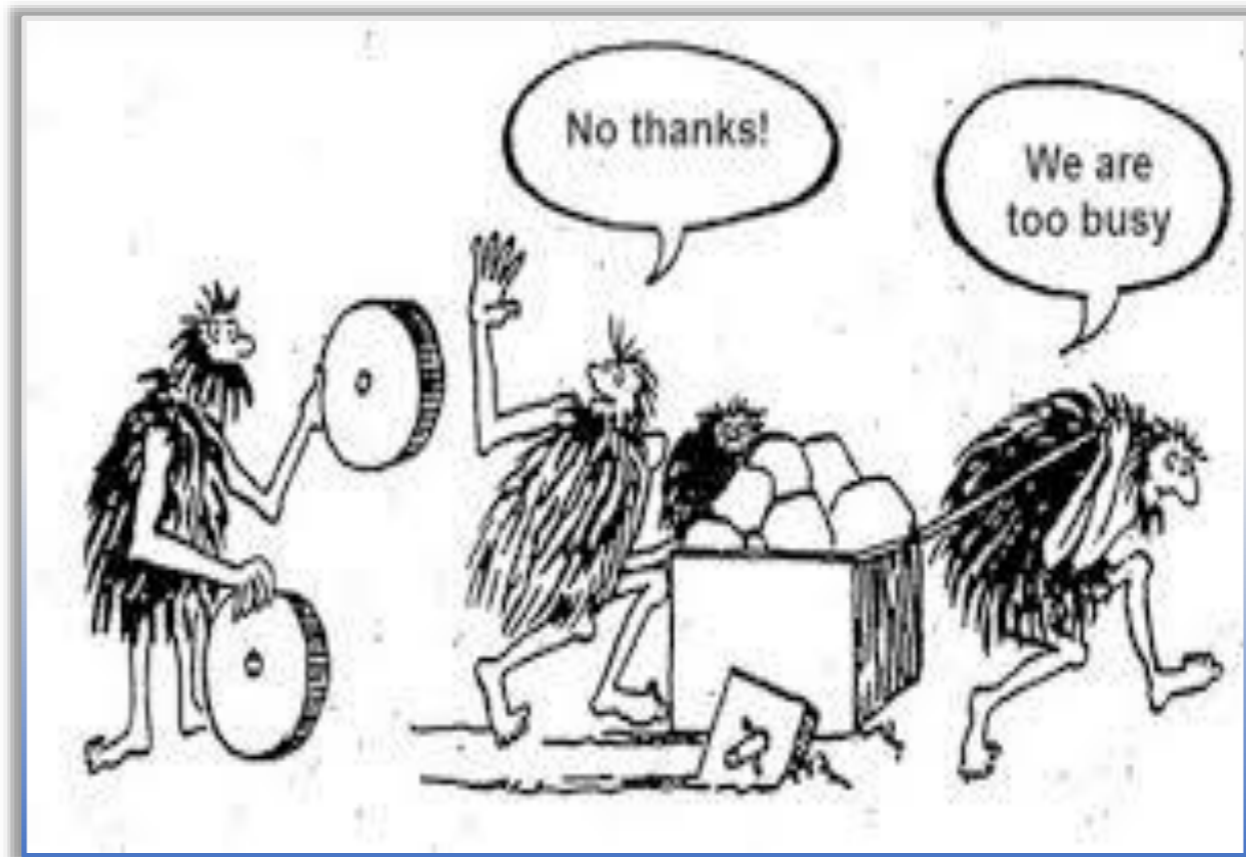
- Authentic data
 - Empirical
 - Relatable contexts
 - Appropriate level
-
- Raw vs. annotated learner corpus files
-
- Enriched corpus files (Götz, 2021, p. 68)
 - Adding mark-ups or meta-data (situational features, comments)
 - POS tagging
 - Error tagging (manual)

Pedagogical Applications

- Learner errors and common areas of difficulty
 - Common patterns of use based on large data
 - Overuse and underuse of linguistic features
 - Characteristics of proficiency levels
 - Comparison between learner and expert data
 - Link between learner language and learner profile
-
- Identify curricular needs
 - Inform the curriculum
 - Create data-driven (Gilquin & Granger, 2010) pedagogically-justified materials

Building a Corpus

- First make sure the corpus you want is not already available.



Learner Corpora

- The Multilingual Academic Corpus of Assignments – Writing & Speech (MACAWS) <https://sites.google.com/email.arizona.edu/macawswebinar/home>
- Corpus & Repository of Writing (CROW) <http://crow.corporaproject.org>
- The International Corpus Network of Asian Learners of English (ICNALE; Ishikawa, 2014). <http://language.sakura.ne.jp/icnale/>
- The Multilingual Corpus of Second Language Speech (MuSSeL) https://l2trec.utah.edu/multi-Lingual_Speech_Corpus.php

Release date: Monday, May 31st, 2021

- The Written Corpus of Utah Dual Language Immersion (CUDLI) <https://l2trec.utah.edu/corpus-of-utah-dual-language-immersion/>

Release date: Monday, Aug 30th, 2021



Corpora with Built-In CL Tools

- The English Corpora (e.g., COCA)
- MICASE & MICUSP
- CROW
- MACAWS
- ICNALE

Concordance Programs

WordSmith MonoConc AntConc

- Only .txt and .xml files are readable by concordancers
- Convert word files to plain texts: [AntFileConverter](#)
- Word lists – frequency & alphabetic
- Concordances – KWICs (**K**ey **W**ord **i**n **C**ontext)
sorting; case; wildcards;
- Distributions
- AntConc tutorial videos by Laurence Anthony
 - https://www.youtube.com/playlist?list=PLiRIDpYmiC0Ta0-Hdvc1D7hG6dmiS_TZj






- <http://crow.corporaproject.org>
- Registration required
- 10,911 samples of college writing (10 m. words)
- Data from Foundations writing courses from 4 US universities
- Filter by country, gender, assignment, date, TOEFL scores
- Search by words, phrases, and word families
- Search across genres
- Search across drafts from an individual learner

- Access through the corpus interface
- Built-in concordancer
- Full texts – Contact
- <https://writecrow.org/crow-for-teachers/>

CROW: Corpus Interface



Corpus & Repository of Writing

[Instructors](#) [Students](#) [Researchers](#) [Developers](#) [Search tips](#) [Source code](#) [Signed in](#) [Sign out](#)

[Report a problem on this page](#)

ⓘ

▶ Institution

▶ Year

▶ Semester

▶ Course

▶ Assignment

▶ Draft

▶ College

▶ Country

▶ First language

▶ Gender

▶ Program

▶ Year in School

▶ TOEFL Score

▶ Advanced Search

Matching texts: **10,911** Word count in search range: **10,154,627**

Showing matching texts 1-20 of 10,911

Display metadata

Should not all the nonnative speaker speak English as native speaker With the world develop, there are many people speak English. There are many people not native speaker but speak English well. With the time gone, English become a world language and fusion with languages around the world. As the ...

Term: Spring 2018 **Course:** ENGL 106 ⓘ **Assignment:** Position Argument ⓘ **Draft:** 2 **Gender:** F **TOEFL:** 73

November 20, 2017 English 106 Academic Argument/Position Native -speaker pronunciation should be taught as a model in English courses If you are given a choice to be a good baker by choosing either Anna Olson or common chef as your teacher, I am pretty sure you will choose Anna Olson because of...

Term: Fall 2017 **Course:** ENGL 106 ⓘ **Assignment:** Position Argument ⓘ **Draft:** 2 **Gender:** F **TOEFL:** 87

English 106 November 1, 2018 In contemporary society, the spread of English whether have a positive influence on worldwide communication has become one of the most controversial topics among the general public. As far as I concerned, the spread of English has already made a positive influence in ...

