

原著論文

健康医学情報を伝える日本語テキストのリーダビリティの改善とその評価： 一般市民向け疾病説明テキストの 読みやすさと内容理解のしやすさの改善実験

Improvement and Evaluation of Readability of Japanese Health Information Texts: An Experiment on the Ease of Reading and Understanding Written Texts on Disease

酒井由紀子
Yukiko SAKAI

Résumé

Purpose: The present study comprehensively examined ways of improving and evaluating the readability of Japanese health information texts.

Methods: The author assessed the improvement and evaluation of a text written by a medical doctor for patients and the general public regarding chronic suppurative otitis media. First, two different Japanese texts were prepared according to three differing readability factors (syntax, vocabulary, and text structure) and the differences between the three texts including the original were verified. Second, 91 college students were assigned one of the three texts and asked to answer a web survey. Readability was assessed using the time taken to read the text, and comprehension was assessed by a multiple-choice test. Comments on the readability of the text were also collected.

Results: Modified Text A with explanations of medical terms in parentheses was more understandable but more difficult to read. Modified Text B, which replaced medical terms with general vocabulary and consolidated the text structure, was more readable but less understandable. Experience of the disease affected the readability. The following conclusions were drawn: 1) syntactic factors contributed to the impression of ease of reading, but conflicted with other factors; 2) medical vocabulary should be replaced with general vocabulary or accompanied by explanations in phrases; 3) the effect of the text's structure on comprehension was not clear due to the problems of improvement and evaluation methods; 4) some human factors should be considered when making the evaluation; and 5) valid comprehension tests are challenging.

酒井由紀子：慶應義塾大学信濃町メディアセンター，160-8582 東京都新宿区信濃町 35

Yukiko Sakai: Shinanomachi Media Center, Keio University, 35 Shinanomachi Shinjuku Tokyo, 160-8582

e-mail: yukiko@lib.keio.ac.jp

受付日：2010年8月11日 改訂稿受付日：2010年12月23日 受理日：2011年4月19日

- I. 健康医学情報を伝えるテキストのリーダビリティ
- II. リーダビリティと関連研究
 - A. 本研究におけるリーダビリティの定義と関連する研究領域
 - B. リーダビリティおよびその改善と評価に関する先行研究
- III. 本研究の枠組み
 - A. リーダビリティに影響する要素、改善とその評価
 - B. 本研究の目的と方法
 - C. 本研究で用いるテキストの改善と評価方法
- IV. 一般市民向け疾病説明テキストの読みやすさと内容理解のしやすさの改善実験
 - A. 実験の実施手順
 - B. テキストの改善とテキスト分析
 - C. 内容理解テストの作成
 - D. Web テストの構築、実施と分析
 - E. Web テストの結果
- V. 一般市民向け疾病説明テキストのリーダビリティの改善とその評価
 - A. 文と段落の長さ
 - B. 医学・医療用語の扱い
 - C. テキスト構造とその改善
 - D. 読み手としての人の特性とその影響
 - E. リーダビリティの評価方法
- VI. 健康医学情報を伝える日本語テキストのリーダビリティの改善とその評価

I. 健康医学情報を伝えるテキストの リーダビリティ

本研究は図書館情報学の立場から、一般市民向けの健康医学情報を伝える日本語テキストのリーダビリティに着目して、科学的な根拠をもとに読みやすさと内容理解のしやすさを改善する手法を確立することを最終的な目的に据えている。図書館情報学の立場で健康医学情報のテキスト（以降、健康医学テキスト）を扱う理由は、一般市民向けに医学・医療に関する情報を提供する健康医学情報サービスが普及し、図書館もその担い手としてサービスを強化し、情報専門家としてその役割が拡大され、関連する問題に対する基盤研究が必要となっているからである。

対処すべき健康医学情報サービスの問題のひとつに、ヘルスリテラシーがある。ヘルスリテラシーとは、健康を保つために、あるいは健康を取り戻すために、必要な情報を得て使いこなすため

