UDC 81`1/`4 DOI https://doi.org/10.24919/2308-4863/42-3-4

Mirjana BORUCINSKY,

orcid.org/0000-0002-1132-9720 Assistant Professor at the English Language Department Faculty of Maritime Studies, University of Rijeka (Rijeka, Croatia) mborucin@pfri.hr **Sandra TOMINAC COSLOVICH,** orcid.org/0000-0002-4320-2863 Associate Professor at the English Language Department Faculty of Maritime Studies, University of Rijeka

(Rijeka, Croatia) stominac@pfri.hr

INTRODUCING DATA-DRIVEN LEARNING INTO THE MARINE ENGINEERING ENGLISH CLASSROOM

Corpora have been extensively used in materials design and development in ELT and to a lesser degree in ESP. However, their application in the ESP classroom is still far from Sinclair's goal envisaged in 1991 of "pushing learning forward by interrogating corpora in an aim of self-directed learning". Hence, this paper will show how data-driven learning (DDL, Johns, 1991; Boulton, 2009), one of the novel approaches to language learning that fosters learner autonomy, inductive and self-directed language learning can be introduced into to the ESP, in particular the marine engineering English classroom. The paper will show: 1. how the maritime English (ME) instructor can compile corpora tailored to their students' needs, 2. how corpus data, in particular frequencies and lexical patterns obtained from word lists, concordances, thesauri and word sketches can help the ESP instructor identify key vocabulary and target structures relevant for their students, and 3. how both students and instructors can identify, classify and generalize linguistic data at various linguistic levels using corpora.

Although the web as a resource for language learning and teaching is frequently used by ESP teachers and ME instructors, web corpora which present linguistic data in a more structured way have been neglected so far. The authors argue that corpus-informed classroom activities will help motivate students, allowing them to become researchers of the language. Such activities can also help teachers bridge the gap from traditional roles to the modern role of coordinator of student work. This is in line with the communicative approach to language learning of the IMO Model course 3.17, in which active learning is encouraged via student involvement. In other words, students are motivated to think about and experiment with language, while the teacher's role is to guide, supervise and encourage them.

Key words: data-driven learning, marine engineering English, IMO Model Course 3.17.

Міряна БОРУЦИНСЬКИ,

огсід.org/0000-0002-1132-9720, асистент кафедри англійської мови морського факультету Рієцького університету (Рієка, Хорватія) mborucin@pfri.hr **Сандра ТОМІНАК КОСЛОВИЧ,** orcid.org/0000-0002-4320-2863

доцент кафедри англійської мови морського факультету Рієцького університету (Рієка, Хорватія) stominac@pfri.hr

ВПРОВАДЖЕННЯ НАВЧАННЯ НА ОСНОВІ ДАНИХ У КЛАС З АНГЛІЙСЬКОЇ МОВИ ДЛЯ СУДНОВИХ МЕХАНІКІВ

Дослідження у сфері корпусної лінгвістики широко використовувалися в розробці матеріалів для викладання англійської мови та меншою мірою англійської мови за професійним спрямуванням. Однак їх застосування під час викладання англійської мови за професійним спрямуванням все ще далеко від мети, поставленої Сінклером у 1991 р,. «просувати навчання шляхом вивчення корпусів із метою самостійного навчання». У цій статті описано, як навчання на основі даних один із нових підходів до вивчення мови, який сприяє автономії студента, індуктивне і самостійне вивчення мови може бути застосоване під час викладання англійської мови за професійним спрямуванням, зокрема англійської мови для суднових механіків. У статті розглянуто: 1) як викладач з морської англійської мови може створити корпус, адаптований під потреби своїх студентів, 2) як дані корпусу,

.....

зокрема частоти і лексичні шаблони, отримані зі списків слів, словників синонімів і статей, можуть допомогти викладачу з англійської мови визначити ключовий словниковий запас і лінгвістичні структури для студентів, 3) як студенти і викладачі можуть ідентифікувати, класифікувати і узагальнювати лінгвістичні дані на різних лінгвістичних рівнях за допомогою корпусів. Хоча Інтернет як ресурс для вивчення і викладання мови часто використовується викладачами з англійської мови за професійним спрямуванням та морської англійської мови, проте вебкорпусу, який представляє лінгвістичні дані у більш структурованому вигляді, досі не приділяється належна увага. Автори стверджують, що заняття, проведені із застосуванням корпусу, допоможуть мотивувати студентів, дозволяючи їм стати дослідниками мови. Такі вправи також можуть допомогти викладачам подолати розрив від традиційних ролей до сучасної ролі координатора роботи студентів. Це відповідає комунікативному підходу до вивчення мови в рамках курсу ІМО Модельного курсу 3.17, в якому активне навчання заохочується через участь студентів. Іншими словами, студентів мотивують думати про мову і експерименти з нею, а роль викладача полягає в тому, щоб направляти, контролювати і заохочувати їх.

Ключові слова: навчання на основі даних, англійська мова для суднових механіків, ІМО Модельний курс 3.17.

Problem statement. An important aspect of datadriven learning (DDL, Johns, 1991; Boulton, 2009) is the use of corpora either in materials design and development or in the classroom. Corpora are, however, not sufficiently implemented in the ESP classroom. Whereas Maritime English (ME) instructors use the web as a (re)source for language learning and teaching (cf. Ungureanu and Varsami, 2019), web corpora, which present linguistic data in a more structured way, have been neglected so far.

