

農学下分野間の語彙的相互関連性

—semi-popularization 記事によるコーパス分析—

山本 佳代

レイヴィン リチャード

Abstract

This paper examines the lexical interrelationships of the six sub-disciplines of the agricultural field to explore the possibilities of a pedagogic approach for EFL (English as a Foreign Language) undergraduates. First, the design of a new corpus consisting of 1 million words of semi-popularization articles in the realm of agriculture is described. In a comparison of the specialized corpus with the COCA (Corpus of Contemporary American English) as a reference corpus, 1,639 keywords were extracted. The keywords were categorized into: 1) the New General Service List, 2) the New Academic Word List, and 3) those not included in the NGSL or NAWL (hereafter referred to as “off-list”). These consisted largely of words such as calf, forage, and alga, which are agriculture-related words that are quite common in English-speaking countries. In a “cluster analysis” of 6 sub-disciplines’ keywords, the resulting dendrogram showed 3 major clusters corresponding roughly to lineages of “animals,” “plants,” and “both.” Based on these results, the pedagogical implications are discussed.

Keywords: AG-UOM corpus, semi-popularization articles, agriculture, undergraduate

はじめに

学術研究・教育機関等で使用することを念頭に、学術領域のテキストを収集し、英語教育を目的にコーパスや語彙リストが作成され、University Word List (Xue & Nation, 1984)、Academic Word List (Coxhead, 2000, 以下 AWL) を

始めとし、Academic Vocabulary List (Gardner & Davies, 2014)、New Academic Word List (Browne, Culligan, & Phillips, 2013) など、言語資料のデジタル化、統計資料の整備により語彙の編纂が行われている。そしてこれらは、複数の下位分野を包括する「一般的学術語彙」として提案されている。

一方、English for Specific Purposes (ESP) の有効性を主張する Hyland & Tse (2007) は、「一般的学術語彙」のリストであることをうたう AWL (Coxhead, 2000) に着目し、それが English for Academic Purposes (EAP) と一般英語 (General English) を区別するという点で、言語学習に関する意味のある議論を提供した点を評価しつつも、分野によっては使用しない語が多く含まれており、実際に各分野の学習者が使う語彙リストには成り得ていない、と指摘した。それよりも学習者自身の研究分野の語彙を直接リストアップし学習する方が良いと主張している。

その後、Hyland & Tse (2007) の主張に支持した形で、農学 (Martínez, Beck, & Panza, 2009)、工学 (Ward, 2009)、医学 (Chen & Ge, 2007; Wang, Liang, & Ge, 2008)、看護 (Yang, 2015) 等の分野固有の語彙リストが ESP として発表されている。これら ESP の分野は、学際性の範囲に差はあるが、さらに多岐にわたる下位分野が包括されている。Gilmore & Millar (2018) は、学際性の度合いの高い学術分野である土木工学 (Civil Engineering) について、東京大学の研究者の助言を得ながら、同大学の同学科の主な 11 の下位分野に含まれるとみなされる 1,100 本の研究論文を集めた約 800 万語のコーパスを構築し、土木工学全体の語彙リストと下位分野の語彙リストを作成するとともに、下位分野間の語彙的関連性をクラスター分析を用いて分析した。その結果、下位分野間の多様性の高さ、一部の語彙の共通性を指摘している。

しかしこのような特定領域コーパスは、しばしば研究論文 (Research Articles) のコーパスに基づいており、そこから得られた結果は、大学院生やキャリア初期の研究者のニーズには合うものの、大学学部生のカリキュラムでは使用が難しいことが指摘されている。学部学生を対象とした ESP リーディングコースでの教育を目的とし Muñoz (2015) は、農学分野の semi-popularization articles のコーパス分析を行い、本格的な研究論文と、新聞や一般雑誌などの記事の中間に位置づけられる、大学のニューズレターなどの記事に用いられる語彙の分析を行った。そして、それらの記事で使用される語彙にも農学の語彙の特徴が顕著に出ているとした。Muñoz (2015) が semi-