の能力を意味する¹⁾。健康医学情報サービスの先進国の米国では、人々のヘルスリテラシーのレベルと、実際に提供される情報の難しさとのギャップが健康政策指針である「Healthy People 2010」²⁾によって指摘され、2000年頃から注目されてきた。このギャップを埋めるためには、①人々のヘルスリテラシーを向上させるための働きかけに加え、②提供する情報を読みやすく、内容理解しやすくするよう適正化するための努力が必要とされている。ここに、どのようなテキストが読みやすく内容が理解しやすいものであるか、そしてどのように改善することが可能かを研究する必要性がある³⁾。

米国ではすでに、健康医学情報提供の担い手として、医学図書館協会（Medical Library Association）および国立医学図書館（National Library of Medicine）が主導し、医学図書館員などの情報専門家が、この①人への働きかけと②情報の適正化の二方向のアプローチから対応を始めて

いる。②の情報の適正化のための具体的な活動には、従来から行われていた役割として、様々なヘルスリテラシーレベルの利用者に合わせた健康医学コレクションの収集と提供がある。拡大された役割としては、医療専門家と協力した、読みやすく内容理解のしやすいテキストの作成や既存のテキストの改善と評価が含まれる⁴⁾ [p.262]。このテキストの作成や改善と評価に関連して、英語には教育学や言語学におけるリーダビリティや関連研究の理論的、実証的な成果の蓄積があり、そこから作成されたチェックリストやガイドラインなどが活用されている⁵⁾。しかし、読みやすく内容理解のしやすいテキストの実践的な提供を確実にするところまで検証が進んでいるわけではない。

日本でもヘルスリテラシーの問題に起因する健康格差が認識され、ヘルスリテラシーと提供される情報の難しさとのギャップの問題解決に向けた研究の必要性が叫ばれている。消費者健康情報学、健康情報学、ヘルスコミュニケーションといった諸領域で取り組みが見られるが、②の情報の適正化のアプローチは口頭でのコミュニケーションに関するものが多く、健康医学情報サービスで扱われるような書かれたテキストに関する研究は少ない。また、日本語を対象とした健康医学情報におけるリーダビリティや関連研究は、わずかに野呂らによる英語のガイドラインに沿って改善したインフォームド・コンセントのアウトカムに関する研究⁶⁾が見られるだけで、読みやすく内容理解のしやすいテキストの作成や改善、評価といった今後の情報専門家にとって必要となる役割に関する基盤研究の蓄積はない。図書館情報学でも Web 上の一般市民向けの健康医学情報の評価は関心の高い領域である。しかし、日本語を対象とした研究では内容の信頼性のみに重点がおかれ、読みやすさと内容理解のしやすさの観点は現在のところ欠けている⁷⁾。健康医学情報サービスという実践の場がある図書館情報学の研究課題として、リーダビリティに着目した日本語テキストの情報の適正化について検討がぜひとも必要である。

本稿では、実際に一般市民向けの日本語の疾病

説明テキストを健康医学テキストの事例としてとりあげ、改善とその評価の手続きを確認するための実験の結果を報告する。次章ではまず、リーダビリティの定義と健康医学情報を伝える日本語テキストのリーダビリティの改善と評価に役立つ先行研究についてレビューを行う。

II. リーダビリティと関連研究

A. 本研究におけるリーダビリティの定義と関連する研究領域

1. 用語の定義と焦点

本研究で着目する英語の Readability を本稿では「リーダビリティ」と呼び、「読みやすさ」と「内容理解のしやすさ」と定義する。文字フォントの種類や大きさ、レイアウトや図表の使用など、可視性の側面を含めて検討する場合もあるが、本研究では純粋なテキスト以外は扱わない。英語のリーダビリティ研究は、教育に科学的成果を取り入れる潮流の中、1920年代に始まった⁸⁾。日本では、1950年代に Readability を測定する公式として開発された英語のリーダビリティフォーミュラが紹介されたが、その際 Readability は「読みやすさと理解のしやすさ」と訳されていた⁹⁾。日本語を対象とした研究では「リーダビリティ」「可読性」「難易度」「わかりやすさ」など用語のバリエーションがあるが、いずれも研究で扱っているのは「読みやすさ」と「内容理解のしやすさ」の2側面が中心である。