Recent research and publications. Major ELT publishers have their state-of-the-art corpora, and corpus-informed research is the basis of modern dictionary and grammar reference books. According to Jones and Waller (2015: 121) corpora have influenced ELT for quite some time in areas such as syllabus design, reference grammars, self-study grammar practice books and textbooks. However, "when it comes to the classroom, Sinclair's 1991 vision of most students and teachers happily interrogating the corpus to push learning forward has not yet been realized" (Tribble, 2012).

Information obtained from corpora, in particular frequencies and lexical patterns, can help the ESP instructor identify target structures relevant for their students. Corpus-informed research is an instance of DDL, one of the novel approaches to language learning that fosters learner autonomy, inductive and self-directed language learning. Above all, it equips the learner with language learning skills that can be used independently. Furthermore, DDL is in line with the holistic view of learners as individuals (*IMO Model Course 3.17*). Within this approach, teachers can direct and facilitate learning, while students themselves have the ultimate responsibility for their own progress.

Through DDL students are exposed to large quantities of authentic data, i.e., the corpus, which allows them to actively explore language and detect linguistic patterns (Boulton, 2009). Boulton (2009) also suggests that the following skills can be refined through corpus use: "predicting, observing, noticing, thinking, reasoning, analyzing, interpreting, reflecting, exploring, making inferences (inductively or deductively), focusing, guessing, comparing, differentiating, theorizing, hypothesizing, and verifying".

Flowerdew (2001: 376) found that science and engineering students accepted DDL quite easily, while business students were far more reluctant in the use of DDL. From the teacher's point of view, DDL has yet to make advances as teachers are either unfamiliar with the possibilities offered by DDL and corpora, unwilling or unable to apply DDL in their (ESP) classroom (cf. Flowerdew, 2001), or as Flowerdew (2012) puts it "DDL has simply not yet penetrated the consciousness of the teaching profession worldwide". For example, in survey carried out in 2004, Mukherjee found that among nearly 250 high school teachers in Germany approximately 80 per cent were unaware of corpus applications in language learning.

The purpose of the article. This paper tries to overcome the issue of the lack of application of DDL in the ESP classroom. We will exemplify how both teachers/instructors and students can make use the vast amount of linguistic data for "hands-on activities in which learners themselves engage in corpus analysis" (Szudarski, 2018: 96). In this paper we will show 1) how the maritime English (ME) instructor can compile corpora tailored to their students' needs, 2) how corpus data, in particular frequencies and lexical patterns obtained from word lists, concordances, thesauri and word sketches can help the ESP instructor identify key vocabulary and target structures relevant for their students, 3) how both students and instructors can identify, classify and generalize linguistic data at various linguistic levels using corpora.

Presentation of the main material. The material will be presented in three steps.

Step 1. Compiling corpora. In this section we will show how the ME instructor can compile a corpus tailored to their students' needs, or even encourage students to compile their own corpus. Finding the right content in ESP, in this case marine engineering is rather difficult. Let us imagine that the

ME instructor wants to introduce new and current content into the classroom, i.e. content that is difficult to find as it presents cutting-edge technology or recent technological advances. For illustration purposes, we shall explore the topic of "dual fuel engines". Instead of spending hours browsing the web, the ME instructor can access the Sketch Engine (SkE) (Kilgarriff et al., 2004) platform (sketchengine.eu1). Once the login has been completed, the ME instructor can start creating a new corpus either by loading his/her own texts or textbooks or using the web. The former method is useful if one already has a larger number of textbooks, articles, instruction manuals and other materials on his/her PC, etc. We shall focus on the latter method, i.e. how a corpus can be compiled using the web, based on BootCaT technology (Baroni and Bernardini, 2004). This is a special feature of *SkE* which allows users to build a specialized corpus from nothing in a matter of minutes. There are two methods to be used: 1. Specific URLs or webpages can be typed from which the corpus will be compiled, or 2. The web can be searched for seed words, i.e. terms that are expected to be typical of the domain of interest. We shall explore the latter option here. For the purposes of our marine engineering course and the topic of "dual fuel engines", we entered the following seed words and phrases: "dual fuel", "engine", "gaseous fuel", "liquid fuel", "Wärtsilä", "FPSO", "tanker", as shown in Figure 1.

Once the seed words have been entered, SkE combines them into random groups of three and submits them to Bing. Bing, in turn, searches the internet and sends pages back, removing ads, navigation menus, etc., i.e. linguistically irrelevant content. Bing normally searches the whole internet (however there is a limit to the number of pages to be downloaded), but the search can also be limited to desired websites only.

