

英語論文コーパスにおける高頻度語連鎖の分析

—英語教育法と言語運用能力の探究をめざすフレイジオロジー—

梅 咲 敦 子

I はじめに

英語による論文作成を苦手と考える非英語母語話者も多い。しかし、論文に含めるべき情報とその順を追った流れには一定の枠組みがあり、各枠組みで用いられる表現にも一定のパターンがある。その情報の流れと表現形式を理解していれば、英語論文の作成は容易になる。この考え方は、Swales (1981, 1990, 2004) のジャンル分析 (genre analysis) に共通し、その中では、情報の推移におけるそれぞれの枠組み (情報推移区分) は move と呼ばれている。大量の論文をコーパスとして、情報推移区分とその中で使用される表現を収集し議論することは、英語教育上意義があると同時に、言語自体の研究にも意味があると考えられる。

従来の言語研究は、言語能力 (competence) と言語運用 (performance) を区別して、Chomsky (1957) 以降、前者の探究が主流となった。なるほど、前者の、無限個の文が生成できる有限個の規則すなわち普遍文法の解明は、重要な言語研究である。しかし、実際の言語使用では、無限個の文の生成というよりも、特定の語の組み合わせが、統語上可能な他の組み合わせよりも高頻度に現れる。この現象を言語運用上の周辺的事象と片付けず、実際の在り様を捉えなおすことも言語の本質に近づく道と考えられるのではないか。その立場から、実際の言語使用に関わる能力、すなわち言語運用能力を追求したい。

筆者は、一定の意味・機能を担う特定の語（句）の組み合わせが人間の記憶に蓄積しており、人はその内から状況に応じた適切な組み合わせを選択して統語規則と合わせて使用すると考える。この主張は、Hoey (2005) のレキシカル・プライミング¹⁾ (lexical priming) や Hunston and Francis (2000) のパターン文法 (pattern grammar) に通じる。さらに、筆者はこの言語運用能力の追求とその英語教育法への応用をフレイジオロジー²⁾ (phraseology) の目標と考える。

本稿では、専門家と日本語を母語とする大学院生の英語論文における高頻度に使用される語連鎖を比較することを通して、語連鎖の言語運用能力に占める重要性に示唆を与える。尚、一定の意味・機能を担う特定の語（句）の組み合わせは、連語の他、語（彙）連鎖 (lexical bundle, Biber et al., 1999) や定型連鎖（表現）(formulaic sequence, Wray, 2002) など、研究者の焦点の当て方や研究手法により用語が異なる。本稿では、自動抽出した n 個の語連鎖 (n-gram) に着目する点で Biber et al. (1999) の研究手法に近く、語連鎖を用いる。

II 論文における語連鎖の抽出方法

コーパスとして、英語論文約1.5億語を独自に収集した Corpus of English Academic Papers (CEAP)³⁾、及び日本語を母語とする言語系大学院生の論文

- 1) Hoey (2005: 11) には “The notion of priming as here outlined assumes that the mind has a mental concordance of every word it has encountered, a concordance that has been richly glossed for social, physical, discursal, generic and interpersonal context.” とあり、priming は人が体験した全ての語の使用状況と文脈に関わる記憶を指す。Lexical priming に定訳はないが、レキシカルプライミングや語彙経験が当てられている。
- 2) 八木 (2013: 51) は、「フレイジオロジーは、言語使用の研究であり、実際の発話資料をデータとし、言語を構成する基本単位である句 (phrase) の意味と用法を研究するのが本旨である。文法規則は句を統合する従属的役割を果たすに過ぎない。このような観点からのフレイジオロジー研究は、言語研究に新しい視点をあたえるとともに、英語教育にも新たな視点を与えることになる。」と述べている。特にこの後半の指摘と筆者の主張は近い。
- 3) Science Direct から機関利用可能であった論文を2005年から2010年にかけて手作業で収集した。出典の詳細は補足資料1参照。

約42万語を独自に収集した Corpus of Learner Papers (COLP) を使用した。語連鎖抽出ソフトとして AntConc⁴⁾、n-gram maker⁵⁾ を用い、パターン検索には Sketch Engine⁶⁾ を利用した。AntConc では抽出できなかった数千万語のコーパスにおける語連鎖抽出が n-gram maker では可能になり、CEAP に保有している全テキストを人文・経済経営・経営社会・自然科学に分類して4語連鎖を抽出した。