Term: Fall 2018 **Course:** ENGL 106 ⓘ **Assignment:** Position Argument ⓘ **Draft:** 2 **Gender:** F **TOEFL:** 84

English 106 04/10/2017 The English accent There are two opinions that Sung discovered about pronunciation preferences in his study. The one is "non-native speakers should be allowed to preserve their lingua-cultural identity by retaining some of their L1 phonological features when speaking Engli...

Term: Spring 2017 **Course:** ENGL 106 ⓘ **Assignment:** Position Argument ⓘ **Draft:** 2 **Gender:** M **TOEFL:** 74

As the world development, English spread to the world have many variable. Some people argue that Standard English better than hybrid English, because they can help human in many different aspects. People in different article to share their idea about advantage of standard English. In article, peop...

Term: Spring 2018 **Course:** ENGL 106 ⓘ **Assignment:** Position Argument ⓘ **Draft:** 2 **Gender:** M **TOEFL:** 73

ICNALE: Spoken Monologues

- The International Corpus Network of Asian Learners of English (Ishikawa, 2014).
- <http://language.sakura.ne.jp/icnale/>
- http://language.sakura.ne.jp/icnale/icnale_online.html

Table 1. The structure of the ICNALE (As of 2021 May)

| Module | Latest Version | Updated | Contents | # of participants | # of samples | # of words |
|------------------------|----------------|-------------|--|-------------------|--------------|------------|
| Spoken Monologues (SM) | V2.0 | 2017 August | 60-seconds monologues about two ICNALE common topics | 1,100 | 4,400 | 500,000 |

ICNALE Spoken Monologue

- KWIC
- Collocation
- Wordlist
- Freq Graph
- Keywords

| | |
|---------------------------------|--|
| Word(s) | <input style="width: 80%;" type="text"/> ? POS |
| Participants | <p>[EFL] : <input type="checkbox"/> CHN <input type="checkbox"/> IDN <input type="checkbox"/> JPN <input type="checkbox"/> KOR <input type="checkbox"/> THA <input type="checkbox"/> TWN</p> <p>[ESL] : <input type="checkbox"/> HKG <input type="checkbox"/> PAK <input type="checkbox"/> PHL <input type="checkbox"/> SIN</p> <p>[ENL] : <input type="checkbox"/> ENS students <input type="checkbox"/> ENS teachers <input type="checkbox"/> ENS</p> |
| Topic | <input checked="" type="checkbox"/> PTJ <input checked="" type="checkbox"/> SMK |
| Production Mode | ICNALE Spoken Monologue |
| Number | <input style="width: 50px;" type="text" value="20"/> ▼ |

← Back
🔍 GO
← Back to Top

Copyright 2015- Dr. Shin Ishikawa, Kobe University. All Rights Reserved.

Hints
✕

| | |
|--------------------------------------|---|
| play | ... "play" |
| play Play | ... "play" and "Play" |
| [play] | ... "play/ Play" "plays/ Plays" "played/ Played" "playing/ Playing" |
| play.[n*] | ... "play" (noun) |
| [play].[n*] | ... "play" "Play" "plays" "Plays" (noun) |
| I [v*] a [n*] | ... "I have a pen" / "I sing a song" / "I played a game" etc |
| He She [v*] a [n*] | ... "He has a pen" / "She sings a song" / "He played a game" etc |

The MuSSeL Corpus



- New release date: Monday, May 31st, 2021
- Five Languages: Mandarin Chinese, French, Portuguese, Russian, Spanish
- Three contexts of learning
 - 3rd to 10th grade students enrolled in Utah's Dual Language Immersion (DLI) program
 - Adult classroom learners
 - Adult post-immersion learners
- Speech samples from
 - ACTFL Assessment of Performance toward Proficiency in Languages (AAPPL)
 - ACTFL Oral Proficiency Interview-Computer (OPIc)

The MuSSeL Corpus



- Corpus Search Filters (Meta-Data)
- Language, Gender, Age group, Learning Context, Proficiency Rating, Associated proficiency ratings
- Task type or language function



THE UNIVERSITY OF UTAH

Section 3

Developing Data-Driven Activities

- a. Example data-driven activities and studies
- b. Interpret corpus search results and create data-driven activities

Checklist for Developing Activities (Reppen, 2010)

- Know what you want to teach!
- Select the best corpus resource for your lesson.
- Explore the corpus **completely** for the point you want to teach.
- Have complete and easy to follow directions
- Provide a variety of ways for interacting with the materials.
- If you are using computers **ALWAYS** have an alternative plan or activity.

Academic Writing: Reporting Verbs

(Shin et al., 2018, p. 45-46)

Semantic Categories of Reporting Verbs

| Category | Verb choices | Student sample writing (verbs altered) |
|--|---|--|
| Argue Introduces information and suggests feeling or stance towards the information. | argue, suggest, assert, note, predict, write , explain, conclude, mention , admit, observe, accept, claim, imply, complain, say , add, hypothesize, insist, maintain, propose, remark, reply, speculate, stress, contend, state, report, postulate, acknowledge, posit, talk about (31 types) | <ul style="list-style-type: none"> → Then the author <i>says</i> why the use and abuse could be a problem. → Then the author <i>postulates</i> why the use and abuse could be a problem. → Then the author <i>proposes</i> why the use and abuse could be a problem. |
| Show Presents an interpretation of action or task performed. | show , illustrate, indicate, demonstrate, confirm, mean, reveal (7 types) | <ul style="list-style-type: none"> → According to the article cited below, some clinical studies <i>showed</i> the safety and advantage of nanorobots. → According to the article cited below, some clinical studies <i>revealed</i> the safety and advantage of nanorobots. → According to the article cited below, some clinical studies <i>indicated</i> the safety and advantage of nanorobots. |
| Find Presents discovery; often used to refer to writers' own actions (e.g., research). | find , realize, observe, discover, establish, infer, recognize, identify, note (9 types) | <ul style="list-style-type: none"> → The peer-reviewed article <i>found</i> Smith & Khawaja, (2011) agrees with Tas (2011). → The peer-reviewed article <i>establishes</i> that (Smith & Khawaja, 2011) agrees with Tas (2011). → The peer-reviewed article <i>identifies</i> agreement between (Smith & Khawaja, 2011) and Tas (2011). |
| Think Presents mental action that conveys feelings and assumptions; <i>often</i> | think , hold, assume, feel, hope, know (5 types) | <ul style="list-style-type: none"> → Some students <i>think</i> it is not necessary to know foreign culture.* → Most international students <i>feel</i> lonely, bored, and they cannot make friends with foreign students.* <p>*Think verbs are much stronger when connected directly to</p> |
| <i>used without evidence.</i> | | evidence; how might you provide support for such statements? |

The bolded verbs above are the ones that are most commonly used by international students, accounting for over three-fourths of all reporting verb use in an earlier study of literature review assignments (Kwon, Staples, & Partridge, 2018).



Reflect: Activity Ideas



Reflect: Activity Ideas

Noticing: Underline the reporting verbs in the following excerpt.

Fill in the blank activity (with/without options).

Negotiation of meaning: Explain your choices to your partner.

Production (pair work):

Report your partners' ideas using appropriate reporting verbs

More scaffolding?

Complete the sentences with the appropriate reporting verb.

(Grammar and Beyond Essentials 4; Bunting et al., 2019, p. 160)

Exercise 2.2 More Reporting Verbs

A Complete the sentences with the appropriate reporting verbs. Remember to write *that* after the verb.

