A Corpus Analysis of Plesionyms *Attribute*, *Characteristic*, and *Trait* in Academic English

*Sadie M. M. Nitta**
Hawaii Pacific University

Abstract
This paper investigates the plesionyms attribute, characteristic, and trait using the Corpus of Contemporary American English (COCA). It explores their frequency and variation across COCA's six registers and sub-registers, then examines their grammatical behaviors and semantic preference and prosody through concordancing in COCA's academic journals (ACAD). The results show that attribute mostly occurs in academic journals related to business, science/technology, and education; characteristic mostly occurs in academic journals related to business, education, and medicine; and trait mostly occurs in academic journals related to science/technology, philosophy/religion, and business. Noun collocations associated with each term suggest a strong preference for standards/quality with positive semantic prosody (*attribute*), demographics/species/property with neutral semantic prosody (*characteristic*), and cognition/behavior/genetics with neutral semantic prosody (*trait*). The paper concludes with a set of materials for one corpus-informed lesson following the Illustration-Induction-Interaction (I-I-I) model (Timmis, 2015).

Introduction

Differentiating between plesionyms, or near-synonyms, that frequently occur in Academic English (AE) is useful for lexical choice in discourse. Although there is a consensus among linguists that true synonymy is "quite rare" (Hirst, 1995, p. 51), dictionaries and thesauri often characterize such words as identical in meaning, which can be "very misleading" to language learners (Biber et al., 2007, p. 43). This paper aims to investigate the plesionyms—more specifically, noun near-synonyms – *attribute, characteristic,* and *trait* using the Corpus of Contemporary American English (COCA). First, it identifies their frequency and variation across COCA's six registers and sub-registers, then assesses their (a) grammatical behaviors and (b) semantic preference and prosody across COCA's academic journals (ACAD) via concordancing. Finally, it provides a set of materials designed for one corpus-informed lesson following the Illustration-Induction-Interaction (I-I-I) model (Timmis, 2015), including English as a Second Language (ESL) teaching implications in academic prose.

The Target Nouns: What We Already Know About Their Form, Meaning, and Use

According to *Longman Student Grammar of Spoken and Written English* (Biber et al., 2002), nouns can be grouped into a small number of classes – concrete, abstract, countable, uncountable, proper, common, and package – which differ in meaning and grammatical behavior (p. 56). In this paper, the target nouns under examination are *attribute, characteristic,* and *trait*.

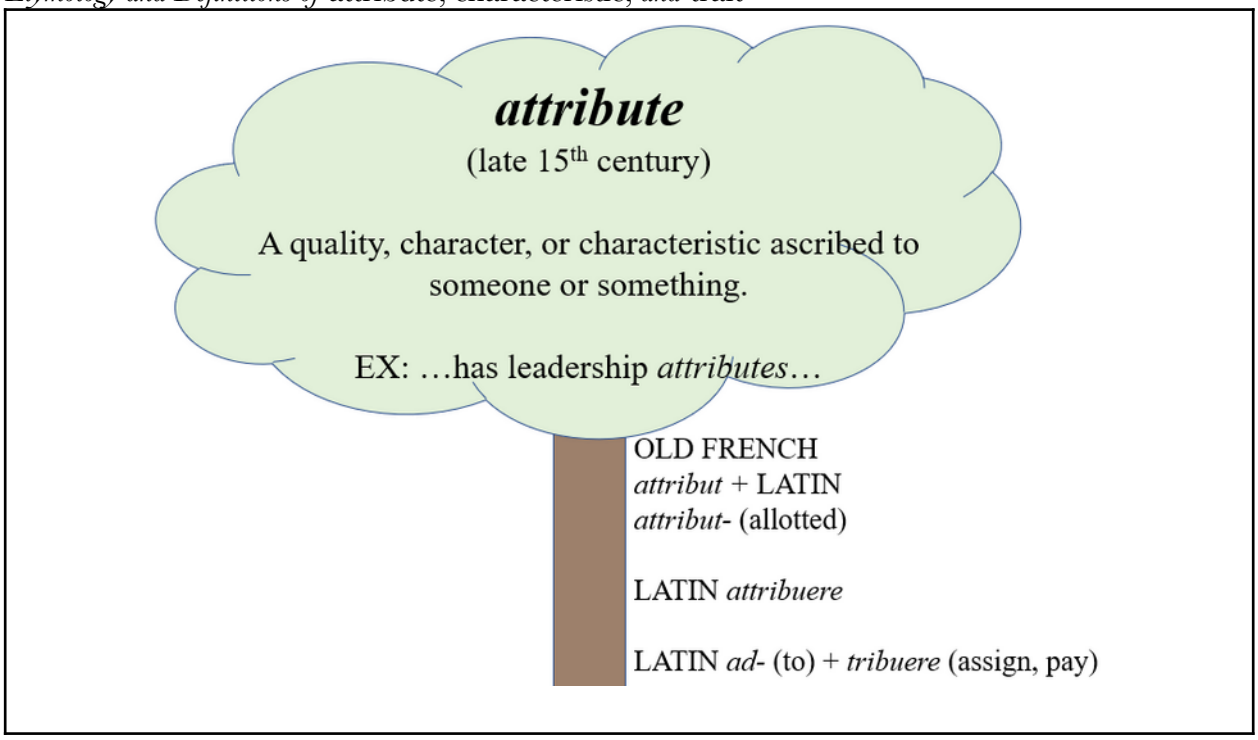
Attribute, characteristic, and *trait* represent common countable nouns, which account for singular and plural forms. Furthermore, the target nouns are considered abstract nouns, which refer to abstractions such as events, states, times, and qualities. According to Biber et al. (2002), "The

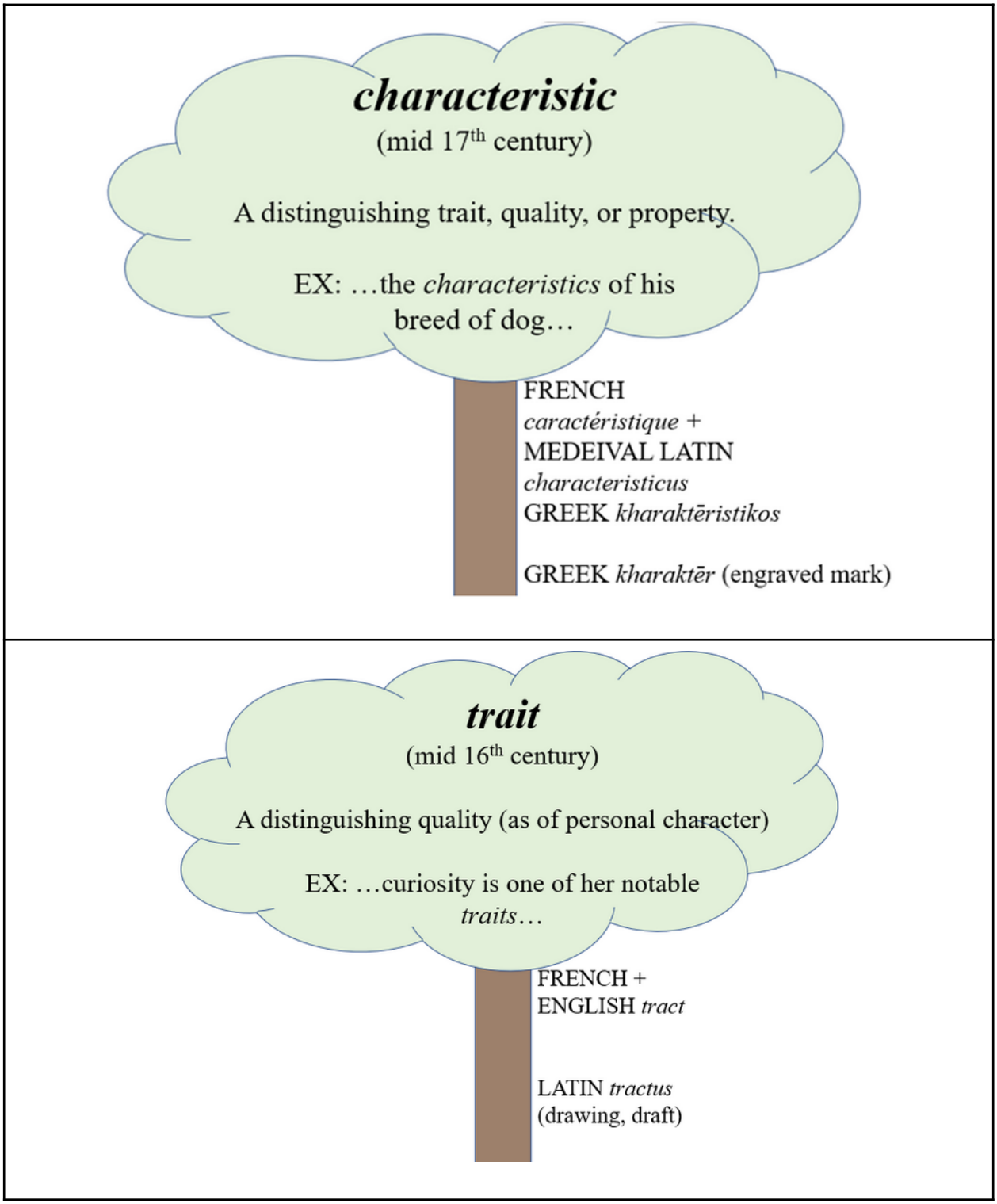
creative commons
Nitta, S. M. M. (2023). A corpus analysis of plesionyms *attribute, characteristic*, and *trait* in academic English. *TESOL Working Paper Series, 21,* 80-120.
Website: Hawaii Pacific University http://www.hpu.edu
*Email: snitta1@my.hpu.edu. Address: 1 Aloha Drive, Honolulu, HI 96813, USA.

distinction between concrete and abstract is purely semantic: it has no real grammatical role, since abstract nouns, like concrete nouns, can be countable, uncountable, common, or proper" (p. 57). So far, there is no strong distinction between the forms of nouns *attribute, characteristic,* and *trait*.

Looking closely at the origins and meanings of nouns *attribute, characteristic,* and *trait* is important for setting up the investigation. Based on *Merriam-Webster.com*, Figure 1 displays their etymology, dictionary definitions, and example phrases ("attribute," 2022; "characteristic," 2022; "trait," 2022). Although *characteristic* derives from Greek and *attribute* and *trait* derive from Latin, all three target nouns entered English through French. Their core meanings refer to the 'quality' of something, and the dictionary definitions are insufficient to differentiate them for language use, which necessitates a corpus analysis. In addition, the *Merriam-Webster.com* examples given for *characteristic* and *trait* show that they both can be used as "species noun" (Biber et al., 2002, p. 64) – often followed by an *of*-phrase to refer to the type of something. This background knowledge on these words will provide greater insight to this paper's investigation and findings.

Figure 1
Etymology and Definitions of attribute, characteristic, *and* trait




characteristic
(mid 17th century)
A distinguishing trait, quality, or property.
EX: …the characteristics of his breed of dog…
FRENCH
caractéristique +
MEDEIVAL LATIN
characteristicus
GREEK kharaktēristikos
GREEK kharaktēr (engraved mark)
trait
(mid 16th century)
A distinguishing quality (as of personal character)
EX: …curiosity is one of her notable traits…
FRENCH +
ENGLISH tract
LATIN tractus
(drawing, draft)

The Phenomenon of Plesionymy (or Near-Synonymy)

Plesionymy, or near-synonymy, refers to two or more words that are close in meaning but "not fully inter-substitutable" (Hirst, 1995, p. 51). Plesionyms can vary in more than one way, such as their connotation, denotation, emphasis, implicature, register, style, or use (Cruse, 1986; DiMarco et al., 1993; Edmonds, 1999). According to Hirst (1995), "It is difficult even for native speakers of a language to command the differences between plesionyms well enough to use them with invariable precision, or to articulate those differences even when they are known" (p. 51). As a result, lexical choice and transfer issues are common for language learners.

Past research can help determine what kinds of differences exist between plesionyms. For example, Gove (1984) analyzed the noun plesionyms *lie, misrepresentation, fib, falsehood,* and *untruth* in context. Gove's (1984) analysis revealed the differences in focus and register amongst these plesionyms (e.g., *misrepresentation* and *untruth* are used euphemistically, whereas *lie* is deliberate and direct; *fib* is used informally and seen as childish, whereas *falsehood* is formal). In another example, Gove (1984) discovered that the difference between the noun plesionyms *enemy* and *foe* is that *enemy* emphasizes antagonism or hatred while *foe* emphasizes active fighting. Room (1985) determined specific differences between the noun plesionyms *forest* and *wood(s)*. According to Room (1985), "A 'wood' is smaller than a 'forest,' is not so primitive, and is usually nearer to civilization. This means that a 'forest' is fairly extensive, is to some extent wild, and on the whole not near large towns or cities. In addition, a 'forest' often has game or wild animals in it, which a 'wood' does not, apart from the standard quota of regular rural denizens such as rabbits, foxes and birds of various kinds…" (p. 270). These differences are context-driven and weigh heavily on one's lexical choice.

Issues related to lexical choice and first language (L1) transfer are quite similar for language learners. Language learners are likely to choose a word in the target language that is closest to a word in their first language. Often, these L1 transfers show examples of plesionyms, not true synonyms (Hirst, 1995). For instance, the German word *Wald* is part of the cross-linguistic plesionym group *forest, wood(s),* and *Wald.* According to Hirst (1995), "the German word *Wald* is close in meaning to the English word *forest*, but *Wald* can denote a rather smaller and more urban area of trees than *forest*" (p. 52). In contrast to Room's (1995) analysis of the plesionyms *forest* and *wood(s)*, it is clear that in some situations, *woods* would be a better translation of *Wald* than *forest.* Meaning differences between plesionyms can be targeted using four broad categories based on Cruse (1986) and Edmond's (1999) research: denotational variation, stylistic variation (including dialect and register), expressive variation (including emotive and attitudinal aspects), and structural variation (including collocational, selectional, and syntactic variations).

Data Driven Learning in Second Language Acquisition and English Language Teaching

The theoretical foundation for Data Driven Learning (DDL) in Second Language Acquisition (SLA) and teaching is the proven value of examining naturally occurring language in context (authenticity) and promoting self-sufficiency in language learning (autonomy) (Timmis, 2015; Friginal, 2018; Green, 2019). One advantage of DDL is the quantity and quality of language evidence, which enables teachers and learners to write comprehensive descriptions and confirm, refute, or refine a language hypothesis (Timmis, 2015). In this light, learners truly become investigators of the language as the teacher guides them through the generalization process towards independent learning.

Through DDL, corpora are recognized as a special type of informant, not a tutor or surrogate teacher (Johns, 1991). Teacher guidance is necessary in the training and initial exploration of the computing tool. Friginal (2018) and other experts suggest that teachers should take on the role of

director and coordinator in DDL. Furthermore, Timmis (2015) states that "Corpus-informed teachers, arguably, will be better placed to make opportunistic decisions in the [ESL] classroom about what vocabulary to highlight from listening or reading texts or from classroom discussions" (p. 53). From this standpoint, DDL can result in memorable learning and teaching moments.

Although there are notable challenges with the implementation of corpus technology, past research has outlined several pedagogical benefits. Most famously, Johns and King (1991) published three primary conclusions — based on exploratory case studies with students (overseas postgraduates in a wide range of subject areas who need to improve their English for instrumental reasons) — about students using concordance output regularly:

1. Using concordancing, learners can develop the ability to see patterning in the target language and to form generalizations to account for that patterning.
2. The teacher has to learn to become a director and coordinator of student-initiated research.
3. Traditional grammar teaching is top-down; DDL is bottom-up.

Concerning the first conclusion, O'Keeffe et al. (2007) recommend creating a lexico-grammatical profile for target vocabulary via concordancing. This corpus strategy leads learners towards making their own generalizations by inviting them to explore a target words' collocates, chunks/idioms, syntactic restrictions, semantic restrictions, prosody, and other relevant or recurring features (pp. 14-15). Due to the overwhelming amount of data, teachers can also use guiding questions with varying degrees of control to help familiarize students with the generalization process. Although there is hesitation amongst teachers to incorporate corpus technology in the classroom, perception studies have shown that "many students enjoy DDL as a novel learning approach" (Vyatkina & Boulton, 2017, p. 1).

Research Questions

1. How frequently are the nouns *attribute, characteristic,* and *trait* used in different registers? How frequently are they used in different sub-registers?
2. What are the differences in the semantic preferences of nouns *attribute, characteristic,* and *trait* in Academic English?
3. What are the differences in the semantic prosody of nouns *attribute, characteristic,* and *trait* in Academic English?

Corpus

Created by Mark Davies (2008-), the Corpus of Contemporary American English (COCA) is a principled collection of natural texts, spanning over 1 billion words — 1,001,610,938, to be exact — and divided into eight registers (or genres): blogs (125 million words [125,496,215]); web pages (130 million words [129,899,426]); TV/movie subtitles (128 million words [128,013,334]); fiction (120 million words [127,352,014]); popular magazines (127 million words [127,352,014]); newspapers (123 million words [122,959,393]); and academic journals (121 million words [120,9388,348]). On the one hand, due to its large database, it is known as a general or reference corpus. On the other hand, due to its ever-increasing database, it is also known as a monitor corpus because users can 'monitor' how the language is used and changes over time. For this paper, the sub-registers of COCA's academic journals (ACAD) or Academic English (AE) were also explored. According to Davies (2008-), COCA's ACAD/AE sub-registers include nearly 100 peer-reviewed journals selected from a balance of history, education, geography/social science, law/political science, humanities, philosophy/religion,

science/technology, business, and miscellaneous. See https://www.english-corpora.org/coca/ for more information.