The question raised here is how many seed words should be entered. As a rule of thumb, fewer seed words produce a smaller but more focused corpus. Depending on the purpose and design of the corpus, 20-60 seeds from a domain are recommended, but the number can be smaller as will be illustrated in our example. It is, however, very important to avoid seed words that might appear in different registers, e.g. bearing 1. (n) (navig) the direction of an object (with reference to you, your ship, another object); the horizontal direction of a line of sight between two objects on the surface of the earth, expressed as the horizontal angle, in degrees, between that direction and north², and 2. (n) (marine engineering) a machine element that constrains relative motion to only the desired motion, and reduces friction between moving parts³. To avoid such ambiguities, the 'allow list' and 'deny list' settings should be explored. The 'allow list' settings allow us to avoid ambiguity of seed words, by making some of the unambiguous seed words compulsory and thus making sure that the document matches the topic. On the other hand, the 'deny list' can be used to reduce ambiguity by restricting the context (e.g. in the case of the noun facility as in offshore facilities, other registers can be limited by introducing words such as medical, telecommunications, etc. as keywords for the 'deny list'). It is only necessary to use the 'deny list' and 'allow list' if irrelevant documents are found. Once the settings had been adjusted (bearing in mind that the default settings will probably produce useful results as well), in a matter of ten minutes, SkE compiled a 1,779 954 token corpus (henceforth referred to as Dual fuel engines corpus) based on the seed words we provided. Once the corpus is compiled, it is advisable to check that the content of the corpus is related to topic or specialized field, and this can be done by using the extract terminology option. Figure 2 shows multi-word terms extracted from the Dual fuel engines corpus.

A brief glance at the terminology helps us decide whether the produced results are satisfactory. Corpus compilation in SkE has many advanced settings, and if the corpus compilation does not produce desired results, one can look into the advanced settings or repeat the corpus compilation procedure several times to enlarge the corpus. In our case, we decided that the results were satisfactory based on keyword extraction and proceeded showing our students techniques of exploring the language on their own by sharing the corpus with each student. However, work on corpora can be done outside the classroom as well, or in distance learning.

Finally, even students can be encouraged to build their own corpus on any (specialized) topic. This will allow them to explore language in their own unique corpus, thus making them more likely to retain new vocabulary.

Step 2. Wordlist, concordances, thesauri and word sketches

Once the corpus has been compiled and the ME deems it satisfactory (based on extracted terminology as described in step 1), corpus tools such as wordlists, concordances, thesauri and word sketches can be further explored.

.....

Source: Wikipedia.

² Source: An English – Croatian Maritime dictionary (Pritchard). URL: https://www.pfri.uniri.hr/bopri/IMEC_Proceedings/Rjecnik_Eng_Hrv.pdf.

ELEXIS funded license available till April 2022.

ø	CREATE CORPUS Dual fuel engines		
	CORPUS: Dual fuel engines (English)		
	1. CREATE CORPUS > 2. ADD TEXTS > 3. COMPILE		
\odot	÷	TEXTS FROM	WEB
•		Input type	Web search URLs Website dual fuel X gasoues X fuel, X liguid X Wärtsilä, X FPSO, X tanker. X engine X
↓≡ N≣		Folder name ⊘	At least 3 words or phrases. Hit ENTER after each one. Multiword phrases in quotes: "wild cat" kitten feed web1
≣5 זייק			Web search settings ← Denylist settings ← Allowiist settings ← Size restrictions ←
			Compile when finished ⑦

Figure 1. Compilation of a specialized corpus from the web

Ø	KEYWORDS	Dual fuel engines Q								
==	SINGLE-WORDS 🗸	MULTI-WORD TERMS 🗸								
	(items: 106,183)	nglish Web 2020 (enTenTen20)								
0			Word	Word		Word		Word	Word	
00			1 dual fuel	 11 fuel oil		21 gas engine		31 dual fuel engine	 41 combustion process	
•≡			2 gaseous fuel	 12 diesel engine		22 fuel system		32 heavy fuel	 42 gas injection	
			3 liquid fuel	 13 engine speed		23 fuel tank		33 methane slip	 43 steam turbine	
:::			4 inert gas	 14 diesel fuel		24 fuel injector		34 flue gas	 44 liquid nitrogen	
Į≡			5 pilot fuel	 15 exhaust gas		25 lubricating oil		35 fuel flexibility	 45 gaseous oxygen	
NE			6 combustion chamber	 16 rocket engine		28 man energy		36 gas system	 46 fuel operation	
			7 gas turbine	 17 gas fuel		27 fuel engine		37 gaseous nitrogen	 47 burner tube	
δΞ			8 non-gaseous fuel	 18 fuel injection		28 floating production		38 natural gas	 48 gas supply	
2			cargo tank	 19 dual-fuel engine		20 cargo vapour		39 diesel oil	 40 cylinder pressure	
			10 fuel gas	 20 gas mode		30 fuel supply		40 heavy fuel oil	 50 internal combustion	
				Rows per page:	50	✓ 1–50 of 1,000	< <	<u>1</u> /20 → >I		

Figure 2. Terminology extraction

The wordlist option allows us to identify most frequent words or word classes such as nouns, verbs, adjectives, etc. in the corpus and decide what vocabulary one should focus on or to identify new vocabulary. From the list of the most frequent nouns we chose the noun *facility* as we believed our students were not entirely familiar with the meaning of this noun in this specific context.

As can be seen in Figure 3, it is possible to navigate directly from word lists to concordances, word sketches as well as the thesaurus. Furthermore, the data can be downloaded in various formats and used in the classroom. The next tool on the corpus dashboard that can be explored are concordances. In Figure 4 concordance lines of the noun *facility* are shown as keywords in context (KWIC). The option *good dictionary examples* (GDEX) will automatically select good examples from which the meaning of a word can be deduced.