popularization articles を研究対象に選んだのは、自身のアルゼンチンの大学学部の ESP コースにて、それらの記事が授業でよく使われることに注目し、「専門的英文を読んだ経験がない学生にもっともアクセスしやすく、一般英語 (General English) とより専門的なテキストとの橋渡しと考えた」ためである (Muñoz, 2015, p. 27)。しかしながら、Muñoz (2015) の研究で使われた semi-popularization articles のコーパスは、トウモロコシ生産 (corn production) のトピックのみを扱った記事で構成されており、農学全般を対象とはしていない。

そこで、本研究では、English for Specific Purposes (ESP) の考えに基づき (Dudley-Evans & St Johns, 1998; Hutchinson & Waters, 1987)、M 大学の農学部生が学習目標とするにふさわしい語彙リスト作成を最終目標に、(1) 農学分野全般の semi-popularization articles からなる特殊コーパスを作成し、統計的基準に基づいてキーワードリストを作成・分析し、また、(2) 学際性の高い農学分野の特徴を捉えるために、日本にある国立大学 (M 大学) の農学部の 6 つの学科の教育・研究内容に沿って作成した上記コーパスの 6 つのサブコーパスを作成し、その語彙的関係性を捉えることを目的とした。リサーチ・クエスションは次の 2 点から成る。

1. 農学分野全般の semi-popularization articles のキーワードはどのようなものか。
2. 農学下位分野間の語彙的関係性はどのようなものか。

これらの問いに対する答えが持つ教育的示唆についても、あわせて議論する。

方法

AG-UOM コーパスの構築

トウモロコシ生産に関する大学のニューズレター記事を集めた Muñoz (2015) にならい、同じくアメリカの 11 大学がウェブ上に掲載しているニューズレターを収集し、M 大学の国際交流協定校 / 機関 (103 機関 2017 年時点) から、英語を母語とする地域に位置し、農学関連のニューズレターを発行

している5機関を加えた(2014年3月～2018年8月)。また、筆者のひとりが担当する農学部1年リーディングクラスで多くの学生が読みたい英文として選ぶ英文記事を発行する *Science Daily* と *Science News for Students* の各ウェブサイトから、2014年～2018年間に掲載された農学分野に関連する記事の収集を行った(*Science Daily* 2016年2月～2018年2月: *Science News for Students* 2014年3月～2017年2月)。これらの記事の発行大学/機関と記事数、発行時期については下記に示す(表1、表2、表3)。

表1 記事収集の対象となった大学・機関: Muñoz (2015) に基づく

Muñozを参照し収集した大学 (USA)	記事数	出版期間 (年/月)
North Dakota State University	168	2014/3 - 2018/8
Kansas State University	129	2014/2 - 2018/6
The Ohio State University	103	2014/3 - 2018/1
The Pennsylvania State University	48	2015/4 - 2016/12
University of Delaware	38	2014/5 - 2018/8
Iowa State University	19	2014/4 - 2017/1
University of Nebraska	17	2014/3 - 2017/1
University of Kentucky	16	2014/3 - 2016/9
Purdue University	15	2017/4 - 2018/8
Oklahoma State University	8	2016/6 - 2016/12
Louisiana State University	7	2015/5 - 2018/8

表2 記事収集の対象となった大学・機関: M 大学協定校に基づく

M 大学協定校・機関	記事数	出版期間 (年/月)
The Pirbright Institute (UK)	26	2015/7 - 2017/1
University of Stirling (UK)	25	2015/2 - 2016/12
The University of Toledo (USA)	22	2015/2 - 2016/11
University of California, Irvine (USA)	13	2015/6 - 2016/11
Canadian Food Inspection Agency (Canada)	10	2016/3 - 2017/1

表3 記事収集の対象となったウェブサイト

ウェブサイト科学雑誌	記事数	出版期間 (年/月)
<i>Science Daily</i>	732	2016/2 - 2018/2
<i>Science News for Students</i>	337	2014/3 - 2017/2
記事総数	1733	