Analysis

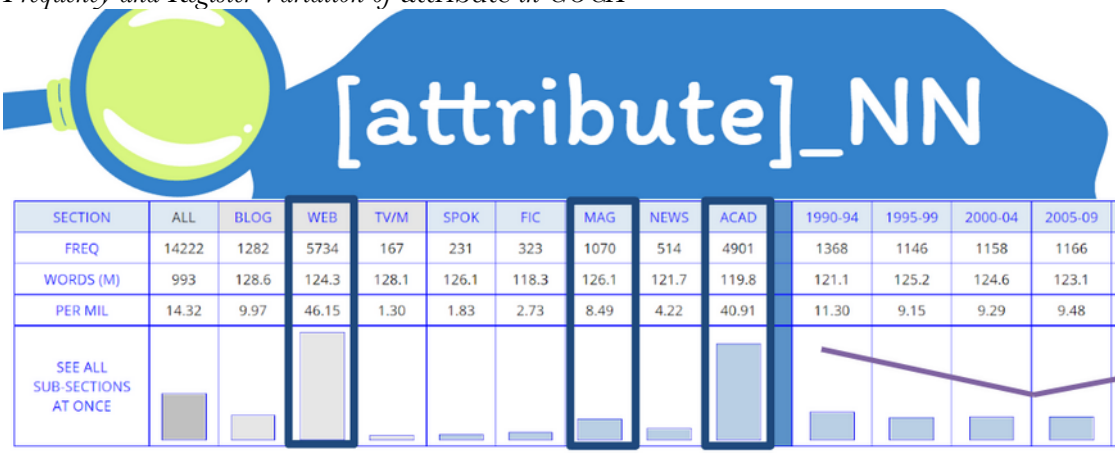
Research Question 1: How frequently are the nouns *attribute, characteristic,* and *trait* used in different registers? How frequently are they used in different sub-registers?

To account for all target forms of the nouns *attribute, characteristic,* and *trait*, the search string utilized was [attribute]_NN, [characteristic]_NN, and [trait]_NN in COCA.

Using the CHART function in COCA, the overall frequency per million of *attribute* is 14.32 (see Figure 2). For *characteristic*, the overall frequency per million is 38.11 (see Figure 3); and for *trait*, 18.63 (see Figure 4). This means that amongst these three nouns, *characteristic* is the most frequent, then *trait* and *attribute*. In addition, Figures 2 through 4 display the variation of *attribute, characteristic,* and *trait* over time (between 1990-2019). It appears that *trait* shows stronger signs of an upward trend (from 15.02 per mil in 1990-94 to 18.88 per mil in 2015-19).

Figure 2

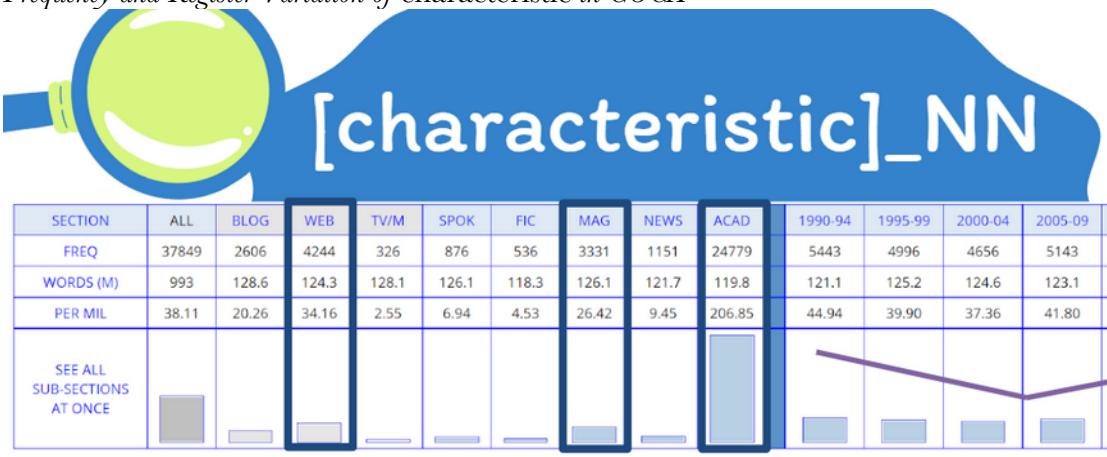
Frequency and Register Variation of attribute *in COCA*



SECTION	ALL	BLOG	WEB	TV/M	SPOK	FIC	MAG	NEWS	ACAD	1990-94	1995-99	2000-04	2005-09	2010-14	2015-19
FREQ	14222	1282	5734	167	231	323	1070	514	4901	1368	1146	1158	1166	1008	1360
WORDS (M)	993	128.6	124.3	128.1	126.1	118.3	126.1	121.7	119.8	121.1	125.2	124.6	123.1	123.3	122.8
PER MIL	14.32	9.97	46.15	1.30	1.83	2.73	8.49	4.22	40.91	11.30	9.15	9.29	9.48	8.17	11.08
SEE ALL SUB-SECTIONS AT ONCE															

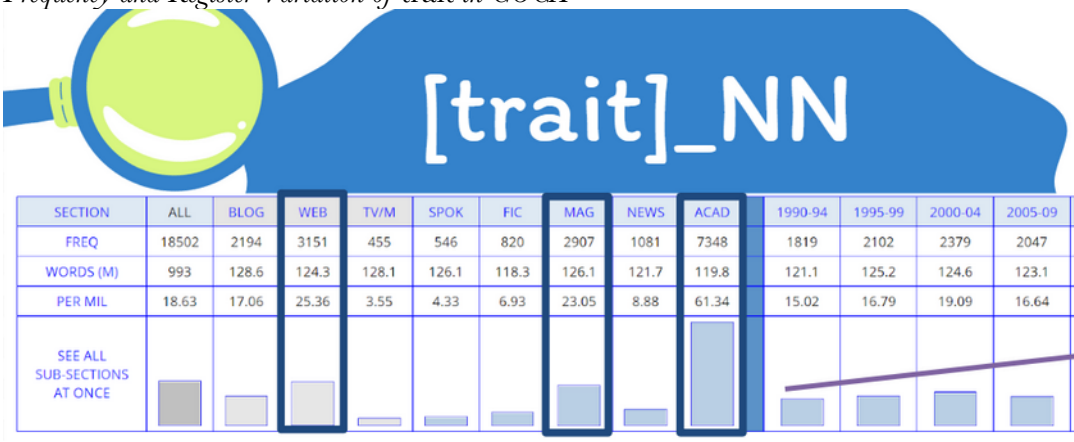
Figure 3

Frequency and Register Variation of characteristic *in COCA*



SECTION	ALL	BLOG	WEB	TV/M	SPOK	FIC	MAG	NEWS	ACAD	1990-94	1995-99	2000-04	2005-09	2010-14	2015-19
FREQ	37849	2606	4244	326	876	536	3331	1151	24779	5443	4996	4656	5143	5706	5055
WORDS (M)	993	128.6	124.3	128.1	126.1	118.3	126.1	121.7	119.8	121.1	125.2	124.6	123.1	123.3	122.8
PER MIL	38.11	20.26	34.16	2.55	6.94	4.53	26.42	9.45	206.85	44.94	39.90	37.36	41.80	46.26	41.18
SEE ALL SUB-SECTIONS AT ONCE															

Figure 4
Frequency and Register Variation of trait *in COCA*



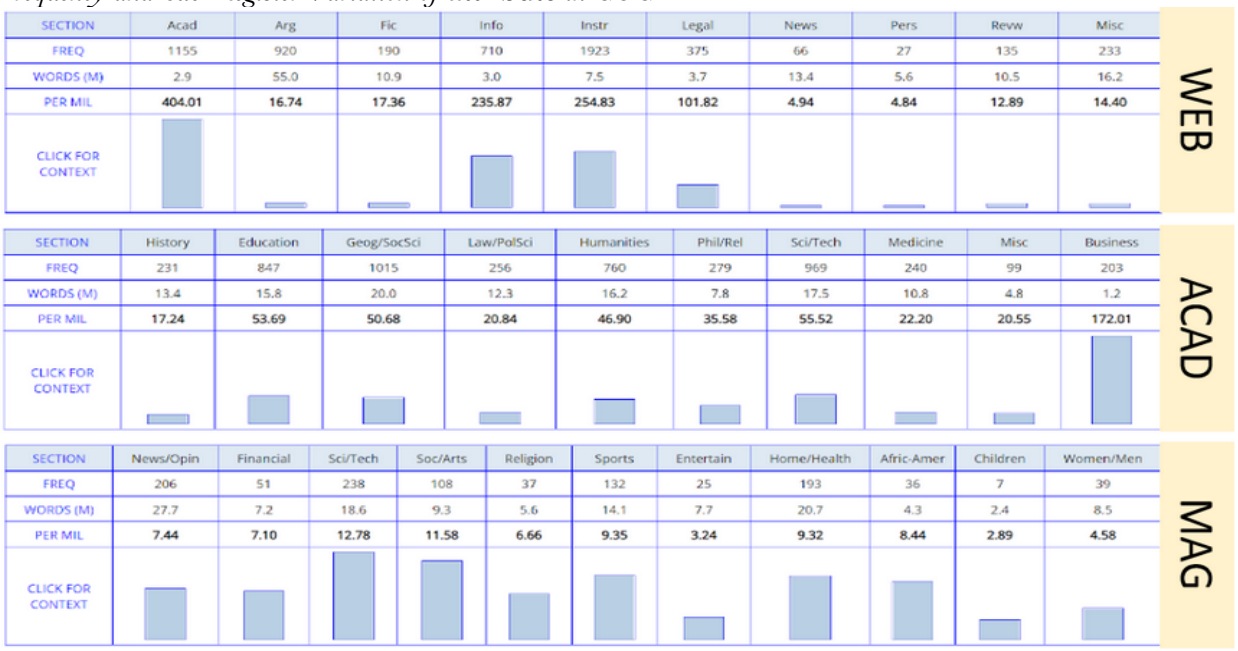
SECTION	ALL	BLOG	WEB	TV/M	SPOK	FIC	MAG	NEWS	ACAD	1990-94	1995-99	2000-04	2005-09	2010-14	2015-19
FREQ	18502	2194	3151	455	546	820	2907	1081	7348	1819	2102	2379	2047	2493	2317
WORDS (M)	993	128.6	124.3	128.1	126.1	118.3	126.1	121.7	119.8	121.1	125.2	124.6	123.1	123.3	122.8
PER MIL	18.63	17.06	25.36	3.55	4.33	6.93	23.05	8.88	61.34	15.02	16.79	19.09	16.64	20.21	18.88
SEE ALL SUB-SECTIONS AT ONCE															

Returning to COCA's initial CHART search, one may also notice the high frequency registers for *attribute, characteristic,* and *trait* ([attribute]_NN, [characteristic]_NN, and [trait]_NN). Interestingly, they share the top three registers but in different order: *Attribute* mostly occurs in web pages (46.15 per mil), academic journals (40.91 per mil), and popular magazines (8.49 per mil). *Characteristic* mostly occurs in academic journals (206.85 per mil), web pages (34.16 per mil), and popular magazines (26.42 per mil). Lastly, *trait* mostly occurs in academic journals (61.34 per mil), web pages (25.36 per mil), and popular magazines (23.05 per mil). Something extremely interesting about this finding is that both *characteristic* and *trait* are more frequently used in academic journals than *attribute*. This result is reflective of the change between Coxhead's (2000) Academic Word List (AWL) and Gardner and Davies' (2013) Academic Vocabulary List (AWL): *Attribute* is included in the AWL but not included in AVL. Coxhead's (2000) AWL is limited to textbooks, whereas Gardner and Davies' (2013) AVL is larger and broader.

Next, the top sub-register of the top three registers were identified. The top sub-registers revealed a general tendency for academic themes (see Figures 5-7). For *attribute*, the top sub-registers include academic web pages (404.01 per mil), business academic journals (172.01 per mil), and science/technology popular magazines (12.78 per mil). For *characteristic*, the top sub-registers include business academic journals (401.63 per mil), academic web pages (190.99 per mil), and science/technology popular magazines (51.88 per mil). For *trait*, the top sub-registers include science/technology academic journals (129.48 per mil), academic web pages (101.09 per mil), and science/technology popular magazines (53.92 per mil). This further illustrates why these nouns are worth noting in Academic English (AE) and how they differ across registers (Cruse, 1986; DiMarco et al., 1993; Edmonds, 1999).

Figure 5

Frequency and Sub-Register Variation of attribute *in COCA*



SECTION	Acad	Arg	Fic	Info	Instr	Legal	News	Pers	Rewv	Misc
FREQ	1155	920	190	710	1923	375	66	27	135	233
WORDS (M)	2.9	55.0	10.9	3.0	7.5	3.7	13.4	5.6	10.5	16.2
PER MIL	404.01	16.74	17.36	235.87	254.83	101.82	4.94	4.84	12.89	14.40
CLICK FOR CONTEXT										

WEB

SECTION	History	Education	Geog/SocSci	Law/PolSci	Humanities	Phil/Rel	Sci/Tech	Medicine	Misc	Business
FREQ	231	847	1015	256	760	279	969	240	99	203
WORDS (M)	13.4	15.8	20.0	12.3	16.2	7.8	17.5	10.8	4.8	1.2
PER MIL	17.24	53.69	50.68	20.84	46.90	35.58	55.52	22.20	20.55	172.01
CLICK FOR CONTEXT										

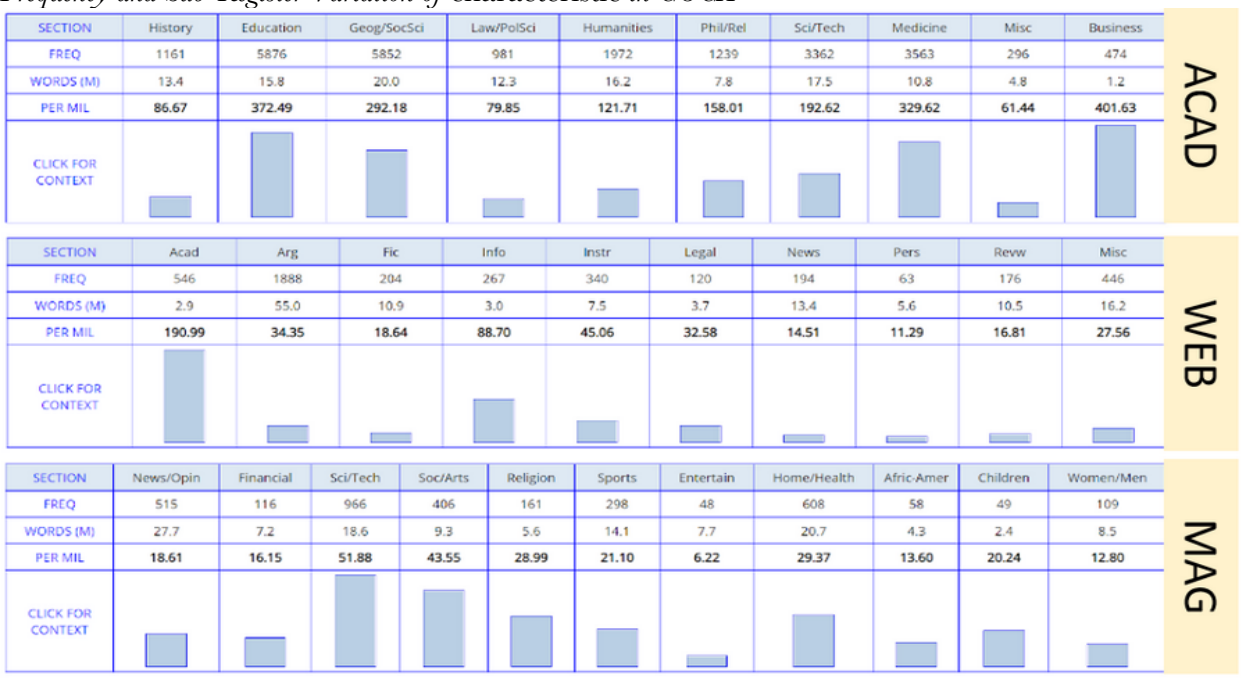
ACAD

SECTION	News/Opin	Financial	Sci/Tech	Soc/Arts	Religion	Sports	Entertain	Home/Health	Afric-Amer	Children	Women/Men
FREQ	206	51	238	108	37	132	25	193	36	7	39
WORDS (M)	27.7	7.2	18.6	9.3	5.6	14.1	7.7	20.7	4.3	2.4	8.5
PER MIL	7.44	7.10	12.78	11.58	6.66	9.35	3.24	9.32	8.44	2.89	4.58
CLICK FOR CONTEXT											

MAG

Figure 6

Frequency and Sub-Register Variation of characteristic *in COCA*



SECTION	History	Education	Geog/SocSci	Law/PolSci	Humanities	Phil/Rel	Sci/Tech	Medicine	Misc	Business
FREQ	1161	5876	5852	981	1972	1239	3362	3563	296	474
WORDS (M)	13.4	15.8	20.0	12.3	16.2	7.8	17.5	10.8	4.8	1.2
PER MIL	86.67	372.49	292.18	79.85	121.71	158.01	192.62	329.62	61.44	401.63
CLICK FOR CONTEXT										

ACAD

SECTION	Acad	Arg	Fic	Info	Instr	Legal	News	Pers	Rewv	Misc
FREQ	546	1888	204	267	340	120	194	63	176	446
WORDS (M)	2.9	55.0	10.9	3.0	7.5	3.7	13.4	5.6	10.5	16.2
PER MIL	190.99	34.35	18.64	88.70	45.06	32.58	14.51	11.29	16.81	27.56
CLICK FOR CONTEXT										

WEB

SECTION	News/Opin	Financial	Sci/Tech	Soc/Arts	Religion	Sports	Entertain	Home/Health	Afric-Amer	Children	Women/Men
FREQ	515	116	966	406	161	298	48	608	58	49	109
WORDS (M)	27.7	7.2	18.6	9.3	5.6	14.1	7.7	20.7	4.3	2.4	8.5
PER MIL	18.61	16.15	51.88	43.55	28.99	21.10	6.22	29.37	13.60	20.24	12.80
CLICK FOR CONTEXT											

MAG

Figure 7
Frequency and Sub-Register Variation of trait *in COCA*

SECTION	History	Education	Geog/SocSci	Law/PolSci	Humanities	Phil/Rel	Sci/Tech	Medicine	Misc	Business
FREQ	288	969	1303	188	818	893	2260	383	121	123
WORDS (M)	13.4	15.8	20.0	12.3	16.2	7.8	17.5	10.8	4.8	1.2
PER MIL	21.50	61.43	65.06	15.30	50.48	113.88	129.48	35.43	25.12	104.22
CLICK FOR CONTEXT										

ACAD

SECTION	Acad	Arg	Fic	Info	Instr	Legal	News	Pers	Revw	Misc
FREQ	289	1458	191	115	97	15	93	75	232	586
WORDS (M)	2.9	55.0	10.9	3.0	7.5	3.7	13.4	5.6	10.5	16.2
PER MIL	101.09	26.52	17.45	38.20	12.85	4.07	6.96	13.44	22.16	36.21
CLICK FOR CONTEXT										

WEB

SECTION	News/Opin	Financial	Sci/Tech	Soc/Arts	Religion	Sports	Entertain	Home/Health	Afric-Amer	Children	Women/Men
FREQ	497	95	1004	212	83	244	76	403	87	47	161
WORDS (M)	27.7	7.2	18.6	9.3	5.6	14.1	7.7	20.7	4.3	2.4	8.5
PER MIL	17.96	13.23	53.92	22.74	14.95	17.28	9.85	19.47	20.40	19.42	18.91
CLICK FOR CONTEXT											

MAG

Focusing on COCA's academic journals (ACAD) register, the most frequent sub-registers of *attribute, characteristic,* and *trait* begin to reveal the nature and differences between these near-synonyms. As shown in Figures 5 through 7, *attribute* mostly occurs in academic journals related to business (172.01 per mil), science/technology (55.52 per mil), and education (53.69 per mil); *characteristic* mostly occurs in academic journals related to business (401.63 per mil), education (372.49 per mil), and medicine (329.62 per mil); *trait* mostly occurs in academic journals related to science/technology (129.48 per mil), philosophy/religion (113.88 per mil), and business (104.22 per mil). In all, the most frequent academic journals across all three are business, science/technology, and education. These are the sub-registers that will be explored in the next set of questions.