比較分析のため、自然科学分野の論文約1700万語を収集した小学館コーパスネットワーク提供の PERC Corpus (PERC)⁷⁾ の他、既存の大規模コーパス (COCA, ukWaC, BNC) を利用した。

Ⅲ 人文・経済経営・経営社会・自然科学の4語連鎖

CEAP の4語連鎖のうち上位50位までを表1と表2に示す。前置詞で始まる語連鎖を抽出すると4分野の50位までに17~24タイプ出現しており、各分野で34~48%にあたる。その他、この結果は梅咲 (2010b) と一致し、4語連鎖の特徴として、中高等学校の教科書では動詞句が高頻度に出現するのに対し、論文では句前置詞を含めて前置詞から成る句 (前置詞句) が多い。従って、大学英語教育が英語の学術的文章の読み書きを目標とするなら、学術的文章を念頭においた英文に多く触れる必要があることを示していると言えよ

入手可能な資料をダウンロードしそのままコーパスとして使用しているため、References に大量に現れるネット特有の語句 (例えば、.... International Journal of Industrial Organization 22 (8-9) (2004), pp. 1237-1263, Article 1 PDF (213 K) | View Record in Scopus | Cited By in Scopus (33) の下線部) は、検索結果から手作業で取り除いた。

- 4) Lawrence Anthony 氏 (早稲田大学) 作成。入手先 <http://www.laurenceanthony.net/software.html>
- 5) 田中良氏 (関西学院大学課程博士) 作成。入手先 http://www7b.biglobe.ne.jp/~ryo_tanaka_corpus_tools/
尚、最初に抽出される語連鎖が大文字であれば、小文字で始まる場合も大文字の語連鎖に集約してカウントされるため、手作業で小文字表示に変更した。
- 6) 搭載コーパスの検索だけでなく、独自作成のコーパスをアップロードして検索することも可能。参照先 <https://www.sketchengine.co.uk/>
- 7) 参照先 http://scn.jkn21.com/~percinfo/index_j.html

表 1 論文 (CEAP) における 4 語連鎖—経済・経営・社会学分野

順位	頻度	Economics & Management	頻度	Management & Society
		総語数：51,420,236		総語数：17,074,910
1	6469	on the other hand	2248	on the other hand
2	4675	in the case of	1903	in the case of
3	4282	in the context of	1388	at the same time
4	4144	are more likely to	1365	as well as the
5	3545	as well as the	1264	in the context of
6	3463	on the basis of	1231	the extent to which
7	3163	at the same time	1138	on the basis of
8	2909	the results of the	1128	in terms of the
9	2774	as a result of	1040	are more likely to
10	2609	in terms of the	1030	as a result of
11	2474	at the end of	936	the results of the
12	2438	the nature of the	756	one of the most
13	2350	it is important to	727	the case of the
14	2182	the degree to which	692	the nature of the
15	2164	the end of the	691	on the one hand
16	2058	with respect to the	679	in the form of
17	2039	in the form of	678	with respect to the
18	1784	for each of the	657	at the end of
19	1747	the value of the	623	the total number of
20	1701	is consistent with the	617	the development of the
21	1680	is likely to be	617	is one of the
22	1646	the size of the	608	the size of the
23	1634	are likely to be	592	the fact that the
24	1525	in this paper we	563	the rest of the
25	1506	in addition to the	557	the end of the
26	1505	the fact that the	555	in the development of
27	1487	can be used to	544	the role of the
28	1480	is based on the	541	in the field of
29	1468	the quality of the	541	a wide range of
30	1451	the relationship between the	524	the purpose of this
31	1423	is positively related to	523	paper presented at the
32	1416	the impact of the	505	the degree to which
33	1395	more likely to be	502	more likely to be
34	1388	as a function of	497	with regard to the
35	1387	the importance of the	494	is based on the
36	1386	to the extent that	485	in addition to the
37	1378	that there is a	479	that there is a
38	1356	is one of the	479	in the number of
39	1348	one of the most	474	the impact of the
40	1343	at the beginning of	456	of the number of
41	1341	at the time of	455	the importance of the
42	1327	of this paper is	443	to the fact that
43	1318	on the one hand	439	the value of the
44	1317	the role of the	438	of this paper is
45	1310	the purpose of this	437	in this paper we
46	1308	the case of the	436	an important role in
47	1308	the use of the	434	a large number of
48	1295	of a firm's	431	can be used to
49	1284	a high level of	431	are likely to be
50	1282	the results of this	429	is likely to be