これらの記事収集後、記事の内容に応じて、M大学農学部 の6つある学科/下位分野(表4)の専門的特徴と合うものに分類した。収集期間の差は、発行記事数の違いによるもので、分野間の語数の開きを期間で調整したことによる。記事の収集と分類にあたっては、M大学の農学研究科の大学院生2名(獣医学、資源環境科学をそれぞれ専攻する留学生)、およびM大学農学研究科の特徴をよく知る元大学職員2名(英語講師経験有り)に作業を依頼した。収集作業は第1期、第2期の2回に分けて行った。1期目は3名で行い、作業前に、分類する上でのポリシーを説明した後、こちらで準備した18本の記事を黙読し、記事の内容がどの学科に最も近いかを記入してもらった。この時の重複率(inter-rater reliability)は60%であったため、再度、各学部のシラバス内容や分類方法について100%になるまで確認を行った。第2期目は、協力者が1名増えたため、改めて収集と分類上のポリシーを説明し、第1期から参加している3名と新たに加わった1名を合わせて4名で再度重複率調査を実施した。重複率は70%と前回よりは上がったが、再度、分類の相違について全員が一致するまで確認と検討を重ねた。収集と分類に関するポリシーの確認は随時行った。このように研究協力者の助けを借りて、6つの農学下位分野のコーパスから成る、農学分野 semi-popularization articles の1,179,064語(約118万語)コーパスを作成した(表4)。Rea Rizzo(2010)によれば、特定の目的のために用いるコーパスとしては、最低100万語の語数を満たしているべきだとされるが、本コーパスは全体でその基準を満たした。大学における英語教育を目的に構築した農学分野全般のコーパスであることから(AGriculture-University of M)、以後AG-UOMコーパスと表記する。

表4 下位分野ごとの記事数と総語数

学科(下位分野)	記事数	Tokens
植物生産環境科学 Agricultural and Environmental Sciences (AE)	408	274,197
森林緑地環境科学 Forest and Environmental Sciences (F)	272	210,286
応用生物科学 Biochemistry and Applied Biosciences (B)	191	126,587
海洋生物環境科学 Marine Biology and Environmental Sciences (M)	192	118,621
畜産草地科学 Animal and Grassland Sciences (AG)	238	175,101
獣医学 Veterinary Sciences (V)	432	273,881
Token 総数		1,179,064

キーワードの抽出

Corpus of Contemporary American English (COCA) を参照コーパスとして、AG-UOM コーパスのキーワード抽出を行った。COCA のサブセクションは、AG-UOM と同時期の、2014 年～2017 年の written fiction, magazine and newspaper texts を用いた（総語数約 4,800 万語）。WordSmith 7.0 (Scott, 2016) を用いて行い、各語が特徴的であることを示す統計的指標には対数尤度 (Log-likelihood) を使用し、有意水準は $p = .0001$ に設定し、1 語につき最低 10 度生じたもののみをキーワードとした（以上は、Gilmore & Millar, 2018 と同じ基準）。また、Gilmore & Millar (2018) では、特定の記事のみに生起するキーワードを除外するための Dispersion の設定を「全体のテキストのうちの最低 5% に生起するもの」としている。しかし Gilmore & Millar (2018) の用いた学術論文のコーパスに比べると、本研究で用いている AG-UOM コーパスはより一般読者向けに書かれた semi-popularization articles を用いており、比較的一般的な語彙が多くを占める一方、各下位分野を代表する専門語の割合がもともと少なく特徴が出にくいと考えられるため、Dispersion を「1% 水準（全体のテキストのうちの最低 1% に生起するもの）」と設定した（全 1,733 記事あるうち、最低 17 本の記事に生起していることに相当する）。

