

# 歯科患者向け英文記事の語彙レベル —特徴語と連語表現に着目して—

吉岡 千由里

## Vocabulary Levels in Dental Articles for Patients: Keywords and Collocations

Chiyuri Yoshioka

### Abstract

This study examines the lexical characteristics of English for Specific Purposes (ESP) vocabulary in a corpus of dental articles for patients. Keywords were extracted from the corpus with AntConc, and vocabulary levels were analyzed with New Word Level Checker using the Scale of English Word Knowledge–Japanese word list. Overall, 61.97% of the keywords were at or below the university entrance exam level including basic dental terms, while 17.95% of the keywords appeared at the level over 10,000 words in the word list. Therefore, when acquiring dental ESP vocabulary, it is effective to focus not only on dental terms containing basic words but also on high-frequency dental terms that do not appear in general English word lists.

**Key words** : ESP, 語彙習得, 歯学用語, 特徴語, コロケーション

### 1. はじめに

本稿は、特定分野に特有の英語使用に焦点をあてる English for Specific Purposes (ESP) の観点から、歯学部生が専門入門期に習得すべき歯学 ESP 語彙の語彙の特徴を分析することを目的とする。具体的には、歯学専門雑誌に掲載された口腔の健康に関する患者向けの英文記事を例にとり、専門家が患者に向けて記述した英文記事で特徴的に使用される語を歯学 ESP 語彙として抽出して語彙レベルを分析し、ESP 語彙習得の観点から考察する。

専門領域では専門家集団がコミュニケーションを図る際に共通語的に使用される特徴的な英語使用がみられ、このような特定目的の英語は ESP と呼ばれる。これまでも ESP 教育の研究においては、それぞれの専門分野のなかで共有される語彙についての分析が積極的に行われ、分野特有の ESP 語彙を効果的に習得するための学習語彙リストが様々な分野で開発されている。さらに、個々の単語に留まらない「語句」の使用実態に着目し、専門分野における特徴的な英語表現を表す Multi-Word Expression (MWE) へと ESP 語彙分析の対象が広がっている (小山・水本, 2010)。

MWE は頻繁に出現する特定の「複数の語の塊り (cluster)」を総称する概念で、ある分野に特

微的な MWE を適切に使用することは、自然でかつ正確な英語表現につながる (小山, 2015)。また、単語の理解を自然な言語使用に近づけて文の理解へとつなげるには共起語 (collocation) の知識が重要であり、特定分野のテキストに頻出する共起パターンの学習は、学術テキストの理解を早める可能性が示唆されている (宮本ら, 2012)。

そこで本稿では、口腔の健康に関する患者向けの英文記事を例にとり、英文記事のなかで特徴的に使用される特徴語と、特徴語を含む複数の語からなる連語表現に着目する。患者向けの英文記事に頻出する特徴語を歯学 ESP 語彙として抽出したうえで、代表的な歯学 ESP 語彙がどのような語と共起しているかを確認し、連語の形で表現される歯学用語について考察する。さらに、特徴語として抽出された ESP 語彙の語彙レベルに着目し、専門入門期に習得すべき歯学 ESP 語彙はどのようなものであるかを探る。

## 2. 特徴語と共起パターン

特定分野テキストで高頻度に使用される ESP 語彙の選定には、石川 (2004) など多くの研究事例で特徴度 (keyness) に着目した特徴語分析の手法がとられている。特徴度とは、対象コーパスに限って特徴的に頻出する語の特徴性の度合いのことで、対象コーパスと参照コーパスにおける頻度差の顕著性の度合いと説明される (石川, 2008)。参照コーパスと比較して特徴度の高い語が、対象コーパスで特徴的に頻出する特徴語として抽出される。

特徴語の抽出にはいくつかの代表的な統計的指標が用いられる。統計的指標を有効に活用することで信頼性の高い特徴語を簡便かつ高精度に抽出することができ、なかでも、対数尤度比は比較的少ない語数の特徴語抽出が可能で、指標値上位ではかなり効果的に特徴語を判別できる (中條・内山, 2004)。

そこで、患者向けの英文記事に特徴的に使用される特徴語の抽出には、統計的指標として対数尤度比を用いることとする。特徴語の抽出作業には、コンコーダンスと呼ばれるコーパス分析用のソフトウェアを使用することで指定した統計的指標を用いた特徴度が自動で計算され、特徴語を抽出することができる。

