

TUMSAT-OACIS Repository - Tokyo

University of Marine Science and Technology

(東京海洋大学)

AntConcを用いた海事英語テキストの語彙分析

メタデータ	言語: jpn 出版者: 公開日: 2015-02-27 キーワード (Ja): キーワード (En): 作成者: 内田, 洋子 メールアドレス: 所属:
URL	https://oacis.repo.nii.ac.jp/records/517

AntConc を用いた海事英語テキストの語彙分析

内田 洋子*

(Accepted October 20, 2014)

A Vocabulary Analysis of Maritime English Text Using AntConc

Yoko UCHIDA *

Abstract: The English proficiency required to prospective deck officers should not be limited to the mastery of SMCP, which only guarantees “a minimum level of communication,” but should extend to the attainment of the ability to comprehend information provided in English texts on Maritime issues. In order to explore the kind of vocabulary necessary for reading Maritime English texts, this study attempted a corpus analysis of Maritime English text *Risk Watch* using AntConc. An overview of the Maritime English corpus was given, followed by a comparison with the so-called “general English,” which was represented by the British National Corpus (BNC) in this study. In the analysis focusing on high-frequency words, it was revealed that the type of content words frequently used in Maritime English differs from that in the BNC. The result highlighted the necessity for deck cadets to learn relatively higher-level words and acquire knowledge on usage peculiar to Maritime English. Further linguistic investigation of the 1,000 most frequently used words in the future will provide hints to establish guidelines for practical vocabulary learning of Maritime English.

Key words: Maritime English, Vocabulary, Corpus Analysis

第一章 はじめに

現在の海事英語教育の中心に据えられているのはIMO標準海事通信用語 (Standard Marine Communication Phrases, SMCP) の習得である。しかし、これは海事英語がカバーする全体像から考えると氷山の一角に過ぎない。本章では、より高い海事英語能力をつけるためには、書きことばの読解訓練が欠かせないものであることを指摘し、そのために必要となる海事英語の語彙は何かを調査する本研究の目的について述べる。

1. 「海事英語を習得する」とは

まず、海事英語を学習する者は、通常、中学・高校段階から英語を学習した経験のある大学生であることを前提とする。

海事英語の習得とは、それまでに培って来た英語力を土台にして、海事という独特のコンテキストの中で必要とされる英語にまつわる能力を獲得することである。船舶運航が多国籍の船員が乗り組む船で、国境を越えて行われることを考えると、この「海事という独特のコンテキスト」での英語使用の機会は思いのほか多い。簡潔明瞭な通

信のためにごく一定の語彙の使用に限定された SMCP にとどまらず、船の設備の使用法について書かれたマニュアル、港湾当局に提出する書類や寄港地から提供される情報、ニューズレター等、書きことばによるものも多い。

書きことばの英語を理解することは船舶職員として非常に大切なことであるにもかかわらず、書きことばの理解に苦勞する学生は多い。その要因としては、話しことばよりも一般的に一文が長く複雑な構造を持つ英語の解釈ができないという部分もあるだろうし、使われる特有の語彙の知識に乏しいケースも多い。いずれにせよ、中学・高校で一定レベルの文法を習得して英語の土台を作った状況で、次の段階として積み重ねていくべきことの一つは、専門語彙を着実に身につけることである。

2. コーパス分析を使った語彙研究

それでは、海事英語を読解する語彙はどのようにして身につければいだろうか？一つの手段は、多くの海事英語文献を読みこなし、そこに出てきた単語を片っ端から覚えていくことであろう。しかし、そのような方法では、たまたま読んだ文献にたまたま一度だけ出現し、他の文献でも滅多に出ないような語をやみくもに覚えることになるやも

* Department of Maritime Systems Engineering, Division of Marine Technology, Graduate School, Tokyo University of Marine Science and Technology, 2-1-6 Etchujima, Koto-ku, Tokyo 135-8533, Japan (東京海洋大学大学院海洋工学系海事システム工学部門)

しれず、非常に非効率的であろう。それよりは、海事英語の理解のために重要と見なされる単語だけをまずは選び、効率的に学びたいと思うのが学習者としての自然な気持ちであろう。コーパス分析はそれを可能にするための一つの方法である。

コーパスとは、なんらかの言語使用場面における言語データを集めて電子化した大量の言語データベースのことである。海事英語のデータベースであれば、海事関係のことが英語で書かれた資料を大量に集めて電子化したものということになる。そして、そこで使われている単語の種類や数、単語同士のまとまり方を調べることをコーパス分析と呼ぶ。コーパス分析は、大量のデータから語彙学習において重要な情報を選び出す手法として、現在、言語教育で広く用いられている。