Based on the *GDEX* option students can deduce the meaning of the word *facility* in the specific context, i.e. they will observe phrases such as *processing facility, power facility, production facility,* etc. Literature review shows that concordances can operate as 'condensed reading' (Gabrielatos, 2005) and it has been established that vocabulary learning comes from

Актуальні питання гуманітарних наук. Вип 42, том 3, 2021

	fuel engines Q (i)					
NOUN (29,382 items 609,916 total	frequency)					
	Lemma	Frequency ?	Lemma Fr	requency ?	Lemma	Frequency ?
	201 general	495	211 mean	480	221 b	467
	202 bar	495	212 limit	479 ***	222 state	465 ***
	203 motor	493	213 aol.com	478 ***	223 location	464
	204 outlet	490	214 connection	475	224 increase	462
	205 rpm	490	915 facility	474 =•=	Concordance	462
	206 e	489	216 compression	473 NE	N-grams	457
	207 stage	485	217 i	473	Word Sketch	453
	208 standard	484	218 model	472	Thesaurus 🔽	452
	209 tube	484	219 quality	471		450
	210 MW	482	220 discharge	470	230 module	449
First 20,000 rows are displayed. Yo	u can download whole list 👤					

Figure 3. Most frequent nouns in the corpus on Dual fuel engines

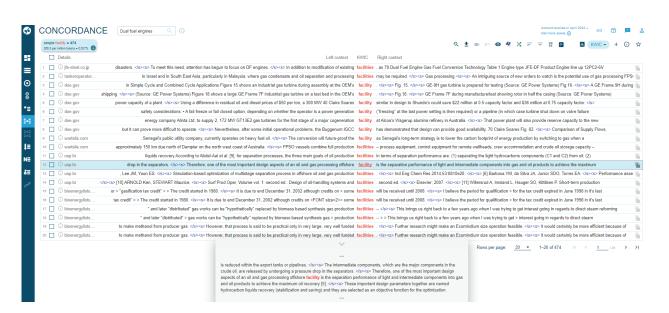


Figure 4. Concordance lines of the noun facility

extensive reading. Hence, by looking at concordances one can focus on target vocabulary. By focusing the students' attention to one item used in various contexts, they can observe the meaning and notice common patterns. In this way students can become lexicographers (cf. Cobb, 1999).

Whereas it is true that most teachers at university level feel reluctant to use corpora and that concordances might be difficult to read and require some skill (cf. Kilgarriff, 2009), we believe that by making and an effort, provided one can find the time in the tightly-packed syllabus, both students and teachers can profit from this activity. Furthermore, not to scare students off from concordances, one can introduce other tools such as the thesaurus or word sketches, and "disguise" them as dictionaries (Kilgarriff, 2009), thus offering students a fuller picture of the meaning and common patterns that words exhibit in language use.

If, at this point the students only have a vague idea of the meaning of the noun *facility*, they can access the thesaurus, which will provide synonyms for the selected query word.

They will notice that nouns such as *plant, infra-structure, equipment,* etc. share some of the core meanings with the noun *facility*. Furthermore, they can analyze which words modify this noun, which verbs collocate with it, etc.

THESAURU	JS	Dual fuel engines
facility as noun 464×		

TH

	Word	Frequency ?	
1	terminal	317	
2	plant	2,098	
3	equipment	1,453	
4	platform	351	
5	unit	1,320	
6	application	1,096	
7	vessel	2,388	
8	infrastructure	147	
9	industry	697	
10	solution	1,411	

Figure 5. Synonyms of the noun facility

facility as noun 464×	•••											
-⇒ ≣≣⊠ ×		Ø×	,	₩Ø×	÷	₩Ø×	÷	₩Ø×	et 10	ØX		•• Ø >
modifiers of "facility"	nouns modifie "facility"	d by		"facility" as ject		h "facility" as ubject	"facility	" and/or	prepositional ph	nrases		al possessor facility"
production	operator		float		mean		reclaim		of "facility"		their	
production facilities	facility operator		recent years , production fac	floating ilities , known as	RECLAIM F facility	ACILITY means a	operate at a former RECL have a	RECLAIM or AIM facility or	"facility" in		their LNG fa	
RECLAIM or former RECLAIM	equipment agricultural machiner	y and	become		become		garage		at "facility"		its its facilities	
facility storage	facilities , constructio equipment , energy s		that the facility former RECLA	becomes a AIM facility . The	time that the a former RE	e facility becomes CLAIM	garages , rep automotive n	pair garages ,	to "facility"			
storage			process		include		dispensing fa parking gara	acilities and	from "facility"			
former			facilities extract and	processing ct contaminants	The facility complete fu	el	capability		for "facility"			
a RECLAIM or former RECLAIM facility			mean		require		capacity acc process facil capability of	ities and a	in "facility"			
offshore			RECLAIM FAC	CILITY means a	facility requi	iles	office		"facility" of			
offshore facility			build		facility oper	ated	offices and re facilities	ecreational	"facility" for			
onshore			facilities were		be		platform		"facility" to			
LNG			supply facilities were	supplied by	facilities are		offshore plat	nore facilities and forms and	"facility" with			
LNG facilities			have		a lavatory fa	acility having a	topsides on		as "facility"			
export			has manufactu	uring facilities	source of		ship availability of and facilities	f bunkering ships	~	≫		
port			If such facilitie they should be	s are provided			equipment					
floating			be				facilities . cor	nachinery and nstruction energy systems				
floating production facility			is a power gen	neration facility			tanker					
test at the engine test facility of ISRO Propulsion			use using the facili	ties provided			facilities and	tankers				
manufacturing												
worldwide and has manufacturing facilities in Germany												
✓												

Figure 6. Word sketch for the noun facility

Based on the word sketch exemplified in Figure 6, the teacher can prepare various activities, such as gap-filling exercises, matching or even grammar consolidation exercises.