Research Question 2: What are the differences in the semantic preferences of nouns *attribute, characteristic,* and *trait* in Academic English?

Using the COLLOCATES function in COCA, the top 30 collocates of *attribute* ([attribute_NN]), *characteristic* ([characteristic]_NN), and *trait* ([trait]_NN) were analyzed in the left (1L) and right (1R) contexts in COCA's academic journals (see Tables 1-2). Then, using the same search string, the top 15 verb, noun[1], adjective, and adverb collocates were identified separately in the left (1L) and right (1R) contexts in COCA's academic journals (see Tables 3-12). Based on this initial search, the semantic preferences among *attribute, characteristic,* and *trait* begin to emerge.

1 Noun collocates required further investigation. Using the same search string, the top 15 noun collocates were identified in the left (1L, 6L) and right (1L, 6R) contexts in COCA's academic journals (see Tables 5-8).

Table 1
Top 30 Collocates (1L) of attribute, characteristic, *and* trait *in COCA's Academic Journals*

#	*attribute*	*characteristic*	*trait*
1	*the* (563)	***the*** (2,300)	*the* (556)
2	*of* (170)	***demographic*** (1,008)	***personality*** (476)
3	*these* (159)	***these*** (623)	*of* (296)
4	*an* (158)	***and*** (481)	*and* (204)
5	*physical* (117)	**.** (386)	*a* (202)
6	*personal* (97)	***personal*** (351)	*these* (196)
7	*and* (95)	***of*** (349)	***character*** (173)
8	*other* (94)	***physical*** (325)	*,* (121)
9	***positive*** (77)	**,** (292)	*this* (119)
10	*important* (57)	***student*** (291)	***type*** (91)
11	***'s*** (53)	*personality* (281)	***cultural*** (89)
12	***each*** (49)	***a*** (253)	***quantitative*** (67)
13	*,* (48)	***other*** (248)	***that*** (65)
14	***those*** (41)	***individual*** (216)	*other* (63)
15	*individual* (40)	***common*** (206)	. (55)
16	***negative*** (40)	***unique*** (203)	*personal* (55)
17	***key*** (39)	***defining*** (195)	*common* (53)
18	***different*** (38)	***important*** (170)	***between*** (51)
19	***three*** (37)	***sociodemographic*** (170)	***for*** (50)
20	*cultural* (36)	***sample*** (160)	*behavioral* (49)
21	*"* (33)	***specific*** (157)	*positive* (49)
22	***its*** (33)	***background*** (145)	*physical* (47)
23	*or* (31)	***essential*** (136)	***quality*** (45)
24	*'* (28)	**'** (130)	***in*** (44)
25	***creative*** (27)	***behavioral*** (130)	***or*** (44)
26	*human* (26)	***baseline*** (127)	**"** (43)
27	***job*** (26)	***certain*** (123)	***human*** (43)
28	***their*** (26)	***similar*** (123)	***with*** (43)
29	*with* (26)	***clinical*** (122)	***phenotypic*** (42)
30	*specific* (25)	***this*** (121)	*certain* (40)

Note. Distinct collocates are in bold, while overlapping collocates occurring at different frequencies are both underlined and bolded in the column with the highest frequency.

As shown in Table 1, *attribute, characteristic,* and *trait* each tend to follow different types of adjectives more than other parts of speech. Given the predicted contexts in which one uses *attribute, characteristic,* and *trait* – when seeking to describe or compare something – this finding makes sense. (The most frequent adjective collocates will be explored further in Tables 9 and 10.)

Table 2
Top 30 Collocates (1R) of attribute, characteristic, *and* trait *in COCA's Academic Journals*

#	*attribute*	*characteristic*	*trait*
1	*of* (1,213)	***of*** (9,068)	*,* (715)
2	. (509)	**,** (2,328)	. (622)
3	*,* (506)	**.** (2,015)	*of* (621)
4	*and* (271)	***and*** (1,577)	*and* (433)
5	*that* (199)	***that*** (965)	*that* (432)
6	*(* (189)	**(** (940)	*in* (356)
7	*are* (131)	***are*** (491)	***anxiety*** (309)
8	*in* (114)	***in*** (432)	*(* (212)
9	*to* (82)	***were*** (290)	*are* (184)
10	*were* (78)	**:** (286)	***is*** (104)
11	*such* (68)	***such*** (279)	*to* (89)
12	*is* (56)	***for*** (236)	*or* (88)
13	*or* (56)	***as*** (232)	*such* (87)
14	*as* (52)	***to*** (223)	***arousability*** (79)
15	*:* (51)	***is*** (187)	*:* (76)
16	*for* (47)	***or*** (181)	*)* (73)
17	*)* (38)	***with*** (174)	*for* (70)
18	*was* (35)	***on*** (171)	*as* (67)
19	*can* (34)	***associated*** (161)	*were* (61)
20	*;* (32)	**)** (111)	*may* (57)
21	*which* (26)	***may*** (108)	*can* (56)
22	*have* (25)	**;** (101)	*with* (53)
23	*"* (22)	**"** (94)	***values*** (49)
24	***agenda*** (21)	***have*** (91)	*associated* (47)
25	*with* (20)	***which*** (90)	*was* (39)
26	*may* (19)	***can*** (86)	***like*** (38)
27	*associated* (18)	***by*** (82)	*;* (37)
28	***used*** (17)	***from*** (77)	*from* (36)
29	*on* (16)	***was*** (72)	***anger*** (31)
30	*by* (15)	***related*** (61)	***dominance*** (30)

Note. Distinct collocates are in bold, while overlapping collocates occurring at different frequencies are both underlined and bolded in the column with the highest frequency.

As shown in Table 2, *attribute, characteristic,* and *trait* each tend to participate in *of*-phrases. This suggests that these target nouns are often associated with the concept or entity being described. It is worth noting this before looking at high frequency noun collocates in the right contexts (Tables 6, 8).

As shown in Tables 3 and 4, *attribute* tends to occur with action verbs (e.g., *assigned*), *characteristic* with communication verbs (e.g., *describe*), and *trait* with mental verbs (e.g., *coping*). This suggests a subtle difference in the real-world processes these nouns engage in respectively.

Table 3
Top 15 Verb Collocates of attribute, characteristic, *and* trait *in the Left Context (1L) in ACAD*

#	*attribute*	*characteristic*	*trait*
1	*identifying* (5)	***defining*** (30)	***yield*** (8)
2	*identify* (5)	***identify*** (29)	***contributing*** (6)
3	***possess*** (3)	***sample*** (23)	***hatching*** (6)
4	***correlated*** (3)	***describe*** (17)	*identify* (6)
5	***stressing*** (3)	***coaching*** (14)	***inherited*** (5)
6	***contained*** (2)	***setting*** (13)	***share*** (5)
7	***contain*** (2)	***examined*** (12)	***assess*** (4)
8	***assigned*** (2)	***identifying*** (12)	*defining* (4)
9	***graduate*** (2)	***learning*** (11)	***measuring*** (4)
10	*defining* (2)	***exhibit*** (10)	***possessed*** (4)
11	***desired*** (2)	***need*** (10)	***preweaning*** (4)
12	***manifest*** (2)	***predisposing*** (10)	***involving*** (3)
13	***emphasize*** (2)	***housing*** (9)	*include* (3)
14	***included*** (2)	***include*** (9)	*identifying* (3)
15	***manipulate*** (2)	***describes*** (7)	***controlled*** (3)

Note. Distinct collocates are in bold, while overlapping collocates occurring at different frequencies are both underlined and bolded in the column with the highest frequency.

Table 4
Top 15 Verb Collocates of attribute, characteristic, *and* trait *in the Right Context (1R) in ACAD*

#	*attribute*	*characteristic*	*trait*
1	*associated* (18)	***associated*** (161)	*associated* (47)
2	*related* (13)	***related*** (61)	*related* (17)
3	***used*** (12)	***include*** (53)	*found* (11)
4	*include* (9)	***included*** (33)	*measured* (11)
5	*needed* (10)	***found*** (29)	***coping*** (10)
6	*included* (9)	***make*** (26)	*include* (10)
7	*identified* (8)	***affect*** (25)	***affecting*** (8)
8	***considered*** (7)	***identified*** (25)	*shared* (8)
9	*found* (5)	***influence*** (21)	***using*** (7)
10	***emphasized*** (5)	***listed*** (21)	***appear*** (6)
11	*listed* (5)	***needed*** (18)	***known*** (6)
12	*affect* (4)	***measured*** (16)	***tend*** (6)
13	*based* (4)	***shared*** (16)	***described*** (5)
14	*make* (4)	***examined*** (15)	***conferring*** (5)
15	***presented*** (4)	***based*** (13)	*considered* (5)

Note. Distinct collocates are in bold, while overlapping collocates occurring at different frequencies are both underlined and bolded in the column with the highest frequency.

Table 5
Top 15 Noun Collocates of attribute, characteristic, *and* trait *in the Left Context (1L) in ACAD*

#	*attribute*	*characteristic*	*trait*
1	*job* (26)	***student*** (291)	***personality*** (476)
2	*personality* (20)	*personality* (281)	***character*** (173)
3	***problem*** (20)	***background*** (145)	***type*** (91)
4	***product*** (20)	***sample*** (137)	***quality*** (45)
5	***parliament*** (15)	***baseline*** (127)	***leadership*** (34)
6	***college*** (14)	***participant*** (110)	***growth*** (30)
7	*student* (13)	***teacher*** (97)	***history*** (25)
8	***cabinet*** (12)	***family*** (90)	***temperament*** (23)
9	*teacher* (12)	***school*** (73)	***conformation*** (21)
10	*quality* (9)	***performance*** (70)	***production*** (18)
11	*type* (9)	***learner*** (62)	***resistance*** (16)
12	***criteria*** (8)	***population*** (61)	***idol*** (13)
13	***service*** (8)	***job*** (59)	***anxiety*** (12)
14	***decision*** (6)	***neighborhood*** (59)	***depression*** (11)
15	***candidate*** (6)	***child*** (55)	*performance* (10)

Note. Distinct collocates are in bold, while overlapping collocates occurring at different frequencies are both underlined and bolded in the column with the highest frequency.

Table 6
Top 15 Noun Collocates of attribute, characteristic, *and* trait *in the Right Context (1R) in ACAD*

#	*attribute*	*characteristic*	*trait*
1	***agenda*** (21)	***curves*** (13)	***anxiety*** (309)
2	***questionnaire*** (12)	***table*** (9)	***arousability*** (79)
3	*values* (9)	*questionnaire* (9)	***values*** (46)
4	***levels*** (8)	***curve*** (7)	***anger*** (31)
5	***weights*** (8)	***inventory*** (6)	***dominance*** (30)
6	***extraction*** (6)	***variables*** (5)	***measures*** (23)
7	***scores*** (6)	***capital*** (4)	***variation*** (22)
8	***set*** (6)	***survey*** (4)	***heritability*** (21)
9	***agenda-setting*** (6)	***subjects*** (3)	***engageability*** (20)
10	***blocks*** (5)	***respondents*** (3)	***value*** (20)
11	***aggregation*** (4)	***ratings*** (3)	***loci*** (18)
12	***information*** (4)	***music*** (3)	***measure*** (18)
13	***preferences*** (4)	***model*** (3)	***evolution*** (17)
14	*variables* (3)	***parents*** (3)	***hope*** (17)
15	***selection*** (3)	***constant*** (2)	***sport*** (16)

Note. Distinct collocates are in bold, while overlapping collocates occurring at different frequencies are both underlined and bolded in the column with the highest frequency.

As shown in Tables 5 and 6, *attribute* tends to occur with abstract nouns related to products of negotiation, such as standards/quality (e.g., *problem, preferences*), *characteristic* with concrete nouns related to demographics/species/property (e.g., *student, respondents*), and *trait* with abstract nouns related to cognition/behavior/genetics (e.g., *depression, anxiety*). This suggests the types of nouns these target words tend to describe: whether they are abstract vs. concrete, observable vs. latent, or living vs. non-living. For more information, an additional search for noun collocates was conducted in the left (6L) and right (6R) contexts.

Interestingly, as shown in Table 7 and 8, *characteristics* is a common collocate for nouns *attribute, characteristic,* and *trait*. Using concordancing, Example Set A provides more context on how the target nouns are used differently.

Example Set A: Concordance Lines of *attribute* and *trait* co-occurring with *characteristic*

(1) …*father-school community is determined by a combination of family characteristics and program attributes*… <ACAD: School of Psychology Review, 2005>

(2) …*a combination of individual factors (e.g., personality traits), home/family characteristics (e.g., roles, responsibilities)*... <ACAD: Journal of Instructional Psychology, 2008>

The first line starts to reveal a meaning difference between *characteristic* and *attribute*. In this case (Line 1), *characteristic* is defining a *family* (a concrete, living thing; more specifically, a human unit), whereas *attribute* is defining a *program* (an abstract, non-living thing). The second line also starts to reveal a meaning difference between *trait* and *characteristic.* Line 2 shows that *trait* tends to occur with abstract nouns related to cognition/behavior/genetics, whereas *characteristic* tends to occur with concrete nouns related to demographics/species/property. This reaffirms the generalization made based on Tables 5 and 6.

Table 7

Top 15 Noun Collocates of attribute, characteristic, *and* trait *in the Left Context (6L) in ACAD*

#	*attribute*	*characteristic*	*trait*
1	***attributes* (48)**	***table*** (1,016)	***personality*** (506)
2	***set*** (43)	***student*** (464)	***character*** (201)
3	***levels*** (38)	***study*** (345)	***selection*** (118)
4	***job*** (35)	*personality* (325)	***type*** (101)
5	***number*** (34)	***students*** (314)	***anxiety*** (78)
6	*students* (33)	***characteristics* (246)**	***trait* (75)**
7	*table* (31)	***school*** (211)	***quality*** (60)
8	*characteristics* (30)	***sample*** (209)	*relationship* (55)
9	***values*** (30)	***data*** (198)	***species*** (55)
10	*personality* (29)	***results*** (197)	***traits* (55)**
11	***product*** (29)	***differences*** (192)	***state*** (54)
12	***problem*** (27)	***information*** (178)	***variation*** (53)
13	*study* (27)	***family*** (173)	***leadership*** (52)
14	***attribute* (25)**	***background*** (171)	*study* (48)
15	***skills*** (25)	***relationship*** (171)	*characteristics* (46)

Note. Distinct collocates are in bold, while overlapping collocates occurring at different frequencies are both underlined and bolded in the column with the highest frequency.

Table 8
Top 15 Noun Collocates of attribute, characteristic, *and* trait *in the Left Context (6R) in ACAD*

#	*attribute*	*characteristic*	*trait*
1	***attributes* (54)**	***students*** (511)	***anxiety*** (387)
2	***levels*** (43)	***table*** (510)	***arousability*** (85)
3	*students* (39)	***children*** (294)	***species*** (77)
4	*table* (36)	***participants*** (270)	***values*** (77)
5	***type*** (34)	***study*** (247)	***scale*** (73)
6	*characteristics* (28)	***characteristics* (245)**	***selection*** (70)
7	***data*** (26)	***sample*** (239)	***trait*** (68)
8	***interest*** (26)	***age*** (232)	***traits*** (62)
9	***model*** (25)	***gender*** (217)	*characteristics* (50)
10	***performance*** (25)	***patients*** (190)	***scores*** (49)
11	*values* (24)	***group*** (173)	***dominance*** (47)
12	***agenda*** (23)	***school*** (168)	***individuals*** (46)
13	***skills*** (23)	***student*** (164)	***measures*** (46)
14	***level*** (22)	***population*** (154)	*students* (45)
15	***research*** (22)	***groups*** (149)	***fitness*** (44)

Note. Distinct collocates are in bold, while overlapping collocates occurring at different frequencies are both underlined and bolded in the column with the highest frequency.