注：総語数は n-gram maker によるため AntConc による数値と多少異なる。

表2 論文 (CEAP) における4語連鎖—人文・自然科学分野

順位	頻度	Arts & Humanities	頻度	Natural Science
		総語数: 61,599,174		総語数: 29,445,207
1	6828	on the other hand	12678	as a function of
2	5472	in the case of	6526	in the case of
3	5141	on the basis of	5457	on the other hand
4	4204	as a function of	5066	as shown in Fig
5	4090	at the same time	4502	are shown in Fig
6	4090	in the context of	4203	a function of the
7	3840	as well as the	3664	is shown in Fig
8	3456	at the end of	2928	the glass transition temperature
9	3125	in terms of the	2925	the temperature dependence of
10	2868	the nature of the	2774	temperature dependence of the
11	2835	the fact that the	2125	as well as the
12	2733	as a result of	2100	with respect to the
13	2619	the results of the	2091	in the range of
14	2600	it is important to	2078	of non crystalline solids
15	2562	with respect to the	2076	is due to the
16	2460	in the present study	1982	the other hand the
17	2409	in the form of	1802	the intensity of the
18	2393	it is possible that	1801	on the basis of
19	2056	at the beginning of	1683	an increase in the
20	2054	on the one hand	1678	the case of the
21	1850	in the absence of	1619	the structure of the
22	1844	for each of the	1616	the size of the
23	1735	can be used to	1592	the fact that the
24	1687	it is possible to	1588	is related to the
25	1631	the total number of	1540	in terms of the
26	1627	to the fact that	1508	in this paper we
27	1623	in addition to the	1481	the thickness of the
28	1615	at the time of	1478	to the formation of
29	1570	the ways in which	1466	a function of temperature
30	1537	paper presented at the	1412	the effect of the
31	1531	that there is a	1401	the increase of the
32	1467	the rest of the	1398	it can be seen
33	1448	a wide range of	1369	in the temperature range
34	1424	in relation to the	1359	it is possible to
35	1407	it is clear that	1355	as a result of
36	1383	a main effect of	1311	in the presence of
37	1361	the degree to which	1287	the position of the
38	1349	the difference between the	1253	to the presence of
39	1347	in the same way	1251	in this case the
40	1336	the role of the	1245	in the present work
41	1321	the way in which	1208	at the same time
42	1311	at the level of	1206	in the form of
43	1260	is consistent with the	1195	to that of the
44	1258	to be able to	1174	in agreement with the
45	1245	are more likely to	1166	a heating rate of
46	1240	it is difficult to	1162	was found to be
47	1236	is one of the	1160	an increase of the
48	1230	[it should be noted]	1140	is consistent with the
49	1218	by the fact that	1139	the formation of the
50	1216	in this paper we	1137	[it should be noted]