To illustrate the nature of words and their behavior in context, one can even compare and contrast two words, such as *facility* and *infrastructure* (see Figure 7).

Words marked green present stronger collocates or modifiers of the noun facility, whereas those marked red are better collocates of the noun infrastructure

and those marked grey present good collocates of both nouns. For instance, verbs such as to provide, to develop and to exist can be combined with both nouns:

(1) On completion of discharge, liquid lines and cargo hoses or loading arms should be drained, purged and depressurized using the facilities provided.

(2) The latest development of the LNG infrastructure provides clear evidence of the increasing importance of LNG fuel to the marine industry.

Step 3. Identifying and generalizing linguistic data at various linguistic levels

How can the data presented above be best exploited be in and outside the classroom? It can be exploited for extending or deepening knowledge of existing language items, for distinguishing synonyms, detecting patterns of usage, collocation, colligation, morphology, and so on (cf. Boulton, 2009)⁴. These can make the students aware of issues of frequency and typicality, register and text type, discourse and style, as well as the fuzzy nature of language itself (Boulton, 2009). If the teacher feels reluctant to perform these activities with their students, corpus data can still be used for materials design, for preparing activities on vocabulary and grammar, at the same time exposing the students to authentic content (cf. Borucinsky and Jelčić Čolakovac, 2020), showing them real language, encouraging them to test hypotheses about how language is used. Language acquired in this way is proven to be likely to be remembered (cf. Johns, 1991).

For students who like to combine computer skills and make observations about language, simple CQL (corpus query language) could be of interest. For instance, the instructor might ask their students to 1. predict which adjectives modify the noun facility and 2. test their hypothesis about language using a simple query language, such as the one illustrated in Figure 8.

The results of the query are shown in Figure 9, a concordance list which students can use to verify whether any of the words they thought of are on the list.

Conclusion. In this paper we have shown how corpus tools and corpora can be used in the ESP classroom to provide students with new and innovative techniques for recording and exploring language. By exposing students to large quantities of authentic data, i.e., the electronic corpus, they can *hone* various skills from guessing to differentiating, hypothesizing and verifying and they are motivated to explore language from a completely different perspective.

Benefits for the ESP instructor are best expressed in bridging the role of a traditional, all-knowing teacher to that of a guide and a mentor. Provided the instructor can find time in the tightly-packed syllabus and possesses basic computer skills (as no knowledge of programming or programming languages is required), he/she can become more confident even with technical materials as the focus remains on language and exploring of linguistic patterns.

VORD SKE	TCH DIF	-ERENC	E Dua	il fuel engi	jines												ecount expire let more space	in April 2022 x	e ()	0
facility 484×		infrastructu	JITO 147×															Q	± 0 (
æ	ėα×	æ		e Ø	×	æ		@ Ø	×	æ	۲	¤ ×	æ	⊛¤×	2		⊕¤>	. e		۲
"facility/infrastruc	ture" and/or		"facility/infra as object	structure	•"	verbs with "facility as sub		tructure	e.	adjective predica "facility/infrastru			"facility/infrastru	ucture" is a	is a "facility	/infrastru	icture"	modifiers of "fa	acility/infra	struc
reclaim	6 0	float	7	0		extract	2	0		unable	1 (performance	1 0	Berhad	1	0	production	45	0
garage	4 0	become	5	0		boast	1	0		ready	2 (plant	1 0	purifier	1	0	reclaim	14	0
capability	4 0	process	3	0		challenge	1	0		vital	1 (challenge	1 0	Europe	1	0	former	12	(
office	3 0	situate	2	0		mean	4	0		subject	2 (Wartsilä	0 1	operator	1	0	offshore	16	(
latform	4 0	fund	2	0		balance	1	0		safe	1 (hull	1	0	storage	23	
iospital	2 0 …	prepare	2	0		put	1	0		similar	1 (installation	1	0	onshore	8	
dministration	0 1	exist	2	3		provide	Concor	rdance	1.0	facility + provide	12				FPSO	1	0	subsea	5	
extent	0 1	bunker	2	12		need	1	-			-				vessel	1	0	local	1	
conomy	0 2	ruin	0	1		develop	1	1	24	infrastructure + provide	2				system	1	0	existing	1	
hrinkage	0 1 …	lack	0	1		exist	1	2	0	facility + provide								pipeline	2	1
ack	0 2	plane	0	1		facilitate	0	1		Advertised as a second de-								community	0	
raud	0 2	render	0	2		grow	0	2	0	infrastructure + provide	2							limited	0	
~			~			~			•=	provide	Ø								~	
2	⊕Ω×	et i		φ¤	×	÷		φ¤	×	æ	۲	¤ ×	÷	⊕¤×	e*		⊕¤>	æ.		۲
nouns mod "facility/infras			ossessors of ty/infrastruct			pronominal po "facility/infra				"facility/infrastructu	ire" of		of "facility/int	frastructure"	"facility/infra	structure	" in	in "facili	ty/infrastru	cture
representative	2 0 …	INEOS	2	0		their	3	0		complex	5 0		shutdown	2 0	Huoma	1	0	shoot	1	
dozen	1 0	OEM	2	0		your	1	0		airplane	1 0		portfolio	2 0	Flow	1	0	uptime	1	
on-shore	1 0	United	1	0		our :	2	1		foot	1 0		insurer	1 0	Trinidad	1	0	accumulation	1	
Frames	1 0	company	1	0		its	3	2		Group	1 0		dozen	1 0 …	River	1	0	release	1	
applicant	1 0	world	0	1						Oil	1 0		sensitivity	1 0	gas-works	1	0	impurity	1	
proponent	1 0									customer	0 1		goal	1 0	Turku	1	0	develop	1	
LP	0 1									production	0 1		lack	2 3 …	Germany	2	0	produce	1	
	0 1									nation	0 1		installation	0 1 …	Finland	1	0	include	1	
Ford	0 1												problem	0 1	place	0	1	gas	1	
Ford zzl																				
	0 1 …												preparation	0 1	port	0	1	use	2	

