

---

Title	Improving disciplinary literacy by developing vocabulary and grammatical profiles: The secondary vocabulary lists
Author(s)	Clarence Green and James Lambert

---

Copyright © 2020 Office of Education Research (OER), NIE

## OER FINAL REPORT SERIES

The OER Final Report series includes final reports from funds managed by Office of Education Research, National Institute of Education, Nanyang Technological University.

Reports are submitted as part of the funding review process and intended for the funding agency, local schools and educators, teacher educators, policymakers, and education scholars. They do not take the place of scholarly, peer-reviewed articles but report on the background, procedures, and major findings of the project.\*

This study was funded by Singapore Ministry of Education (MOE) under the Education Research Funding Programme (DEV 01/16 CGG) and administered by National Institute of Education (NIE), Nanyang Technological University, Singapore. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the Singapore MOE and NIE.

\*In some cases reports show coloured font or highlights. These are an artifact of the review process and not intended to have any special weight or meaning within the report itself.

EDUCATION RESEARCH FUNDING PROGRAMME

**PROJECT CLOSURE REPORT**



**Improving Disciplinary Literacy by Developing Vocabulary  
and Grammatical Profiles:  
The Secondary Vocabulary Lists**

By

Clarence Green  
James Lambert

National Institute of Education  
Singapore

## **EXECUTIVE SUMMARY (NO MORE THAN 5 PAGES)**

### **INTRODUCTION/BACKGROUND**

The knowledge of academic vocabulary is crucial for educational success, and recently there has been a push amongst teachers and researchers to assist students at the pretertiary level to develop their disciplinary literacy and understanding of how academic English varies across disciplines.

### **STATEMENT OF PROBLEMS**

English for Academic Purposes research has developed advanced methods for producing corpus-informed vocabulary resources, but these have yet to be fully leveraged to promote disciplinary literacy within the secondary school context. For example, the focus of most previous wordlists has been on general academic vocabulary or the discipline specific vocabulary needed in tertiary education.

### **PURPOSE OF STUDY**

The current research contributes a series of discipline-specific wordlists for secondary school education, the Secondary School Vocabulary Lists (SVL),

### **METHODOLOGY / DESIGN**

A corpus was developed of textbooks covering eight core subjects: Biology, Chemistry, Economics, English, Geology, History, Mathematics, and Physics. Key vocabulary was extracted from the materials and developed into pedagogical wordlists.

### **FINDINGS / RESULTS**

The Secondary Vocabulary List (SVL) captures all the essential vocabulary students need to read well in eight subjects and should be taught in schools. The SVL goes beyond wordlists alone in developing accompanying word family and word association (i.e. collocation) lists for the disciplinary lexis. The SVL thus provides secondary education teachers with an unprecedented set of resources covering key vocabulary for the eight core disciplines informed by innovative corpus methods.

### **CONTRIBUTIONS**

Theory: We have published two articles in high impact journal regarding the vocabulary of secondary school subject areas.

Practice: We have developed multiple resources for disciplinary literacy pedagogy.

NIE Programmes: Dr Linda HANINGTON has incorporated our Secondary Vocabulary Lists into the Practical Pronunciation for Teacher course, that all PGDE students have to take. Thus our work has impacted all PGDE students.

Others: I have had several requests about the SVL. Dr Paul Doyle and ELIS in particular have plans to use the work, which is very exciting, to form the basis for a pronunciation dictionary.

### **CONCLUSION**

The SVL is a very rich resource for vocabulary and could be extremely useful for students' English language development in Singapore if promoted to teachers.

### **ACKNOWLEDGEMENTS**

This study was funded by the Education Research Funding Programme, National Institute of Education (NIE), Nanyang Technological University, Singapore, project no. OER DEV 01/16

Project Number: DEV 01/16 CGG  
Name of PI: Clarence Green

Form RC3b | **2016**

---

CGG. The views expressed in this paper are the author's and do not necessarily represent the views of NIE.