3. 研究の目的

本研究の目的は、コーパス分析を通して書きことばの海事英語を構成する語彙特徴を明らかにすることにある。その特徴が高校段階までで学習する「一般英語」で学習する内容とどのように類似し、また異なっているかを記述する。本研究では、調査する海事英語資料の概観および高頻度語の中身の一部の検証に限定して行う。

第二章 分析の手順

この章では、本研究で用いた言語資料および分析の手順について説明する。

1. 言語資料

本研究で用いた言語資料は、*Risk Watch* という海事情報誌である¹⁾。*Risk Watch* は船主のための相互保険組合である *Britannia* によって毎年ほぼ 3 ~ 4 回のペースで発行されている。*Britannia* のメンバーである船主やその船の航海士・乗組員にリスク管理情報を伝達する手段として、毎号、8 ページ程度の各種情報（例：衝突事故の原因分析、特定の国や港における貨物の検査体制の変化、コンテナの積み付けの注意点、密航者への対応策）を記事にして提供している。

本研究では、*Britannia* の HP 上にアーカイブとして掲載されている 2000 年 7 巻 1 号から 2013 年 20 巻 3 号までの 46 号分を言語資料として用いた。掲載されている記事数は 315 を数え、その内容の下位分類と記事の数は次の通りである：

- (1) 貨物の扱い
 - a. コンテナ (19)
 - b. その他の貨物 (53)
- (2) 人命救助 (18)
- (3) 航海

- a. 事故 (24)
 - b. 操船 (62)
 - (4) 規則 (19)
 - (5) 各種作業
 - a. 危険地帯 (21)
 - b. 燻蒸 (3)
 - c. メインテナンス (12)
 - d. 係留 (9)
 - e. その他 (12)
 - (6) 密輸及び密航者の問題 (12)
 - (7) その他（船員の問題、海賊、条約改正など）(58)
- 一つの記事を一つのカテゴリーに分類できない場合は二つにまたがる形で分類したため、括弧内の合計は 315 にはならない。

発行元の業種柄、保険が関わってくる貨物や事故における過失の判断に関する記事が多く見受けられるが、いずれの記事も航海士や乗組員の円滑な運航のために有益と思われ、これを難なく読みこなす程度の語彙力を有することは職業上、きわめて重要であると考えられる。このことから、本言語資料を将来海技士になることを目指す学生が必要とする読解力及び語彙力を培う英語資料としてふさわしい海事英語コーパスである、と判断した。

このコーパスを分析するために、HP 上で .pdf 形式であった本資料を .txt ファイルに変換し、記事一つにつきファイル一つという形で保管した。

2. 分析方法

海事英語コーパスの分析は、代表的コンコーダンスである AntConc (Anthony 2014)²⁾ に .txt ファイルを読み込んで遂行した。

1) 総語数のカウント

AntConc の Concordance 機能を使用して、異なり語数と延べ語数をカウントした。

2) 単語頻度数の調査

AntConc の Wordlist 機能に e_lemma_no_hyphen.txt (Someya 1998)³⁾ によるレマ化も併せて施した。この際、“be → am, are, is, was, were, being, been, 'm, m” は停止するなど微調整を加えた。

3) 一般英語との比較

本コーパスで用いられている語彙が一般英語と比較してどのような特徴を有しているかを知るために、British National Corpus (BNC)⁴⁾ との比較を行った。BNC は 20 世紀後半のイギリス英語 1 億語（うち 90% が Written Part、10% が Spoken Part を構成）を収集・記録したコーパスである。今回は、語彙の頻度数に関する海事英語コーパスの特徴を記述するために、Kilgarrieff の lemma.num のファイ

ル⁵⁾ を利用し、高頻度語の比較を試みた。

第三章 結果

AntConc を用いて、本海事英語コーパスに対する様々な側面からの計量を行うことにより、次に挙げるような結果が明らかとなった。

1. 海事英語コーパス概観

まずはコーパスの全体像を掴むべく、基本データを以下に提示する。

1) 総語数

異なり語数 = 6,764、延べ語数 = 155,533 であった。

ここでは、レマ化を行ったことにより、名詞の単複の違いや動詞の活用形などは、異なり語数には反映されていない。Table 1 の「内訳」部で明らかのように、例えば、名詞の単複を表す cargo と cargoes、動詞の load, loaded, loading といった形は、それぞれ併せて一つと数えられている。また、以下の分析に関わってくる注意点として、ship (名詞と動詞) や have (本動詞と助動詞) のように二つ (以上) の品詞にまたがって用いられる場合も一つと数えられている。