キーワードにおける NGSL と NAWL、Off-list Words のカバー率

その次に、各キーワードのうち、(1) New General Service List (NGSL) (Browne et al., 2013) に含まれるもの（英語で最も頻繁に使われる 2801 語のリストにある語）、(2) New Academic Word List (NAWL) (Browne et al., 2013) に含まれるもの（多くの学術分野で頻繁に用いられるが、NGSL には含まれていない 963 語のリストに含まれるもの）、そして (3) (1) と (2) どちらにも含まれない “off-list” のキーワードに分類し、各下位分野の (1)～(3) のカバー率を算出し、Gilmore & Miller (2018) の SCCERA とのコーパス全体の比較を行った。下位分野間については、(1) と (2)、off-list keywords のカバー率比較を行った。抽出した “off-list” キーワードについては、下位分野間で比較し、それぞれの独自性と下位分野間の共通性について観察した。

キーワードに関する下位分野間の共通性の分析（クラスター分析）

さらに、本研究で仮定している農学の6つの下位分野に含まれるキーワードの分布をデータとして、「階層的凝集クラスター分析 (hierarchical agglomerative cluster analysis) Ward 法」を行い、下位分野間のキーワードに関する共通性から、下位分野間の「近さ」を析出した。「凝集型階層クラスター分析」とは、「あるグループ（分野）のメンバー（語）が、お互いにどれくらい似ているか、また他のグループのメンバーとはどれほど似ていないかという基準で、グループ間のクラスターを作る（凝集する）方法」である (Gries, 2013; Gilmore & Millar, 2018)。

結果と考察

AG-UOM コーパスのキーワード

COCA との比較により、統計的に有意な 1,639 語のキーワードが抽出された (対数尤度比がプラスのもののみ使用した $p < .0001$)。さらにこれらのキーワードにおける既存の語彙リストのカバー率をみたところ、NGSL に属する語が 836 語 (全体の 51.0%)、NAWL に属する語が 269 語 (全体の 16.4%)、そしてどちらにも属さない語が 534 語 (“off-list” キーワード, 全体の 32.6%) あった (表5)。この “off-list “キーワード” が、農学分野特有の語だと考えることができる。表5は、AG-UOM コーパスと Gilmore & Millar (2008) の土木工学学術コーパス (SCCERA) との比較である。off-list keywords が占める割合が、学術コーパスの SCCERA よりも AG-UOM コーパスが高いことが分かった。また、NAWL は殆ど同じであるが、NGSL は SCCERA の方が全体を占める割合が高い。このことから、AG-UOM コーパスには特徴的な語彙が多く含まれていることが窺える。

表5 NGSL、NAWL、off-list keywords の被覆率

コーパス	被覆率(%)			被覆率合計(%)
	NGSL	NAWL	Off-list keywords	
AG-UOM	51.0	16.4	32.6	100
SCCERA	60.4	16.9	22.8	100

Note: それぞれの off-list keywords は異なる。(AG-UOM = 534 words、SCCERA = 675 words)

“Off-list” キーワードの各農学下位分野コーパスにおける分布

農学分野の特徴を表わすと考えられる “off-list” キーワードの各下位分野トップ20を俯瞰する。表6は、COCAを参照コーパスとし、各下位分野のキーワードリストを作った上で、さらに “off-list” キーワードの各下位分野トップ20の語を示したものである。複数の下位分野に生じている語は太字にしてある。