1 A report from the Dietary Guidelines for Americans suggests that people should "eat fewer calories, be more active, and make wiser food choices."

a suggests **b** complains **c** demonstrates



2 The Centers for Disease Control and Prevention (CDC) _____ one way to fight childhood obesity is to remove high-calorie, low-nutrition snacks.

a emphasizes **b** illustrates **c** warns

3 Former President Bill Clinton _____ many parents are overworked and don't have time to prepare healthy food for their children.

a describes **b** believes **c** estimates

4 A report by the Centers for Disease Control and Prevention _____ child whose body mass index is above the 85th percentile is overweight.

a states **b** recommends **c** warns

5 The Dietary Guidelines for Americans _____ people can reduce the risk of many chronic diseases, including cancer, through better nutrition.

a assume **b** predict **c** point out

Production (pair work)

(Grammar and Beyond Essentials 4; Bunting et al., 2019, p. 161)

Exercise 2.3 Using Reporting Verbs

Pair Work Talk with a partner about one of the topics below or a topic approved by your teacher. Take notes on your partner's answers, and then report your partner's ideas to the class. Use reporting verbs in your report.

- At what age should children have cell phones?
- At what age should children be allowed to join social networks?
- At what age should children have their own computers?
- At what age should children have TVs in their rooms?

Monica argues that children under 10 should not have cell phones. She points out that cell phones are expensive and young children could easily break them. She also believes that children under 10 should always have an adult present, so they wouldn't need their own phones.

Using Corpora to Inform EAP instruction

- Academic Speech
- Classroom Discussions

Paired speaking task

Situation: You and your partner are roommates sharing an apartment. You only have one small refrigerator so you need to plan your meals, and do your shopping together. You need to make a grocery list for the week.

Directions: You will have **3 minutes** to prepare with your partner. Please do the following things:

- Decide who will be partner A and who will be partner B
- Talk about why you want to buy **each** food or drink in your list
- Talk about why it's a good idea or bad idea to buy the foods/drinks your partner suggests
- Together, you should **agree on any 2 items** you will buy at the store

After preparing, you and your partner will have **2 minutes** to record your conversation.

Items to put on the grocery list

PARTNER A

- cookies
- fruit
- energy drinks
- milk

PARTNER B

- bread
- vegetables
- coffee
- yoghurt

Useful Phrases for Paired Speaking Tasks

| Agree | Disagree | Show understanding | Ask for a clarification | Give clarification | Comprehension check |
|--|---|--|---|---|-------------------------|
| Yeah, maybe you are right. Oh, good idea! I like the idea! Sure, we can do that! Nice! That sounds good! Yeah, why not. We can ... Sounds like a plan! | I see your point but... I like the idea but... I don't think it is very important, Let's ... Well, we can think about it. Hmmm, I see but I think it is better to ... I don't think that's gonna work. | Oh, I see your point. Oh, I see now. Got it! Oh, now I understand! I see. Yeah! Right! | What exactly do you mean? Do you mean... What is ... Could you explain more? I still don't understand. Why do you think so? Can you be more specific? | Let me explain... What I mean is... Let me say it another way... It's like this... What I'm trying to say is... | Do you see what I mean? |

Functions of High-Frequency Expressions in MICASE (Simpson, et al., 2002) and T2K-SWAL (Biber, et al. 2002)

| # | Expressions | Functions |
|----|----------------------------|---|
| 1 | Yeah, but | Disagreement, (clarification) |
| 2 | What I mean | Clarification, (checking understanding) |
| 3 | OK, but | Disagreement |
| 4 | You're saying | Asking for / giving clarification |
| 5 | That's right | Agreement |
| 6 | You're right | Agreement |
| 7 | That's true | Agreement |
| 8 | I agree (, but) | Agreement, (disagreement) |
| 9 | Do you mean | Asking for clarification |
| 10 | I have a question | Asking for clarification/information |
| 11 | (I see) what you're saying | Asking for clarification, (showing understanding) |
| 12 | I don't think so | Disagreement |
| 13 | What do you mean? | Asking for clarification |
| 14 | Good point | Agreement |
| 15 | (I see) your point (, but) | Agreement, (disagreement) |
| 16 | I agree with ... (, but) | Agreement, (disagreement) |
| 17 | Sounds good | Agreement |
| 18 | In other words | Asking for / giving clarification |
| 19 | Are you saying | Asking for clarification |
| 20 | Me too (, but) | Agreement, (disagreement) |
| 21 | I agree ... but | Disagreement |

Noticing from Reading

(Bogorevich & Kia, forthcoming)

Activity 1. Noticing from reading

Directions: Form groups of 2. Read the dialogues. What expressions do interlocutors use to *agree*, *disagree*, and give or ask for a *clarification*.

Transcript 1: SGR385SU057

S1: it's part of the real numbers...?

S2: A-D minus B-C otherwise it would be orientation-reversing.

S1: yeah but just we're also in the real numbers so you can't have a complex conjugate.

S2: right.

S1: but

S2: oh i see.

S1: but that doesn't help us though. like, because it's an inequality.

Transcript 2: OFC270MG048

S9: okay so, what, do you mean by path by paths? i mean, (i was) talking about (xx)

S2: well, what i mean by pa- let's- let's look at one that is dependent. let's say node four is given as evidence. and you're trying to see if node two and node five, are independent of each other.

okay. so you start at node two and you say well, okay is node two, dependent on node four, yes? is node two dependent on node one? yes, because rule three doesn't apply anymore.

Transcript 3: LES425JG077

S1: why wouldn't you account for the probability of the egg staying as an egg?

S2: ah, good point good point that's right. if your if your uh time unit was sufficiently small then there might, there might be uh, a loop here as well. that's a good point right.

Transcript 4: OFC195SU116

S1: okay so if you divide that flux by area again

S18: yeah so i don't want that there if i ha- say Q is like that. yeah

S1: so you don't ~~wanna~~ do this. now|

S18: i didn't really mean that. <LAUGH>

S1: okay

S18: ~~cuz~~ i haven't (written it yet.)

S1: so, and you're saying like energy over time right?

S18: ~~mhm~~

Agree _____

Disagree _____

Give a clarification _____

Ask for a clarification _____

Noticing from Listening

(Bogorevich & Kia, forthcoming)

Noticing from Listening

Directions: Work individually. Listen to three short dialogs. Identify what expression(s) the speakers use to agree, disagree, or ask/give a clarification. Each dialog can have one or more expression types. Next, check in pairs.

Dialog 1: (MICASE SGR175SU123, 10:14-11:18)

<http://media.talkbank.org/TalkBank/CABank/MICASE/study%20groups/sgr175su123.mp3>

Agree _____

Disagree _____

Give a clarification _____

Ask for a clarification _____

Dialog 2 (MICASE SGR999MX115, 1:16:50-1:18:17)

<http://media.talkbank.org/TalkBank/CABank/MICASE/study%20groups/sgr999mx115.mp3>

Agree _____

Disagree _____

Give a clarification _____

Ask for a clarification _____

Dialog 3 (MICASE SGR175SU123, 1:08:30-1:09:55)

<http://media.talkbank.org/TalkBank/CABank/MICASE/study%20groups/sgr175su123.mp3>

Agree _____

Disagree _____

Give a clarification _____

Ask for a clarification _____

¹All three dialogs have been extracted from MICASE.