2) 単語頻度分析

次ページの Table 1 は、コーパスを構成する単語のうち、出現頻度が高いものの順位 1 位から 100 位までをリストアップしたものである。

いわゆる機能語は黒、内容語は色付き (特に海事英語としての使用が顕著であると判断した単語は赤、それ以外の単語は青) で示してある。

3) 出現頻度と語彙のカバー率

出現頻度を具体的に見ていく。Table 1 の「頻度 %」は、各単語のテキスト内での出現頻度をパーセントで示したもののだが、ランキング 1 位の the は全体の 8.7% 近くを占めている。つまり the は 10 語に 1 語弱の頻度で出現していることになる。また、順位 10 位までの高頻度語の頻度 % を足し合わせると、高頻度語上位 10 語だけで、海事英語テキスト全体の 29.62%、すなわち 30% 近くの出現語をカバーしていることが分かる。さらに、上位 20 語、30 語、…と拡げていくと、それぞれ次のようなカバー率となる：

上位 20 語 : 36.31%
 上位 30 語 : 40.29%
 上位 40 語 : 42.67%
 上位 50 語 : 44.71%
 上位 60 語 : 46.45%
 上位 70 語 : 47.96%

上位 80 語 : 49.31%
 上位 90 語 : 50.54%
 上位 100 語 : 51.66%

これは、高頻度の単語 100 語を知っていれば、海事英語テキストに出現する半分の語は分かることを示唆する。(むしろ、ここでの「分かる」とは単語レベルの話であり、その知識をもって文章を正しく理解できる保証はない。)

次に、Table 1 には単語例が示されていない順位 100 位以降の単語も含めた語彙カバー率がどのようになっているかを示したのが Figure 1 である。1,000 語知っていれば、80% を超える単語がカバーできていることに、また 2,000 語になると、92.42% ものカバー率が見込める。

一方で、低頻度語に関しては、出現回数が 2 の単語は 926 語、1 の単語は 2,206 語におよび、異なり語数、すなわち出現する単語種類数 (6764 語) のそれぞれ 13.69%、32.61% を占めている。

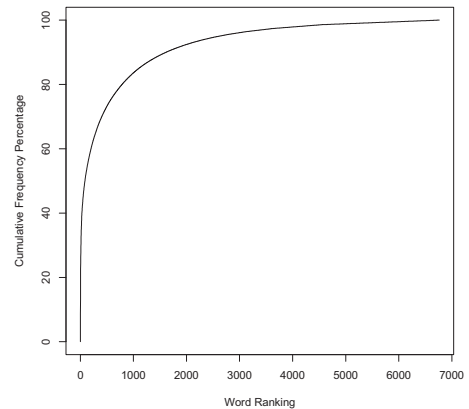


Figure 1 : 単語の頻度ランキングと語彙カバー率

4) 高頻度語の傾向：機能語と内容語

Table 1 に戻って、上位 100 語の種類について詳しく見ていくと、上位 20 語ぐらいまでは、機能語 (文の骨格を作っている語、黒字で示されている) が圧倒的に多いことが分かる。最も基本的な内容語である be 動詞と have 以外、トップ 10 はすべて機能語、それに続く 10 語ではまさに海事用語と言ふべき ship と cargo が入っているものの、それ以外はすべて機能語である。次の 20 語は内容語 (海事用語が多い) と機能語が半々程度になり、それ以降 (41 位～) は徐々に内容語が多くを占めるようになっていく。

なお、本コーパスは品詞タグ付けの作業を行っていないため、複数の品詞にまたがりうる語も 1 語とカウントされている。例えば、can は助動詞と名詞になり得るが、cans という複数形にならない限り区別がつかず、本研究で 33