表6 農学下位分野における off-list キーワードトップ20

AE	AG	B	F	M	V
CORN	BEEF	PIRBRIGHT	DEFORESTATION	CORAL	MAMMALS
BEES	CALVES	VACCINE	DROUGHT	REEFS	VETERINARY
SOYBEAN	FORAGE	INSTITUTE	INVASIVE	MARINE	DNA
WEEDS	CALF	GERMS	EROSION	CORALS	MOSQUITOES
SOYBEANS	GRAZING	MICROBES	PERMAFROST	REEF	CHIMPANZEES
GLYPHOSATE	CALVING	DNA	FORESTRY	ALGAE	PRIMATES
KOCHIA	HERD	MEALWORMS	SERICEA	WHALES	EXTINCT
WEED	HAY	FMD	ECOSYSTEM	SHARKS	CALVING
POLLEN	LIVESTOCK	MICROCYSTIN	ASH	ECOSYSTEMS	LEMURS
ATRAZINE	PASTURE	SPIDERS	TIMBER	BLEACHING	EXTINCTION
DICAMBA	WEANING	VACCINES	WILDLIFE	WHALE	BUGS
BEE	BULLS	BIOFILMS	CFIA	PHOSPHORUS	PRIMATE
ANTS	SILAGE	ATMOSPHERIC	CANOPY	MAMMALS	ELEPHANTS
KANSAS	BULL	BIOMASS	ECOSYSTEMS	AQUACULTURE	CALVES
FERTILIZER	HEIFERS	DR	MOISTURE	EEL	GORILLAS
GROWERS	WEANED	AVIAN	PEAT	DOLPHINES	DINOSAURS

SORGHUM	ACRE	FMDV	BIOMASS	UNDERWATER	POULTRY
ALFALFA	KRIS	INFLUENZA	ERUPTION	FISHERIES	AVIAN
MOESTURE	PASTURES	PCR	MICROBES	ZOOPLANKTON	GORILLA
ACRE	CORN	FLU	RAINFOREST	ECOSYSTEM	FEATHERS

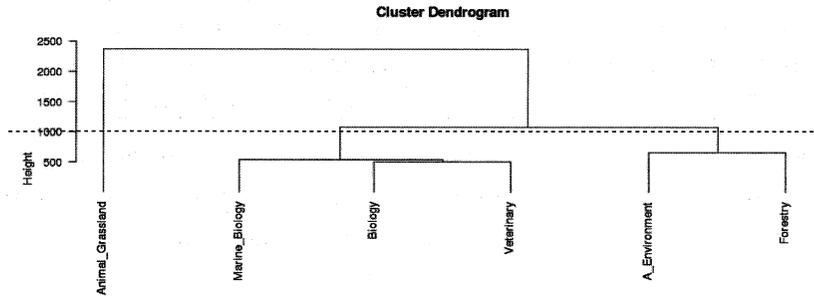
Note: AE = Agricultural Environmental Sciences, AG = Animal Grasslands Sciences, B = Biochemistry and Applied Biology, F = Forest and Environmental Sciences, M = Marine Biology and Environmental Sciences, V = Veterinary Sciences

このように見ると、各下位分野は独自の特徴語を持ちながら、下位分野間で共通に存在する語があることが分かる。植物環境科学科 (AE) で主に作物として言及される CORN は、畜産草地科学では飼料としての CORN としてトップ 20 にランク入りしていたり、畜産草地科学でしばしば言及される CALF, CALVES などが、獣医学でもトップ 20 にランク入りしている。また、特定の метод論や手法を示すような語が、複数の下位分野間で現れていることもある。たとえば応用生物科学に現れている DNA は、獣医学でも上位に入っているが、これは両分野が遺伝子学や遺伝子工学をも応用していることを示していると思われる。

その一方で、1つの分野に顕著に表れる語も少なくない。例えば、海洋科学に特徴的な CORAL, REEF, ALGAE や、植物生産科学で特徴的な BEE, POLLEN、FERTILIZER は、他のどのトップ 20 にも現れていない。

キーワードに関する下位分野間の共通性の分析 (クラスター分析)

このような下位分野間のキーワードの共通性をデータとして、下位分野間の「距離」を凝集型階層クラスター分析を用いて統計的に算出し、グラフ化した (図 1)。距離の算出には「ユークリッド距離」を用いた。



Note. A_Environment = Agricultural Environmental Sciences, Animal_Grassland = Animal Grasslands Sciences, Biology = Biochemistry and Applied Biology, Forestry = Forest and Environmental Sciences, Marine_Biology = Marine Biology and Environmental Sciences, Veterinary = Veterinary Sciences