Table 9
Top 15 Adjective Collocates of attribute, characteristic, *and* trait *in the Left Context (1L) in ACAD*

#	*attribute*	*characteristic*	*trait*
1	*physical* (117)	***demographic*** (1,008)	***cultural*** (89)
2	*personal* (97)	***personal*** (351)	***quantitative*** (67)
3	*other* (94)	***physical*** (235)	*other* (63)
4	***positive*** (77)	***other*** (248)	*personal* (55)
5	*important* (57)	***individual*** (216)	*common* (53)
6	*individual* (40)	***common*** (206)	*behavioral* (49)
7	***negative*** (40)	***unique*** (203)	*positive* (49)
8	*different* (38)	***important*** (170)	*physical* (47)
9	***key*** (37)	***sociodemographic*** (170)	***human*** (43)
10	*cultural* (36)	***defining*** (165)	***phenotype*** (42)
11	***creative*** (27)	***specific*** (157)	*certain* (40)
12	*human* (26)	***essential*** (136)	***different*** (39)
13	*specific* (25)	***behavioral*** (130)	***psychological*** (36)
14	*unique* (25)	***certain*** (123)	***desirable*** (35)
15	*certain* (24)	***similar*** (123)	***life-history*** (35)

Note. Distinct collocates are in bold, while overlapping collocates occurring at different frequencies are both underlined and bolded in the column with the highest frequency.

Table 10
Top 15 Adjective Collocates of attribute, characteristic, *and* trait *in the Right Context (1R) in ACAD*

#	*attribute*	*characteristic*	*trait*
1	*necessary* (10)	***similar*** (24)	***anxious*** (15)
2	***inherent*** (3)	***necessary*** (19)	***cognitive*** (10)
3	*common* (2)	***common*** (18)	*typical* (8)
4	*essential* (2)	***unique*** (10)	***verbal*** (7)
5	***constant*** (2)	***typical*** (10)	***somatic*** (7)
6	*unique* (2)	***relevant*** (8)	***involved*** (6)
7	***independent*** (1)	***important*** (7)	*important* (5)
8	***unaffected*** (1)	***indicative*** (6)	***similar*** (4)
9	*specific* (1)	***specific*** (5)	***young*** (3)
10	*similar* (1)	***suitable*** (4)	*common* (3)
11	***responsible*** (1)	***peculiar*** (4)	***likely*** (3)
12	***physical*** (1)	***pertinent*** (3)	***distinctive*** (2)
13	***organizational*** (1)	***essential*** (3)	***familiar*** (2)
14	***central*** (1)	***attributable*** (2)	*essential* (2)
15	*involved* (1)	***critical*** (2)	*peculiar* (2)

Note. Distinct collocates are in bold, while overlapping collocates occurring at different frequencies are both underlined and bolded in the column with the highest frequency.

Table 11
Top 15 Adverb Collocates of attribute, characteristic, *and* trait *in the Left Context (1L) in ACAD*

#	*attribute*	*characteristic*	*trait*
1	*how* (2)	***how*** (7)	***both*** (6)
2	***away*** (1)	***either*** (2)	*how* (5)
3		*on* (2)	*on* (2)
4		***yet*** (2)	***highly*** (2)
5		***unfortunately*** (1)	***when*** (1)
6		***together*** (1)	***thus*** (1)
7		***simply*** (1)	***then*** (1)
8		***rather*** (1)	***recently*** (1)
9		*perhaps* (1)	*perhaps* (1)
10		***out*** (1)	***higher*** (1)
11		***now*** (1)	***hardly*** (1)
12		***mainly*** (1)	***example*** (1)
13		***indeed*** (1)	
14		***in*** (1)	
15		***i.e.*** (1)	

Note. Distinct collocates are in bold, while overlapping collocates occurring at different frequencies are both underlined and bolded in the column with the highest frequency.

Table 12
Top 15 Adverbs Collocates of attribute, characteristic, *and* trait *in the Right Context (1R) in ACAD*

#	***attribute***	***characteristic***	***trait***
1	*also* (5)	***also*** (28)	*also* (8)
2	*more* (4)	***in*** (19)	*in* (8)
3	***when*** (3)	***often*** (15)	*most* (8)
4	*often* (3)	***as*** (14)	*often* (7)
5	***best*** (2)	***most*** (13)	*more* (5)
6	***at*** (2)	***alone*** (10)	***usually*** (3)
7	*directly* (2)	***more*** (9)	*traditionally* (3)
8	*in* (2)	***commonly*** (8)	***together*** (3)
9	*most* (2)	***frequently*** (6)	***simultaneously*** (3)
10	***only*** (2)	***significantly*** (5)	*frequently* (3)
11	***wholly*** (1)	***differently*** (4)	*as* (3)
12	*usually* (1)	***typically*** (4)	*commonly* (3)
13	*traditionally* (1)	***traditionally*** (4)	***already*** (2)
14	***too*** (1)	***that*** (3)	*directly* (2)
15	*together* (1)	***strongly*** (3)	***even*** (2)

Note. Distinct collocates are in bold, while overlapping collocates occurring at different frequencies are both underlined and bolded in the column with the highest frequency.

As shown in Tables 9 and 10, *attribute* tends to occur with evaluative adjectives (e.g., *positive, important*), *characteristic* with topical adjectives (e.g., *personal, physical*), and *trait* with relational adjectives (e.g., *cultural, life-history*). In addition, for *characteristic*, the high frequency of *defining* suggests that a *characteristic* is fundamental to the thing being described and helps distinguish it from other things that may otherwise share some similarities or traits.

As shown in Tables 11 and 12, *attribute* tends to occur with adverbs of degree (e.g., *most*), *characteristic* with adverbs of degree (e.g., *significantly*), and *trait* with adverbs of frequency (e.g., *often*). This suggests that *attribute* and *characteristic* are fundamental to the thing being described and helps distinguish it from other things that may otherwise share similarities or traits. On the other hand, *trait* seems to emphasize things that are commonly observable.

Next, using the COMPARE function in COCA, the the top 50 collocates of *attribute* ([attribute_NN]), *characteristic* ([characteristic]_NN), and *trait* ([trait]_NN) were analyzed in the left (3L) and right (3R) contexts in COCA's academic journals (see Figures 8-13). Based on this next search, the semantic preferences among *attribute, characteristic,* and *trait* continue to emerge.

Figure 8

[attribute]_NN VS [characteristic]_NN 3L Top 50 Collocates using Compare (sorted by FREQ)

WORD 1 (W1): **ATTRIBUTE** (0.38)

	WORD	W1	W2	W1/W2	SCORE
1	THE	4740	9610	0.5	1.3
2	OF	2044	4273	0.5	1.3
3	AND	1190	4861	0.2	0.7
4	,	1155	3325	0.3	0.9
5	.	918	3372	0.3	0.7
6	AN	853	286	3.0	7.9
7	A	774	1899	0.4	1.1
8	TO	537	1670	0.3	0.9
9	THESE	535	1320	0.4	1.1
10	'S	469	669	0.7	1.9
11	WITH	461	1291	0.4	1.0
12	FOR	376	541	0.7	1.8
13	PHYSICAL	372	767	0.5	1.3
14	OTHER	369	856	0.4	1.1
15	"	354	616	0.6	1.5
16	THAT	351	977	0.4	1.0
17	ALL	341	472	0.7	1.9
18	AS	341	565	0.6	1.6
19	IS	334	668	0.5	1.3
20	IN	319	660	0.5	1.3
21	OR	307	642	0.5	1.3
22	HAS	299	436	0.7	1.8
23	ON	297	1182	0.3	0.7
24	THIS	280	399	0.7	1.9
25	IF	256	60	4.3	11.4
26	HAVE	249	718	0.3	0.9
27	THEIR	228	703	0.3	0.9
28	ITS	219	510	0.4	1.1
29	BY	216	505	0.4	1.1
30	HIS	212	124	1.7	4.5
31	ARE	211	620	0.3	0.9
32	POSITIVE	198	101	2.0	5.2
33	THOSE	198	374	0.5	1.4
34	SOME	187	600	0.3	0.8
35	MOST	174	483	0.4	1.0
36	PERSONAL	170	548	0.3	0.8
37	IMPORTANT	168	357	0.5	1.3
38	VALUE	166	19	8.7	23.3
39	ONE	153	452	0.3	0.9
40	)	150	878	0.2	0.5
41	'	148	370	0.4	1.1
42	DOM	144	0	288.0	766.5
43	NAME	135	12	11.3	29.9
44	ELEMENT	134	3	44.7	118.9
45	CONTENT	134	21	6.4	17.0
46	NOT	133	199	0.7	1.8
47	MANY	130	319	0.4	1.1
48	EACH	121	128	0.9	2.5
49	ANY	119	186	0.6	1.7
50	SUCH	115	221	0.5	1.4

WORD 2 (W2): **CHARACTERISTIC** (2.66)

	WORD	W2	W1	W2/W1	SCORE
1	THE	9610	4740	2.0	0.8
2	AND	4861	1190	4.1	1.5
3	OF	4273	2044	2.1	0.8
4	.	3372	918	3.7	1.4
5	,	3325	1155	2.9	1.1
6	A	1899	774	2.5	0.9
7	TO	1670	537	3.1	1.2
8	THESE	1320	535	2.5	0.9
9	WITH	1291	461	2.8	1.1
10	DEMOGRAPHIC	1278	14	91.3	34.3
11	ON	1182	297	4.0	1.5
12	THAT	977	351	2.8	1.0
13	)	878	150	5.9	2.2
14	OTHER	856	369	2.3	0.9
15	PHYSICAL	767	372	2.1	0.8
16	HAVE	718	249	2.9	1.1
17	THEIR	703	228	3.1	1.2
18	'S	669	469	1.4	0.5
19	IS	668	334	2.0	0.8
20	IN	660	319	2.1	0.8
21	OR	642	307	2.1	0.8
22	ARE	620	211	2.9	1.1
23	"	616	354	1.7	0.7
24	SOME	600	187	3.2	1.2
25	AS	565	341	1.7	0.6
26	PERSONAL	548	170	3.2	1.2
27	FOR	541	376	1.4	0.5
28	1	520	13	40.0	15.0
29	DEFINING	519	29	17.9	6.7
30	ITS	510	219	2.3	0.9
31	BY	505	216	2.3	0.9
32	MOST	483	174	2.8	1.0
33	ALL	472	341	1.4	0.5
34	CERTAIN	462	78	5.9	2.2
35	ONE	452	153	3.0	1.1
36	UNIQUE	445	80	5.6	2.1
37	HAS	436	299	1.5	0.5
38	STUDENT	426	19	22.4	8.4
39	INDIVIDUAL	423	59	7.2	2.7
40	COMMON	415	59	7.0	2.6
41	THIS	399	280	1.4	0.5
42	:	391	105	3.7	1.4
43	TABLE	381	13	29.3	11.0
44	THOSE	374	198	1.9	0.7
45	PERSONALITY	373	37	10.1	3.8
46	SPECIFIC	373	92	4.1	1.5
47	'	370	148	2.5	0.9
48	IMPORTANT	357	168	2.1	0.8
49	DIFFERENT	341	101	3.4	1.3
50	SIMILAR	323	45	7.2	2.7

Figure 9

[attribute]_NN VS [characteristic]_NN 3R Top 50 Collocates using Compare (sorted by FREQ)

WORD 1 (W1): **ATTRIBUTE** (0.38)

	WORD	W1	W2	W1/W2	SCORE
1	OF	2859	14491	0.2	0.5
2	THE	2486	6913	0.4	1.0
3	,	2435	5523	0.4	1.2
4	.	2174	4532	0.5	1.3
5	AND	1142	3523	0.3	0.9
6	IS	992	797	1.2	3.3
7	TO	898	1521	0.6	1.6
8	A	859	1737	0.5	1.3
9	THAT	796	2431	0.3	0.9
10	IN	742	1696	0.4	1.2
11	ARE	558	1200	0.5	1.2
12	AS	520	1392	0.4	1.0
13	(	500	1507	0.3	0.9
14	BE	408	450	0.9	2.4
15	FOR	356	750	0.5	1.3
16	"	323	719	0.4	1.2
17	MUST	311	56	5.6	14.8
18	OR	299	484	0.6	1.6
19	ON	298	397	0.8	2.0
20	VALUE	278	17	16.4	43.5
21	WITH	248	884	0.3	0.7
22	AN	246	396	0.6	1.7
23	:	243	748	0.3	0.9
24	IT	223	411	0.5	1.4
25	SUCH	222	736	0.3	0.8
26	WHICH	209	361	0.6	1.5
27	NOT	203	269	0.8	2.0
28	)	201	343	0.6	1.6
29	CAN	181	299	0.6	1.6
30	BY	180	417	0.4	1.1
31	THIS	180	476	0.4	1.0
32	IF	171	72	2.4	6.3
33	SET	170	36	4.7	12.6
34	WERE	163	547	0.3	0.8
35	HAVE	161	341	0.5	1.3
36	HAS	159	124	1.3	3.4
37	'S	147	270	0.5	1.4
38	I	143	244	0.6	1.6
39	BUT	141	234	0.6	1.6
40	YOU	128	195	0.7	1.7
41	MAY	127	293	0.4	1.2
42	GOD	123	23	5.3	14.2
43	SPECIFIED	122	8	15.3	40.6
44	PRESENT	120	60	2.0	5.3
45	USED	120	95	1.3	3.4
46	;	117	196	0.6	1.6
47	THEY	114	368	0.3	0.8
48	WILL	113	135	0.8	2.2
49	WAS	111	214	0.5	1.4
50	RETURN	108	1	108.0	287.4

WORD 2 (W2): **CHARACTERISTIC** (2.66)

	WORD	W2	W1	W2/W1	SCORE
1	OF	14491	2859	5.1	1.9
2	THE	6913	2486	2.8	1.0
3	,	5523	2435	2.3	0.9
4	.	4532	2174	2.1	0.8
5	AND	3523	1142	3.1	1.2
6	THAT	2431	796	3.1	1.1
7	A	1737	859	2.0	0.8
8	IN	1696	742	2.3	0.9
9	TO	1521	898	1.7	0.6
10	(	1507	500	3.0	1.1
11	AS	1392	520	2.7	1.0
12	ARE	1200	558	2.2	0.8
13	WITH	884	248	3.6	1.3
14	IS	797	992	0.8	0.3
15	FOR	750	356	2.1	0.8
16	:	748	243	3.1	1.2
17	SUCH	736	222	3.3	1.2
18	"	719	323	2.2	0.8
19	WERE	547	163	3.4	1.3
20	OR	484	299	1.6	0.6
21	THIS	476	180	2.6	1.0
22	BE	450	408	1.1	0.4
23	BY	417	180	2.3	0.9
24	IT	411	223	1.8	0.7
25	THEIR	399	107	3.7	1.4
26	ON	397	298	1.3	0.5
27	AN	396	246	1.6	0.6
28	THEY	368	114	3.2	1.2
29	WHICH	361	209	1.7	0.6
30	)	343	201	1.7	0.6
31	THESE	342	55	6.2	2.3
32	HAVE	341	161	2.1	0.8
33	FROM	332	95	3.5	1.3
34	STUDENTS	321	28	11.5	4.3
35	ASSOCIATED	299	47	6.4	2.4
36	CAN	299	181	1.7	0.6
37	WE	295	107	2.8	1.0
38	MAY	293	127	2.3	0.9
39	'S	270	147	1.8	0.7
40	NOT	269	203	1.3	0.5
41	I	244	143	1.7	0.6
42	EACH	236	64	3.7	1.4
43	BUT	234	141	1.7	0.6
44	MAKE	230	66	3.5	1.3
45	ALL	222	76	2.9	1.1
46	THOSE	217	41	5.3	2.0
47	WAS	214	111	1.9	0.7
48	?	203	68	3.0	1.1
49	CHILDREN	198	18	11.0	4.1
50	PARTICIPANTS	197	3	65.7	24.7

As shown in Figures 8 and 9, *attribute* continues to show a strong preference for standards/quality (e.g., *value*), whereas *characteristic* continues to show a strong preference for demographics/species/property (e.g., *demographic*).