Table 1 海事英語コーパスにおいて頻度数の高い
単語トップ 100

順位	単語	出現頻度数	頻度%	内 訳
1	the	13464	8.66	
2	be	7008	4.51	am (1), are (996), be (1799)
3	of	5504	3.54	been (502), is (1714),
4	to	5077	3.26	been (502), m (73)
5	and	4109	2.64	was (1450), were (473)
6	in	2967	1.91	
7	a	2849	1.83	
8	that	2149	1.38	that (2055), those (94)
9	have	1612	1.04	had (486), has (455), have (622),
10	on	1328	0.85	having (48), ve (1)
11	ship	1325	0.85	ship (935), shipped (21),
12	for	1253	0.81	shipping (38), ships (331)
13	with	1041	0.67	
14	cargo	1038	0.67	cargo (865), cargoes (173)
15	as	1036	0.67	
16	by	1022	0.66	
17	this	952	0.61	these (240), this (712)
18	not	925	0.59	
19	or	910	0.59	
20	at	906	0.58	
21	it	903	0.58	
22	from	758	0.49	
23	should	727	0.47	
24	vessel	648	0.42	vessel (455), vessels (193)
25	master	601	0.39	master (462), masters (139)
26	which	574	0.37	
27	an	549	0.35	
28	port	515	0.33	port (421), ports (94)
29	load	457	0.29	load (112), loaded (129), loading (199), loads (17)
30	use	454	0.29	use (184), used (178), uses (3),
31	when	400	0.26	using (89)
32	container	398	0.26	container (197), containers (201)
33	can	392	0.25	
34	if	379	0.24	
35	any	372	0.24	
36	all	358	0.23	
37	crew	356	0.23	crew (320), crews (36)
38	there	354	0.23	
39	may	349	0.22	
40	report	347	0.22	report (174), reported (81), reporting (55), reports (37)
41	no	346	0.22	no (343), nos (3)
42	their	342	0.22	
43	they	327	0.21	
44	will	326	0.21	
45	member	323	0.21	member (94), members (229)
46	officer	315	0.20	officer (205), officers (110)
47	but	303	0.19	
48	such	299	0.19	
49	being	295	0.19	
50	take	294	0.19	take (96), taken (113), takes (7), taking (51), took (27)
51	water	294	0.19	water (248), waters (46)
52	incident	282	0.18	incident (201), incidents (81)
53	time	280	0.18	time (224), timed (1), times (53), timing (2)
54	oil	277	0.18	oil (255), oils (22)
55	also	272	0.17	
56	do	266	0.17	did (117), do (94), doing (23), done (32)
57	make	263	0.17	made (178), make (71), makes (14)
58	risk	263	0.17	risk (222), risked (2), risking (1), risks (38)
59	other	262	0.17	other (255), others (7)
60	hold	255	0.16	held (21), hold (154), holds (80)
61	would	250	0.16	
62	into	247	0.16	
63	case	243	0.16	case (165), cases (78)
64	sample	241	0.15	sample (52), sampled (2), samples (102), sampling (85)
65	board	239	0.15	board (216), boarded (22), boards (1)
66	require	228	0.15	require (41), required (133),
67	during	227	0.15	requires (36), requiring (18)
68	provide	227	0.15	provide (82), provided (73), provides (57), providing (15)
69	tank	227	0.15	tank (136), tanks (91)
70	discharge	222	0.14	discharge (157), discharged (33), discharges (19), discharging (13)
71	result	221	0.14	result (108), resulted (47), resulting (23), results (43)
72	one	220	0.14	one (218), ones (2)
73	ensure	215	0.14	ensure (186), ensured (5),
74	where	213	0.14	ensuring (24)
75	association	212	0.14	association (209), associations (3)
76	he	206	0.13	
77	only	205	0.13	
78	authority	202	0.13	authorities (143), authority (59)
79	pilot	201	0.13	pilot (172), piloted (1), pilots (28)
80	watch	200	0.13	watch (195), watched (1), watches (4)
81	damage	195	0.13	damage (131), damaged (48),
82	out	195	0.13	damages (14), damaging (2)
83	operation	194	0.12	operation (109), operations (85)
84	sea	194	0.12	sea (180), seas (14)
85	however	192	0.12	
86	safety	192	0.12	
87	cause	190	0.12	cause (69), caused (61), causes (26), causing (34)
88	carry	188	0.12	carried (103), carries (3),
89	more	187	0.12	carry (39), carrying (43)
90	some	185	0.12	
91	test	183	0.12	test (75), tested (26),
92	two	181	0.12	testing (58), tests (24)
93	must	180	0.12	
94	give	176	0.11	gave (19), give (49), given (79), gives (12), giving (17)
95	state	174	0.11	state (80), stated (20),
96	then	174	0.11	states (63), stating (11)
97	system	171	0.11	system (113), systems (58)
98	up	171	0.11	
99	anchor	169	0.11	anchor (104), anchored (31), anchoring (25), anchors (9)
100	could	169	0.11	

(参考) トップ 100 からは外れたものの、101 位から **note**, **include**, **condition**, **high**, **about**, **fuel**, **order**, **area**, **so**, **than**, **after**, **within**, **before**, **fine**, **new**, **under**, **record**, **collision**, **regulation**, **both**, **hatch**, **course**, **deck**, …と続いていく。

位として挙げられている 392 例も助動詞と名詞の両方が混在しているのが現状である。

ただし、一つ一つの例を確認していくことにより、品詞を認定することは可能である。AntConc の Concordance 機能 (Figure 2 参照) を使って確認したところ、15 例が名詞として使われており、あとは助動詞の例であった。15 の名詞例のうち 9 例は can test (試験の一種) という句で使われており、あとの 6 例はそれ以外の形で用いられていた。例 : The accident victim, the 1st Oiler, went to the engine room workshop to make a garbage can from an empty steel drum. (11-1: Personal Injury Warning)