注：総語数は n-gram maker によるため AntConc による数値と多少異なる。

表3 大学院生論文 (COLP) における4語連鎖

順位	頻度	COLP		順位	頻度	COLP 続き
		総語数: 424,502				
1	210	on the other hand		26	41	native speakers of English
2	113	senior high school students		27	41	in the current study
3	95	the results of the		28	40	high school students'
4	87	in the present study		29	39	there was no significant
5	82	come to and get		30	38	the students In the
6	81	to and get to		31	38	the meaning of the
7	73	junior high school students		32	38	in terms of the
8	68	the Textbooks of three		33	37	This story talks about
9	68	textbooks of three countries		34	37	in the English class
10	63	in the textbooks of		35	37	explanation This story talks
11	58	of the present study		36	36	pre and post tests
12	49	the result of the		37	36	English as a Foreign
13	49	that is to say		38	35	the end of the
14	49	is one of the		39	35	students' attitudes toward
15	48	can be said that		40	35	in the context of
16	47	it can be said		41	34	proper nouns and adjectives
17	44	the Pre and the		42	33	students in the lower
18	44	pre and the post		43	33	as a result of
19	44	for senior high school		44	32	the purpose of the
20	43	the other hand the		45	32	in Junior High School
21	43	on the basis of		46	31	the total number of
22	42	in the case of		47	31	the purpose of this
23	42	at the same time		48	31	the junior high school
24	42	as one of the		49	31	of junior high school
25	42	as a foreign language		50	30	was no significant difference
				51	30	at the end of

う。

IV 専門家英語論文と大学院生英語論文の4語連鎖

表3に日本語を母語とする言語系大学院生の論文の4語連鎖50位までを示した。表1、2の専門家英語論文の場合と比較する。

1. on the one hand と on the other hand

内容を対照させる副詞の機能を持つ合接詞 (conjunct) on the other hand は、表1、2から分かるように、3分野で出現頻度1位、残りの1分野で3位と高頻度で使用されている。同様に大学院生論文でも第1位である。日本語母語話者にも、比較的専門家と同様の頻度で使用できる語連鎖と言えよう。

ここで注目すべきは、専門家論文では on the other hand が1位にある人

文・経営・経済・社会学分野では、on the one hand も15位から43位までの間で出現している。しかし、言語系大学院生論文では、on the one hand は100位（22回出現）までにも上がらず、COLP 全体で6回の出現に留まっている。すなわち、(1) のように On the one hand, ～. On the other hand, ～. で対比させる構文としては日本語を母語とする大学院生には過少使用と言えよう。

- (1) Clearly, the conservative approach taken in the present study has pros and cons. On the one hand, a conservative approach minimizes concerns about sensationalizing matters related to workplace same-sex harassment. Even marginally sensationalizing such matters may compromise the seriousness of the debate over same-sex harassment. On the other hand, a conservative approach may underestimate the true scope of the costs and lead organizations to avoid dealing constructively with same-sex sexual harassment in their workplaces. [CEAP: IJoIR]

2. as a function of

自然科学分野の論文では as a function of の頻度が第1位であり、表4に示すように人文・経済・経営・社会学分野でも100位以内の頻度があるが、言語系大学院生の論文では極端に頻度が少ない。自然科学分野では「～の関数として」を表す句として使用頻度が高いことは容易に考えつくが、その他の分野ではどのような意味に使用されるのか、また大学院生論文に極端に少ない要因を探る。

表4 as a function of の CEAP と COLP における頻度

	順位	頻度	100万語換算
Arts & Humanities	4	4,204	68.2
Economics & Management	34	1,388	27.0
Management & Society	98	297	17.4
Natural Science	1	12,678	430.6
COLP	—	3	7.1

用例 (2)～(4) では、as a function of が前置詞の機能を果たす句前置詞の

ように機能し、「～次第で、～によって、～ごとの、～に対応した、～に応じて、～ごとに」を表すのに対し、(5)～(7)では、as a function of は前置詞＋名詞句からなり、explain や view と共起し、「～の産物・作用・機能・働き・関数として、～との相関で、～に関連づけて、～の（関係しあった）結果として生じるものとして」「説明できる、捉える」を表す。すなわち、「動詞＋O₁＋as＋O₂」のパターンと捉えるべきであり、as a function of の連鎖が句前置詞として機能しているのではない。

- (2) If words are unconditionally semantically processed, irrespective of their frequency of occurrence, then performance in the arrow task should vary as a function of whether the standard sound is congruent or incongruent with the upcoming arrow. [CEAP: Cog]
- (3) Skilled workers were much more likely to belong to a cooperative association as well as receiving weekly benefits in case of sickness or accident than unskilled workers. The parameters of age and the square of age are both statistically significant. The estimated probability as a function of age is bell-shaped : it is increasing for workers younger than about forty and decreasing for older workers (see first column of Table 5 and Fig. A.1) [CEAP EiEH]
- (4) ... and we expect its data transfer performance to be similar to TCPs. To test this we measured the time needed to transfer a 6 MB file as a function of the window size (TCP) and number of outstanding Interests (CCN). [COCA]
- (5) Two distinct but related issues are at stake: the reasons for variation in the incidence of polygynous polygamy and monogamy, and the motivation for the social imposition of universal monogamy regardless of status and resources.
- 7.1. Female choice
- The overall incidence of polygyny may be explained as a function of female mate choice. [CEAP: THotF]
- (6) Day (2002) states that while both production and service delivery activities are aimed at creating value for customers, customer expectations affect satisfaction, which in turn influences the assessments of value, rather than that value plays a