また、特定分野で特徴的に用いられる MWE の中でも特に専門用語に着目した場合、有用な共起パターンは名詞句であるという指摘がある (中島, 2021)。歯学用語においても多くが名詞句である。そこで、特徴語を含む複数語のつながりで表現される歯学用語については名詞句に着目し、名詞+名詞の共起パターンと形容詞+名詞の共起パターンで表現される連語表現について抽出することとする。

## 3. 方法

分析の素材とするのは、American Dental Association が発行する *The Journal of the American Dental Association* に掲載された患者向けの英文記事である。乳歯の重要性、齲蝕の予防、歯の痛み、様々な症状とその治療法など口腔の健康に関するテーマが幅広く取り上げられている印象を受ける。このような患者向け英文記事のうち、2018年から2020年にかけて掲載された21本が1冊にま

とめられている (American Dental Association, 2020)。本稿ではこの21本の英文記事を分析素材とした。各英文記事のタイトル部から Conclusion 部までを分析の対象とし、引用リストや脚注は分析の対象外とした。また、分析に使用するコンコーダンスでの処理が適切になされるようにタイトルやサブタイトルなどの大文字の記載を小文字に修正した。以下、この分析対象データを「英文記事テキスト」と呼ぶ。

分析には AntConc (Anthony, 2024) を使用し、統計指標には対数尤度比を指定した。また、特徴語を抽出する際の参照コーパスには AntConc で提供されている American English 2006 (AmE06) を使用した。AmE06は複数のジャンルから集められた約100万語の書き言葉コーパスである。AmE06については Anthony (2024) を参照されたい。

まず AntConc の keyword 機能を使い、分析対象とした英文記事テキストに特徴的に現れる語を対数尤度比に基づいて検出した。さらに、検出された語のなかから固有名詞、代名詞、略称、単位などを削除して残った内容語を英文記事テキストの特徴語として抽出した。

続いて、抽出された特徴語について、語彙的傾向を確認するために語彙レベルを解析した。語彙レベルの解析には New Word Level Checker (NWLC) (Mizumoto, 2021) を使用し、語彙レベルの指標となるリストには大学生レベルの日本人英語学習者が知っている可能性の高い順に約 75,000語がリストされた Scale of English Word Knowledge - Japanese (SEWK-J) を指定した。なお、SEWK-J については Mizumoto (2021) および水本 (2022) を参照されたい。

さらに、抽出された特徴語のうち特徴度の高い上位語について、特徴語の直前または直後に出現する語を AntConc で検出し、コンコーダンスラインと文脈を参照して歯学用語として使用されていることが確認できたものを連語表現として抽出した。

#### 4. 特徴語と連語表現の抽出結果

分析の対象とした英文記事テキストについて AntConc で分析した結果、総語数は延べ語数 (Tokens) が10,505語、異なり語数 (Types) が1,623語となった。英文記事テキストに特徴的に現れる語として異なり語257語が検出された。そこから、固有名詞、代名詞、略称、単位などを削除した結果、234語の特徴語が抽出された。なお、語のレマ化は行われず、*teeth* と *tooth*、*gum* と *gums*、*brush* と *brushing* などのように語の表記の形が異なれば別の語としてカウントされている。抽出された特徴語について特徴度の高い順に上位30語を表1に示す。

表 1 特徴語の抽出結果

Rank	Type	Keyness (Likelihood)	Rank	Type	Keyness (Likelihood)
1	teeth	876.31	16	rays	181.46
2	dentist	863.28	17	jaw	177.53
3	tooth	631.42	18	antibiotics	175.37
4	dental	591.33	19	risk	173.94
5	fluoride	458.63	20	cancer	161.90
6	gum	336.13	21	brush	156.71
7	oral	308.59	22	disease	156.66
8	mouth	287.51	23	fillings	155.88
9	pain	261.87	24	cause	153.39
10	cavities	257.20	25	gums	147.89
11	decay	233.17	26	erosion	143.13
12	medications	223.84	27	diabetes	141.07
13	help	192.63	28	treatment	130.07
14	toothpaste	183.45	29	seal	127.87
15	sealants	183.39	30	cavity	127.05