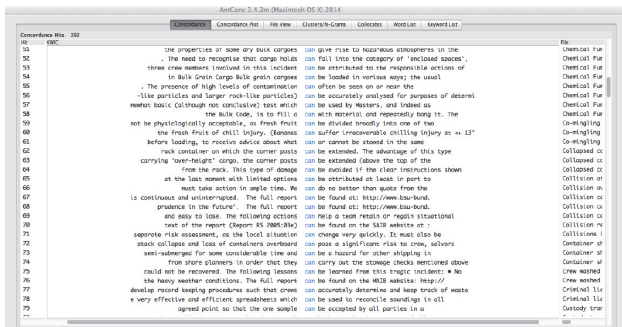


Figure 2 : AntConc の Concordance 機能画面
can が現れる文を前後の文脈とともに列挙している。

5) 高頻度語の傾向 : 内容語の種類

内容語の中心を成すのは動詞、名詞、形容詞、副詞であるが、海事英語コーパスのトップ 100 では各品詞に対して次のような語が見られた :

動詞 : be, board*, carry, cause*, damage*, discharge*, ensure, give, have, hold*, load*, make, provide, report*, require, result*, risk*, ship*, state*, take, test*, use*, watch* (23 語)

名詞 : anchor, association, authority, board, cargo, case, cause, container, crew, damage, discharge, hold, incident, master, member, officer, oil, operation, pilot, port, report, result, risk, safety, sample, sea, ship, state, system, tank, test, two, time, use, vessel, watch, water (37 語)

形容詞 : more, only, other (3 語)

副詞 : however, more (2 語)

前項で can が 2 つの品詞にまたがる例を示したのと同様のケースが、ここでは動詞と名詞で多く見られる。そこで、そのような単語については、動詞と名詞の両方に載せ、動詞の方にアスタリスク (*) を付した。なお、形容詞と副詞の 2 つにまたがる例として、more があった。

挙げられた内容語は動詞と名詞、特に名詞が圧倒的に多い。また、海事英語らしい語も目立っている。

2. 一般英語との比較

本海事コーパスと、現代イギリス英語の傾向を示すことを目的として収集された BNC の書きことばの語彙データを比較することにより、海事英語を読むにあたって特に知っておきたい語彙特徴は何かの記述を試みる。

1) BNC において高頻度で出現する単語

Kilgarriff の lemma.num に基づいた BNC における高頻度トップ 100 の単語は次のとおりである :

順位 1 ~ 10 : the, be (動詞), of, and, a, in, to (不定詞), have (動詞), it, to (前置詞)

11 ~ 20 : for (前置詞), I, that (接続詞), you, he, on, with, do, at, by

21 ~ 30 : not, this (限定詞), but, from, they, his (限定詞), that (限定詞), she, or, which (限定詞)

31 ~ 40 : as (接続詞), we, an, say, will (助動詞), would, can (助動詞), if, their, go

41 ~ 50 : what (限定詞), there, all (限定詞), get, her (限定詞), make, who (代名詞), as (前置詞), out, up

51 ~ 60 : see, know, time, take, them, some (限定詞), could, so, him, year

61 ~ 70 : into, its, then, think, my, come, than, more, about, now

71 ~ 80 : last (形容詞), your, me, no, other (形容詞), give, just, should, these (限定詞), people

81 ~ 90 : also, well (副詞), any (限定詞), only, new, very, when (接続詞), may, way, look (動詞)

91 ~ 100 : like (前置詞), use, her (代名詞), such, how, because, when (副詞), as (副詞), good, find, man

ここでも機能語は黒字、内容語は青字で示してある。併記されているラベルから明らかのように、海事英語コーパスと異なり、BNC は品詞を区別して数えている。このことに留意しながら、2 つのコーパスを比較する。

2) 海事英語コーパスと BNC の比較 : 順位 1-10 位

2 つのコーパスで極めて類似した単語がランクインしている。1 位 = the, 2 位 = be, 3 位 = of というのは 2 つのコーパスでまったく同様であるし、4 位について、海事英語は to になっているのは、不定詞の to (BNC では 7 位) と前置詞の to (同 10 位) が区別なく合算されているからであろう。BNC で 4 位となっている and は、海事英語では to に続いて 5 位となっている。

順位の多少の違いがあるものの、それ以外の in, a, have についてもいずれのコーパスにも見られる。両コーパスとも内容語は基本同士の be と have で、それ以外はすべて機能語である。