role in satisfaction assessments of customers. Along these lines, quality could be viewed as a function of, rather than an antecedent to customer satisfaction. Despite these contrasting views, the debate regarding the relationship between the two constructs appears to be reaching some form of consensus (Bitner and Hubbert, 1994; Caruana and Pitt, 1997). [CEAP: JoPaSM]

- (7) There are many issues in implementation; however, of central importance is capability. Capability can be viewed as a function of education and experience.

[CEAP: IJoPM]

量的にも経済・経営・社会学分野の論文で as a function of の直前に来る動詞を調べると、図 1 に示す通り vary, change, increase, decrease のような変化を表わす動詞、または view, express, see, define, determine のような捉え方を表わす動詞と共起することが多いことが分かる。

COLP では as a function of は refer to, explain, decrease に後続する 3 例しかないため、過少使用の要因は特定し難いが、ひとつには as a function of を例 (7) のような「～の産物として」を表わす場合の使用が難しいためと推測できる。もうひとつには、COLP には vary according to が 26 例あることから、「～ごとに」を表わす場合 as a function of を according to で代用している可能性が高いからと考えられる。

この代用の要因を探るために、類似表現として「動詞 + according to」を ukWaC で検索した結果を図 2 に示す。動詞には vary など変化を表わす語が多く見られる。しかし、図 2 で according to と共起頻度の高い classify が as a function of と共起する例が図 1 で使用したコーパスにも PERC にも見られなかった。このことから、as a function of は「連続する変化」または「より具体的な変化」の基準を示す際に用いるのが普通だが、according to との使い分けが日本語を母語とする大学院生には難しいため、according to で代用していると言えるかもしれない。

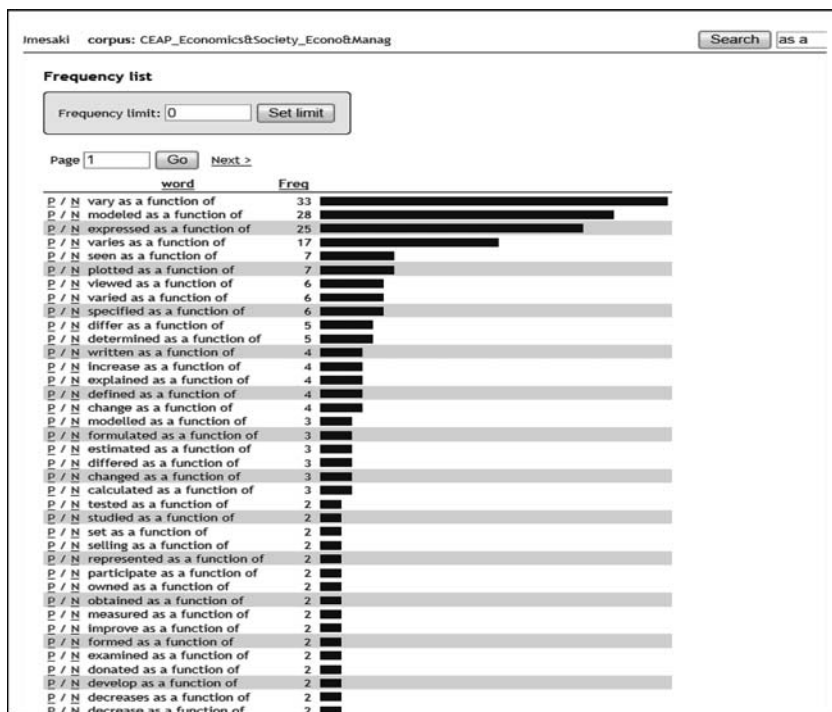


図1 経済・経営分野における as a function of が後続する動詞

注：Sketch Engine による検索 [tag="V.*"] "as" "a" "function" "of"