Figure 7. Word sketch difference for the noun facility and infrastructure

⁴ More on how to use corpora in the classroom can be found in Firginal (2018), Thomas (2016) and Kilgarriff (2009) and O'Keefe, McCarthy and Carter (2007).

Ø	CONCORDANCE	Dual fuel engines 🔍 🕢		
	BASIC ADVANCED	ABOUT		
##	Query type ③ simple	CQL [tag="J.*"] [word="facility"] 	COMMON TAGS	Ð
\odot	lemma phrase	CQL BUILDER []	noun	N.* 🗈
8	word	Default attribute ?	verb	V.* 🗈
•=	character	lemma	adjective	J.* 🗈
≣•≣	CQL		adverb	RB.? 🗈
			pronoun	PP.? 🗈
	Subcorpus ⑦ none (the whole corpus) 👻 🔒	+	conjunction	CC 🖻
t≡			preposition	IN 🖷
NE	Filter context 💿 🗸		determiner	DT 🖷
β≣	Text types ? 🗸	_	numeral	CD 🚡
12		GO	particle	RP 🖺
			view help for ALL TAGS	

Figure 8. Advanced search of language structures using CQL

CONCORDAN	Dual fuel engines Q ()		Account expires in April 2022 * 🕢 😁 🧭
CQL [tag="J."] [word="facility"] = 11.24 per million tokens = 0.0011%	20 0		옥 🛓 📰 🗠 🕢 🤻 X 📰 포 🎇 🖪 🛛 🖪 (KWC 🗸 🕂 (
Details	Left context	KWIC	Right context
1 🔲 🛈 usp.br	a pressure drop in the separators. <s> Therefore, one of the most important design aspects of an oil and gas processing</s>	offshore facility	is the separation performance of light and intermediate components into gas and oil products to achieve the maximum
2 🔲 🛈 bioenergylists	for over 100 years that can meet those specifics with an off the shelf system built with off the shelf parts in an ISO	certified facility	. <s> I have been barraged with wishes, excuses, timelines of nearly a year or more and ZERO that is ZERO examples and</s>
з 🔲 🛈 setrac.in	, based on calculations, shall assist the Master to determine how fast the ship can safely load a particular cargo at a	particular facility	, taking into account the vessel's design parameters and the cargo involved. <s> The Chief Officer should indicate, in the</s>
4 🔲 🛈 intechopen.com	(CaO) is produced by heating calcium carbonate (CaCO3) about at 900°C. Figure 1 shows a schematic representation of an	integrated facility	for the production of calcium carbide [8], <s> Moreover, the processes are seen in Eqs. </s> <s> (1) and (2) [8, 9, 10], </s>
5 🔲 🛈 slideshare.net	stored for shipment to another location for processing. <s> 5. </s> <s> Offloading – transfer of produced hydrocarbons from an</s>	offshore facility	into shuttle tankers or barges for transport to terminals or Deepwater ports. <s> 6. </s> <s> An FPSO relies on subsea tech</s>
в 🔲 🛈 investopedia.co	offshore oil-drilling rig alone is approximately \$650 million. <s> This amount does not include well completion costs,</s>	ongoing facility	production maintenance costs, and platform decommissioning costs (the cost of removing the platform at the end of its
7 🔲 🛈 almsgt.com	the catalysts for an extensive overhaul and upgrade of the gas detection system on the Terra Nova FPSO.	remote facility	with limited egress. <s> Therefore, any hazardous gas release in the facility requires complete production shutdown,</s>
8 🔲 🛈 aimsgt.com	. <s> The team involved personnel from safety, risk analysis, operations, instrumentation and controls engineering. </s> <s> A</s>	detailed facility	analysis, based on computational fluid dynamics modeling, was performed. <s> In aggregate, more than 1,400 gas leak</s>
o 🔲 🛈 icheme.org	the extraction hood. <s> Due to the uncertainty over the flow behaviour and the potential for butane accumulation in the</s>	experimental facility	, this condition was not examined and only tests with the higher flow rate were conducted.
10 🔲 🛈 wikipedia.org	gas. <s> As such, the removal of the sulfuret of hydrogen was given the highest level of priority in the gas-works. </s> <s> A</s>	special facility	existed to extract the sulfuret of hydrogen - known as the purifier. <s> The purifier was arguably the most important</s>
11 📋 🛈 wikipedia.org	facility existed to extract the sulfuret of hydrogen – known as the purifier.	important facility	in the gas-works, if the retort-bench itself is not included. <s> Originally, purifiers were simple tanks of lime-water,</s>