Figure 10

[attribute]_NN VS [trait]_NN 3L Top 50 Collocates using Compare (sorted by FREQ)

WORD 1 (W1): **ATTRIBUTE** (0.77)

	WORD	W1	W2	W1/W2	SCORE
1	THE	4740	3423	1.4	1.8
2	OF	2044	2612	0.8	1.0
3	AND	1190	1496	0.8	1.0
4	,	1155	1634	0.7	0.9
5	.	918	1134	0.8	1.1
6	AN	853	328	2.6	3.4
7	A	774	2347	0.3	0.4
8	TO	537	696	0.8	1.0
9	THESE	535	790	0.7	0.9
10	'S	469	446	1.1	1.4
11	WITH	461	635	0.7	0.9
12	FOR	376	704	0.5	0.7
13	PHYSICAL	372	235	1.6	2.1
14	OTHER	369	446	0.8	1.1
15	"	354	461	0.8	1.0
16	THAT	351	742	0.5	0.6
17	AS	341	296	1.2	1.5
18	ALL	341	311	1.1	1.4
19	IS	334	468	0.7	0.9
20	IN	319	340	0.9	1.2
21	OR	307	332	0.9	1.2
22	HAS	299	138	2.2	2.8
23	ON	297	420	0.7	0.9
24	THIS	280	427	0.7	0.9
25	IF	256	61	4.2	5.5
26	HAVE	249	326	0.8	1.0
27	THEIR	228	241	0.9	1.2
28	ITS	219	74	3.0	3.9
29	BY	216	163	1.3	1.7
30	HIS	212	149	1.4	1.9
31	ARE	211	337	0.6	0.8
32	POSITIVE	198	197	1.0	1.3
33	THOSE	198	399	0.5	0.6
34	SOME	187	326	0.6	0.7
35	MOST	174	248	0.7	0.9
36	PERSONAL	170	126	1.3	1.8
37	IMPORTANT	168	136	1.2	1.6
38	VALUE	166	31	5.4	7.0
39	ONE	153	284	0.5	0.7
40	)	150	247	0.6	0.8
41	'	148	97	1.5	2.0
42	DOM	144	0	288.0	374.7
43	NAME	135	6	22.5	29.3
44	ELEMENT	134	1	134.0	174.3
45	CONTENT	134	2	67.0	87.2
46	NOT	133	179	0.7	1.0
47	MANY	130	212	0.6	0.8
48	EACH	121	88	1.4	1.8
49	ANY	119	149	0.8	1.0
50	KEY	115	72	1.6	2.1

WORD 2 (W2): **TRAIT** (1.30)

	WORD	W2	W1	W2/W1	SCORE
1	THE	3423	4740	0.7	0.6
2	OF	2612	2044	1.3	1.0
3	A	2347	774	3.0	2.3
4	,	1634	1155	1.4	1.1
5	AND	1496	1190	1.3	1.0
6	PERSONALITY	1170	37	31.6	24.3
7	.	1134	918	1.2	0.9
8	THESE	790	535	1.5	1.1
9	THAT	742	351	2.1	1.6
10	FOR	704	376	1.9	1.4
11	TO	696	537	1.3	1.0
12	WITH	635	461	1.4	1.1
13	CHARACTER	577	22	26.2	20.2
14	IS	468	334	1.4	1.1
15	"	461	354	1.3	1.0
16	OTHER	446	369	1.2	0.9
17	'S	446	469	1.0	0.7
18	THIS	427	280	1.5	1.2
19	ON	420	297	1.4	1.1
20	THOSE	399	198	2.0	1.5
21	IN	340	319	1.1	0.8
22	ARE	337	211	1.6	1.2
23	OR	332	307	1.1	0.8
24	AN	328	853	0.4	0.3
25	SOME	326	187	1.7	1.3
26	HAVE	326	249	1.3	1.0
27	ALL	311	341	0.9	0.7
28	AS	296	341	0.9	0.7
29	ONE	284	153	1.9	1.4
30	HUMAN	258	83	3.1	2.4
31	COMMON	248	59	4.2	3.2
32	MOST	248	174	1.4	1.1
33	)	247	150	1.6	1.3
34	THEIR	241	228	1.1	0.8
35	PHYSICAL	235	372	0.6	0.5
36	CERTAIN	232	78	3.0	2.3
37	MANY	212	130	1.6	1.3
38	POSITIVE	197	198	1.0	0.8
39	GENETIC	190	3	63.3	48.7
40	SUCH	189	115	1.6	1.3
41	SAME	182	93	2.0	1.5
42	NOT	179	133	1.3	1.0
43	BY	163	216	0.8	0.6
44	CULTURAL	157	59	2.7	2.0
45	(	155	66	2.3	1.8
46	SHARE	151	38	4.0	3.1
47	ANY	149	119	1.3	1.0
48	HIS	149	212	0.7	0.5
49	BETWEEN	146	67	2.2	1.7
50	MORE	146	81	1.8	1.4

Figure 11

[attribute]_NN VS [trait]_NN 3R Top 50 Collocates using Compare (sorted by FREQ)

WORD 1 (W1): **ATTRIBUTE** (0.77)

	WORD	W1	W2	W1/W2	SCORE
1	OF	2859	2021	1.4	1.8
2	THE	2486	1849	1.3	1.7
3	,	2435	3136	0.8	1.0
4	.	2174	2928	0.7	1.0
5	AND	1142	1731	0.7	0.9
6	IS	992	715	1.4	1.8
7	TO	898	1043	0.9	1.1
8	A	859	876	1.0	1.3
9	THAT	796	1885	0.4	0.5
10	IN	742	1360	0.5	0.7
11	ARE	558	854	0.7	0.9
12	AS	520	717	0.7	0.9
13	(	500	625	0.8	1.0
14	BE	408	391	1.0	1.4
15	FOR	356	475	0.7	1.0
16	"	323	440	0.7	1.0
17	MUST	311	26	12.0	15.6
18	OR	299	334	0.9	1.2
19	ON	298	230	1.3	1.7
20	VALUE	278	56	5.0	6.5
21	WITH	248	438	0.6	0.7
22	AN	246	158	1.6	2.0
23	:	243	393	0.6	0.8
24	IT	223	264	0.8	1.1
25	SUCH	222	335	0.7	0.9
26	WHICH	209	222	0.9	1.2
27	NOT	203	250	0.8	1.1
28	)	201	225	0.9	1.2
29	CAN	181	276	0.7	0.9
30	THIS	180	165	1.1	1.4
31	BY	180	247	0.7	0.9
32	IF	171	98	1.7	2.3
33	SET	170	25	6.8	8.8
34	WERE	163	221	0.7	1.0
35	HAVE	161	292	0.6	0.7
36	HAS	159	173	0.9	1.2
37	'S	147	165	0.9	1.2
38	I	143	323	0.4	0.6
39	BUT	141	229	0.6	0.8
40	YOU	128	289	0.4	0.6
41	MAY	127	184	0.7	0.9
42	GOD	123	9	13.7	17.8
43	SPECIFIED	122	3	40.7	52.9
44	PRESENT	120	30	4.0	5.2
45	USED	120	33	3.6	4.7
46	;	117	112	1.0	1.4
47	THEY	114	289	0.4	0.5
48	WILL	113	149	0.8	1.0
49	WAS	111	205	0.5	0.7
50	RETURN	108	0	216.0	281.0

WORD 2 (W2): **TRAIT** (1.30)

	WORD	W2	W1	W2/W1	SCORE
1	,	3136	2435	1.3	1.0
2	.	2928	2174	1.3	1.0
3	OF	2021	2859	0.7	0.5
4	THAT	1885	796	2.4	1.8
5	THE	1849	2486	0.7	0.6
6	AND	1731	1142	1.5	1.2
7	IN	1360	742	1.8	1.4
8	TO	1043	898	1.2	0.9
9	A	876	859	1.0	0.8
10	ARE	854	558	1.5	1.2
11	AS	717	520	1.4	1.1
12	IS	715	992	0.7	0.6
13	(	625	500	1.3	1.0
14	FOR	475	356	1.3	1.0
15	"	440	323	1.4	1.0
16	WITH	438	248	1.8	1.4
17	:	393	243	1.6	1.2
18	BE	391	408	1.0	0.7
19	ANXIETY	366	0	732.0	562.7
20	SUCH	335	222	1.5	1.2
21	OR	334	299	1.1	0.9
22	I	323	143	2.3	1.7
23	FROM	295	95	3.1	2.4
24	HAVE	292	161	1.8	1.4
25	THEY	289	114	2.5	1.9
26	YOU	289	128	2.3	1.7
27	CAN	276	181	1.5	1.2
28	IT	264	223	1.2	0.9
29	NOT	250	203	1.2	0.9
30	BY	247	180	1.4	1.1
31	WE	237	107	2.2	1.7
32	ON	230	298	0.8	0.6
33	BUT	229	141	1.6	1.2
34	)	225	201	1.1	0.9
35	HE	222	65	3.4	2.6
36	WHICH	222	209	1.1	0.8
37	WERE	221	163	1.4	1.0
38	WAS	205	111	1.8	1.4
39	?	204	68	3.0	2.3
40	LIKE	198	68	2.9	2.2
41	THEIR	187	107	1.7	1.3
42	MAY	184	127	1.4	1.1
43	HAS	173	159	1.1	0.8
44	'S	165	147	1.1	0.9
45	THIS	165	180	0.9	0.7
46	AN	158	246	0.6	0.5
47	ONE	149	75	2.0	1.5
48	WILL	149	113	1.3	1.0
49	WOULD	144	63	2.3	1.8
50	ASSOCIATED	133	47	2.8	2.2

As shown in Figures 10 and 11, *attribute* continues to show a strong preference for standards/quality (e.g., *must, return*), whereas *trait* continues to show a strong preference for cognition/behavior/genetics (e.g., *genetic, anxiety*).

Figure 12

[characteristic]_NN VS [trait]_NN 3L Top 50 Collocates using Compare (sorted by FREQ)

WORD 1 (W1): **CHARACTERISTIC** (2.05)

	WORD	W1	W2	W1/W2	SCORE
1	THE	9610	3423	2.8	1.4
2	AND	4861	1496	3.2	1.6
3	OF	4273	2612	1.6	0.8
4	.	3372	1134	3.0	1.5
5	,	3325	1634	2.0	1.0
6	A	1899	2347	0.8	0.4
7	TO	1670	696	2.4	1.2
8	THESE	1320	790	1.7	0.8
9	WITH	1291	635	2.0	1.0
10	DEMOGRAPHIC	1278	21	60.9	29.7
11	ON	1182	420	2.8	1.4
12	THAT	977	742	1.3	0.6
13	)	878	247	3.6	1.7
14	OTHER	856	446	1.9	0.9
15	PHYSICAL	767	235	3.3	1.6
16	HAVE	718	326	2.2	1.1
17	THEIR	703	241	2.9	1.4
18	'S	669	446	1.5	0.7
19	IS	668	468	1.4	0.7
20	IN	660	340	1.9	0.9
21	OR	642	332	1.9	0.9
22	ARE	620	337	1.8	0.9
23	"	616	461	1.3	0.7
24	SOME	600	326	1.8	0.9
25	AS	565	296	1.9	0.9
26	PERSONAL	548	126	4.3	2.1
27	FOR	541	704	0.8	0.4
28	1	520	5	104.0	50.8
29	DEFINING	519	82	6.3	3.1
30	ITS	510	74	6.9	3.4
31	BY	505	163	3.1	1.5
32	MOST	483	248	1.9	1.0
33	ALL	472	311	1.5	0.7
34	CERTAIN	462	232	2.0	1.0
35	ONE	452	284	1.6	0.8
36	UNIQUE	445	80	5.6	2.7
37	HAS	436	138	3.2	1.5
38	STUDENT	426	3	142.0	69.4
39	INDIVIDUAL	421	94	4.5	2.2
40	COMMON	415	248	1.7	0.8
41	THIS	399	427	0.9	0.5
42	:	391	99	3.9	1.9
43	TABLE	381	1	381.0	186.2
44	THOSE	374	399	0.9	0.5
45	SPECIFIC	373	130	2.9	1.4
46	PERSONALITY	373	1170	0.3	0.2
47	'	370	97	3.8	1.9
48	IMPORTANT	357	136	2.6	1.3
49	DIFFERENT	341	136	2.5	1.2
50	SIMILAR	323	82	3.9	1.9

WORD 2 (W2): **TRAIT** (0.49)

	WORD	W2	W1	W2/W1	SCORE
1	THE	3423	9610	0.4	0.7
2	OF	2612	4273	0.6	1.3
3	A	2347	1899	1.2	2.5
4	,	1634	3325	0.5	1.0
5	AND	1496	4861	0.3	0.6
6	PERSONALITY	1170	373	3.1	6.4
7	.	1134	3372	0.3	0.7
8	THESE	790	1320	0.6	1.2
9	THAT	742	977	0.8	1.6
10	FOR	704	541	1.3	2.7
11	TO	696	1670	0.4	0.9
12	WITH	635	1291	0.5	1.0
13	CHARACTER	577	6	96.2	196.7
14	IS	468	668	0.7	1.4
15	"	461	616	0.7	1.5
16	'S	446	669	0.7	1.4
17	OTHER	446	856	0.5	1.1
18	THIS	427	399	1.1	2.2
19	ON	420	1182	0.4	0.7
20	THOSE	399	374	1.1	2.2
21	IN	340	660	0.5	1.1
22	ARE	337	620	0.5	1.1
23	OR	332	642	0.5	1.1
24	AN	328	286	1.1	2.3
25	SOME	326	600	0.5	1.1
26	HAVE	326	718	0.5	0.9
27	ALL	311	472	0.7	1.3
28	AS	296	565	0.5	1.1
29	ONE	284	452	0.6	1.3
30	HUMAN	258	176	1.5	3.0
31	COMMON	248	415	0.6	1.2
32	MOST	248	483	0.5	1.1
33	)	247	878	0.3	0.6
34	THEIR	241	703	0.3	0.7
35	PHYSICAL	235	767	0.3	0.6
36	CERTAIN	232	462	0.5	1.0
37	MANY	212	319	0.7	1.4
38	POSITIVE	197	101	2.0	4.0
39	GENETIC	190	79	2.4	4.9
40	SUCH	189	221	0.9	1.7
41	SAME	182	272	0.7	1.4
42	NOT	179	199	0.9	1.8
43	BY	163	505	0.3	0.7
44	CULTURAL	157	161	1.0	2.0
45	(	155	204	0.8	1.6
46	SHARE	151	303	0.5	1.0
47	HIS	149	124	1.2	2.5
48	ANY	149	186	0.8	1.6
49	MORE	146	210	0.7	1.4
50	BETWEEN	146	280	0.5	1.1

Figure 13

[characteristic]_NN VS [trait]_NN 3R Top 50 Collocates using Compare (sorted by FREQ)

WORD 1 (W1): **CHARACTERISTIC** (2.05)

	WORD	W1	W2	W1/W2	SCORE
1	OF	14491	2021	7.2	3.5
2	THE	6912	1849	3.7	1.8
3	,	5523	3136	1.8	0.9
4	.	4532	2928	1.5	0.8
5	AND	3523	1731	2.0	1.0
6	THAT	2431	1885	1.3	0.6
7	A	1737	876	2.0	1.0
8	IN	1696	1360	1.2	0.6
9	TO	1521	1043	1.5	0.7
10	(	1507	625	2.4	1.2
11	AS	1392	717	1.9	0.9
12	ARE	1200	854	1.4	0.7
13	WITH	884	438	2.0	1.0
14	IS	797	715	1.1	0.5
15	FOR	750	475	1.6	0.8
16	:	748	393	1.9	0.9
17	SUCH	736	335	2.2	1.1
18	"	719	440	1.6	0.8
19	WERE	547	221	2.5	1.2
20	OR	484	334	1.4	0.7
21	THIS	476	165	2.9	1.4
22	BE	450	391	1.2	0.6
23	BY	417	247	1.7	0.8
24	IT	411	264	1.6	0.8
25	THEIR	399	187	2.1	1.0
26	ON	397	230	1.7	0.8
27	AN	396	158	2.5	1.2
28	THEY	368	289	1.3	0.6
29	WHICH	361	222	1.6	0.8
30	)	343	225	1.5	0.7
31	THESE	342	71	4.8	2.4
32	HAVE	341	292	1.2	0.6
33	FROM	332	295	1.1	0.6
34	STUDENTS	321	26	12.3	6.0
35	ASSOCIATED	299	133	2.2	1.1
36	CAN	299	276	1.1	0.5
37	WE	295	237	1.2	0.6
38	MAY	293	184	1.6	0.8
39	'S	270	165	1.6	0.8
40	NOT	269	250	1.1	0.5
41	I	244	323	0.8	0.4
42	EACH	236	63	3.7	1.8
43	BUT	234	229	1.0	0.5
44	MAKE	230	107	2.1	1.1
45	ALL	222	123	1.8	0.9
46	THOSE	217	81	2.7	1.3
47	WAS	214	205	1.0	0.5
48	?	203	204	1.0	0.5
49	CHILDREN	198	33	6.0	2.9
50	PARTICIPANTS	197	9	21.9	10.7

WORD 2 (W2): **TRAIT** (0.49)

	WORD	W2	W1	W2/W1	SCORE
1	,	3136	5523	0.6	1.2
2	.	2928	4532	0.6	1.3
3	OF	2021	14491	0.1	0.3
4	THAT	1885	2431	0.8	1.6
5	THE	1849	6912	0.3	0.5
6	AND	1731	3523	0.5	1.0
7	IN	1360	1696	0.8	1.6
8	TO	1043	1521	0.7	1.4
9	A	876	1737	0.5	1.0
10	ARE	854	1200	0.7	1.5
11	AS	717	1392	0.5	1.1
12	IS	715	797	0.9	1.8
13	(	625	1507	0.4	0.8
14	FOR	475	750	0.6	1.3
15	"	440	719	0.6	1.3
16	WITH	438	884	0.5	1.0
17	:	393	748	0.5	1.1
18	BE	391	450	0.9	1.8
19	ANXIETY	366	7	52.3	107.0
20	SUCH	335	736	0.5	0.9
21	OR	334	484	0.7	1.4
22	I	323	244	1.3	2.7
23	FROM	295	332	0.9	1.8
24	HAVE	292	341	0.9	1.8
25	YOU	289	195	1.5	3.0
26	THEY	289	368	0.8	1.6
27	CAN	276	299	0.9	1.9
28	IT	264	411	0.6	1.3
29	NOT	250	269	0.9	1.9
30	BY	247	417	0.6	1.2
31	WE	237	295	0.8	1.6
32	ON	230	397	0.6	1.2
33	BUT	229	234	1.0	2.0
34	)	225	343	0.7	1.3
35	HE	222	118	1.9	3.8
36	WHICH	222	361	0.6	1.3
37	WERE	221	547	0.4	0.8
38	WAS	205	214	1.0	2.0
39	?	204	203	1.0	2.1
40	LIKE	198	140	1.4	2.9
41	THEIR	187	399	0.5	1.0
42	MAY	184	293	0.6	1.3
43	HAS	173	124	1.4	2.9
44	'S	165	270	0.6	1.3
45	THIS	165	476	0.3	0.7
46	AN	158	396	0.4	0.8
47	WILL	149	135	1.1	2.3
48	ONE	149	175	0.9	1.7
49	WOULD	144	183	0.8	1.6
50	ASSOCIATED	133	299	0.4	0.9

As shown in Figures 12 and 13, *characteristic* continues to show a strong preference for demographics/species/property (e.g., *demographic*, *participants*), whereas *trait* continues to show a strong

preference for cognition/behavior/genetics (e.g., *character, anxiety*). While traits are inherited, characteristics can be acquired or changed.

In summary, it seems that *attribute* tends to occur with action verbs (e.g., *assigned*), abstract nouns related to standards/quality (e.g., *value*), evaluative adjectives (e.g., *important*), and adverbs of degree (e.g., *most*). *Characteristic* tends to occur with communication verbs (e.g., *describe*), concrete nouns related to demographics/species/property (e.g., *student*), topical adjectives (e.g., *physical*), and adverbs of degree (e.g., *significantly*). *Trait* tends to occur with mental verbs (e.g., *coping*), abstract nouns related to cognition/behavior/genetics (e.g., *anxiety*), relational adjectives (e.g., *life-history*), and adverbs of frequency (e.g., *often*). Lastly, *traits* tend to describe things that are inherited, whereas *characteristics* can describe things that are acquired.