3) 海事英語コーパスと BNC の比較 : 順位 11 位以降

11 位から 50 位までの単語を見てみると、海事英語では内容語が徐々に増えていって 40 語中 13 語を数えるのに

対し、BNC では5語にとどまる。また、51位以降では海事英語では非常に多く(39語)が内容語であるのに対し、BNC では半分の25語である。

この原因としては、繰り返し述べているように、BNC では品詞ごとに単語の分類がされているのに対し、海事英語ではそれがなされていないからという可能性がある。事実、BNC のトップ100の中では複数の品詞にまたがっている単語で、海事英語では1単語にまとめられているような機能語は: as, her, that, to, when がある。また、同じ限定詞の this, these も海事英語では1単語にまとめられている。

このような分類の違いは存在するものの、極めて高頻度な単語は機能語が多く占め、頻度順位が下がっていくにつれて内容語が増えていく傾向は海事英語とBNCのいずれでも同じである。機能語の数には限りがあるので、順位200位まで範囲を広げれば、いずれのコーパスも内容語ばかりということになるであろう。

4) 海事英語コーパスとBNCの比較: 内容語

2つのコーパスの最も大きな違いはトップ100以内に出現する内容語の質にある。BNCの内容語32語はいずれも非常に基本的で、すべて中学レベルの段階で学ぶような単語から構成されている。

海事英語コーパスの場合はその限りではない。いわゆる「単語の難易度」を測定する方法や基準には様々なものがあるが、学習者向け辞書として本学学生が高校までの段階で最も使用者が多い辞書の一つである『ジーニアス英和辞典』(第4版)⁶⁾で用いられているランク付け(Aランク: 中学基本語約1100語; Bランク: 高校学習語約3400語; Cランク: 大学生・社会人に必要な語約5100語)を参考にして海事英語のトップ100に出現する内容語53語を調べると、Bランクの単語が20、Cランクの単語が2(vessel, anchor)あった。

第四章 考察

本章ではここまでで明らかとなった多くの事実を考察し、海事英語学習者が語彙学習にあたって持つべき視点は何かということに焦点を当てて論じていく。

1. 単語の出現頻度

単語の出現頻度の調査結果より、(1)高頻度トップ10を圧倒的に機能語が占め、そのレパートリーは大規模な一般的イギリス英語を代表するBNCと変わらない、(2)全体の異なり語数の約3分の1で90%以上の出現単語をカバーしている、(3)頻度が1や2しかない低頻度な単語が極めて多い、の3点が明らかとなった。

これは、石川(2008: 63)⁷⁾が述べる「自然言語における語彙は、機能語中心の一握りの高頻度語と、圧倒的多数の低頻度語から構成されている」という傾向とも一致している。

2. 高頻度の内容語

一般的に、高頻度な内容語は基本動詞に圧倒的に多く、コミュニケーション動詞(say)、アクション動詞(go, come, get)、メンタル動詞(see, know, think)から構成されると言われる(投野2006: 40-41)⁸⁾。また、名詞や形容詞は、文の骨組みを構成する動詞と比べると入れ替え可能な語であるため、圧倒的に少ないとも言われる(投野2006: 49-50)。BNCにおいて順位1~100位に入っている単語は、いずれもこれに該当するものばかりである。

ところが、海事英語コーパスにおける高頻度の内容語では事情が大きく異なっており、頻度の高い内容語の名詞は動詞よりも数が大きく上回っている。また、動詞も必ずしも基本動詞ばかりではない。さらに、難易度の点でもBNCと海事英語には大きな違いが見られ、中学レベルの内容語ばかりのBNCとは対照的に、海事英語には高校レベルやそれ以上の単語が含まれている。その上、海事に偏った語彙も高頻度の単語として多く出現している(Table 1では赤字で示している)。

BNCでは難易度が低い単語が高頻度語として多くランクインしている理由として一つ考えられるのは、BNCには全体の10%とはいえ、話しことば(Spoken Part)が含まれていることが挙げられるかもしれない。話しことばでは難易度が低い単語が多く用いられる傾向があるため、中学レベルで学習する基本語が多く含まれる結果となったということである。

高頻度の内容語の調査結果から、海事英語テキストを読み解く上では、「基本語」以上の語彙力が要求されることは明らかである。しばしば「英語は中学レベルでもなんとかなる」などと言われるが、日常的な会話であればともかく、海事英語という専門性を持つ内容の文章読解においては、中学レベルの単語だけでは到底太刀打ちできず、高校レベルを超える単語も登場する可能性があることを肝に銘じておかなければならない。

3. 内容語の用法

ここまでの説明はまだ表面的なものに過ぎない。高校レベル以上の単語や海事英語に分類される単語で高頻度に出現する単語が文章内でどのように使われているかを知ることが「どのような海事英語語彙を学ぶべきか」を論じる上で、必要不可欠である。網羅的な説明は別の機会に譲るとして、いくつかのヒントとなる着眼点を(単純なものから多少複雑なものまで)挙げていくことにする。