**KEYWORDS**

Vocabulary; Literacy; Corpus Linguistics; Pedagogical Resources

# **Improving Disciplinary Literacy by Developing Vocabulary and Grammatical Profiles: The Secondary Vocabulary Lists**

**Clarence Green**

**James Lambert**

National Institute of Education

## **INTRODUCTION/BACKGROUND**

With an increasing number of school students continuing to tertiary education, developing students' academic literacy during secondary school is essential. As Wingate (2011) notes, "in today's mass higher education system, many students are not fully prepared for the demands of academic study" (p. 66). Thus, research-based resources with the potential to assist teachers prepare secondary students for tertiary education are essential. Vocabulary is a vital component of educational success in both first and second language contexts (Webb & Nation, 2017), and in English for Academic Purposes (EAP) pedagogical materials to facilitate instructed vocabulary acquisition have been the focus of much research; in particular, advanced methods have developed for producing corpus-informed wordlists (Gardner & Davies, 2014; Lei & Liu, 2016). However, such wordlists have largely been developed in the context of tertiary level ESL education. The current advances in EAP for developing such educational resources have yet to be fully leveraged in other educational contexts.

Recent trends in this direction are exemplified by the Middle School Vocabulary Lists (Greene & Coxhead, 2015). Our project further advances the trend by reporting on a large corpus project that releases to the research and teaching community the Secondary School Vocabulary Lists (SVL), a series of lemma-based discipline-specific academic vocabulary lists.

These lists cover important vocabulary from eight core subjects that students need to master during secondary education: Biology, Chemistry, Economics, English, Geology, History, Mathematics, and Physics. Furthermore, the current study goes beyond previous wordlist development by contributing accompanying pedagogical lists containing word associations (i.e. collocations) and word families.

The SVL is designed for secondary school as a resource to assist teachers in the development of disciplinary literacy. Disciplinary literacy is a pedagogical approach and research area of increasing significance in secondary education internationally.

### **STATEMENT OF PROBLEMS (I.E., JUSTIFICATION FOR THE STUDY)**

The value of pedagogical material informed by objective methodological procedures developed in corpus linguistics is widely recognized (Brezina & Gablasova, 2017). The insight underlying such wordlists is that frequency, combined with metrics such as range and dispersion, profiles for teachers and students the relative usefulness of words (Brezina & Gablasova, 2017). Nation (2006, p. 79), for example, calculated that for any general text, comprehension requires about 98% understanding of its vocabulary, and that remarkably a teacher can provide this coverage by targeting the most frequent 8000-9000 word families of English.

Vocabulary provides a foundation from which grammar, phonology, and morphology emerge, and in a subject area it provides access to conceptual knowledge (Coxhead, 2018). Vocabulary selection for pedagogical purposes is therefore crucial.

### **PURPOSE OF STUDY (INCLUDING RESEARCH QUESTIONS AND/OR OBJECTIVES)**

To connect the SVL to current secondary school needs, the corpus was based on secondary school textbooks; the majority (82%) published within the past five years. Texts were taken from the ATL textbook lists of the Ministry of Education (Singapore), and United

Kingdom and Singapore A-level/O-level syllabi documents (the UK and Singapore exams being aligned). All textbooks were published in these two countries. When recommended texts were exhausted, texts specifically marketed at O-Level or A-level were sampled to reach the desired word count. In total, the corpus consisted of 206 textbooks across 8 disciplines.

## **METHODOLOGY/DESIGN**

Texts were scanned using Omnipage 18 and converted via optical character recognition to plain text. Indexes, references, front matter, and contents pages were removed. Following Gardner and Davies (2014) and Lei and Liu (2016), the SVL is based on lemmas rather than word families. A lemma had to meet 6 criteria within each discipline to be included in the SVL. The methods replicate the current state-of-the-art methods developed by Davis and Gardener (2014) for their Academic Vocabulary List and extended to discipline-specific material by Lei and Liu (2016). This procedure is based on lemmas rather than word families. A lemma had to meet each of the following criteria to be included in the pedagogical list.

1. Minimum Frequency: A lemma was required to have a minimum 28.57 occurrences per million words in the target discipline subcorpus. This metric was added by Lei and Liu (2016) to the methods of Davies and Gardener (2014), and taken from Coxhead (2000), in order to ensure a lemma was of relatively high frequency in the discipline.

2. Range: A lemma was required to occur in more than 50% of texts for a discipline, following Davis and Gardner (2014). This ensured that a word was often encountered in the subject's reading material.