AntConcによる検出結果から固有名詞等を除いた上位30語

次に、英文記事テキストから抽出された特徴語について、NWLCを使用して語彙レベルを分析した。特徴語の語彙レベルを表2に示す。なお、NWLCでは語をレマ化して計量する（水本, 2022）ため、異なり語数は196語となった。また、SEWK-JではL1からL6までの各レベルに500語、L7とL8の各レベルに1,000語、L9とL10の各レベルに2,500語がリストされ、L10（1万語レベル）を超える語はOver10にリストされている（Mizumoto, 2021）。

表 2 特徴語の語彙レベル

Word Level	Type	Frequency	%	Cumulative
PropN/Num	0	0	0.00%	0.00%
L1	44	59	25.21%	25.21%
L2	37	45	19.23%	44.44%
L3	23	29	12.39%	56.84%
L4	10	12	5.13%	61.97%
L5	16	18	7.69%	69.66%
L6	5	6	2.56%	72.22%
L7	5	6	2.56%	74.79%
L8	6	6	2.56%	77.35%
L9	1	1	0.43%	77.78%
L10	2	2	0.85%	78.63%
Over10	39	42	17.95%	96.58%
N/A	8	8	3.42%	100.00%
ALL	196	234	100.00%	100.00%

NWLCによる解析結果。Word LevelはSEWK-Jに基づく語彙レベル、Typeは異なり語数、Frequencyは総語数、%は総語数の出現割合、Cumulativeはその累積値。

さらに、特徴語がどのような語と共起して複数語から成る用語として用いられているかを調査するために、特徴度の上位10語に着目し、連語の形で意味のある歯学用語となる連語表現を抽出した。出現頻度が比較的高い連語表現が確認された4語 *tooth*, *dental*, *gum*, *oral* について、分析対象の英文記事テキストから抽出された連語表現を出現頻度の高い順に表3に示す。

表3 英文記事テキストに使用されている連語表現の例

特徴語	連語表現
tooth	tooth decay (28), tooth pain (8), tooth root(s) (4), tooth surface(s) (3), tooth appearance (1), tooth loss (1)
dental	dental erosion (15), dental x-ray (8), dental treatment(s) (6), dental sealants (5), dental product(s) (4), dental care (3), dental office (3), dental pain (3), dental visits (3), dental floss (2), dental filling(s) (2), dental amalgam (1), dental experts (1), dental health (1), dental hygienist (1), dental infections (1), dental professionals (1), dental radiographs (1), dental sensitivity (1)
gum	gum disease (27), gum line (4), gum problems (4)
oral	oral health (18), oral sores (5), oral care (4), oral hygiene (3), oral cancer (2), oral pain (2), oral side effects (2), oral examination (1), oral self-care (1), oral appliance (1), oral bacteria (1), oral development (1), oral diseases (1), oral rinse (1)

( ) 内の数字は英文記事テキストでの出現頻度

最後に、英文記事テキストの中に使用されている特徴語のうち語彙レベルが Over10となる語を抽出した。出現頻度の高い順に表4に示す。

表4 英文記事テキストから抽出された Over10レベルの特徴語

語彙レベル	特徴語
Over10	fluoride (50), sealants (20), erosion (17), enamel (15), denture(s) (13), acetaminophen (12), acidic (11), clench(ing) (9), dentin (9), fluoridated (9), opioid(s) (9), oropharyngeal (9), gingivitis (8), thrush (8), interdental (7), periodontitis (7), amalgam (6), cementum (6), resin (6), composites (5), ionomers (4), obstetricians (4), reflux (4), teething (4), chemotherapy (3), gynecologists (3), ibuprofen (3), moistened (3), radiographs (3), stiffness (3) trimester (3), vaccinated (3), vaccination (3), benzocaine (2), caries (2), connective (2), fluorosis (2), sippy (2), temporomandibular (2)

( ) 内の数字は英文記事テキストでの出現頻度

## 5. 考察

英文記事テキストから特徴語として234語が抽出された。そのほとんどは、表1に示す通り歯科に関連する語と言えそうである。語彙レベルを確認すると、表2に示す通りL3レベルまでの比率が比較的高く、大学入学共通テストレベルとされるL4レベルまでで特徴語全体の61.97%がカバーされる。このL4レベルまでを既習レベルと捉えると、残る38.03%が未習語と考えられるが、なかでも1万語を超えるOver10レベルのカバー率は17.95%でありL3レベルを抜いてL2レベルと同等の比率を占める。