Scaffolded Practice

(Bogorevich & Kia, forthcoming)

Scaffolded Practice: Clarification

Directions: Use the cards to role-play the dialogs.

| | | | |
|----|--|----|--|
| #1 | <p>Partner A</p> <ul style="list-style-type: none"> - You say: "There are no active drug ingredients in placebos." - Provide a clarification. Use: "Well, you know, what I mean is..." - Check understanding. Use: "You know what I mean?" | #1 | <p>Partner B</p> <ul style="list-style-type: none"> - Ask for a clarification. Use: "What do you mean?" - Show understanding. Use: "Oh, I see what you are saying." |
| #2 | <p>Partner B</p> <ul style="list-style-type: none"> - Ask for a clarification. Use: "Double major?" - Show understanding. Use: "Oh, ok." | #2 | <p>Partner A</p> <ul style="list-style-type: none"> - You say: "You can do a double major." - Provide a clarification. Use: "You know, in other words..." - Check understanding. Use: "You see what I'm saying?" |
| #3 | <p>Partner A</p> <ul style="list-style-type: none"> - You say: "You need to construct a figure." - Provide a clarification. Use: "Well, what I am saying..." | #3 | <p>Partner B</p> <ul style="list-style-type: none"> - Ask for a clarification. Use: "What do you mean by that?" - Show understanding. Use: "Ok, got it." |
| #4 | <p>Partner B</p> <ul style="list-style-type: none"> - Ask for a clarification. Use: "Poster sessions?" - Show understanding. Use: "That's right." | #4 | <p>Partner A</p> <ul style="list-style-type: none"> - You say: "We will have poster sessions." - Provide a clarification, use: "Well, I mean..." |
| #5 | <p>Partner A</p> <ul style="list-style-type: none"> - You say: "I need to see a doctor to have my leg x-rayed." - Provide a clarification, use: "Well, I mean..." - Check understanding, use: "You see what I'm saying?" | #5 | <p>Partner B</p> <ul style="list-style-type: none"> - Ask for a clarification, use: "What do you mean x-rayed?" - Show understanding, use: "Oh, I see." |

Group Academic Discussion: Paired-Speaking Task

(Bogorevich & Kia, forthcoming)

Paired Speaking Task

Directions: You will have **3 minutes** to prepare with your partner and **2 minutes** to record your dialog for the teacher to grade.

Situation: You and your partner have a presentation next week. You are going to give advice to your classmates on how to conserve energy.

Preparation:

- Decide who will be partner A and who will be partner B.
- Decide on only two topics that you will cover in your presentation.
- Choose at least 6 phrases from the list of useful expressions. Each partner needs to use 3 of them.
- Discuss different ways to conserve energy at home and use the expressions.

| PARTNER A | PARTNER B |
|---|--|
| <ul style="list-style-type: none"> • Replace light bulbs with energy efficient LED bulbs. • Install solar panels. | <ul style="list-style-type: none"> • Take shorter showers. • Replace all appliances with new efficient ones. |

| Useful Expressions: | Use this space to write notes: |
|--|--------------------------------|
| <p>Ask for/give clarification: Do you mean...? What do you mean? You are saying? What I mean... In other <u>words</u>...</p> <p>Show/check understanding: I see (what you're saying) You see what I mean? You know what I mean?</p> <p>To agree/disagree: Yeah, (but...) Ok, (but...) That's right (but...) You're right (but...) Sounds good! (but...) Good point/idea! (but...)</p> | |



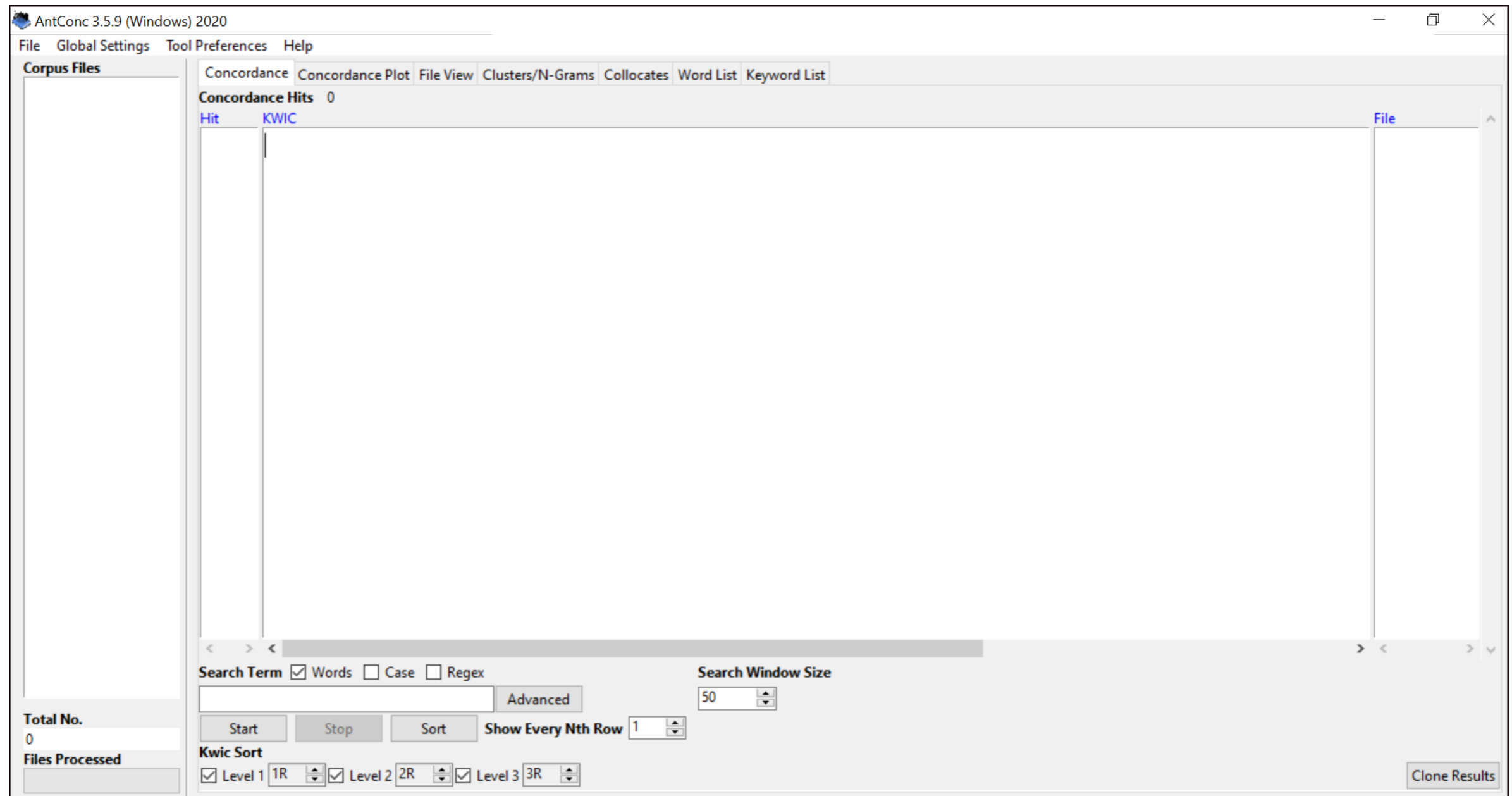
DISCUSSION TIME

AntConc

- Only .txt and .xml files are readable by concordancers
- Convert word files to plain texts: [AntFileConverter](#)
- AntConc tutorial videos by Laurence Anthony
 - https://www.youtube.com/playlist?list=PLiRIDpYmiC0Ta0-Hdvc1D7hG6dmiS_TZj



AntConc: User Interface



CL Tools

- Word lists – frequency & alphabetic
- Concordances – KWICs (**K**ey **W**ord **i**n **C**ontext)
- Distributions
- N-grams/Clusters/Lexical Bundles

CL Tools: Word Lists

(Götz, 2021, p. 78)

- Initial idea
- Counts different word forms
- Sorted by frequency ascending and descending
- Sort by word
- Download the wordlist (File → save output to text file)
- Word usage based on frequency
- Find spelling variants/errors
 - Ask students to mark unfamiliar words in the list and make vocabulary lessons based on those vocabulary terms
 - Use the most frequent words to make vocab activities such as sentence scramble, fill in the blank