1) 上位100位までの単語だけでも類義語が学べる

Table 1の青字と赤字の単語を見て、その例文での使われ方をAntConcのConcordance機能で確認していくと、英語では多くの類義語が存在し、それぞれの用法があることが見えてくる。

「船」は ship だけでなく vessel も出てくるが、前者の出現数が後者の2倍多い。「船長」は日本語であれば captain (キャプテン) がすぐに思い浮かべられるが、本海事英語コーパスはイギリス英語がベースなので、master が圧倒的に多い (captain は頻度数 10 回で、順位 1731 位)。「事故」といえば accident がまずは出てくるが、こちらは頻度数 140 回で、順位 129 位。順位 52 位の incident の方が 282 回の頻度数で、ずっと多い。これ以外にも sea と ocean、risk と danger などが類義語として挙げられる。各単語がどのようなコンテキストやコロケーションで用いられるか調べると、単語に対する知識が増し、スムーズな読解にも役立つことであろう。

2) 「中学レベル」の単語も侮れない

高頻度語の使われ方を詳しく見ていくと、中学レベルの単語だからといって、分かった気になってはいけないうケースもある。

頻度順位 51 番の water を例に取る。高校段階までの英語では、この単語は「水」という不可算名詞で用いられるのが一般的だが、AntConc の Sort 機能も活用してつぶさに見ていくと、海事英語テキストで頻出している語義・用法があることがある。Water で 248 回、waters で 46 回の出現回数が見られるが、前者では ballast water (16 回)、bilge water (7 回)、fresh water (5 回)、oily water (39 回) などのように、特定の形容詞と結びついて複合語を作っている。

また、(不可算名詞のはずの water に -s がついた形であるということで驚くかもしれない) waters で見ていくと、Australian waters, coastal waters, Columbian waters, confined waters, European waters, internal waters, international waters, Japanese waters, restricted waters, territorial waters, …のように、国やエリアを表す形容詞が waters を限定している用法ばかりになっていることが分かる。

ここで『ジーニアス英和辞典』(第4版)で調べると、4番目の語義として「[~s]水域、海域；領海、近海」という意味を見つかることができる。つまり、海事のコンテキストでは、coastal waters (沿岸水域)のように「海域」という意味で water が用いられることが多々あるのである。学習者の中には、その特別な語義がある可能性を考えずに water = 「水」と短絡的に考える者も多いし、辞書を引いても相当する語義は後の方に出てくるので、正しい意味にたどり着けない者も多い。

3) 機能語も手強い

機能語の用法はすべての英語で一律である、と決めつけるのもいけない。ここで、助動詞に目を向けてみる。助動詞のうち、海事英語コーパスで最も頻度が高いのは should (23 位、727 回) であり、後に続く can (33 位、392 回) や may (39 回、349 回) と比べても図抜けている。また、BNC の場合は will, would, can が 30 位台、could が 50 位台であるのに対して should は 70 位台である、というのと比

較しても、海事英語コーパスの頻度は高い傾向があるように見受けられる。そこで、should の用法を観察し、顕著な特徴はないかを見てみる。

助動詞 should といえば、「~すべきである」というのが真っ先に頭に思い浮かぶであろう。海事のコンテキストであれば法令遵守等、「すべき」ことは多いはずで、実際にこの意味の用法が圧倒的に多いのだが、それでは具体的にどのようなことを「すべき」か、should の後に結びつきやすい動詞を調べてみる。すると、能動の形の動詞の場合 (should be + 動詞の過去分詞形ではない形の場合)、consider, contact, ensure, include, seek, note、などが多いことが分かる。次のような例文単位で代表的な結びつきを確認することは学習者にとって非常に大切であろう：

例1 : Ideally shippers should ensure that such cargoes are adequately covered but by the nature of the trade this is unlikely to occur. (16-1: Steel Carried in Open-Top Containers – The Need for Inspection)

例2 : Information on labels should include date and time, vessel name, port, location (e.g. cargo tank, manifold), description of cargo, type of sample (e.g. first foot, composite, running, ‘Upper, Middle, Lower’), identity of sampler, and seal number. (18-1: Sampling Risks)

なお、さらに詳しく見ていくと、if ~ should が倒置を起こして “Should …” で始まる文例、“should have 動詞の過去分詞” という形なども出てくる。学習者にとって用例の宝庫と言えよう。また、こういった用法は、SMCP のような最小限度レベルの英語では出てこない文法項目であることも指摘しておきたい。

第五章 まとめ

本研究は、海事英語テキスト *Risk Watch* を言語データとして、書きことばの海事英語を構成する語彙特徴を明らかにするために、AntConc を用いる分析を試みた。言語データを概観すると共に、主に高頻度語の特徴に着目し、一般英語との類似点と相違点を示した。