2. レビューの対象とする研究領域

リーダビリティに影響を与える要素には多様なものがあり、研究初期の1935年にはすでに288種類が集められているが¹⁰⁾、リーダビリティを予測するためにテキストで測定可能な指標として、それらは①構文的要素、②語彙的要素、③テキスト構造の3種類に分類することができる。しかし、既存の健康医学テキストのリーダビリティに関する研究では、主に構文的要素と語彙的要素のみが集中してとりあげられていて、読みやすさや内容理解にかかわると言われているテキスト構造は検討されていない。

第1表 レビューで取り上げる研究とその位置づけ

1-1. 英語を対象とした研究

	健康医学テキストにかかわる リーダビリティ関連研究	リーダビリティ研究	その他の分野の研究
①構文的要素	1. リーダビリティフォーミュラによる 健康医学テキストの分析		
②語彙的要素	2. Consumer Health Vocabulary の開発と応用研究		
③テキスト構造		5. テキスト構造とリーダビリティ	

1-2. 日本語を対象とした研究

	健康医学テキストにかかわる リーダビリティ関連研究	リーダビリティ研究	その他の分野の研究
①構文的要素		3. 日本語リーダビリティ フォーミュラの基盤研究と 自動処理を用いた研究	
②語彙的要素	4. 日本語の一般市民向け医 学・医療分野の語彙研究		
③テキスト構造			5. テキスト構造とリーダビ リティ

注) 表中のアラビア数字は、第Ⅱ章B節中の項番号を表す

そこで、次節では、日本語の健康医学テキストのリーダビリティの改善のために検討すべき要素とその改善方法および評価方法を具体的に見極めるため、健康医学テキストを対象としたリーダビリティ研究を中心に、他の関連分野も含めてレビューを行う。第1表は、レビューでとりあげる五つの研究領域や主題について、英語を対象としたものと日本語を対象としたものとに分け、それぞれの研究の属性（健康医学テキストのリーダビリティ研究、リーダビリティ研究、その他の分野の研究のいずれか）と、扱われている要素の種類（①構文的要素、②語彙的要素、③テキスト構造のいずれか）を示したものである。レビューの最後には、読みやすさと内容理解の側面からの改善に対する人による評価方法についても、先行研究の中から取り上げ確認する。

B. リーダビリティおよびその改善と評価に関する先行研究

1. 英語のリーダビリティフォーミュラによる健康医学テキストの分析

リーダビリティフォーミュラはテキストの特性をもとに、客観的に読みやすさを測定するツールで、英語のフォーミュラは教育学の成果として1930年代から開発されていた。そのため、米国ではヘルスリテラシーが提唱される以前の1970年代から、フォーミュラを用いた健康医学テキストの分析が行われていた。2004年に医学研究所(Institute of Medicine)がまとめた最初の包括的なヘルスリテラシーの報告書によると、同種の研究がすでに300件を超えている¹¹⁾。そのほとんどは、テキストのリーダビリティが、米国民の平均的なヘルスリテラシーに対して難しすぎることを指摘している¹²⁾。