3. Dispersion: Following Lei and Liu (2016), a dispersion metric of 0.5 was set. The Oakes Dispersion test was used in Concord (Scott, 2016), which divided the corpus for a discipline into 8 equal parts and computed homogeneity of occurrence across these equal parts. A lemma with low dispersion, i.e. below 0.5, was deemed to be unevenly distributed across the corpus and thus, like range, this indicated it was not of common utility across the discipline.

4. Range Ratio: As with the two previous studies, a further requirement was that a lemma occur in 50% of the texts in the target discipline with at least 20% of its overall expected frequency, i.e., 20% the overall frequency of the lemma in the discipline. This metric ensures that a word is at a relatively high frequency consistently in the discipline.

5. Discipline-specific frequency Ratio and Keyness: It was decided that a lemma needed to be a keyword in the discipline at a ratio 3 times higher than the rest of the corpus. Keyness was computed by the Keyword tool of Wordsmith (Scott, 2016) against the rest of the corpus, i.e., all other disciplines excluding the target discipline. This departs slightly from both Davies and Gardner (2014) and Lei and Lui (2016) who used a lemma inclusion ratio of 1.5 times higher than a general corpus, e.g. COCA. In fact, Gardner and Davies (2014) excluded words that were 3 times higher because they felt that this ratio would exclude lemmas that were too discipline-specific. Therefore, it was the latter ratio that was adopted to target discipline-specific lemmas and by using the rest of the academic corpus rather than a general corpus as a reference corpus for computation, this eliminated lemmas that were unlikely to be the focus of instruction, i.e., had low 'teachability' (Martinez & Schmitt, 2012; Simpson-Vlach & Ellis, 2012).

6. Major Part of Speech: A lemma was required to be amongst the four major parts of speech, namely: noun, verb, adjective, or adverb. Other word classes such as modifiers, modal verbs, auxiliaries, prepositions, and so forth, were excluded even if they met other statistical benchmarks as they were deemed to have low teachability.

## **FINDINGS / RESULTS**

The SVL fifty most frequent lemmas for each discipline are given in rank order in Table 1, the complete lists have been published in Green and Lambert (2018, 2019) and are freely available. As is clear from Table 1, even an intuitive check suggests the methods and procedures have produced a series of lists that reflect important vocabulary across the eight different subject areas. In total, the procedures produced 897 target lemmas for Biology, 524

for Chemistry, 480 for Economics, 686 for English, 720 for Geography, 728 for History, 559 for Physics, and 257 for Mathematics. Together, there are 4,856 lemmas that can be targeted by teachers representing the core of the academic language in the reading material of secondary school students.