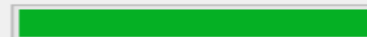
Corpus Files

- SM_CHN_PTJ_A2_0.t
- SM_CHN_PTJ_B1_1.t
- SM_CHN_PTJ_B1_2.t
- SM_CHN_PTJ_B2_0.t
- SM_CHN_SMK_A2_0
- SM_CHN_SMK_B1_1
- SM_CHN_SMK_B1_2
- SM_CHN_SMK_B2_0
- SM_ENS_PTJ_XX1.tx
- SM_ENS_PTJ_XX2.tx
- SM_ENS_PTJ_XX3.tx
- SM_ENS_SMK_XX1:
- SM_ENS_SMK_XX2:
- SM_ENS_SMK_XX3:
- SM_HKG_PTJ_B1_1.t
- SM_HKG_PTJ_B1_2.t
- SM_HKG_PTJ_B2_0.t
- SM_HKG_SMK_B1_1
- SM_HKG_SMK_B1_2
- SM_HKG_SMK_B2_0
- SM_IDN_PTJ_A2_0.t
- SM_IDN_PTJ_B1_1.t
- SM_IDN_PTJ_B1_2.t
- SM_IDN_PTJ_B2_0.t
- SM_IDN_SMK_A2_0
- SM_IDN_SMK_B1_1
- SM_IDN_SMK_B1_2
- SM_IDN_SMK_B2_0
- SM_JPN_PTJ_A2_0.t
- SM_JPN_PTJ_B1_1.t
- SM_JPN_PTJ_B1_2.t
- SM_JPN_PTJ_B2_0.t
- SM_JPN_SMK_A2_0
- SM_JPN_SMK_B1_1
- SM_JPN_SMK_B1_2
- SM_JPN_SMK_B2_0
- SM_KOR_PTJ_A2_0.t
- SM_KOR_PTJ_B1_1.t
- SM_KOR_PTJ_B1_2.t
- SM_KOR_PTJ_B2_0.t
- SM_KOR_SMK_A2_0
- SM_KOR_SMK_B1_1
- SM_KOR_SMK_B1_2
- SM_KOR_SMK_B2_0
- SM_PAK_PTJ_A2_0.t
- SM_PAK_PTJ_B1_1.t
- SM_PAK_PTJ_B1_2.t
- SM_PAK_PTJ_B2_0.t
- SM_PAK_SMK_A2_0
- SM_PAK_SMK_B1_1
- SM_PAK_SMK_B1_2
- SM_PAK_SMK_B2_0
- SM_PHL_PTJ_B1_1.t
- SM_PHL_PTJ_B1_2.t
- SM_PHL_PTJ_B2_0.t
- SM_PHL_SMK_B1_1
- SM_PHL_SMK_B1_2
- SM_PHL_SMK_B2_0

Total No.

78

Files Processed



Concordance Concordance Plot File View Clusters/N-Grams Collocates Word List Keyword List

Word Types: 6566 Word Tokens: 489832 Search Hits: 0

| Rank | Freq | Word |
|------|-------|----------|
| 1 | 17958 | the |
| 2 | 15481 | to |
| 3 | 15205 | and |
| 4 | 13116 | i |
| 5 | 10908 | it |
| 6 | 10160 | is |
| 7 | 10010 | a |
| 8 | 8979 | in |
| 9 | 8326 | that |
| 10 | 8077 | for |
| 11 | 7931 | smoking |
| 12 | 7639 | they |
| 13 | 7441 | time |
| 14 | 6735 | of |
| 15 | 6463 | have |
| 16 | 5678 | part |
| 17 | 5659 | job |
| 18 | 5573 | can |
| 19 | 5482 | you |
| 20 | 5187 | be |
| 21 | 5141 | so |
| 22 | 5050 | people |
| 23 | 4987 | think |
| 24 | 4896 | because |
| 25 | 4625 | not |
| 26 | 4428 | uh |
| 27 | 4236 | their |
| 28 | 4219 | s |
| 29 | 4032 | we |
| 30 | 3995 | smoke |
| 31 | 3953 | should |
| 32 | 3829 | are |
| 33 | 3495 | t |
| 34 | 3174 | will |
| 35 | 3146 | if |
| 36 | 3077 | with |
| 37 | 2994 | students |

Search Term Words Case Regex

Advanced

Hit Location

Search Only 0

Start Stop Sort

Lemma List Loaded

Word List Loaded

Sort by Invert Order

Sort by Freq

CL Tools: Concordances

AntConc

- See the distribution across texts and within texts
- Example: Passives (were *ed)

The screenshot shows the AntConc 3.5.6 (Windows) 2018 interface. The main window displays a list of concordance hits for the search term "were *ed". The hits are numbered 1 through 18 and show various contexts where the word "were" is used in a passive construction. The search term is entered in the "Search Term" field, and the "Kwic Sort" options are checked for Level 1 (2L), Level 2 (1L), and Level 3 (1R). The "Search Window Size" is set to 100. The "Total No." of hits is 36, and the "Files Processed" is shown as a green bar.

AntConc 3.5.6 (Windows) 2018

File Global Settings Tool Preferences Help

Concordance Concordance Plot File View Clusters/N-Grams Collocates Word List Keyword List

Concordance Hits 25

| Hit | KWIC | File |
|-----|--|-----------|
| 1 | sure about that but I'd say the majority of businesses that are started were started [unclear words] offices , that's pretty typical 2: [unclear v | busmgle |
| 2 | actories. [274] Uh what they call factory farms. But before regulations were released, the E.P.A. cut a deal, allowing the pork industry to hire | busbalel |
| 3 | ous ? and it's interesting when students were asked college students were asked about monogamy . seventy percent of them yes they wer | busmgle |
| 4 | turning this in last week was telling me the opposite story? 2: cos we were confused- 1: cos you were confused. don't think too much. OK w | busbalel |
| 5 | as telling me the opposite story? 2: cos we were confused- 1: cos you were confused. don't think too much. OK we set out we take the origin | busbalel |
| 6 | fluence of alcohol and an accident resulted, and the criminal charges were filed, uh, it was a six person, so a court not of record, kind of jur | busbalel |
| 7 | g paid something ? they're getting paid everything . everything they were promised in order for the equity holders to get anything , the de | busfilegr |
| 8 | ot, the introduce the widows of the two, guards, security guards who were killed in the White House remember a couple months ago in tha | busbmle |
| 9 | both in part A. One A. and one B. There, unfortunately, R.O.A. is, you were asked forty percent, but it needs to be? 2: Fifty. 1: Fifty percent. S | busfileuc |
| 10 | of both people and the friends involved . if you were a friend or you were involved in this kind of a situation . ok I want you to identify fir | busmgle |
| 11 | ost-audit. uh anybody here ever [250] have a job where they're they were evaluated against [2 syll] something that they had uh were part | busbalec |
| 12 | out with that dishonesty that was occurring what kind of risk factors were involved . .ok I want you to look at I'm gonna ask you to look at | busmgle |
| 13 | lass of shares . so if you wanted more than one class of shares , you were forced to [unclear words] use the AMEX or NASDAQ . uh 2: [un | busfilegr |
| 14 | we just used the formulas that we had and solved for the one that we were interested in. these two equations are equal to each other. so I : | busmgle |
| 15 | e about this. would it affect your judgment in a case, if the attorneys were characterized as follows, and I'm not saying this is bad. I'm just | busbalel |
| 16 | right marketing activities were done. And how well those activities were organized in the company. Some companies, you know don't en | busmkle |
| 17 | hat I mean any other g- thoughts on this poem uh beyond those that were presented by the group here. I guess Lisa you saw it as more sh | busmgle |
| 18 | a team building day. we went out to stone mountain? we had to we were assigned as groups at the beginning of the day? and we had di | busmgle |