3. is likely to be

動詞句を検討すると、表1、2から専門家は likely を含む4語連鎖を多用していることが分かる一方、表3から大学院生にその傾向は見られない。(8)(9)のように「～である(～する)可能性が高い」ことを表現するのに適しているが、日本語母語話者には使用し難い表現と言えよう。

- (8) In addition, the frequency of pro forma reporting is much higher in 2000 than in 1998 and 1999. Thus, the estimate we obtain during this time period is likely to be closer to the upper bound of all non-GAAP nomenclatures used by companies in any quarter during our sample period. [CEAP: JoAaE]

word	Frequency
P N vary according to	3,086
P N varies according to	1,130
P N classified according to	469
P N is according to	437
P N grouped according to	328
P N varied according to	320
P N arranged according to	298
P N made according to	297
P N go according to	294
P N calculated according to	286
P N differ according to	277
P N live according to	267
P N selected according to	265
P N act according to	250
P N be according to	222
P N change according to	210
P N assessed according to	200
P N set according to	197
P N paid according to	188
P N goes according to	185
P N organised according to	179
P N used according to	178
P N judged according to	161
P N graded according to	158
P N ranked according to	155
P N allocated according to	155
P N determined according to	152
P N went according to	151
P N defined according to	147
P N conducted according to	147
P N adjusted according to	145
P N chosen according to	142
P N operate according to	140
P N treated according to	138
P N distributed according to	137

図2 ukWaCにおける動詞+according to

注：Sketch Engine による検索 [tag="V.*"] "according" "to"

- (9) In as much as service recovery performance is likely to be influenced by job demand stressors, it is logical to posit a direct positive effect of service recovery performance on organisational commitment. [CEAP: JoRaCS]

4. it can be said

他方、表3では大学院生による it can be said that の多用が見られる。専門家論文コーパス分野別の数値は、表5に示す。(10)の例に示す通り、it can be said は、結果に基づく主張の客観的まとめ、事実としての断定を避け

る downtoner としての機能を果たしている。この過剰使用の要因は、日本語の「～と言えよう」の英語への置き換えであろう。

表5 it can be said の CEAP と COLP における頻度

	順位	頻度	100万語換算
Arts & Humanities	—	86	1.4
Economics & Management	—	64	1.2
Management & Society	—	54	3.2
Natural Science	—	59	2.0
COLP	16	47	110.7

- (10) It is interesting to notice, however, that when this condition is referred to, the examples presented usually involve non-causative verbs. Therefore, it can be said that the distinction between causative verbs and non-causative verbs is not relevant. [CEAP: LS]

V ディスコース上の機能と4語連鎖

論文に必要な情報区分と関連づけて4語連鎖を分析する。例えば、論文で得た結果について自身の論文の限界を指摘しておく必要がある。以下にその表現を扱う。

1. it should be noted と it must be noted

表6は it should be noted の分野別と大学院生論文コーパスにおける頻度を示している。自然科学分野の成果では、当該実験や分析の限界に触れることが重要なため、この表現の使用頻度が他分野よりも多いものと推測できる。類似する表現に、it must be noted があるが、図3に示す通り、経済・経営分野における使用頻度は10倍以上の差で it should be noted の方が高い。

用例(11)(12)(13)と(14)(15)を比較すると、機能として it should be note that は「但書、ことわり書き」「論文の結果の限界の指摘」を持つことが多いのに対し、it must be noted that は「焦点化、注意喚起」のために使

表6 it should be noted の CEAP と COLP における頻度

	順位	頻度	100万語換算
Arts & Humanities	48	1230	20.0
Economics & Management	—	869	16.9
Management & Society	—	325	19.0
Natural Science	50	1137	43.0
COLP	—	10	23.6