**Table 1**

Top 50 most frequent lemmas for eight core subjects with part-of-speech label

Biology			Chemistry		Economics		Geography	
Rank	Lemma	Freq. p/m	Lemma	Freq. p/m	Lemma	Freq. p/m	Lemma	Freq. p/m
1	cell.n	11982	reaction.n	9319	price.n	8204	area.n	4746
2	blood.n	6125	acid.n	7469	cost.n	4906	country.n	3613
3	plant.n	3809	ion.n	5692	demand.n	4273	population.n	2465
4	enzyme.n	2971	atom.n	5533	rate.n	4147	food.n	2302
5	molecule.n	2844	form.v	5268	firm.n	4004	river.n	1985
6	gene.n	2762	solution.n	5107	income.n	3688	land.n	1900
7	dna.n	2700	electron.n	4942	good.n	3546	development.n	1658
8	protein.n	2515	gas.n	4479	market.n	3514	city.n	1645
9	body.n	2410	bond.n	4140	government.n	3166	place.n	1627
10	glucose.n	2332	metal.n	4051	supply.n	3014	cent.n	1432
11	concentration.n	2279	carbon.n	3922	business.n	2934	rock.n	1390
12	organism.n	2070	molecule.n	3884	tax.n	2875	soil.n	1380
13	contain.v	2062	compound.n	3696	bank.n	2861	global.adj	1336
14	chromosome.n	2059	hydrogen.n	3690	profit.n	2846	sea.n	1336
15	oxygen.n	2057	element.n	3492	curve.n	2781	impact.n	1229
16	membrane.n	2039	group.n	3418	economy.n	2733	tourist.n	1220
17	structure.n	2020	sodium.n	2889	money.n	2728	local.adj	1183
18	carbon.n	1984	mass.n	2813	output.n	2636	map.n	1173
19	muscle.n	1629	react.v	2812	product.n	2627	million.n	1172
20	allele.n	1577	oxide.n	2614	total.adj	2364	wind.n	1166
21	leaf.n	1523	formula.n	2487	worker.n	2363	climate.n	1160
22	dioxide.n	1516	chloride.n	2394	service.n	2292	activity.n	1159
23	tissue.n	1448	oxygen.n	2184	trade.n	2247	tourism.n	1139
24	substance.n	1433	mole.n	2183	quantity.n	2215	urban.adj	1137
25	amino.n	1364	table.n	2090	account.n	2156	environment.n	1135
26	wall.n	1342	structure.n	2060	consumer.n	2104	earthquake.n	1093
27	carry.v	1312	contain.v	2031	pay.v	2028	natural.adj	1018
28	photosynthesis.n	1278	substance.n	1979	economic.adj	1978	erosion.n	942
29	animal.n	1193	copper.n	1854	interest.n	1956	south.n	937
30	tube.n	1183	volume.n	1818	increase.n	1929	crop.n	920
31	species.n	1146	concentration.n	1795	production.n	1876	live.v	908
32	potential.n	1143	dioxide.n	1793	capital.n	1838	farm.n	887
33	bacterium.n	1131	chemical.adj	1698	balance.n	1794	plate.n	885

34	genetic.adj	1105	iron.n	1698	sale.n	1734	weather.n	880
35	respiration.n	1046	oxidation.n	1668	revenue.n	1731	beach.n	849
36	heart.n	1043	salt.n	1571	wage.n	1697	rainfall.n	846
37	light.adj	1037	particle.n	1565	resource.n	1609	study.n	842
38	nucleus.n	1030	hydroxide.n	1529	buy.v	1597	international.adj	817
39	human.adj	995	aqueous.adj	1499	inflation.n	1585	location.n	803
40	neurone.n	975	mixture.n	1440	sell.v	1519	tropical.adj	794
41	root.n	960	molecular.adj	1407	policy.n	1515	over.adv	794
42	disease.n	914	nitrogen.n	1381	company.n	1504	health.n	787
43	chain.n	912	property.n	1380	gdp.n	1369	coastal.adj	786
44	hormone.n	908	ionic.adj	1368	marginal.adj	1343	north.n	784
45	active.adj	905	magnesium.n	1267	asset.n	1326	building.n	779
46	transport.n	902	ammonia.n	1243	unemployment.n	1300	grow.v	757
47	sugar.n	880	chlorine.n	1240	real.adj	1289	disease.n	755
48	vessel.n	878	proton.n	1229	industry.n	1257	flow.n	755
49	release.v	857	strong.adj	1200	growth.n	1256	desert.n	743
50	gamete.n	838	calcium.n	1174	investment.n	1225	develop.v	741