分析の結果、一般英語と海事英語共に、言語の基盤を成す機能語の使われ方は類似していたものの、伝えたい内容の中身を構成する内容語の高頻度語の種類には大きな違いが見られた。一般英語では主に中学までで学ぶ語が頻度の上位を占めるのに対し、海事英語の場合は高校学習語レベルの単語が多く出現、少数だが大学生・社会人レベルに必要な語も見られた。

さらに、海事英語の内容語のテキスト内での使われ方も特異で、例えば、中学レベルの単語であっても、中学レベルでは (おそらく高校レベルでも) 普通は扱われない海事特有の語義が表されているケースも見られた。この事実は、海事英語テキストの内容を理解するための語彙として、高校レベルにプラス α の知識が必要とされることを示唆している。

本研究で得られたデータは海事英語テキストにおける言語用法について一定の事実を明らかにした。今後、これを足がかりとして更なる言語学的な分析、すなわち言語的に意味のある質的解釈を加えることにより、特に高校レベルまでの英語学習を済ませた学習者が更に ESP (English for Specific Purposes) として海事英語を身につけて行く際の語彙学習の指針となるような事実を明らかにすることが期待される。今後、語彙カバー率が 80% を超える上位 1,000 語に現れる内容語の用法について、AntConc を活用して調べ上げることが求められる。

なお、用いた海事英語資料は、タグ付けをすることにより、さらに精巧な分析を行うことが可能になるかもしれない。特に、一般英語と比較しながら Keyword List 機能等を用いて特徴後を抽出するような場合、現在の BNC に歩み寄った形でのタグ付けを行うことにより、厳密な統計的解釈も可能になるであろう。Constituent Likelihood Automatic Word-tagging System (CLAWS4)⁹⁾ のような自動タグ付けシステムを利用するのも一考であろう。

将来の海技士の SMCP を超えた英語力の強化に向けて、今後の研究の進展が強く望まれる。

謝辞

本稿は、本学卒業研究における言語分析手法と研究結果のまとめ方をデモンストレーションするために書かれました。データとして用いた海事情報誌 *Risk Watch* のテキスト

ファイル化及び分類作業に協力して下さった海事システム工学科 4 年の真田一騎君と成川宏海君に御礼申し上げます。

参考文献

- 1) Risk Watch. <http://www.britanniapandi.com>
- 2) Anthony, L. AntConc (Version 3.4.2m) [Computer Software]. Tokyo, Japan: Waseda University, 2014. Available from <http://www.antlab.sci.waseda.ac.jp/>
- 3) Someya, Y. English lemma list, 1st ver., 1998, Retrieved Aug. 26, 2014, from http://www.antlab.sci.waseda.ac.jp/antconc_index.html
- 4) British National Corpus. <http://www.natcorp.ox.ac.uk>
- 5) Adam Kilgarriff, BNC database and word frequency lists. Retrieved Aug. 26, 2014, from <http://www.kilgarriff.co.uk/bnc-readme.html>
- 6) 小西友七, 南出康世(編集主幹). ジーニアス英和辞典, 第 4 版. 大修館書店, 2007, 2249p.
- 7) 石川慎一郎. 英語コーパスと言語教育: データとしてのテキスト. 大修館書店, 2008, 265p.
- 8) 投野由紀夫. コーパス超入門: コーパスでわかる英語学習のコツ. 小学館, 2006, 159p.
- 9) Constituent Likelihood Automatic Word-tagging System (CLAWS4). <http://ucrel.lancs.ac.uk/claws/>

AntConc を用いた海事英語テキストの語彙分析

内田 洋子

(東京海洋大学大学院海洋工学系海事システム工学部門)

要旨: 将来の海技士に求められる英語力は「必要最低限のコミュニケーション」を目的とする SMCP の習得にとどまらず、必要な海事情報を得るための書きことばを読解する力にまで拡充すべきである。それでは、それを達成するためには、どのような語彙を習得すればよいのか? この疑問に答える一手段として、本研究では AntConc を用いて海事英語テキスト *Risk Watch* を構成する語彙特徴を明らかにする分析を試みた。言語データを概観すると共に、主に高頻度語の特徴に着目し、BNC に代表される一般英語との類似点と相違点を示した。分析の結果、一般英語と海事英語の機能語の使われ方は類似していたが、内容語の高頻度語の種類には大きな違いが見られた。海事英語テキストの内容を理解するための語彙として、高校レベルにプラス α の知識が必要とされることを示唆している。今後、頻度上位 1,000 語を構成する内容語の用法を綿密に記述し、高校レベルまでの英語学習を済ませた学習者が更に海事英語を身につけて行く際の語彙学習の指針となるような事実を明らかにすることが期待される。

キーワード: 海事英語、語彙、コーパス分析