健康医学テキストに最もよく用いられるリーダビリティフォーミュラは、Flesch Reading Ease,

Flesch-Kincaid, Fog Index, SMOG (Simple Measure of Gobbledygook), Fry の 5 種類で、いずれも 1940 年代から 70 年代にかけて開発された。同時代のほかの伝統的フォーミュラと同様に、文の長さと言語の長さの 2 変数のみから成る¹³⁾。変数が 2 種類しかないのは、コンピュータ普及以前に開発され、手作業を想定してなるべく数える変数が少ない方が望ましかったためである。文の長さは構文の複雑さを、語の音節数は語彙の難しさを代替しており、言語や読解の発達に従ってより複雑な文や、難しい語彙を多く理解できるようになるという教育発達理論に依拠している¹⁴⁾。また、教育用の学年別標準読解テスト集の分析から開発されたため、測定結果は学年レベルで示される。したがって、数字が大きいほど、高学年の教科書と特徴が似ていて難しいことを表す。現在では、コンピュータによる測定が可能である。1974 年以降、多様な変数を追加した複雑なリーダビリティフォーミュラも考案されているが、伝統的リーダビリティフォーミュラを用いた測定結果と、いずれも相関が高く差がなかったとされている¹⁴⁾。

リーダビリティフォーミュラを用いる長所は、既存ツールによって簡易にかつ客観的に評価ができる点と、人のヘルスリテラシー能力との比較が学年レベルで容易にできる点である。しかし、批判もある。その一つめは、教育分野で開発されたツールをそのまま適用し、人を使った実験で有用性を検証していないという指摘である。二つめは、構文的な複雑さのみを取り上げていて、リーダビリティに影響する他の重要な要素を排除しているのではないかという点である。これらの批判に対して、人に対する内容理解テストを実施し、構文的要素による伝統的なフォーミュラでも一定程度の有効性は認められることを検証した研究が 2004 年に発表されている¹⁵⁾。

2. Consumer Health Vocabulary の開発と応用研究

伝統的なフォーミュラが表面的な構文的要素のみを指標としているという批判に対し、健康医学テキストで用いられる特徴的な語彙、つまり医学・医療用語に着目した研究が 2005 年以降集中

して行われている。背景には Web 上の健康医学情報サイトが普及し、一般市民が用いる語彙の収集が容易になったことがある。特に米国国立医学図書館などの助成金を受け、Zeng が中心となって進めている Consumer Health Vocabulary (以下、CHV)¹⁶⁾ というプロジェクトでは、すでに一般市民の用いる健康医学語彙集を開発してオープンソースとして公開し、語彙集はリーダビリティの測定のほか、言い換え提案機能を組み込んだ情報検索システムなどに応用されている。CHV の構築にあたっては、米国国立医学図書館の MedlinePlus¹⁷⁾ などに一般市民が検索のために入力した用語が集められ、様々な健康医学サイトのテキストから抽出された医学・医療用語とともに、医学文献データベース MEDLINE の抄録から収集された専門用語や、報道機関のサイトの医学記事ページから収集された中間的な用語との比較分析などが、コーパスを構築して実施されている^{18), 19)}。さらに、専門家の用語集である医学・医療用語シソーラスを複合的に集積し、用語の意味タイプや用語同士の関係など豊富な語彙情報を付与した Unified Medical Language System (以下、UMLS)²⁰⁾ とのマッピングも行われ²¹⁾、専門用語から一般市民が理解できる用語への置き換えに応用が可能になっている。

3. 日本語リーダビリティフォーミュラの基盤研究と自動処理を用いた研究

日本にも 2000 年頃にヘルスリテラシーの概念が紹介されたが、健康医学テキストのリーダビリティをとりあげた実証的研究は見当たらない。これは、日本語には確立したリーダビリティフォーミュラが普及していなかったためと思われる。しかし、1950～1960 年代には英語のフォーミュラに触発され、日本語のフォーミュラの開発を目指し、リーダビリティに影響を与える要素の抽出を追求した、教育学分野等の一連の研究が見られる。森岡と庵造は、文の長さと言語の含有率を二大要素として抽出している^{22), 23)}。また、阪本は文の長さに加え語彙の難しさを測定するための難易度別の語彙リスト『教育基本語彙』²⁴⁾ を作成