English			History		Mathematics		Physics	
Rank	Lemma	Freq. p/m	Lemma	Freq. p/m	Lemma	Freq. p/m	Lemma	Freq. P/m
1	word.n	3968	war.n	5894	find.v	13421	energy.n	7273
2	language.n	2408	soviet.adj	4037	value.n	7406	force.n	6680
3	say.v	2117	source.n	3790	equation.n	7327	wave.n	4168
4	text.n	2010	state.n	3430	point.n	6818	object.n	3753
5	write.v	1977	world.n	2629	line.n	5323	field.n	3686
6	writer.n	1915	government.n	2556	graph.n	4672	current.adj	3646
7	think.v	1867	power.n	2360	give.v	4307	speed.n	3516
8	read.adv	1665	union.n	2123	solution.n	3956	direction.n	2964
9	go.v	1643	party.n	1659	example.n	3808	mass.n	2939
10	life.n	1436	military.adj	1624	area.n	3795	temperature.n	2866
11	child.n	1397	policy.n	1578	curve.n	3587	magnetic.adj	2558
12	man.n	1354	communist.adj	1539	term.n	3497	resistance.n	2557
13	paragraph.n	1299	american.adj	1493	function.n	2999	move.v	2543
14	idea.n	1264	political.adj	1479	diagram.n	2830	charge.n	2540
15	look.v	1245	nation.n	1446	angle.n	2827	distance.n	2468
16	student.n	1244	leader.n	1251	axis.n	2553	circuit.n	2367
17	come.v	1234	treaty.n	1158	let.v	2448	wire.n	2365
18	sentence.n	1210	support.v	1053	hence.adv	2267	diagram.n	2314
19	essay.n	1074	army.n	1018	circle.n	2121	particle.n	2305
20	just.adv	1017	president.n	1015	solve.v	2093	air.n	2091
21	feel.v	977	west.n	971	gradient.n	2089	electric.adj	2028
22	thing.n	959	control.n	937	length.n	2051	calculate.v	2026
23	school.n	953	east.n	930	form.n	1963	unit.n	1944
24	passage.n	929	cold.adj	927	coordinate.n	1899	ray.n	1939
25	story.n	870	nuclear.adj	923	side.n	1895	velocity.n	1923

26	phrase.n	834	foreign.adj	923	probability.n	1784	pressure.n	1920
27	writing.n	829	missile.n	875	vector.n	1777	heat.n	1865
28	reader.n	817	crisis.n	874	draw.v	1679	light.n	1792
29	here.adv	793	agree.v	823	triangle.n	1658	constant.adj	1778
30	learn.v	765	peace.n	801	root.n	1654	potential.adj	1721
31	english.n	757	begin.v	792	sum.n	1557	frequency.n	1670
32	tell.v	754	support.n	781	step.n	1456	acceleration.n	1664
33	society.n	745	member.n	777	distance.n	1401	motion.n	1648
34	ask.v	745	western.adj	686	method.n	1319	surface.n	1647
35	view.n	737	plan.n	682	positive.adj	1290	coil.n	1616
36	understand.v	712	ally.n	666	range.n	1278	travel.v	1508
37	old.adj	703	eastern.adj	659	calculator.n	1259	length.n	1455
38	character.n	665	german.n	635	set.n	1251	magnet.n	1435
39	woman.n	664	communism.n	632	represent.v	1223	difference.n	1422
40	speaker.n	654	election.n	624	tangent.n	1215	earth.n	1418
41	sense.n	642	independence.n	617	constant.adj	1178	angle.n	1409
42	meaning.n	635	agreement.n	616	height.n	1176	resistor.n	1396
43	speak.v	634	revolution.n	574	centre.n	1166	sound.n	1396
44	today.adv	591	security.n	567	obtain.v	1136	act.v	1376
45	purpose.n	582	weapon.n	567	normal.adj	1130	weight.n	1302
46	seem.v	575	nazi.n	563	radius.n	1100	wavelength.n	1253
47	argument.n	565	africa.n	547	real.adj	1080	kinetic.adj	1244
48	topic.n	549	establish.v	543	expression.n	1078	car.n	1235
49	young.adj	548	conflict.n	538	equal.adj	1030	electrical.adj	1190
50	speech.n	512	invasion.n	531	substitute.v	1019	work.n	1177

Table 2 provides an illustration of the information in the discipline-specific word association lists (with the 5 strongest collocates) that accompany the target vocabulary on the SVL. A word is often defined by the ‘company it keeps’, and a target word’s associations can therefore help in the understanding of its meaning and use. Thus, once the lemma list for each subject had been finalized, we took an extra step not present in many previously developed academic wordlists and created lists of each word’s discipline-specific collocates. Such collocations can be used in pedagogical materials development, such as mind-maps, word-searches, crossword activities based on weekly spelling lists, etc., in order to promote disciplinary literacy.