Search Term Words Case Regex Advanced Search Window Size 100

Start Stop Sort Show Every Nth Row 1

Kwic Sort Level 1 2L Level 2 1L Level 3 1R

Clone Results

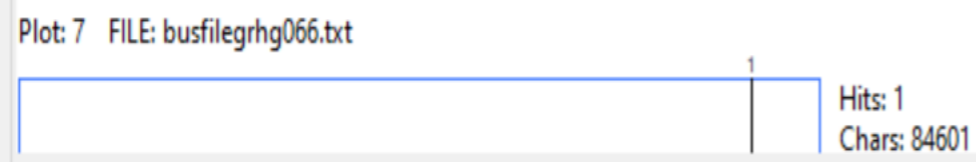
Corpus Files

- busacleldln052.txt
- busacleldmn048.txt
- busacleudhs057.txt
- busacleudms056.txt
- busacleudms058.txt
- busatlegrmg077.txt
- busatlegrmg084.txt
- busatleudhg013.txt
- busatleudhg016.txt
- busbalegrhn218.txt
- busbalegrhn219.txt
- busbaleldhn045.txt
- busbaleldhn046.txt
- busbaleldmn044.txt
- busbaleldmn054.txt
- busbaleudhn152.txt
- busbaleudhn154.txt
- busbaleudhn170.txt
- busbaleudln049.txt
- busbmleudhg019.txt
- busfilegrhg066.txt
- busfilegrmg002.TXT
- busfileudhg024.txt
- busmglegrhg105.TXT
- busmgleldhg174.txt
- busmgleudhg096.txt
- busmgleudhg104.txt
- busmgleudhg145.txt
- busmgleudhg147.txt
- busmqleudmq114.txt

Total No.
36
Files Processed

Concordance Concordance Plot File View Clusters/N-Grams Collocates Word List Keyword List

Concordance Hits 25 Total Plots (with hits) 15



Search Term Words Case Regex Plot Zoom

were *ed Advanced x1

Start Stop Show Every Nth Row 1



Reflect: *Activity Ideas*

Multi-word vs. single word verbs

- Single word verbs are usually preferred over multi-word verbs in formal writing.
- CROW
- Searching for a word/word family
- *look into* → *investigate*
- *put up with* → *tolerate*

Search by word family (e.g., eats/eating/eaten/ate)

When you aren't looking for an exact word form, use the 'word family' radio button:

▼ Advanced Search

Tip: To search by phrase, place multiple words in quotation marks

Word form handling: Exact word Word family (a search for 'eat' returns results for 'eat/eats/eating/eaten/ate/eaten')

Multiple word handling: Contains any word Contains all words

Filter by ID..

- › Institution
- › Year
- › Semester
- › Course
- › Assignment
- › Draft
- › College
- › Country
- › First language
- › Gender
- › Program
- › Year in School
- › TOEFL Score

"look into" ⓘ

▼ Advanced Search

Tip: To search by phrase, place multiple words in quotation marks

Word form handling: Exact word Word family (a search for 'eat' returns results for 'eat/eats/eating/eaten/ate/eaten')

Multiple word handling: Contains any word Contains all words

Matching texts: 126 Word count in search range: 10,154,627

Showing matching texts 1-20 of 126

| Search query | Instances in matching texts ¹ | Normed (per 1 million) ² | Texts containing term |
|--------------|--|-------------------------------------|-----------------------|
| "look into" | 141 | 13.89 | 126 |

1: The total number of instances within texts that match all search criteria

2: The number of instances per 1 million words in the corpus

Display metadata

...hat "The studies shows that #sociolinguistics also **look into** how #language combine with the #society in the past." I abbreviated the 'By studying writ.....

Term: Fall 2016 **Course:** ENGL 106 ⓘ **Assignment:** Reflection ⓘ **Draft:** F **Gender:** F **TOEFL:** 86

...psychology and #anthropology. 4.#Sociolinguistics **look into** the interaction of #language and #society, withe language as the beginning point. 5. Man.....

Term: Fall 2016 **Course:** ENGL 106 ⓘ **Assignment:** Register Rewrite ⓘ **Draft:** 2 **Gender:** F **TOEFL:** 86

...psychology and #anthropology. 4.#Sociolinguistics **look into** the interaction of #language and #society, withe language as the beginning point. 5. Man.....

...hat "The studies shows that #sociolinguistics also look into how #language combine with the #society in the past." I abbreviated the 'By studying writ.....

Term: Fall 2016 **Course:** ENGL 106 ⓘ **Assignment:** Reflection ⓘ **Draft:** F **Gender:** F **TOEFL:** 86

...psychology and #anthropology. 4.#Sociolinguistics look into the interaction of #language and #society, withe language as the beginning point. 5. Man.....

Term: Fall 2016 **Course:** ENGL 106 ⓘ **Assignment:** Register Rewrite ⓘ **Draft:** 2 **Gender:** F **TOEFL:** 86

...psychology and #anthropology. 4.#Sociolinguistics look into the interaction of #language and #society, withe language as the beginning point. 5. Man.....

Term: Fall 2016 **Course:** ENGL 106 ⓘ **Assignment:** Register Rewrite ⓘ **Draft:** 1 **Gender:** F **TOEFL:** 86

...ere has been a fall in literacy. He says we should look into this new way of texting to understand it and he gives us many examples of the "new langua.....

Term: Fall 2017 **Course:** ENGL 106 ⓘ **Assignment:** Register Rewrite ⓘ **Draft:** 1 **Gender:** M **TOEFL:** N/A

...psychology and #anthropology. 4.#Sociolinguistics look into the interaction of #language and #society, withe language as the beginning point. 5. Man.....

Term: Fall 2016 **Course:** ENGL 106 ⓘ **Assignment:** Register Rewrite ⓘ **Draft:** F **Gender:** F **TOEFL:** 86

...psychology and #anthropology. 4.#Sociolinguistics look into the interaction of #language and #society, withe language as the beginning point. 5. Man.....

Term: Fall 2016 **Course:** ENGL 106 ⓘ **Assignment:** Register Rewrite ⓘ **Draft:** 4 **Gender:** F **TOEFL:** 86

...psychology and #anthropology. 4.#Sociolinguistics look into the interaction of #language and #society, withe language as the beginning point. 5. Man.....

Term: Fall 2016 **Course:** ENGL 106 ⓘ **Assignment:** Register Rewrite ⓘ **Draft:** 3 **Gender:** F **TOEFL:** 86

...hat "The studies shows that #sociolinguistics also look into how #language combine with the #society in the past." I abbreviated the 'By studying writ.....

Term: Fall 2016 **Course:** ENGL 106 ⓘ **Assignment:** Description and Explanation ⓘ **Draft:** 2 **Gender:** F **TOEFL:** 86

...hat "The studies shows that #sociolinguistics also look into how #language combine with the #society in the past." I abbreviated the 'By studying writ.....

Term: Fall 2016 **Course:** ENGL 106 ⓘ **Assignment:** Description and Explanation ⓘ **Draft:** 1 **Gender:** F **TOEFL:** 86

...hat "The studies shows that #sociolinguistics also look into how #language combine with the #society in the past." I abbreviated the 'By studying writ.....

Term: Fall 2016 **Course:** ENGL 106 ⓘ **Assignment:** Description and Explanation ⓘ **Draft:** F **Gender:** F **TOEFL:** 86

...te the credibility of sources, which brought me to look into the founding parties of some of my sources in the annotated bibliography. For example, I

Term: Spring 2017 **Course:** ENGL 108 ⓘ **Assignment:** Reflection ⓘ **Draft:** F **Gender:** M **TOEFL:** 80

...tudent which is not fair and it makes the students look into other places just to be able to pay the school. In addition, internationals' fees are way.....

Term: Fall 2017 **Course:** ENGL 108 ⓘ **Assignment:** Public Argument ⓘ **Draft:** F **Gender:** M **TOEFL:** N/A

...kers, Any speakers of English. and he respectively **investigate** to the people and which pronunciation model they will choice. Most English learner sele.....

Term: Fall 2016 **Course:** ENGL 106 ⓘ **Assignment:** Position Argument ⓘ **Draft:** 1 **Gender:** M **TOEFL:** 77

...s inherent in Gulf English. In a study designed to **investigate** the pronunciation of English monophthongs amongst L1 Arabic-speaking university graduat.....