To illustrate, consider the following examples for their semantic preferences:

Example Set 1: Concordance Lines of *attribute*

(1) *These attributes were identified by a panel of experts composed of recognized members of the art education community.* <ACAD: Arts Education Policy Review, 1996>
(2) *The beliefs about the nature of power, and the values and attributes associated with warriors and hunters…* <ACAD: Africa Today, 1994>
(3) *What are the important attributes that you consider while buying an apartment?* <ACAD: The Journal of Real Estate Research, 2016>
(4) *In both classes, the enjoyability of classes ranks as one of the most important attributes…* <ACAD: Economic Inquiry, 2018>

Example Set 2: Concordance Lines of *characteristic*

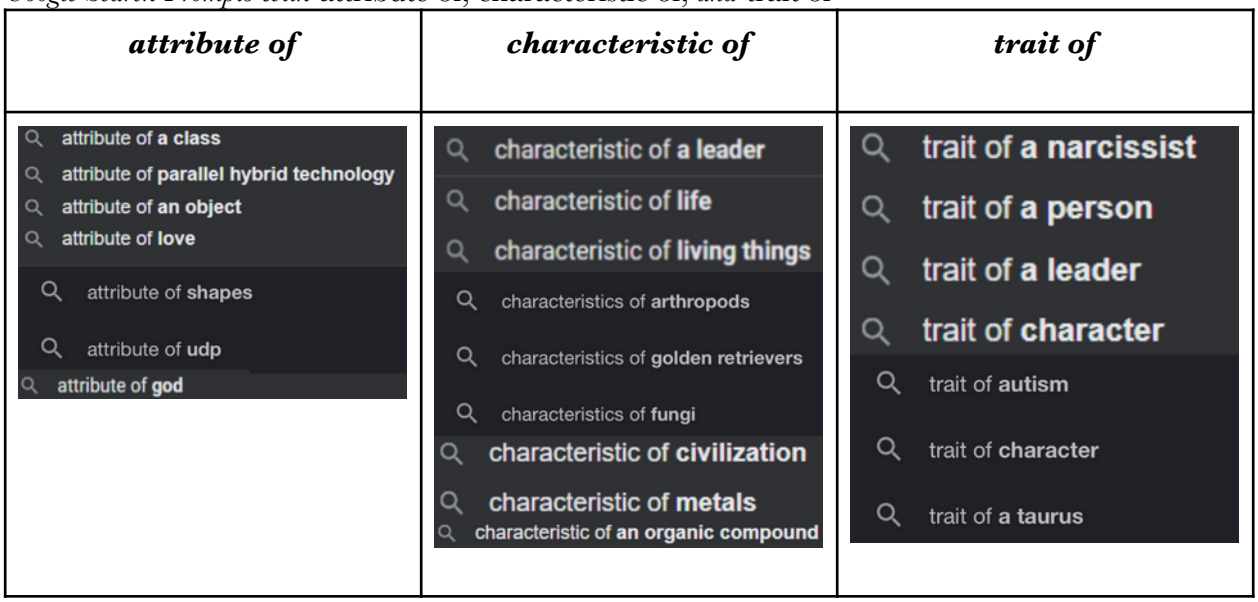
(1) *However, participant characteristics are reported separately for lead teachers and teaching assistants.* <ACAD: Language, Speech & Hearing Services in Schools, 2014>
(2) *The most typical student characteristics indicated a U.S.-born (65.7%) female (51.5%) majoring in accounting (29.8%) who started college at this university (52.2%) and who was working part-time (63.8%).* <ACAD: Academy of Business Research Journal, 2016>
(3) *Currently, the demographic characteristics of tattooed persons show wide variations in gender, age, social class, race, political party and occupation.* <ACAD: College Student Journal, 2015>
(4) *However, these characteristics are disproportionately found in the lower socioeconomic, high-crime, older neighborhoods that racial minorities often call home.* <ACAD: Vanderbilt Law Review, 2017>

Example Set 3: Concordance Lines of *trait*

(1) *Because each human's genes are unique, they are a personal map for that person's biological past and future — the traits inherited from parents and the ones to be passed on to children.* <ACAD: ABA Journal, 1998>
(2) *…for generalized anxiety disorder or panic attack for anxiety traits, and for major depressive disorder for depression traits.* <ACAD: Journal of Sport Behavior, 2000>
(3) *Unusual behavioral traits allow the fire ant to form large colonies and to dominate the territories it invades.* <ACAD: Bioscience, 1994>
(4) *Traits can be thought of as falling along a continuum that ranges from traits considered by many to be desirable (e.g., happy, self-confident, considerate) to traits that are often considered to be undesirable (e.g., aggressive, spiteful, crude).* <ACAD: Journal of Instructional Psychology, 1999>

Furthermore, a Google search using the common construction (noun + *of*-phrase) with *attribute, characteristic,* and *trait* reveals similar semantic preferences based on the search engine's predicted text (see Figure 14).

Figure 14
Google Search Prompts with attribute of, characteristic of, *and* trait of



Research Question 3: What are the differences in the semantic prosody of nouns *attribute, characteristic,* and *trait* in Academic English?

Using the COLLOCATES and COMPARE functions in COCA, the semantic prosody of *attribute* ([attribute_NN]), *characteristic* ([characteristic]_NN), and *trait* ([trait]_NN) was determined. To summarize the findings, consider the following examples for their semantic prosody:

Example Set 4: Concordance Lines of *attribute, characteristic,* and *trait*

(1) *Security is an important attribute of any good or service…*
(2) *The most typical student characteristics indicated a U.S.-born (65.7%) female (51.5%)...*
(3) *In fact, risk-taking is one personality trait that is most of the time connected with entrepreneurs.*

For *attribute* (Line 1), the noun often occurs in a context with positive meaning and an authoritative tone. For *characteristic* (Line 2), the noun often occurs in a context with neutral meaning and an informative tone. Lastly, for *trait* (Line 3), the noun often occurs in a context with neutral meaning and an authoritative tone. One other interesting feature of *attribute* is its tendency to describe things that are non-living (Line 1), whereas *trait* tends to describe things that are living (Line 3). This semantic restriction is similar to Room's (1985) explanation of a *wood* and *forest* and could potentially impact lexical choice and L1 transfer in translation (Hirst, 1995). When looking at the closest translation of *trait* in Japanese, 特性 (とくせい, *tokusei*), a quick survey from my Japanese students revealed that it is also primarily used to describe living things. In that case, the cross-linguistic plesionym group *attribute, trait,* and *tokusei* may be a non-issue for Japanese learners of English.

Application

Target Population

The target students for this application are adults (from over 70 countries) studying Academic English (AE) at university. They are currently enrolled in an advanced level (C1) writing course. Students can successfully discuss in detail issues related to success, including building a motivated team, and understand various communication styles (or registers). By the end of the term, the students' goal is to move up to C2, which enables students to explore science and technology and promotes creativity in writing. These themes and goals are highlighted by the materials in this section of the paper. According to the Common European Framework's (2015) English Vocabulary Profile Online - American English, the noun *attribute* is a C2 level word, *characteristic* is B2, and *trait* is C2. Although two of the three target nouns are technically above the students' level, the goal is to move them from C1 to C2 by the end of the term, so it is acceptable to start introducing all three target nouns to students.

Materials

The following set of materials designed for one corpus-informed lesson that follows the Illustration-Induction-Interaction (I-I-I) cycle. First, Figure 15 displays a sample handout, representative of the Illustration and Induction steps.

Figure 15
Handout 1: Illustration and Induction

Instructions: In this task, examine the following concordance lines with *attribute (n.)*, *characteristic (n.)*, and *trait (n.)* from COCA's Academic register and answer questions about them. What patterns do you see in the use of *attribute (n.)*, *characteristic (n.)*, and *trait (n.)*? Some guiding questions have been provided to help you guess the patterns. You may work with a partner or in a small group.

1. *What are the important attributes that you consider while buying an apartment?* <ACAD: The Journal of Real Estate Research, 2016>
2. *Security is an important attribute of any good or service…* <ACAD: Energy Journal, 2000>
3. *…there was strong agreement across respondent groups on which two attributes (Usability and Performance/Audit Trail) were considered the most important…* <ACAD: Journal of Information Systems, 1991>
4. *Senge (1990) contends that a critical attribute to leadership is the ability to bring about a shared picture of the future.* <ACAD: Education, 1992>
5. *Our closeness to patiences and the intelligence this provides is an important attribute in assessing health needs, and making informed decisions about the development of new services or improving existing ones.* <ACAD: Practice Nurse, 2011>
6. *The most typical student characteristics indicated a U.S.-born (65.7%) female (51.5%) majoring in accounting (29.8%) who started college at this university (52.2%) and who was working part-time (63.8%).* <ACAD: Academy of Business Research Journal, 2016>
7. *…the result of the physical and chemical properties of a particular element and the biological characteristics of the exposed tissue.* <ACAD: Natural History, 1990>
8. *Two demographic characteristics, parental absence and socioeconomic status (SES), were…* <ACAD: School Psychology Review, 1990>
9. *The results are listed as percentages, showing the percent of teachers who identified a specific characteristic as being associated with the potential school dropout (see chart B).* <ACAD: Education, 1990>
10. *Currently, the demographic characteristics of tattooed persons show wide variations in gender, age, social class, race, political party and occupation.* <ACAD: College Student Journal, 2015>
11. *Optimism is the trait that keeps him confident and motivated in the pursuit of the desired goal.* <ACAD: Academy of Entrepreneurship Journal, 2019>
12. *What can you do to develop innovative thinking traits?* <ACAD: Technology & Engineering Teacher, 2012>
13. *Unusual behavioral traits allow the fire ant to form large colonies and to dominate the territories it invades.* <ACAD: Bioscience, 1994>
14. *Possession of a sense of humor at any age is a socially desirable trait and children who possess it are more likely to be attractive to others.* <ACAD: Roeper Review, 1991>
15. *These behavioral traits predict later engagement in substance abuse, violence, criminal behavior, and poor post-school outcomes that adversely impact the student's future, their family and their community.* <ACAD: Education & Treatment of Children, 2014>

Guiding Questions:

A. What are the most frequent adjectives that occur before or after *attribute (n.)*?
    a. What types of adjectives are they?
        i. descriptive vs. evaluative
        ii. positive vs. neutral
        iii. observable vs. latent
        iv. other

B. What nouns occur after *attribute (n.)*?
    a. What types of nouns are they?
        i. abstract vs. concrete
        ii. observable vs. latent
        iii. living vs. non-living
        iv. other

C. What are the most frequent adjectives that occur before or after *characteristic (n.)*?
    a. What types of adjectives are they?
        i. descriptive vs. evaluative
        ii. positive vs. neutral
        iii. observable vs. latent
        iv. other

D. What nouns occur after *characteristic (n.)*?
    a. What types of nouns are they?
        i. abstract vs. concrete
        ii. observable vs. latent
        iii. living vs. non-living
        iv. other

E. What are the most frequent adjectives that occur before *trait (n.)*?
    a. What types of adjectives are they?
        i. descriptive vs. evaluative
        ii. positive vs. neutral
        iii. observable vs. latent
        iv. other

F. What nouns occur after *trait (n.)*?
    a. What types of nouns are they?
        i. abstract vs. concrete
        ii. observable vs. latent
        iii. living vs. non-living
        iv. other

G. What else do you notice about the use of *attribute (n.)*, *characteristic (n.)*, and *trait (n.)*? Based on the data, what questions do you have about their use?

Note. Adapted from Green (2019).

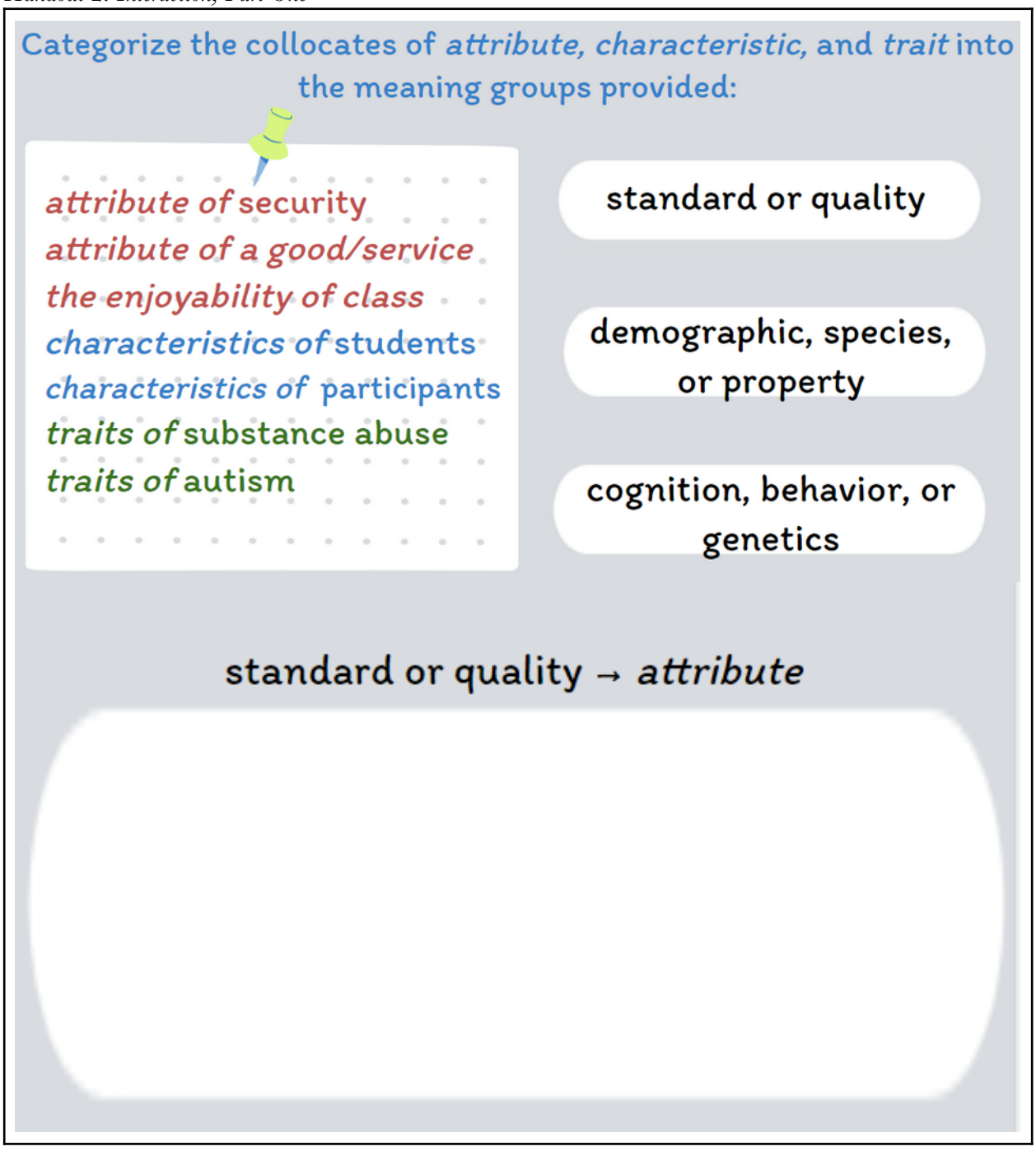
The structure of this handout was adapted from Green (2019) and contains a minimum of five concordance lines per target noun (*attribute, characteristic,* and *trait*) alongside guided questions with varying degrees of control. As students progress through the questions, the number of acceptable answers or interpretations is not limited to one. However, given the limited data set, the teacher can anticipate what types of responses students will give and continue guiding them through the generalization process by pointing to corpus data (e.g., the specific concordance line or number). This type of structure helps introduce students to the I-I-I cycle and to begin learning how to approach corpus data in general.