**Table 2**

Extracts from discipline-specific word association lists ordered by MI score

<b>Biology</b>	<b>Chemistry</b>	<b>Physics</b>	<b>Geography</b>
<u>ADENINE.n:</u>	<u>POLYMER.n:</u>	<u>URANIUM.n:</u>	<u>PRECIPITATION.n:</u>

nicotinamide.n (12.43)	polyvinyl.n (9.45)	plutonium.n (13.75)	hail.n (10.39)
thymine.n (11.99)	kevlar.n (9.37)	fission.n (10.84)	evapotranspiration.n (9.98)
cytosine.n (11.86)	repeat.v (9.3)	split.v (10.81)	input.n (8.84)
guanine.n (11.78)	perspex.n (9.28)	fuel.n (10.01)	exceeds.v (8.46)
uracil.n (11.56)	monomer.n (9.19)	neutrons.n (9.17)	falls.v (8.36)
<b>English</b>	<b>Mathematics</b>	<b>Economics</b>	<b>History</b>
<u>DESCRIPTION.n:</u>	<u>ESTIMATED.n:</u>	<u>OLIGOPOLY.n:</u>	<u>DISARMAMENT.n:</u>
physical.adj (7.25)	unbiased.adj (11.88)	interdependence.n (10.94)	conference.n (8.91)
scene.n (6.89)	mean.v (9.86)	monopolistic.adj (9.85)	commission.n (8.81)
character.n (6.44)	variance.n (9.8)	oligopolistic.adj (9.66)	failure.n (8.6)
writer.n (5.54)	population.n (9.63)	collusion.n (9.17)	collective.adj (7.89)
story.n (5.39)	testing.v (9.58)	game.n (8.45)	nuclear.adj (7.54)

## CONTRIBUTIONS OF STUDY

The value of the SVL can be demonstrated by considering the words at work in secondary texts themselves. Consider the extract in (2), analysed by *Lextutor* (Cobb, 2012):

(2) **Codominance** occurs when both alleles *controlling a trait* are fully expressed in the heterozygous condition. *Incomplete dominance occurs when neither of the two alleles* controlling a trait is completely dominant over the other and both express themselves to result in an intermediate phenotype in the hybrids. (Perfect Guide: O Level Biology, Lee & Sim, 2014)

In (2), underlined words are in the most frequent 1000 GSL words, those in italics in the AWL, and those in bold in the SVL biology. Therefore, if a student has mastered the most frequent 1000 words of English, they would know 71.74% of the words in the paragraph, a further 8.7% of words are covered by the AWL, while 19.57% of the text is entirely covered by the SVL, i.e. *alleles, dominant, codominance, heterozygous, hybrids, phenotype, trait*. A long term curriculum to enhance disciplinary literacy that followed Nation's (2016) progression from general high frequency words, to general academic

words, to discipline-specific words could effectively employ the GSL, AWL and SVL at progressive year levels from primary through middle and secondary school to achieve remarkable vocabulary coverage for students (in the case (2), 100% coverage). Of course, 100% coverage will not always be the case, though in such cases, the words not covered will be those of very low frequency in the discipline, e.g. *pinnae*.

Offering extensive classroom activities is beyond the scope of this paper, but recent book length treatments on the use of wordlists for pedagogy by Nation (2016) and Greene and Coxhead (2015) are recommended. Nevertheless, let us consider some examples to further demonstrate the value of the SVL. One traditional teaching practice for vocabulary is to select study words for a period of time, followed by a test on spelling, definitions, and productive use in sentences (Nagy et al. 2012). This activity can be enhanced for disciplinary literacy easily with the SVL, e.g., a teacher can target all words from Biology, or all words from Physics, etc., knowing that these words will be relevant to their students content courses (Römer, 2011). Vocabulary items can be drawn from the lemma lists, word association lists, or word families. Other activities can be developed using the supplementary resources; for example, though we noted *set* is in fact in the GSL, though only in mathematics does it relate to words such as *empty*, *universal*, and *ordered*, i.e. the word association lists capture the subtypes of the mathematical sets. Concept maps are popular pedagogically (Lewis & Wray, 2012, p. 32), and a productive activity would be to provide some word associates of *set* to students as a starting point, asking students to map these to the target word and further elaborate on other subtypes that occur to them. Alternatively, guessing from context tasks (Webb & Nation, 2017, p. 637) would help bind these word associations together, reinforcing meaning and lexical priming (Hoey, 2005), e.g. an *ordered* \_\_\_\_ is a [definition]; A genealogical tree is a type of *empty/ordered/universal* (circle correct) set.