し、語彙比重法を提唱している。堀川は、文の長短、構文が簡易か複雑か、漢字の多少で組み合わせを作り、わかりやすい組み合わせを提示した²⁵⁾。その他、北尾による漢字交じり文の方がひらがな文より読みやすいことを確認した実験研究がある²⁶⁾。

日本語のリーダビリティ研究はその後盛んにはならなかったが、1980年代以降、情報技術の発達に伴って、機械処理を用いた研究が自然言語処理や日本語教育などの分野で再びなされるようになった。現在では、その成果としてWeb上に測定ツールがいくつか公開されている。佐藤らが開発し2008年5月に公開した『日本語テキストの難易度を測る』²⁷⁾は文字の出現確率と教科書コーパスとの類似度から難易度を推定し、学年レベルで表示するツールである。柴崎らが開発した『日本語リーダビリティ測定 Version 0.5.0-UD』²⁸⁾は、ひらがなの割合、1文の平均述語数、文字数、文節数を変数としたリーダビリティフォーミュラから、学年レベルに準じたリーダビリティスコアを算出する。このフォーミュラも、教科書の読解教材で作成したコーパス分析に基づいて構築されている。『チュウ太の道具箱』²⁹⁾では、日本語学習者のための読解教材の難しさを、複数の要素について別々に判定するツール群が公開されている。2010年4月現在公開されているのは「語彙チェッカー」と「漢字チェッカー」の2種類で、いずれも日本語能力試験の級別の出題基準に準拠したツールである。

リーダビリティの改善に関連する工学的領域には、2000年以降盛んになった言い換え研究がある。応用範囲は機械翻訳や文書要約まで幅広いが、その研究課題の一つに「テキスト簡単化」がある。この領域では、語、文のほかに文章のレベルの言い換えも視野に入れた研究がすすめられているが、これまでの多くの成果は構文的・語彙的言い換えに関するものである。言い換え研究の方向性には、“大規模コーパスを用いた統計的手法の高度化を追求する方向と、意味の問題に踏み込んで古くて新しい問題に再挑戦する方向がある”³⁰⁾。テキスト簡単化に評価を取り入れた研究

には、聾者の読解支援のための「文可読性基準のモデル化」³¹⁾と博物館などの解説文の「リライト支援」³²⁾の研究がある。評価者は読み手ではなく、前者は聾学校の教諭、後者は子供向けの解説文の書き手である。

4. 日本語の一般市民向け医学・医療分野の語彙研究

日本語の一般市民にわかりやすい医学・医療用語に関する研究は、端緒についたばかりである。特定トピックを対象とした研究に、久保による副作用医薬品情報の研究がある。その成果として、副作用用語を自覚症状に置き換えた患者用語集が作成され公開されている³³⁾。国立国語研究所の「病院の言葉」プロジェクトは、医療現場の口頭コミュニケーションで使われる用語を広く集め、わかりやすくする工夫を類型別にまとめた研究である³⁴⁾。2007年10月から2009年3月にかけて実施された一連の調査研究に基づいているが、最終報告書には代表的な57語のみが掲載されている。専門家と一般市民の用いる医学・医療用語の研究用コーパスもそれぞれ作成されているが、マッピングが行われているかは不明で、公開もされていない。

5. テキスト構造とリーダビリティ

テキスト構造は、内容理解に影響を与える要素の一つとされ、リーダビリティ研究では検討されてきたものの、英語での健康医学テキストを対象とした研究でもとりあげられていない。そこで、関連する分野の諸領域の成果を言語を問わず概観する。なお、研究分野により用語が異なるが、本稿ではまとまった文章の、文や語を越える単位を包含した構成を示すのに「テキスト構造」という用語を用いて代表させる。

まず、図書館情報学でも取り組まれているディスコースアナリシスとジャンル分析³⁵⁾のアプローチについて確認する。ディスコースは日本語では「談話」と訳され、“いくつかの文が連続し、まとまりのある内容をもった言語表現のことで、話されたもの、書かれたもの両者を含む”³⁶⁾。