Term: Fall 2017 **Course:** ENGL 106 ⓘ **Assignment:** Position Argument ⓘ **Draft:** 1 **Gender:** F **TOEFL:** 77

...kers, Any speakers of English. and he respectively **investigate** to the people and which pronunciation model they will choice. Most English learner sele.....

Term: Fall 2016 **Course:** ENGL 106 ⓘ **Assignment:** Position Argument ⓘ **Draft:** F **Gender:** M **TOEFL:** 77

...osts which is rhetorical questions. First, when we **investigate** Finegan academic reading, we will find rhetorical questions only in the introduction pa.....

Term: Fall 2016 **Course:** ENGL 106 ⓘ **Assignment:** Description and Explanation ⓘ **Draft:** 2 **Gender:** M **TOEFL:** N/A

...osts which is rhetorical questions. First, when we **investigate** Finegan academic reading, we will find rhetorical questions only in the introduction pa.....

Term: Fall 2016 **Course:** ENGL 106 ⓘ **Assignment:** Description and Explanation ⓘ **Draft:** 1 **Gender:** M **TOEFL:** N/A

...osts which is rhetorical questions. First, when we **investigate** Finegan's academic reading, we will find rhetorical questions only in the introduction

Term: Fall 2016 **Course:** ENGL 106 ⓘ **Assignment:** Description and Explanation ⓘ **Draft:** F **Gender:** M **TOEFL:** N/A

...o real letter, we both should talk as friends, and **investigate** details clearly to remind you in the future of moving forward. Together, we can implant.....

Term: Spring 2018 **Course:** ENGL 107 ⓘ **Assignment:** Reflection ⓘ **Draft:** 2 **Gender:** F **TOEFL:** N/A

...o real letter, we both should talk as friends, and **investigate** details clearly to remind you in the future of moving forward. Together, we can implant.....

Term: Spring 2018 **Course:** ENGL 107 ⓘ **Assignment:** Reflection ⓘ **Draft:** 1 **Gender:** F **TOEFL:** N/A

...nian Academy, take it with a smile. Now, we should **investigate** details to remind you in the future a better map and tricks. "I; first person pronoun".....

Term: Spring 2018 **Course:** ENGL 107 ⓘ **Assignment:** Reflection ⓘ **Draft:** F **Gender:** F **TOEFL:** N/A

...t seems too vague, what exactly does that company "**investigate**"? Secondly the visuals, things such as videos and photos and how to properly take them.....

Term: Fall 2017 **Course:** ENGL 107 ⓘ **Assignment:** Genre Analysis ⓘ **Draft:** 2 **Gender:** M **TOEFL:** 94

...t seems too vague, what exactly does that company "**investigate**"? Secondly the visuals, things such as videos and photos and how to properly take them.....

Term: Fall 2017 **Course:** ENGL 107 ⓘ **Assignment:** Genre Analysis ⓘ **Draft:** 1 **Gender:** M **TOEFL:** 94

...t seems too vague, what exactly does that company "**investigate**"? Secondly the visuals, things such as videos and photos and how to properly take them.....

Term: Fall 2017 **Course:** ENGL 107 ⓘ **Assignment:** Genre Analysis ⓘ **Draft:** F **Gender:** M **TOEFL:** 94

...opeans and many other classifications of people to **investigate** if there is a correlation towards their Global Warming viewpoint. My research will prov.....

Term: Spring 2017 **Course:** ENGL 108 ⓘ **Assignment:** Research Proposal ⓘ **Draft:** 1 **Gender:** M **TOEFL:** 96

...echnology. For this research proposal, I intend to **investigate** the uncharted territory of whether or not artificial intelligence will eventually be ab.....



Reflect: *Activity Ideas*



Reflect: Activity Ideas

- Create activities for target vocabulary
- Example sentences or excerpts
- Matching or chart exercise for students
- Comparison with expert data

SECTION: ACADEMIC (5,179)
 FIND SAMPLE: [100](#) [200](#) [500](#) [1000](#)
 PAGE: << < 1 / 52 > >>

| CLICK FOR MORE CONTEXT | | | | <input type="checkbox"/> [?] | SAVE LIST | CHOOSE LIST | CREATE NEW LIST | [?] | SHOW DUPLICATES |
|------------------------|------|------|-------------------------------------|------------------------------|-----------|-------------|---|-----|-----------------|
| 1 | 2019 | ACAD | Business and Economic Horizons | A | B | C | attitudes to development aid, in the article we reconsider the impact of aid and investigate whether the declarations to increase aid effectiveness | | |
| 2 | 2019 | ACAD | PLoS ONE | A | B | C | . They can be compared with the space syntax centrality measures that were proposed to investigate the performance of the introduced metr | | |
| 3 | 2019 | ACAD | PLoS ONE | A | B | C | bias caused by the scarceness of the data. Future research with larger data-sets could investigate the use of shorter time intervals to produce | | |
| 4 | 2019 | ACAD | PeerJ | A | B | C | way different from younger adults is not entirely clear. The present study intended to investigate the age differences in the use of animacy cu | | |
| 5 | 2019 | ACAD | ... Pacific Journal of Reproduction | A | B | C | Nwannenna. Ifeyinwa, Ubah. Azubuike, Ogwu. David # Objective: To investigate the effect of different fractions from the aqueous methanolic | | |
| 6 | 2019 | ACAD | Saudi J Anaesthesia | A | B | C | three key methods for creating a list of potential journal names, and strategies to investigate the details of each. # Your Reference List # The f | | |
| 7 | 2019 | ACAD | Saudi J Anaesthesia | A | B | C | and Scope # Once you have a list of journal titles you will need to investigate further to determine how suitable these journals really are for yc | | |
| 8 | 2019 | ACAD | Media and Communication | A | B | C | in home and host countries. # Amanda Alencar and Vasilki Tsagkroni (2019) investigate integration as a two-way process. In their work " Prosp | | |
| 9 | 2019 | ACAD | ...uter Applications in Archaeology | A | B | C | of 20 experiments were conducted. Several parameters were varied between experiments in order to investigate their influence on the perfor | | |
| 10 | 2019 | ACAD | Construction Building Mat | A | B | C | # 39 JA Canfield and IG Clator Development of a scaling law and techniques to investigate penetration in concrete. NEL Report No 2057, U.S. N | | |
| 11 | 2019 | ACAD | PeerJ | A | B | C | EVI (plant productivity) in conjunction with land cover data will allow us to investigate the productivity-diversity relationship at large spatial sca | | |
| 12 | 2019 | ACAD | PeerJ | A | B | C | the observable impact of the accumulation of local scale effects. This will let us investigate how human modified landscapes might modulate a | | |
| 13 | 2019 | ACAD | PeerJ | A | B | C | a covariate within the model as recommended by Wood (2006). To further investigate the mechanisms underlying the impact of land cover on | | |
| 14 | 2019 | ACAD | PeerJ | A | B | C | pasture across the island of Ireland provides a useful, large-scale opportunity with which to investigate this impact as high levels of nutrient ac | | |
| 15 | 2019 | ACAD | Construction Building Mat | A | B | C | concrete 4,6,9,18-20. # In the last few decades, studies have been carried out to investigate the inhibiting effectiveness of different organic an | | |
| 16 | 2019 | ACAD | Construction Building Mat | A | B | C | solution such as saturated calcium hydroxide solution. Therefore, there is a need to investigate the effect of cation type of chloride salts on co | | |
| 17 | 2019 | ACAD | ...ter Science & Technology | A | B | C | 4. Solution Approach # 4.1 Overview # Our objective in this research is to investigate the suitability of employing semi-supervised learning tec | | |
| 18 | 2019 | ACAD | ...ter Science & Technology | A | B | C | datasets with different distributions would result in the same classification accuracy. We could not investigate this matter further because the | | |
| 19 | 2019 | ACAD | ...ter Science & Technology | A | B | C | sample adequate enough to classify the overall input data. Moreover, future work could investigate developing cluster-based technique for pa | | |
| 20 | 2019 | ACAD | Landscape Architecture Frontiers | A | B | C | daily activity patterns. Amber L. Pearson et al. 26 used wearable cameras to investigate children's daily access to water areas. By analyzing the | | |
| 21 | 2019 | ACAD | Health Services Research | A | B | C | In what follows, we use a difference-in-difference approach and a Poisson count model to investigate the effects of GS policies on drug overdc | | |
| 22 | 2019 | ACAD | Health Services Research | A | B | C | , Table 2). Since GS policies target controlled substance possession, we specifically investigate controlled substances (all and prescription, as ill | | |
| 23 | 2019 | ACAD | European Research Studies | A | B | C | effect among markets has been extensively studied in recent years. Some studies try to investigate the spillover effect across developed mark | | |
| 24 | 2019 | ACAD | European Research Studies | A | B | C | and Politis 2011). Christopoulos et al. (2014) will try to investigate the long-term spillover effect in the European markets represented by Portu | | |
| 25 | 2019 | ACAD | European Research Studies | A | B | C | at all these studies, this research will try to complete the literature review by investigate the relationship of Islamic stock markets in the short | | |
| 26 | 2019 | ACAD | European Research Studies | A | B | C | . # 5. Conclusion # The main objective of this study has been to investigate the spillover effect and correlation among Islamic stock indices in A | | |