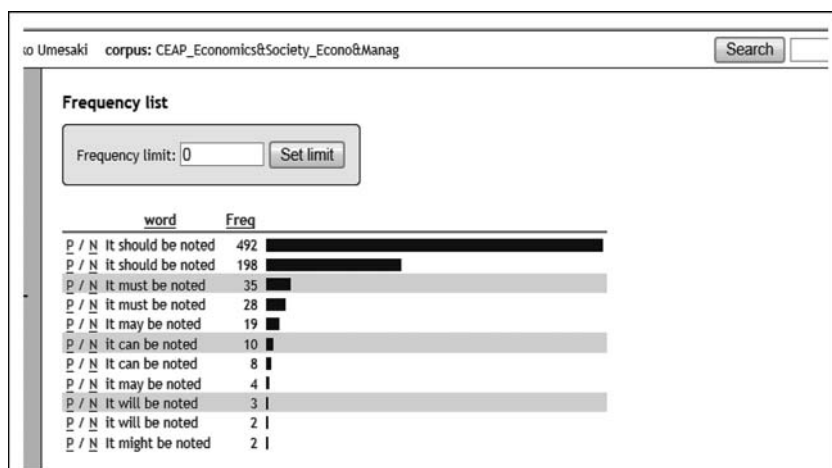


図3 経済・経営分野における it ~ be noted の頻度

注：Sketch Engine による検索 "it|It" [] |1,1| "be" "noted"

用されることが多い。つまり、前者は「但し～には注意を要する」、後者は「注目すべきは～」とまとめられる。

- (11) In addition, a detailed investigation of practitioners' attitudes towards multiple bankers, both with the same and different institutions, would yield further useful insights. Finally, it should be noted that the data was collected only in the UK. As such, the results and discussion are applicable particularly to the UK context, but may also provide insights into banking choice in other countries. However, as

cultural and product differences may be apparent, further research is recommended in alternative cultural contexts to establish the degree of generalisability of the current findings. In conclusion, this study presented a detailed analysis of choice criteria and their relative importance in the case of a main and secondary banking relationship. Results showed that,...

[CEAP: JoRaCS] => 結果の限界の指摘

- (12) This possibility suggests that there might be important differences in how multiple meanings of ambiguous words are activated depending on whether they are across or within languages. However, it should be noted that our dependent measure, naming errors, may not have been sensitive enough to pick up on the time-course differences with which the multiple meanings were activated. Future studies, using more sensitive dependent measures such as eye-tracking should address this issue further. [CEAP: JoMaL] => 論文の結果の限界の指摘

- (13) Methodology: Journals were allocated to subject categories by two independent researchers and differences (which occurred in about 20% of cases) were resolved by discussion. It should be noted that allocating journals to subject category on the basis of personal experience or title alone is obviously still a very subjective process. [CEAP: TBar] => 留意点の但書

- (14) They argue that project changes also is an important cause of underestimation, which lead to a conclusion that the estimates in their sample were rather reliable. The cost estimation process, which can be viewed as a factor more closely related to the project organization, has also received some attention among authors in this field of research. Flyvbjerg et al. deliver an interesting contribution on this issue. First of all, it must be noted that their study focuses on a certain category of projects, so-called transportation infrastructure projects, and secondly, the data comes from projects all around the world from a period of time stretching from 1910 to 1998. They study the differences between actual and estimated cost and argue that “the cost estimates used to decide whether such projects should be built are highly and systematically misleading.” This underestimation could, according to Flyvbjerg et al., best be explained by strategic misrepresentation. [CEAP: IJoPM] => 焦点化、注意を向けさせる

- (15) The generalizability of the results is also limited as a result of the small number of cases used in the research. However, it must be noted that the aim of the

research was not to provide generalizable results, but to explain the process of internationalization as experienced in a particular context with the purpose of theory building. Future research that tests the framework may place greater importance on generalizability. [CEAP: JoWB] => 焦点化、注意を向けさせる

この相違は、should と must の意味機能と関連し、should は一般的義務を、must は人間の内面からの義務感を表わすとされ、must のほうが義務感が強いと考えられていることと一致する。

VI 考察

自作専門家論文コーパス (CEAP) と大学院生論文コーパス (COLP) を主として使用して、論文における4語連鎖の量的質的分析を試みた。量的分析では、すでに梅咲 (2010b) で指摘済みの中等・高等学校の教科書における動詞句中心の4語連鎖とは異なり、学術論文における句前置詞を含む前置詞句の高頻度という特徴を経済・経営・社会科学・自然科学分野にも広げて提示し確認した。