12 🔲 🛈 spe.org	required approximately 25 hours per week of operator attention, or approximately 3 hours per day, on the continuously	manned facility	. <s> The unit consistently exceeds the H2S-removal requirement (typically less than 20 ppm in the treated gas) and has</s>
13 🔲 🛈 wordpress.com	62 LNG Bunkering: Shore, Ships & Barges Infrastructure availability to support LNG as a Marine Fuel From Shore	Dedicated facility	Truck on jetty Containerized fuel From Sea LNG bunker barge or vessel Mooring dolphin Ship-to-ship Courtesy: Jensen
14 🔲 🛈 shipsbusiness.c	, based on calculations, shall assist the Master in determining how fast the ship can safely load a particular cargo at a	specific facility	, taking into account the vessel's design parameters and the shipment involved. <s> The Chief Officer should indicate, in</s>
15 🔲 🛈 wikipedia.org	aircraft carrier[7] It is the world's largest floating liquefied natural gas platform as well as the largest	offshore facility	ever constructed. <s> Contents </s> <s> The main double-hulled structure was built by the Technip Samsung Consortium in</s>
18 🔲 🛈 wikipedia.org	the need for long pipelines to land-based LNG processing plants. <s> However, fitting all the equipment onto a single</s>	floating facility	was a significant challenge. <s> On 25 July 2017, after a journey of 5,800 km (3,600 mi) from its construction site in South</s>
17 🔲 🛈 aqmd.gov	OCS Waters and transport of such crude oil and gas in Southern California Coastal or OCS Waters shall be included in the	same facility	which is under the same ownership or use entitlement as the crude oil and gas production facility on-shore. <s> FORMER</s>
18 🔲 🛈 aqmd.gov	usage by the biogas engine is 90% or more, based on the higher heating value of the fuels used. <s> The calculation of the</s>	monthly facility	biogas use percentage may exclude natural gas fired during: any electrical outage at the facility; a Stage 2 or higher
19 🔲 🛈 eagle.org	the Complex World of Shuttle Tankers <s> Shuttle Tankers are a specialized type of tanker, designed to load cargo from an</s>	offshore facility	or a larger vessel that is not able to transit to the destination due to draft restrictions.
20 🔲 🛈 bluewater.com	ia the offloading system. <s> What is an FPSO? </s> <s> A Floating Production Storage and Offloading (FPSO) installation is a</s>	floating facility	, usually based on a (converted) oil tanker hull. <s> It is equipped with hydrocarbon processing equipment for separation</s>

Figure 9. Results of the query [tag="J.*"] [word="facility"]

BIBLIOGRAPHY

1. Baroni, M. and S. Bernardini. BootCaT: Bootstrapping corpora and terms from the web. *Proceedings of LREC*. Lisbon : ELDA. 2004. pp. 1313–1316.

2. Borucińsky, M. and J. Jelčić-Čolakovac. Promoting Authenticity in the ESP Classroom: The Impact of ICT and Use of Authentic Materials on Reading Comprehension. In Petkova, T. V. and V. S. Chukov (Eds.). *5th International e-Conference on Studies in Humanities and Social Sciences: Conference Proceedings*. Belgrade: Center for Open Access in Science. 2020. pp. 31–44. DOI: 10.32591/coas.e-conf.05.03031b.

3. Boulton, A. Data-driven learning: reasonable fears and rational reassurance. *Indian Journal of Applied Linguistics*. 2009. 35 (1). pp. 81–106.

4. Cobb, T. Breadth and depth of vocabulary acquisition with hands-on concordancing. *Computer Assisted Language Learning* 12. 1999. pp. 345–360.

5. Flowerdew, L. The exploitation of small learner corpora in EAP materials design. In Ghadessy, M., Henry, A. and R. Roseberry (Eds.). *Small Corpus Studies and ELT: Theory and practice*. Amsterdam: John Benjamins. 2001. pp. 363–379.

6. Flowerdew, L. Corpora and language education. Basingstoke : Palgrave Macmillan, 2012.

.....

.....

7. Friginal, E. Corpus linguistics for English teachers: Tools, online resources, and classroom activities. Abingdon/New York : Routledge, 2018.

8. Gabrielatos, C. Corpora and language teaching: Just a fling, or wedding bells? TESL-EJ, 2005. 8 (4). pp. 1–37.

9. IMO Model Course 3.17. Maritime English. London: International Maritime Organization, 2015.

10. Johns, T. Should you be persuaded: Two examples of data-driven learning. In Johns, T. and P. Kingn (Eds.). Classroom concordancing. English Language Research Journal. 1991, 4. pp. 1–16.