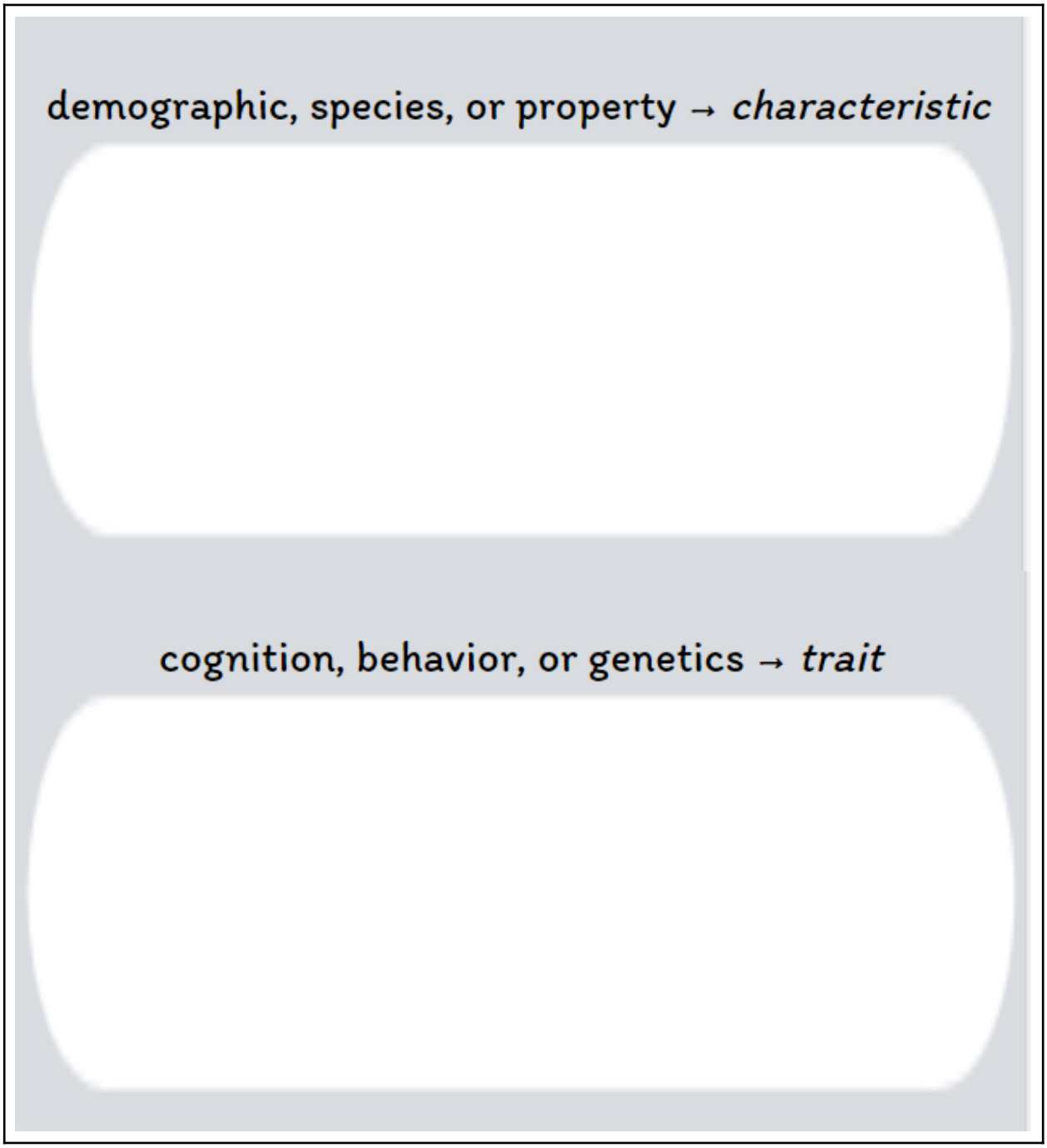
Next, Figure 16 displays a sample handout to promote Interaction with the target nouns and corpus tools. The structure of this handout was adapted from Nelson's (2006) Business English teaching materials based on semantic associations. In this case, students are asked to categorize the collocates of *attribute, characteristic,* and *trait* into the meaning groups (semantic categories) provided. As students deduce which collocate to match with each meaning group, the teacher can encourage them to use COCA's LIST, COLLOCATE, and COMPARE functions to confirm, refute, or refine their language hypotheses (Timmis, 2015).

Last but not least, Figure 17 displays a sample handout to promote further Interaction with the target nouns. It is vitally important that this writing activity comes after the Illustration and Induction steps. Without sufficient exposure to the forms, meanings, and uses of *attribute, characteristic,* and *trait*, students may find it very difficult to approach the task. The theme and goal of this handout are connected with the C1 and C2 skill set: 'characteristics of a successful student.' First, students form small groups and draft survey questions based on the theme. Second, the students will interview other students on campus and collect demographic data. Once each group has identified the top three *characteristics* of a successful student mentioned by participants, students will individually write a paragraph summary (150 words) to highlight the results. In the writing process, they will be asked to underline the target nouns that are used[2]. Lastly, students will report their results to the class in a class presentation or poster session format. The benefit of ending the I-I-I cycle with this activity is that it can function as a formative assessment for the vocabulary unit.

2 In this case, not using *trait* shows the understanding that it is not a feature that can be developed or created.

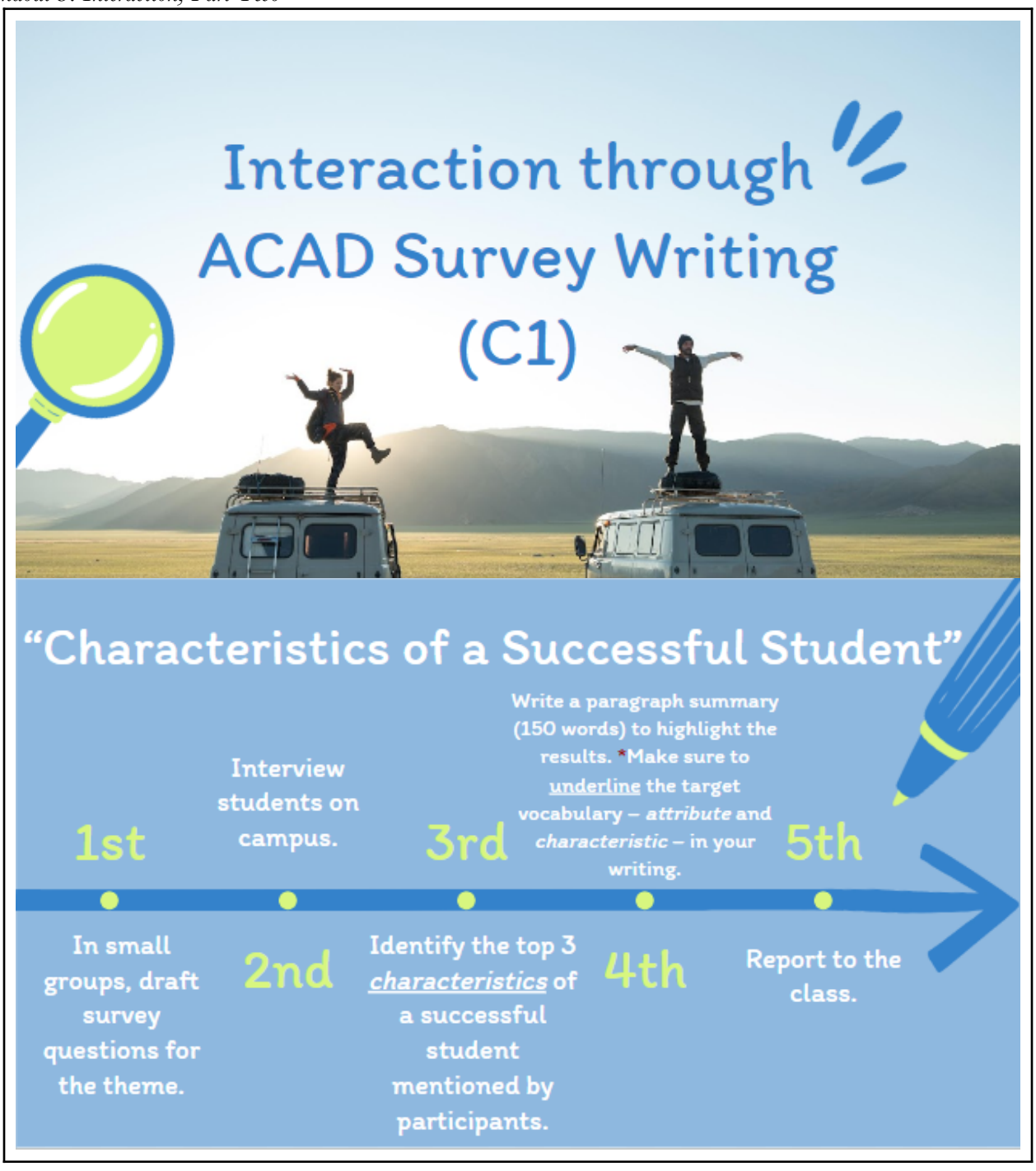
Figure 16
Handout 2: Interaction, Part One

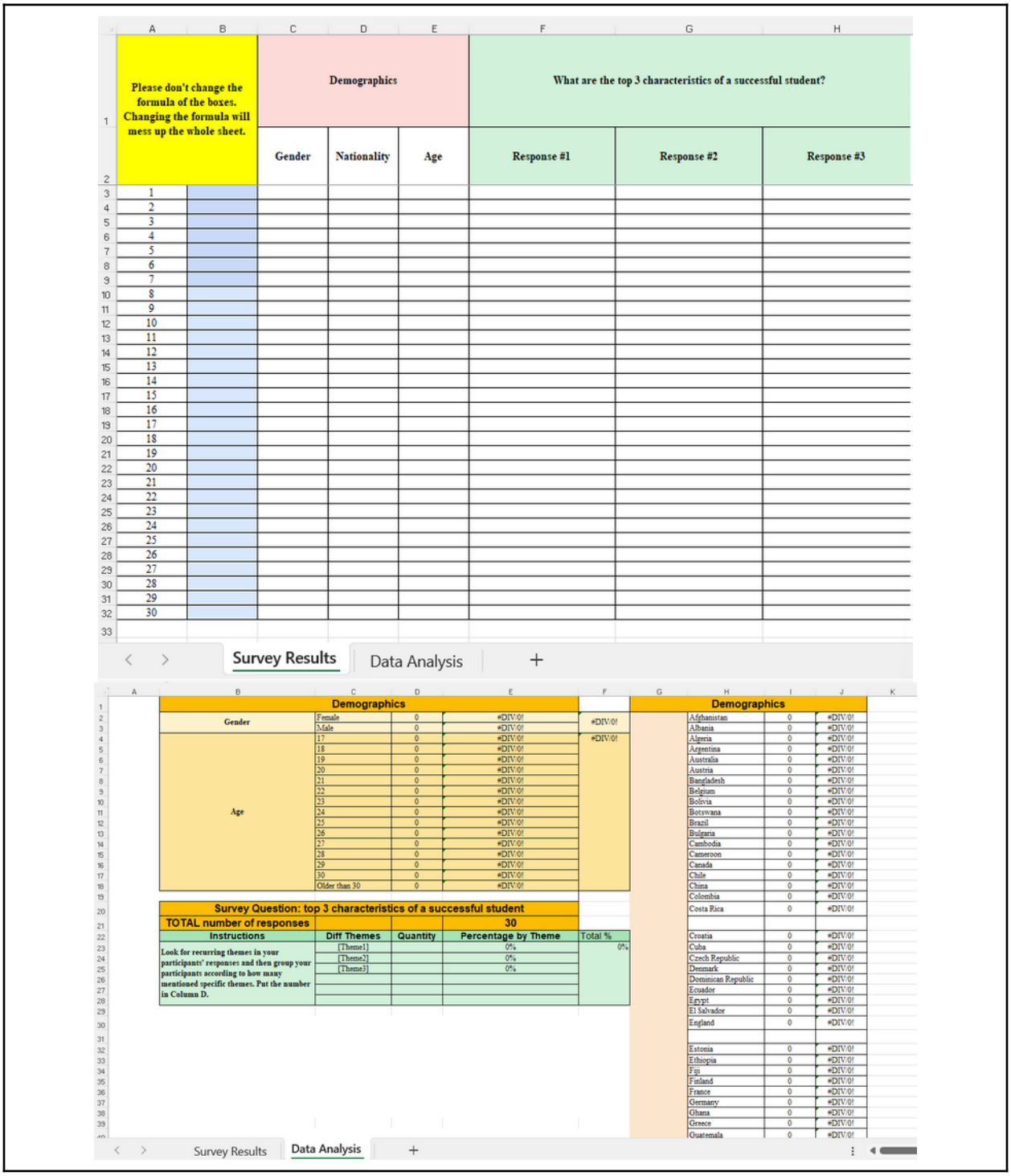




Note. Adapted from Nelson (2006).

Figure 17
Handout 3: Interaction, Part Two



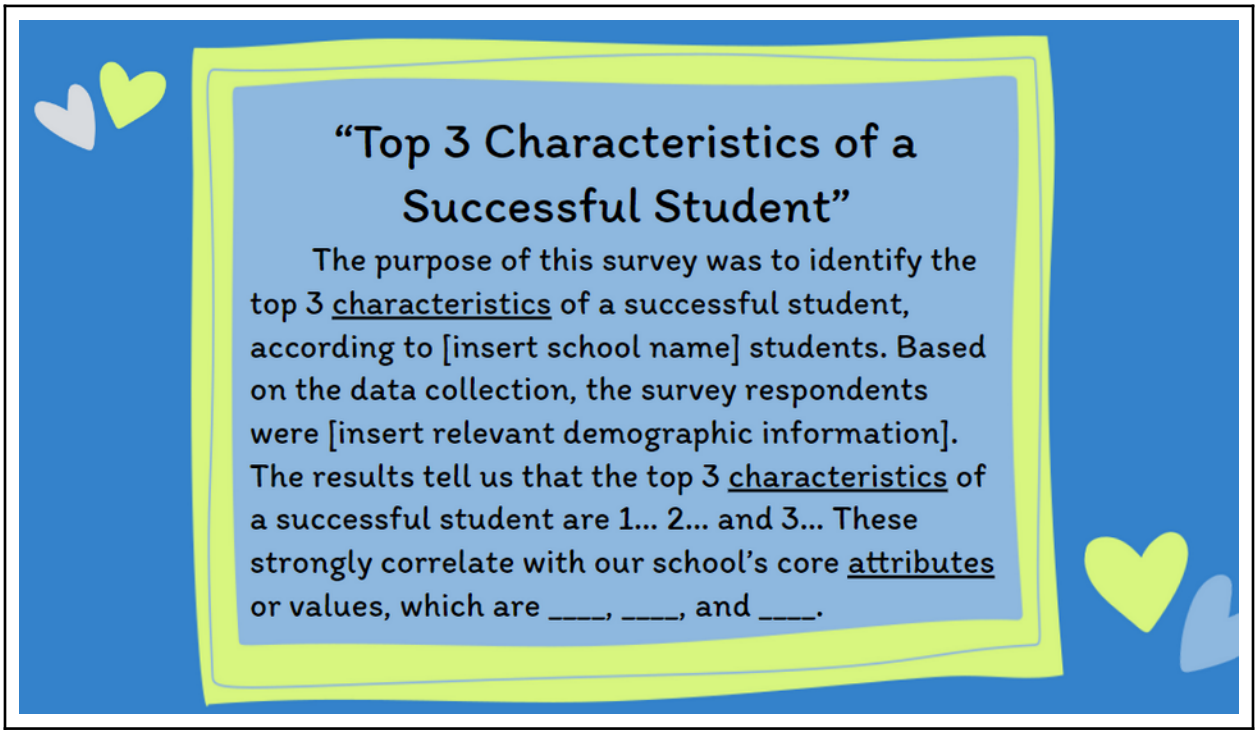


Please don't change the formula of the boxes. Changing the formula will mess up the whole sheet.		Demographics			What are the top 3 characteristics of a successful student?		
		Gender	Nationality	Age	Response #1	Response #2	Response #3
1							
2							
3							
4							
5							
6							
7							
8							
9							
10							
11							
12							
13							
14							
15							
16							
17							
18							
19							
20							
21							
22							
23							
24							
25							
26							
27							
28							
29							
30							

Survey Results | Data Analysis | +

Demographics					Demographics		
Gender	Female	0	#DIV/0!	#DIV/0!	Afghanistan	0	#DIV/0!
	Male	0	#DIV/0!		Albania	0	#DIV/0!
Age	17	0	#DIV/0!	#DIV/0!	Algeria	0	#DIV/0!
	18	0	#DIV/0!		Argentina	0	#DIV/0!
	19	0	#DIV/0!		Australia	0	#DIV/0!
	20	0	#DIV/0!		Austria	0	#DIV/0!
	21	0	#DIV/0!		Bangladesh	0	#DIV/0!
	22	0	#DIV/0!		Belgium	0	#DIV/0!
	23	0	#DIV/0!		Bolivia	0	#DIV/0!
	24	0	#DIV/0!		Botswana	0	#DIV/0!
	25	0	#DIV/0!		Brazil	0	#DIV/0!
	26	0	#DIV/0!		Bulgaria	0	#DIV/0!
	27	0	#DIV/0!		Cambodia	0	#DIV/0!
	28	0	#DIV/0!		Cameroon	0	#DIV/0!
	29	0	#DIV/0!		Canada	0	#DIV/0!
	30	0	#DIV/0!		Chile	0	#DIV/0!
	Older than 30	0	#DIV/0!		China	0	#DIV/0!
					Colombia	0	#DIV/0!
					Costa Rica	0	#DIV/0!
Survey Question: top 3 characteristics of a successful student							
TOTAL number of responses			30				
Instructions	Diff Themes	Quantity	Percentage by Theme	Total %			
Look for recurring themes in your participants' responses and then group your participants according to how many mentioned specific themes. Put the number in Column D.	[Theme1]		0%	0%			
	[Theme2]		0%				
	[Theme3]		0%				
					Croatia	0	#DIV/0!
					Cuba	0	#DIV/0!
					Czech Republic	0	#DIV/0!
					Denmark	0	#DIV/0!
					Dominican Republic	0	#DIV/0!
					Ecuador	0	#DIV/0!
					Egypt	0	#DIV/0!
					El Salvador	0	#DIV/0!
					England	0	#DIV/0!
					Estonia	0	#DIV/0!
					Ethiopia	0	#DIV/0!
					Fiji	0	#DIV/0!
					Finland	0	#DIV/0!
					France	0	#DIV/0!
					Germany	0	#DIV/0!
					Ghana	0	#DIV/0!
					Greece	0	#DIV/0!
					Guatemala	0	#DIV/0!

Survey Results | Data Analysis | +


"Top 3 Characteristics of a Successful Student"
The purpose of this survey was to identify the top 3 characteristics of a successful student, according to [insert school name] students. Based on the data collection, the survey respondents were [insert relevant demographic information]. The results tell us that the top 3 characteristics of a successful student are 1... 2... and 3... These strongly correlate with our school's core attributes or values, which are ____, ____, and ____.