CL Tools: Cluster Analysis

- Prefabricated multi-word units
- N-grams/Clusters/Lexical Bundles: words that co-occur

N-grams/Clusters/Lexical Bundles

AntConc 3.5.6 (Windows) 2018

File Global Settings Tool Preferences Help

Corpus Files

- busacleldln052.txt
- busacleldmn048.txt
- busacleudhs057.txt
- busacleudms056.txt
- busacleudms058.txt
- busatlegrmg077.txt
- busatlegrmg084.txt
- busatleudhg013.txt
- busatleudhg016.txt
- busbalegrhn218.txt
- busbalegrhn219.txt
- busbaleldhn045.txt
- busbaleldhn046.txt
- busbaleldmn044.txt
- busbaleldmn054.txt
- busbaleudhn152.txt
- busbaleudhn154.txt
- busbaleudhn170.txt
- busbaleudln049.txt
- busbmleudhg019.txt
- busfilegrhg066.txt
- busfilegrmg002.TXT
- busfileudhg024.txt
- busmglegrhg105.TXT
- busmgleldhg174.txt
- busmgleudhg096.txt
- busmgleudhg104.txt
- busmgleudhg145.txt
- busmgleudhg147.txt
- busmqleudmq114.txt

Concordance Concordance Plot File View Clusters/N-Grams Collocates Word List Keyword List

Total No. of Cluster Types 19 Total No. of Cluster Tokens 25

| Rank | Freq | Range | Cluster |
|------|------|-------|--------------------|
| 1 | 3 | 1 | were asked |
| 2 | 3 | 1 | were married |
| 3 | 2 | 1 | were involved |
| 4 | 2 | 1 | were confused |
| 5 | 1 | 1 | were asked |
| 6 | 1 | 1 | were killed |
| 7 | 1 | 1 | were released |
| 8 | 1 | 1 | were assigned |
| 9 | 1 | 1 | were characterized |
| 10 | 1 | 1 | were evaluated |
| 11 | 1 | 1 | were filed |
| 12 | 1 | 1 | were forced |
| 13 | 1 | 1 | were interested |
| 14 | 1 | 1 | were organized |
| 15 | 1 | 1 | were presented |
| 16 | 1 | 1 | were promised |
| 17 | 1 | 1 | were raised |

Search Term Words Case Regex N-Grams Cluster Size Min. 2 Max. 2

Min. Freq. 1 Min. Range 1

Sort by Invert Order Search Term Position On Left On Right

Sort by Freq

Total No. 36

Files Processed

Clone Results

Corpus Files

- busacleldln052.txt
- busacleldmn048.txt
- busacleudhs057.txt
- busacleudms056.txt
- busacleudms058.txt
- busatlegrmg077.txt
- busatlegrmg084.txt
- busatleudhg013.txt
- busatleudhg016.txt
- busbalegrhn218.txt
- busbalegrhn219.txt
- busbaleldhn045.txt
- busbaleldhn046.txt
- busbaleldmn044.txt
- busbaleldmn054.txt
- busbaleudhn152.txt
- busbaleudhn154.txt
- busbaleudhn170.txt
- busbaleudln049.txt
- busbmleudhg019.txt
- busfilegrhg066.txt
- busfilegrmg002.TXT
- busfileudhg024.txt
- busmglegrhg105.TXT
- busmgleldhg174.txt
- busmgleudhg096.txt
- busmgleudhg104.txt
- busmgleudhg145.txt
- busmgleudhg147.txt
- busmqleudmq114.txt

Total No.
36

Files Processed

Concordance Concordance Plot File View Clusters/N-Grams Collocates Word List Keyword List

Total No. of Cluster Types 19 Total No. of Cluster Tokens 25

| Rank | Freq | Range | Cluster |
|------|------|-------|--------------------|
| 1 | 3 | 1 | were asked |
| 2 | 3 | 1 | were married |
| 3 | 2 | 1 | were involved |
| 4 | 2 | 1 | were confused |
| 5 | 1 | 1 | were asked |
| 6 | 1 | 1 | were killed |
| 7 | 1 | 1 | were released |
| 8 | 1 | 1 | were assigned |
| 9 | 1 | 1 | were characterized |
| 10 | 1 | 1 | were evaluated |
| 11 | 1 | 1 | were filed |
| 12 | 1 | 1 | were forced |
| 13 | 1 | 1 | were interested |
| 14 | 1 | 1 | were organized |
| 15 | 1 | 1 | were presented |
| 16 | 1 | 1 | were promised |
| 17 | 1 | 1 | were raised |

Search Term Words Case Regex N-Grams

were *ed

Advanced

Cluster Size

Min. 2 Max. 2

Start Stop Sort

Min. Freq. Min. Range

1 1

Sort by Invert Order Search Term Position

Sort by Freq On Left On Right

Clone Results

English Spoken Corpora

- London-Lund corpus (Conversation, interviews, news broadcasts)
- Longman corpus (American and British conversation)
- BNC (British conversation and planned speech)
- Hong Kong Corpus of Spoken English (various registers)
- T2K-SWAL (university classroom teaching, office hours)
- MICASE (teaching, office hours, presentations, study groups, service encounters)
- LINDSEI (learner corpus of interviews)
- Switchboard (telephone conversations)
- Santa Barbara Corpus

English Written Corpora

- [Corpus of Journal Articles \(CJA\)](#)
- [Corpus of Historical American English \(COHA\)](#)
- [British Academic Written English Corpus \(BAWE\)](#)
- [British National Corpus \(BNC\)](#)-Spoken and Written
- [Business Letter Corpus](#)
- [Michigan Corpus of Upper-Level Student Papers \(MICUSP\)](#)
- [State of the Union Corpus](#)
- [Time Magazine Corpus](#)
- [Wikipedia Corpus](#)

Corpora in Other Languages

- <https://www.sketchengine.eu/corpora-and-languages/> (Links to an external site.)
- <https://www ldc.upenn.edu/new-corpora> (Links to an external site.)
- <http://fa.persianspeechcorpus.com/>



QUESTIONS?



THE UNIVERSITY OF UTAH

L2TReC Webinar Series

Creating Data-Driven Pedagogical Materials Using Learner Corpora: A Guide for Language Teachers

Speaker: Dr. Elnaz Kia

Thursday, May 6, 2021, 4:00-6:00 PM MST

Email: elnaz.kia@utah.edu



L2TReC

**SECOND LANGUAGE
TEACHING & RESEARCH CENTER**

THE UNIVERSITY OF UTAH