図1 6つの下位分野キーワードによるクラスター分析結果

図1のように線を引いた場合、まず、主に植物を対象とする科学である植物生産環境科学 (AE) と森林緑地生産科学 (F) が一まとまりとなり、その一方で比較的動物を対象とすることの多い、海洋生物環境科学 (M)、応用生物科学 (B)、獣医学 (V) がひとまとまりとなっている。例えば、DNA という語が共通して上位にあった応用生物科学と獣医学は「距離が近い」との分析結果となった。また、一方で、畜産草地科学 (AG) は、これらの植物科学と動物科学のグループを超越して、単独のクラスターとなった。この点については、畜産草地科学は、実際に植物科学 (「草地」) と動物科学 (「畜産」) の両方の特徴を色濃く併せ持つ学際性の高い分野であると考えられる。畜産草地科学の off-list トップ 60 では、家畜の繁殖や生育課程を表す語が多く (TRIMESTER, CALVE, CALVING, HEIFER, STEER, SIRES)、それらに伴う飼料としての植物の種類 (BROMEGRASS, ALFALFA, OATS, CORN, SOYBEANS) や栄養に関する語 (NUTRITIONAL, NUTRITIONIST) が並ぶのがこの分野の特徴と言える (表7)。

表7 畜産草地科学 off-list トップ 60

BEEF	WEANING	MOISTURE	HERDS	RATION	TRIMESTER
CALVES	BULLS	DICKINSON	PONDER	TETANY	LACTATION
FORAGE	SILAGE	FORAGES	SORGHUM	GRASSLANDS	NITRATE
CALF	BULL	CHAPS	BENCHMARK	MANAGERIAL	FEEDLOT
GRAZING	HEIFERS	STEERS	BALES	BROMEGRASS	UD
CALVING	WEANED	CARCASS	DAKOTA	YEARLING	BARLEY
HERD	ACRE	DROUGHT	CALVE	GESTATION	SIRES
HAY	KRIS	NUTRITIONAL	INTAKE	TURNOUT	NUTRITIONIST
LIVESTOCK	PASTURES	ALFALFA	GRAZE	OATS	GRAZED
PASTURE	CORN	PHOSPHORUS	INVENTORY	UTILIZATION	SOYBEANS

結論

6つの農学下位分野に関連するトピックの Semi-popularization Articles を集め、100万語レベルの AG-UOM コーパスを作成した。そのコーパスから、COCA の非学術的テキストを参照コーパスとして用いて、キーワードを抽出し、農学分野のキーワードリストを作成した。また、やはり COCA を参照コーパスとして、M 大学の農学部6学科に対応する6つのサブコーパスにおけるキーワードを別途抽出し、下位分野間の比較を行った。その結果、下位分野間の語彙的関連性が浮かび上がった。

まず、11の下位分野を持つ土木工学分野全体の学術論文で構成された Gilmore & Millar (2018) の SCCERA コーパスと、6つの下位分野を持つ農学分野 AG-UOM コーパスのキーワードにおける NGSL と NAWL、off-list キーワードの被覆率から見ていきたい。NAWL では、両コーパス間に大きな被覆率の差は見られなかった (SCCERA 16.9%、AG-UOM 16.4%)。しかし、NGSL は約 10% という違いで AG-UOM コーパスの被覆率が低く算出される (SCCERA 60.4%、AG-UOM 51.0%)、off-list キーワードでは逆に約 10% 高い被覆率となった (SCCERA 22.8%、AG-UOM 32.6%)。この結果により、AG-UOM には土木工学分野学術コーパスよりも一般的な語の比率が低いことがわかった。つまり、より特徴的な語が多く含まれることが示唆された。語の種類としては、calf や forage, graze, hay, microbe, alga, fertilizer など、英語