Conclusion

When it comes to the application of corpus linguistics to teach plesionyms (or near-synonyms), there are many benefits and challenges. Some of the benefits include exposure to authentic, naturally occurring language in context and learner progression towards independent learning (Timmis, 2015; Friginal, 2018; Green, 2019). Using COCA, learners can build vocabulary banks dedicated to plesionyms and even create their own corpus based on the course text. Some of the challenges include the amount of corpus data, data randomization, and unfamiliarity with corpus query functions and techniques. All of these challenges are common reasons why teachers shy away from direct corpus use. Despite these familiar difficulties, plesionyms are context-dependent and require corpus analysis. The materials presented here, which present corpus data with added structure, are purposed to enable learners to discover the usage of the target phenomenon in context. Additional ways to develop students' ability to understand and use language in context include consulting experts within the target's high frequency registers (e.g., academic journals, web pages, popular magazines) and sub-registers (e.g., business, education, science/technology); exploring other corpora, such as parallel or learner corpora for cross-linguistic plesionym groups (e.g., L1 transfer in English to Japanese translation and vice versa); and implementation of student-initiated research (e.g., learners choose their own set of plesionyms to investigate) (Nelson, 2006; Hirst, 1995). One limitation of this study is its focus on a single corpus (COCA) and register (ACAD). There is potential to explore broader categories and corpora, such as parallel or learner corpora. Another limitation is insufficient data presented to represent patterns. Generally, the top 50 collocates and 10 concordance lines minimum should be presented for each pattern. The main searches were also limited to COCA's CHART, COLLOCATE, and COMPARE functions. In the end, there is always room for improvement and deeper exploration.

References

attribute. 2022. In *Merriam-Webster.com.* Retrieved November 2, 2022, from https://www.merriam-webster.com/dictionary/attribute

Biber, D., Leech, G. & Conrad, S. (2002). *Longman Student Grammar of Spoken and Written English*. Longman.

Biber, D., Reppen. R. & Conrad, S. (2007). *Corpus linguistics: Investigating language structure and use*. Cambridge University Press.

characteristic. 2022. In *Merriam-Webster.com.* Retrieved November 2, 2022, from https://www.merriam-webster.com/dictionary/characteristic

Common European Framework (2015). *English Vocabulary Profile Online - American English* [Website]. Available online at https://www.englishprofile.org/american-english.

Coxhead, A. (2000). A New Academic Word List. *TESOL Quarterly*, *34*, 213-238.

Cruse, D. A. (1986). *Lexical semantics.* Cambridge University Press.

Davies, M. (2008-). *The Corpus of Contemporary American English (COCA): 560 million words, 1990-present* [Website]. Available online at https://corpus.byu.edu/coca/.

DiMarco, C., Hirst, G., & Stede, M. (1993). The semantic and stylistic differentiation of synonyms and near-synonyms. In *AAAI Spring Symposium on Building Lexicons for Machine Translation* (pp. 114-121).

Edmonds, P. (1999). *Semantic Representations of Near-Synonyms for Automatic Lexical Choice*. University of Toronto.

Friginal, E. (2018). *Corpus linguistics for English teachers: New tools, online resources, and classroom activities*. Routledge.

Gardner, D., & Davies, M. (2013). A new academic vocabulary list. *Applied linguistics*, *35*(3), 305-327.

Gove, P. (1984). *Webster's new dictionary of synonyms*. Merriam-Webster.

Green, B. A. (2019). Teaching and assessing multiword expressions using an open educational resources academic English corpus. In E. Hinkel (Ed.), *Teaching essential units of language: Beyond single-word vocabulary* (pp. 134-164). Routledge.

Hirst, G. (1995). Near-synonymy and the structure of lexical knowledge. In *AAAI Symposium on Representation and Acquisition of Lexical Knowledge: Polysemy, Ambiguity, and Generativity* (pp. 51-56).

Johns, T. (1991). Should you be persuaded: Two samples of data-driven learning materials. *English Language Research Journal, 4*, 1-16.

Johns, T., & King, P. (Eds.). (1991). Classroom concordancing. *English Language Research Journal*, *4*.

Nelson, M. (2006). Semantic associations in Business English: A corpus-based analysis. *English for Specific Purposes*, *25*(2), 217-234.

O'Keeffe, A., McCarthy, M., & Carter, R. (2007). *From corpus to classroom: Language use and language teaching*. Cambridge University Press.

Room, A. (1985). *Dictionary of confusing words and meanings.* Routledge.

Timmis, I. (2015). *Corpus linguistics for ELT: Research and practice.* New York: Routledge.

trait. 2022. In *Merriam-Webster.com.* Retrieved November 2, 2022, from https://www.merriam-webster.com/dictionary/trait

Vyatkina, N., & Boulton, A. (2017). Corpora in language teaching and learning. *Language Learning and Technology*, *21*(3), 66-89.

Appendix A

Answer Key for "Handout 1: Illustration and Induction"

Guiding Questions:

A. *"Important" occurs in Lines 1, 2, 3, and 5.*

a. *Evaluative, positive, latent.*

B. *"Apartment" in Line 1, "security" "good or service" in Line 2, "usability and performance/audit trail" in Line 3, "leadership" in Line 4, and "health needs" in Line 5.*

a. *Abstract, latent, non-living.*

C. *"Demographic" in Lines 3 and 10. Also, "typical" and "biological" in Lines 6 and 7.*

a. *Descriptive, neutral, observable, topical.*

D. *"Student" in Line 6, "school dropout" in Line 9, and "tattooed persons" in Line 10.*

a. *Concrete, observable, living.*

E. *"Behavioral" in Lines 13 and 15. Also, "confident" "motivated" in Line 1, "innovative thinking" in Line 2,*

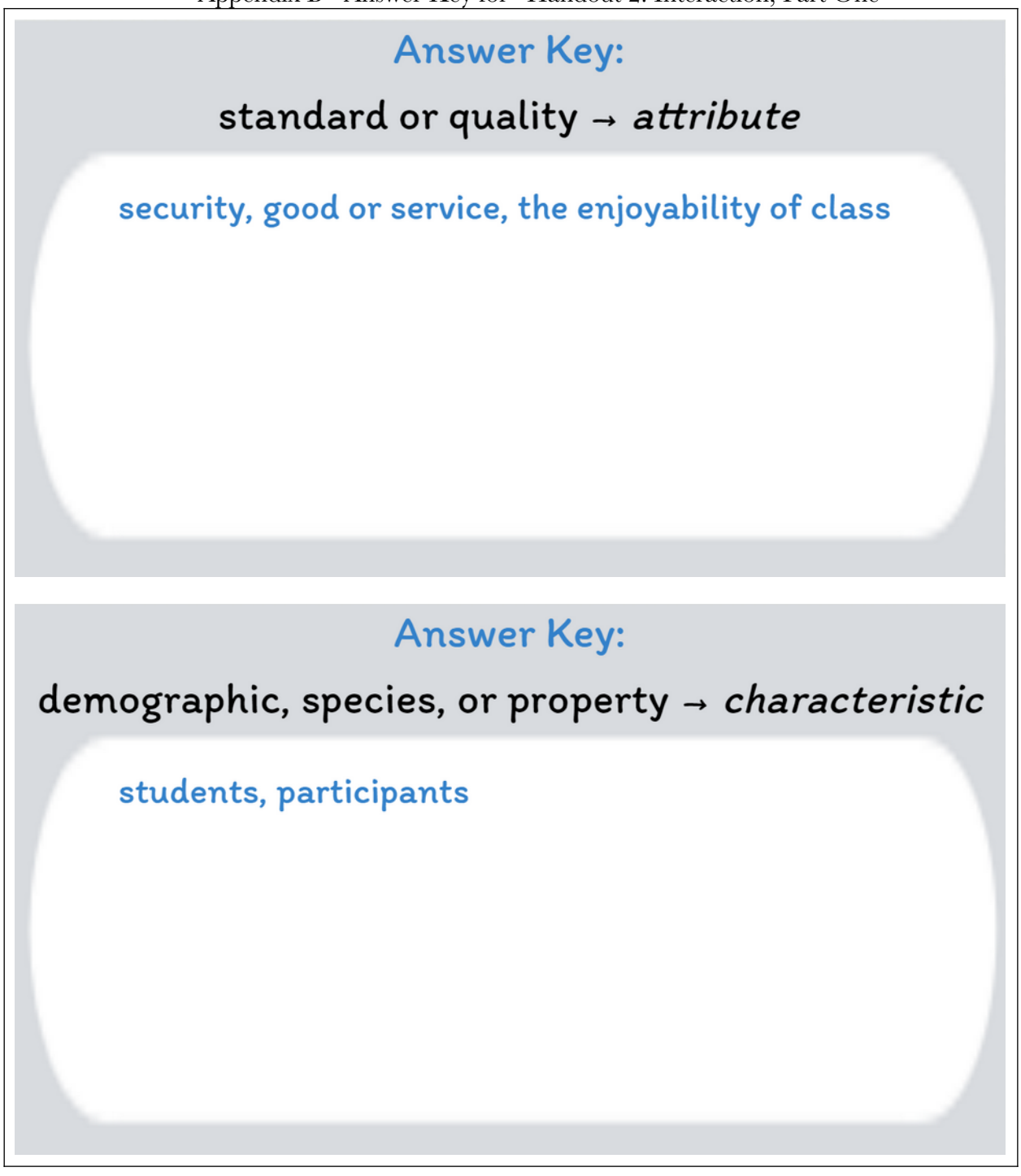
a. *Descriptive, neutral, observable, relational.*

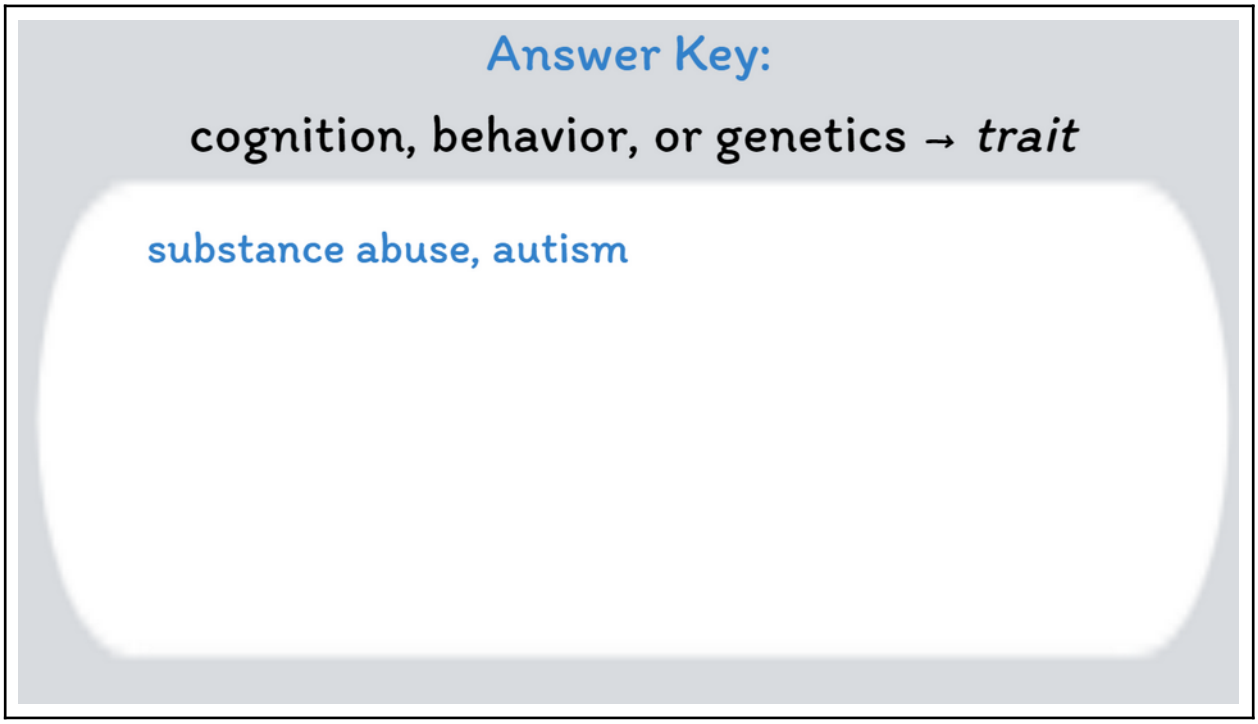
F. *"Him" in Line 11, "you" in Line 12, "fire ant" in Line 13, "children" in Line 14, "substance abuse, violence, criminal behavior…" in Line 15.*

a. *Concrete, observable, living,*

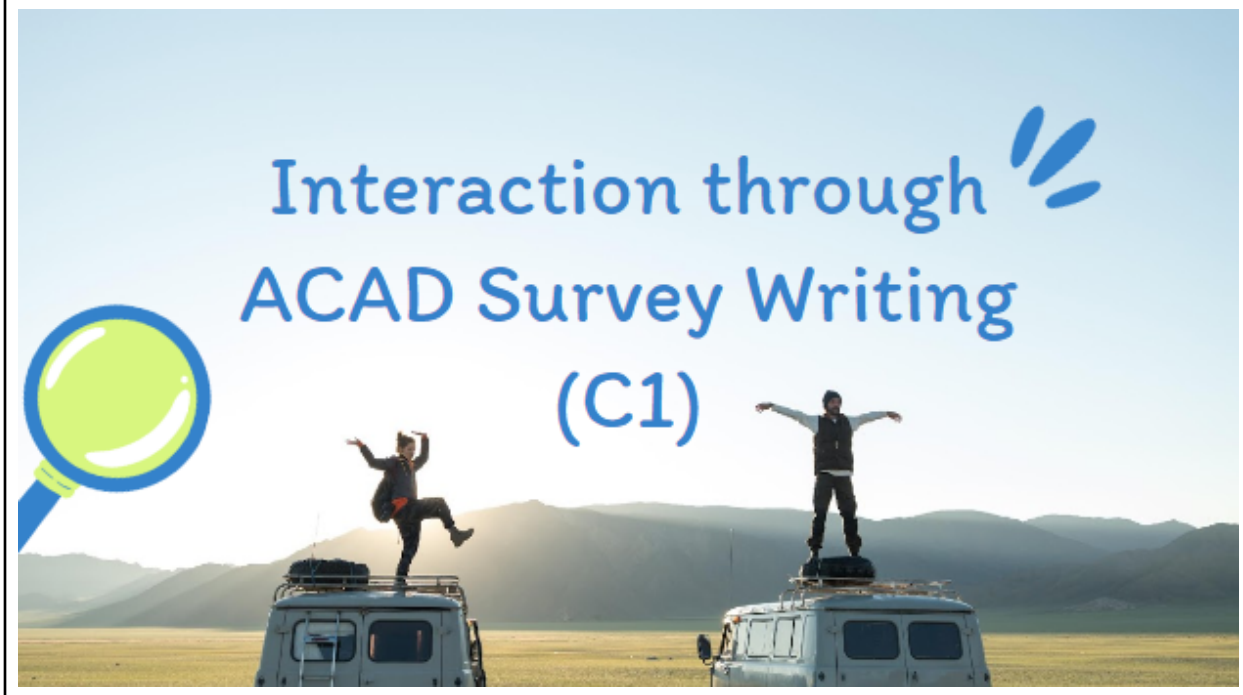
G. *'Attribute' and 'characteristic' both show examples of of-phrases. Of-phrases are used with species nouns. Using a corpus, maybe I want to investigate of-phrases for 'trait' too?*

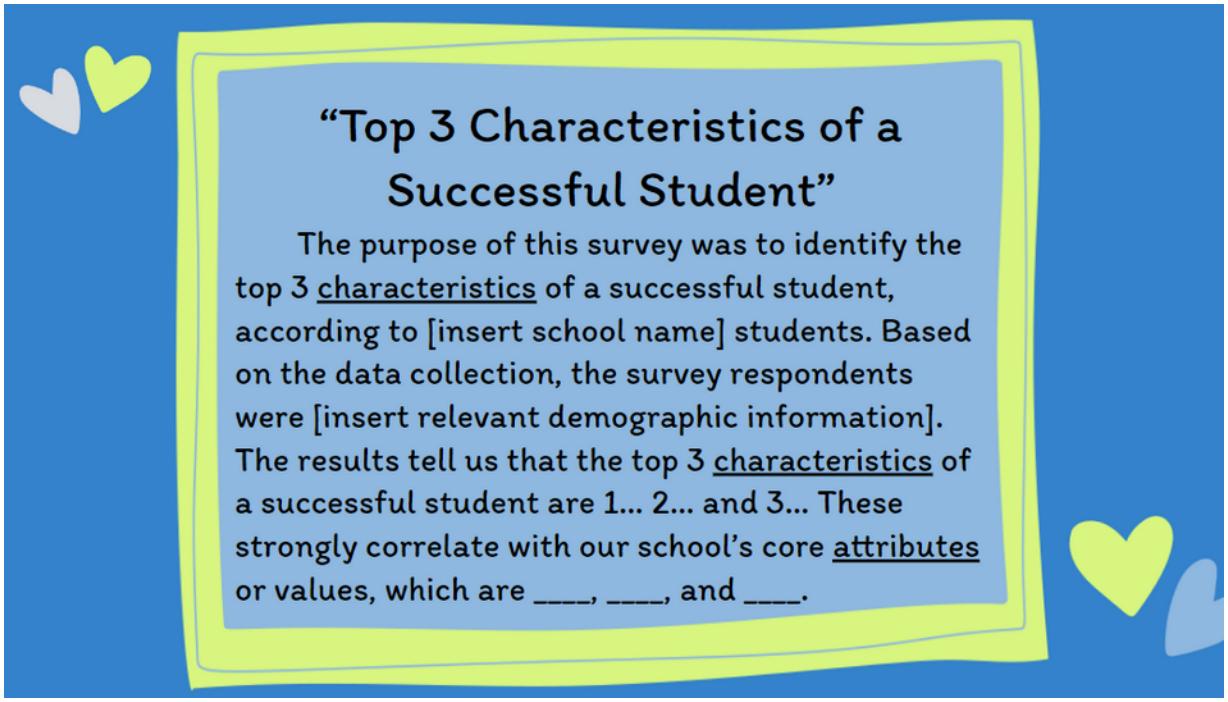
Appendix B - Answer Key for "Handout 2: Interaction, Part One"




Answer Key:
cognition, behavior, or genetics → *trait*
substance abuse, autism

Appendix C - Answer Key for "Handout 3: Interaction, Part Two"





About the Author

Sadie M. M. Nitta holds a BA TESOL degree from Brigham Young University – Hawai'i and an MA TESOL graduate from Hawai'i Pacific University. She has taught English in Hawai'i and Japan as well as contributed to corpus projects. Her main research interests are corpus linguistics, sociolinguistics, and peace linguistics.