これまでも指摘されているように、特定の分野に焦点を当てるESP教育では、日常語として使用される基本語が分野内では特有の意味を持つ専門用語として用いられていることに注意が必要である(Coxhead, 2017; スミス山下ら, 2015)。一方で、初級レベルと同等の割合を占めている1万語を超えるレベルにも意識を向けた語彙指導が必要であろう。つまり、既習レベルではあるが改めて専門用語としての学習が必要な語と、1万語を超えるレベルであるにも関わらず専門用語としては使用頻度の高い語の両面に注意する必要があることが見えてくる。

次に、英文記事テキストで歯学用語として用いられている連語表現に着目する。表3は、英文記事テキストのなかで出現頻度が比較的高い連語表現が確認された *tooth, dental, gum, oral* の例である。*tooth, dental, gum, oral* と共起して連語表現を作る語には、*pain, root, surface, treatment, care, health, line* など初級レベルの語が確認できる。表3にある共起語の語彙レベルを確認したところ、大学入学共通テストレベルとされるL4レベルまでの語が約75%を占める。一方で、*decay, erosion, sealants, amalgam, sensitivity* などL5レベル以上の語も確認できる。L5レベル以上を高校での学習機会がなかった未習レベルと捉えると、表3にある共起語の約25%が未習語と考えられる。

このような語彙レベルの傾向を示す患者向け英文記事を題材に歯学ESP語彙を習得するとした場合、専門入門期のESP教育ではどのような語彙の習得を目標にすべきであろうか。L4レベルまでの既習レベルと、未習レベルのなかでも語彙出現率の高いOver10レベルに着目して考察を続ける。

上述のように、表3に例示した共起語の約75%がL4までの既習レベルである。連語表現の理解には、高校までに習得した既習語の知識を援用できそうである。既習知識を援用しながら専門用語として改めて導入し、既習の基本語を使って表現できる患者向けの歯学用語としての習得を図ることがひとつの鍵となるであろう。また、分析対象とした英文記事テキストは患者に向けて書かれた英文記事であるため、一般の人が理解できるように日常的な語彙を意識して使用している面も窺える。一般向けに書かれた英文記事を学習教材として活用することは、基本的な既習語の歯学ESPでの用いられ方を知るとともに、患者に合わせた適切な歯学ESP語彙の用い方を習得することにも繋がる可能性が示唆される。

もう一方の課題は、未習レベルのなかでも出現率の高いOver10レベルを含めた未知語の習得であろう。英文記事テキストから抽出した特徴語に占めるOver10レベルの割合は約18%と未習語レベルでは突出している。では、Over10レベルの特徴語とはどのような歯学用語であろうか。表4に英文記事テキストから抽出されたOver10レベルの語を示した。Over10レベルとは、日本人大学生が知っている可能性が高い順に並べた語彙リストにおいて1万語を超えるレベルを意味する。即ち、表4に示した特徴語は日本人大学生が知っている可能性は非常に低いとされるレベルにある。

さらに、特定分野に特化しない一般的な学習語彙リストでは、ここに示した語の多くがリスト対象外となる。従って、これらの語は日本人大学生にとって未知語であるにもかかわらず一般的な語彙学習では習得する機会を得る可能性が低い語であると考えられる。つまり、意識的に導入する必要のある歯学 ESP 語彙であるといえよう。

また、表 4 には借用語（カタカナ語）として用いられる用語が少なからず見受けられる。薬学生を対象にした調査ではカタカナ語で理解している英単語は習得されやすいことが示唆されている（スミス山下ら, 2020）。英文記事テキストの例では, *sealants, enamel, denture, amalgam, resin, composite, caries* などの Over10 レベルの語が見られるが、日本語の用語としてはカタカナ語で用いられることの多いこれらの語の難度は歯学生にとってはそれほど高くはない可能性がある。また、連語の形で学習することで、使用実態に即した歯学表現として効率的に習得することも期待できる。

このように、専門入門期に習得目標とすべき歯学 ESP 語彙は、既習レベルの基本的な語を用いて表現可能な歯学用語と、大学生が知っている可能性が非常に低いとされるレベルにありながら使用頻度の高い歯学用語であると考えられる。