圏では比較的日常的に使われ、農学分野で高頻度な反面、いわゆる AWL のような従来の学術語彙リストから抜け落ちている語が多く抽出された。Martinez et al. (2009) による農学分野の研究論文コーパスに基づくキーワードリストよりも、対象学生である大学学部生の持つ学術的スキルや英語力にふさわしい語が抽出できたと考えられる。これらは、初年次の農学部生でも比較的習得し易い語であり、かつ専門分野でも知っておきたい語と言える。また、農学下位分野間の“off-list”キーワード比較により、各分野の独自性が明らかになると共に、下位分野間の類似性の特徴から、横断的要素が下位分野間に存在することが分かった。例えば、応用生物科学 (B) と獣医学 (V) では遺伝学が大きな位置を占める点で共通しており、語彙面でもその傾向が見える。

クラスター分析からは、これら6つの下位分野には動物科学としてのまとまりと植物科学としてのまとまりの2つの流れがあることが分かった。そしてそれらをまたぐ形の畜産草地科学 (AG) の強い学際性を指摘することが可能となった。

これらの結果から、「農学単語リスト」または、縦割り式の下位分野ごとの単語リストではなく、下位分野ごとに共通性と独自性の両方を備えた「単語リスト」が、農学部および農学部生のニーズに最も合った教材の形であると考えられる。

謝辞

本論文においてデータ分析に協力頂いた宮崎大学語学教育センター荒木瑞夫氏に感謝の意を表す。

本研究は、JSPS 科研費 JP16K02844、19K00913 の助成を受けたものである。

参考文献

- Browne, C., Culligan, B., & Phillips, J. (2013). The new general service list. <http://www.new-generalservicelist.org>. (Accessed 29 May 2019).
- Chen, Q., & Ge, G. (2007). A corpus-based lexical study on frequency and distribution of Coxhead's AWL word families in medical research articles (RAs). *English for Specific*

- Purposes*, 26(4), 502–514. <https://doi.org/10.1016/j.esp.2007.04.003>
- Coxhead, A. (2000). A New Academic Word List. *TESOL Quarterly*, 34(2), 213–238. <https://doi.org/10.2307/3587951>
- Dudley-Evans, T., & St John, M. (1998). *Developments in ESP: A multi-disciplinary approach*. Cambridge: Cambridge University Press.
- Gardner, D., & Davies, M. (2014). A new academic vocabulary list. *Applied Linguistics*, 35(3), 305–327. <https://doi.org/10.1093/applin/amt015>
- Gilmore, A., & Millar, N. (2018). The language of civil engineering research articles: A corpus-based approach. *English for Specific Purposes*, 51, 1–17. <https://doi.org/10.1016/j.esp.2018.02.002>
- Hutchinson, T., & Waters, A. (1987). *English for Specific Purposes: A learner-centered approach*. Cambridge: Cambridge University Press.
- Hyland, K., & Tse, P. (2007). Is there an “academic vocabulary”? *TESOL Quarterly*, 41(2), 235–253.
- Martínez, I. A., Beck, S. C., & Panza, C. B. (2009). Academic vocabulary in agriculture research articles: A corpus-based study. *English for Specific Purposes*, 28(3), 183–198. <https://doi.org/10.1016/j.esp.2009.04.003>
- Muñoz, V. L. (2015). The vocabulary of agriculture semi-popularization articles in English: A corpus-based study. *English for Specific Purposes*, 39, 26–44.
- Rea-Rizzo, C. (2010). Getting on with corpus compilation: From theory to practice. *ESP World*, 27, 1–23.
- Wang, J., Liang, S., & Ge, G. (2008). Establishment of a Medical Academic Word List. *English for Specific Purposes*, 27(4), 442–458. <https://doi.org/10.1016/j.esp.2008.05.003>
- Ward, J. (2009). A basic engineering English word list for less proficient foundation engineering undergraduates. *English for Specific Purposes*, 28(3), 170–182. <https://doi.org/10.1016/j.esp.2009.04.001>
- Xue, G., & Nation, P. (1984). A University Word List. *Language Learning and Communication*, 3(2), 215–229.
- Yang, M. N. (2015). A nursing academic word list. *English for Specific Purposes*, 37(1), 27–38. <https://doi.org/10.1016/j.esp.2014.05.003>