英語学習の観点から、日本語母語話者の過剰使用として it can be said を取り上げ、過少使用として On the one hand～. と On the other hand～. の対比構文、および as a function of～ を取り上げた。これらの過剰・過少使用の要因は、基本的には母語の影響と考えるのが妥当であろう。すなわち、「～と言えよう」を英語に置き換えた it can be said のような語連鎖は日本語母語話者に使用頻度が高く、逆に as a function of は「～の機能・関数として」という日本語に置き換え易く、「～ごとに、～次第で、～に呼応して(変わる)」または「～の産物として、～に関連受けて(説明できる)」のような日本語には置き換え難い。日本語母語話者は、「関数として」の置き換えが可能な文脈でのみ as a function of を使用するため、過少使用に繋がると考えられよう。また、同形の語連鎖でも句前置詞や副詞句のように機能が異なる、すなわち別の語連鎖(パターン)として捉える必要のある場合もある。

言語分析の観点から、it should be noted と it must be noted の語連鎖の論

文における機能が但書と焦点化であることは、should と must の相違に起因する。すなわち、その起源は個々の語の結び付きと関連があると言える。同時に、言語運用能力を考える場合には、これらの4語連鎖が適格な状況で使用できるかどうかが重要であり、その力こそ能力であろう。すなわち、言語学的には、語連鎖はその塊そのものが、特定の意味機能をもつ形式と捉える必要があるのではなかろうか。

VII おわりに

大学英語教育に役立つであろう、専門家の英語論文1.5億語から分野別に高頻度4語連鎖を抽出した。英語論文においては、専門用語に加えて、個々の単語ではなく語の連鎖として表現を習得する必要があることを本調査を通して示した。論旨の展開と結びつけて表現を示すことが重要で、その抽出を今後も試みたい。さらに、日本語母語・英語非母語話者に過剰使用される表現については、英語母語話者は、その表現に変えてどのような英語表現を使用しているかを特定して行く必要があろう。

(筆者は関西学院大学商学部教授)

付記

本研究は科学研究費助成金(基盤研究(c) 課題番号2352077「コーパスを活用したフレジオロジーに基づく英語指導法の開発と検証」平成23~25年度、研究代表者梅咲敦子)の一部および関西学院大学2014年度個人特別研究費を使用させていただいた。

参考文献

- Biber, D., S. Johansson, G. Leech, S. Conrad and E. Finegan (1999) *Longman Grammar of Spoken and Written English*. Harlow: Pearson Education.
- Chomsky, N. (1957) *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.
- Hoey, M. (2005) *Lexical Priming*. London: Routledge.
- Hunston, S. and G. Francis (2000) *Pattern Grammar*. Amsterdam: John Benjamins.
- Swales, J. M. (1981) *Aspects of Article Introductions*. (Aston ESP Research Reports No. 1) Birmingham: Language Studies Unit, University of Aston.
- Swales, J. M. (1990) *Genre Analysis: English in Academic and Research Settings*. Cambridge: Cambridge University Press.

Swales, J. M. (2004) *Research Genres*. Cambridge: Cambridge University Press.

梅咲敦子 (2010a) 「英語論文作成にみるフレイジオロジー」(フレイジオロジー研究会第1回例会2010.3.7.ハンドアウト)

梅咲敦子 (2010b) 「教科書コーパスの意義—連語分析を例に—」『アジアの英語教科書コーパスの構築と相互比較—社会・文化的コンテキストと関連させて—』(平成19年度～21年度科学研究費補助金基盤研究B成果報告書(研究代表者中村純作)) pp. 49-58.

梅咲敦子 (2014) 「英語論文における n-gram とフレイジオロジー」(フレイジオロジー研究会第9回例会2014.3.1ハンドアウト)

八木克正 (2013) 「フレイジオロジーと実証性」『関西学院大学社会学部紀要』116, 45-60.

Wray, A (2002) *Formulaic Language and the Lexicon*. Cambridge: Cambridge University Press.