11. Jones, C. and Waller, D. Corpus linguistics for Grammar. A Guide for research. Routledge Corpus Linguistics Guides. London and New York: Routledge, 2015.

12. Kilgarriff, A., Rychlý, P., Smrž, P and D. Tugwell. Itri-04-08 the sketch engine. Information Technology. 2004. pp. 105-116.

13. Kilgarriff, A. Corpora in the classroom without scaring the students. Proc. 18th Int. Symposium on English Teaching. Taipei. 2009.

14. O'Keefe, M. McCarthy, A. and R. Carter. From Corpus to Classroom: Language Use and Language Teaching. Cambridge University Press, 2007.

15. Pritchard, B. An English-Croatian Maritime Dictionary. URL: https://www.pfri.uniri.hr/bopri/documents/Tlex-novi. pdf (Accessed on July 16, 2021).

16. Sketch Engine. Sketchengine.eu (Accessed on July 1, 2021).

17. Szudarski, P. Corpus linguistics for vocabulary. Routledge Corpus Linguistics Guides. London and New York: Routledge, 2018.

 Thomas, J. *Discovering English with Sketch Engine*. Versatile, 2016.
 Tribble, Ch. Corpora in the Language-Teaching Classroom. https://doi.org/10.1002/9781405198431.wbeal0226 2012 (Accessed on July 1, 2021).

20. Ungureanu, C. and C. Varsami. Teaching Maritime English based on web resources. Young Scientist 71.2. 2009. pp. 83-85.

21. Wikipedia. URL: https://en.wikipedia.org/wiki/Bearing (mechanical) (accessed July 16, 2021).

REFERENCES

1. Baroni, M. and S. Bernardini. BootCaT: Bootstrapping corpora and terms from the web. Proceedings of LREC. Lisbon: ELDA. 2004. pp. 1313-1316.

2. Borucinsky, M. and J. Jelčić-Čolakovac. Promoting Authenticity in the ESP Classroom: The Impact of ICT and Use of Authentic Materials on Reading Comprehension. In Petkova, T. V. and V. S. Chukov (Eds.). 5th International e-Conference on Studies in Humanities and Social Sciences: Conference Proceedings. Belgrade: Center for Open Access in Science. 2020. pp. 31-44. DOI: 10.32591/coas.e-conf.05.03031b.

3. Boulton, A. Data-driven learning: reasonable fears and rational reassurance. *Indian Journal of Applied Linguistics*. 2009. 35 (1). pp. 81-106.

4. Cobb, T. Breadth and depth of vocabulary acquisition with hands-on concordancing. Computer Assisted Language Learning 12. 1999. pp. 345-360.

5. Flowerdew, L. The exploitation of small learner corpora in EAP materials design. In Ghadessy, M., Henry, A. and R. Roseberry (Eds.). Small Corpus Studies and ELT: Theory and practice. 2001. pp. 363–379. Amsterdam: John Benjamins.

6. Flowerdew, L. Corpora and language education. Basingstoke: Palgrave Macmillan, 2012. 7. Friginal, E. Corpus linguistics for English teachers: Tools, online resources, and classroom activities. Abingdon/New York: Routledge, 2018.

8. Gabrielatos, C. Corpora and language teaching: Just a fling, or wedding bells? TESL-EJ, 2005. 8 (4), pp. 1–37.

9. IMO Model Course 3.17. Maritime English. London: International Maritime Organization, 2015.

10. Johns, T. Should you be persuaded: Two examples of data-driven learning. In Johns, T. and P. Kingn (Eds.). Classroom concordancing. English Language Research Journal. 1991, 4. pp. 1–16.

11. Jones, C. and Waller, D. Corpus linguistics for Grammar. A Guide for research. Routledge Corpus Linguistics Guides. London and New York: Routledge, 2015.

12. Kilgarriff, A., Rychlý, P., Smrž, P and D. Tugwell. Itri-04-08 the sketch engine. Information Technology. 2004. pp. 105-116.

13. Kilgarriff, A. Corpora in the classroom without scaring the students. Proc. 18th Intrnl Symposium on English Teaching. Taipei. 2009.

14. O'Keefe, M. McCarthy, A. and R. Carter. From Corpus to Classroom: Language Use and Language Teaching. Cambridge University Press, 2007.

15. Pritchard, B. An English-Croatian Maritime Dictionary. https://www.pfri.uniri.hr/bopri/documents/Tlex-novi.pdf (Accessed on July 16, 2021).

16. Sketch Engine. Sketchengine.eu (Accessed on July 1, 2021).

17. Szudarski, P. Corpus linguistics for vocabulary. Routledge Corpus Linguistics Guides. London and New York: Routledge, 2018.

18. Thomas, J. Discovering English with Sketch Engine. Versatile, 2016.

19. Tribble, Ch. Corpora in the Language-Teaching Classroom. https://doi.org/10.1002/9781405198431.wbeal0226 2012 (Accessed on July 1, 2021).

20. Ungureanu, C. and C. Varsami. Teaching Maritime English based on web resources. Young Scientist 71.2. 2009. pp. 83-85.

21. Wikipedia. URL: https://en.wikipedia.org/wiki/Bearing (mechanical) (accessed July 16, 2021).

.....