## 6. おわりに

患者向けの英文記事を例に、一般の人に向けて使用される歯学 ESP 語彙について、特徴語と連語表現に着目して分析した。分析対象とした英文記事は一般向けではあるが、1 万語レベルを超える語の出現率は低いとはいえない。歯科医向けの専門誌や歯学論文となればさらに高い語彙レベルが要求されることは想像に難くない。他方、患者や一般の人に向けて発信する際には、学術的な用語の使用を避け、一般の人が理解しやすい歯学 ESP 語彙を用いることが求められるであろう。

専門入門期には、基本的な語を用いて表現される歯学用語を習得するとともに、大学生が知っている可能性が非常に低いとされるレベルにありながら患者に対して頻繁に使用される歯学用語を習得することが望まれる。専門入門期に取り組む歯学 ESP の教材として、患者や一般の人に向けて書かれた英文記事を活用することには意義がありそうである。

本稿で対象とした英文記事テキストは約 1 万語にすぎない。歯学 ESP の使用実態を分析するためには、もっと多くの素材を広く対象にすることが必要と考える。また、将来の論文購読を意識した準備段階として習得すべき歯学 ESP 語彙についても調査分析することが必要であろう。今後の課題としたい。

【参考文献】

- American Dental Association (2020) *JADA+ Tips for Caring for Your Teeth and Mouth: For the Patient Pages*. Retrieved September 26, 2025, from <https://jada.ada.org/pb/assets/raw/Health%20Advance/journals/adaj/onlineissueversion2-1590154327693.pdf>
- Anthony, L. (2024) AntConc (Version 4.3.1) [Computer Software] Waseda University. Retrieved September 26, 2025, from <https://www.laurenceanthony.net/software/AntConc>
- 中條清美・内山将夫 (2004) 「統計的指標を利用した特徴語抽出に関する研究」『関東甲信越英語教育学会研究紀要』 18, 99-108.
- Coxhead, A. (2017) Specialized vocabulary: What the research shows. *ALIS Newsletter*. Retrieved June 14, 2023, from <http://newsmanager.commpartners.com/tesolalis/issues/2017-08-18/3.html>
- 石川慎一郎 (2004) 「司法英語 ESP 語彙表構築の試み：FROWN コーパスと米国司法文献コーパスの比較に基づく特徴語の抽出」『神戸大学国際コミュニケーションセンター論集』 1, 13-27. <https://doi.org/10.24546/00517984>
- 石川慎一郎 (2008) 『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
- 小山由紀江 (2015) 「コーパスと EAP/ESP 教育」 投野由紀夫 (編) 『コーパスと英語教育』 131-156. ひつじ書房
- 小山由紀江・水本篤 (2010) 「単語連鎖にみる科学技術分野と他分野の英語表現比較」『統計数理研究所共同研究レポート239 ESP コーパスからの特徴表現の抽出』 1-12. <http://hdl.handle.net/10112/12992>
- 宮本祥子・五百蔵高浩・宮本謙三・宅間豊・井上佳和・竹林秀晃・岡部孝生・滝本幸治 (2012) 「理学療法分野における英語表現 - 共起パターン - の特性」『理学療法学』 39(2), 90-101. <https://doi.org/10.15063/rigaku.KJ00008046837>
- Mizumoto, A. (2021) New Word Level Checker [Web application]. Retrieved September 26, 2025, from <https://nwlc.pythonanywhere.com/>
- 水本篤 (2022) 「New Word Level Checker の概要」『メソドロジー研究部会報告論集』 12, 1-24. <https://doi.org/10.69194/methodologysig.12>
- 中島和郎 (2021) 「コーパスデータからのソーシャルワーク ESP 連語表現抽出の試み」『立正大学社会福祉研究所年報』 23, 47-62. <http://hdl.handle.net/11266/00009901>
- スミス山下朋子・天ヶ瀬葉子・野口ジュディー (2015) 「薬学生を対象とした専門用語の理解度調査：種類別対応の必要性」『大阪薬科大学紀要』 9, 13-17.
- スミス山下朋子・天ヶ瀬葉子・山下直子・野口ジュディー (2020) 「意味と発音に焦点を当てた薬学生の専門用語の語彙習得について」『大阪薬科大学紀要』 14, 77-86.