## **CONCLUSION**

The technical advances in recent EAP corpus-informed wordlists had not yet been fully leveraged in educational contexts such as secondary school, where there is a need for resources to facilitate disciplinary literacy. Further, in the contemporary landscape of secondary education Singapore, as well as with growing numbers of second language speakers and international schools, resources informed by EAP research that has developed in the context of ESL education are increasingly relevant. The current project has responded by producing the Secondary School Vocabulary Lists (SVL), a set of academic vocabulary resources freely available in the supplementary materials of this journal, including discipline-specific lemma lists, word association lists, and word families for eight subjects: Biology, Chemistry, Economics, English, Geology, History, Mathematics, and Physics.

## **ACKNOWLEDGEMENTS**

This study was funded by the Education Research Funding Programme, National Institute of Education (NIE), Nanyang Technological University, Singapore, project no. OER DEV 01/16

## NOTES

1. A more extensive version of this report is published as Green, C. & Lambert, J. (2018). Advancing disciplinary literacy through English for academic purposes: Discipline-specific wordlists, collocations and word families for eight secondary subjects. *Journal of English for Academic Purposes*, 35, 105-115

## REFERENCES

- Brezina, V., & Gablasova, D. (2013). Is there a core general vocabulary? Introducing the New General Service List. *Applied Linguistics*, 36(1): 1–22.
- Brezina, V., & Gablasova, D. (2017). How to produce vocabulary lists? Issues of definition, selection and pedagogical aims. A response to Gabriele Stein. *Applied Linguistics*, 38(5), 764-767.
- Cobb, T. (2012). *The compleat lexical tutor for data driven learning on the web*. Montreal University of Quebec, available at <http://lexutor.ca/>
- Coxhead, A. (2000). A new academic word list. *TESOL Quarterly*, 34: 213–238.
- Coxhead, A. (2018). *Vocabulary and English for Specific Purposes Research: Quantitative and Qualitative Perspectives*. New York: Routledge..
- Gardner, D., & Davies, M. (2014). A new academic vocabulary list. *Applied Linguistics*, 35(3): 305–327.
- Greene, J., & Coxhead, A. (2015). *Academic vocabulary for middle school students*. Baltimore: Brookes.
- Hoey, M. (2005). *Lexical priming: A new theory of words and language*. London: Routledge.
- Lei, L., & Liu, D. (2016). A new medical academic word list: A corpus-based study with enhanced methodology. *Journal of English for Academic Purposes*, 22: 42–53.
- Nation, P. (2006). How large a vocabulary is needed for reading and listening?. *Canadian Modern Language Review*, 63(1): 59-82.

Nation, P. (2016). *Making and using word lists for language learning and testing*.

Amsterdam: John Benjamins.

Römer, U. (2011). Corpus research applications in second language teaching. *Annual Review of Applied Linguistics*, 31: 205–225.

Scott, M. (2016). *WordSmith Tools* (Version 6 and 7) [Computer software]. Liverpool.

Shanahan, T., & Shanahan, C. (2017). Disciplinary Literacy: Just the FAQs. *Educational Leadership*, 74(5): 18-22.

Simpson-Vlach, R., & Ellis, N. (2010). An academic formulas list: New methods in phraseology research. *Applied Linguistics*, 31(4): 487–512.

Webb, S., & Nation, P. (2017). *How vocabulary is learned*. Oxford: OUP.

Wingate, U. (2011). A comparison of 'additional' and 'embedded' approaches to teaching writing in the disciplines. In M. Deane & P. O'Neil (Eds.), *Writing in the disciplines* (pp. 65–87). Basingstoke: Palgrave Macmillan.

## **APPENDIX**

Please see the two main articles from this research, with links to the resources:

Green, C. & Lambert, J. (2018). [Advancing disciplinary literacy through English for academic purposes: Discipline-specific wordlists, collocations and word families for eight secondary subjects](#). *Journal of English for Academic Purposes*, 35, 105-115

Green, C., & Lambert, J. (2019). [Position vectors, homologous chromosomes and gamma rays: Promoting disciplinary literacy through secondary phrase lists](#). *English for Specific Purposes*, 53, 1-12.