ディスコースアナリシスでは、テキストの構造を明らかにするために、結束性 (Cohesion)³⁷⁾ や一貫性 (Coherence)³⁸⁾ といった概念を、文と文の意味のつながりをとらえるために用いている。前者が文法や語彙に基づく表面的な連鎖を指すのに対し、後者はテキスト内では明示されない概念的な連鎖に基づいているとされる³⁹⁾。これらの概念は、まとまった文章の単位のテキストの構造を明らかにする、テキスト言語学などの分野でも用いられている。

ジャンル分析は、ディスコースアナリシスを用いた一研究領域で、特定コミュニティや状況に存在するテキスト構造を含む“典型的なディスコースの形式”⁴⁰⁾ を把握するものである。この特定ジャンルに見られるテキスト構造に関する知識は、理解の促進につながる事が検証されている^{41), 42), 43)}。その理由は、記憶作業領域の限られた認知プロセスを助けているためと説明されている⁴⁴⁾。本研究でとりあげる一般市民向けの健康医学テキストについても、特定の構造が、執筆する医師や医療の専門家と読者である患者を含む一般市民に共有されていれば、それに従って執筆することで理解が促進されるはずである。しかし、医学分野のディスコースアナリシスでは、医師・患者間の会話や患者の疾病談話など個別の患者を交えた話し言葉はよく取り上げられているが、書き言葉を対象として特定構造を明らかにした研究は、学術論文や構造化抄録など専門家同士のものに限定されている。一般市民向けのテキストは、まだ典型的なディスコースの形式が確立していないものと推測される。

英語のリーダビリティ研究でも、1970年代後半から、認知科学の視点を持ち込みテキスト構造に重点をおいた新たなリーダビリティの要素の提案が複数なされている。Kintschらは、伝統的なフォーミュラは表面的なテキストの特徴のみにとらわれていると批判し、やはり認知プロセスにおいて記憶作業領域が限られていることを理由に、負担が少なく処理しやすいテキスト構造にかかわる変数を考案し、リーダビリティの指標として提案している。しかし、それらの変数は、実際には

伝統的フォーミュラでとりあげたテキストの要素とも相関関係があることがわかっている^{43), 45)}。

読みやすく内容理解しやすい日本語のテキスト構造を最も具体的に示したものに、1981年出版の『理科系の作文技術』⁴⁶⁾ に代表される「科学論文の書き方」の解説書がある。表現は多様であるが、いずれの解説書においても、パラグラフを形成することとトピックセンテンスを置くこと、そして重点先行主義とすることを中心とした技術的要素が推奨されている。パラグラフは“内容的に連結された、いくつかの文の集まりで、全体として一つのトピックについてある一つのことを言うもの”、トピックセンテンスは“パラグラフで何を言おうとするのか糸口に概論的に述べた文のこと”と定義されている。これらの技術的要素の根拠はやはり認知科学の成果で、人間の脳の記憶作業領域が限られていることがここでも主張されている^{47), 48)}。ただし、修飾句や節が前置される日本語の特性として、トピックセンテンスをパラグラフの冒頭に置くのが難しいこと、論理が枝葉から幹にいたるわかりにくい展開になりがちなことも解説されている⁴⁶⁾。

国語学の分野でも、それまでの文の単位の文法をテキストに相当する文章の単位に拡大した文章論があり、日本語のテキスト構造の役割に言及がある。ディスコースアナリシスやテキスト言語学における「結束性」や「一貫性」は、文章論では内容的な文と文の連結を表す「連接関係」という用語で説明されている⁴⁹⁾。また、佐久間は1987年に、日本語の文章における形式的な段落と異なる、パラグラフに相当する「文段」という単位を提唱し、トピックセンテンスに相当する「中心文」とともにそれらの役割を実証データで検証している^{50), 51), 52)}。ここでも、日本語の「中心文」は、実態としてその位置や数が多様であることが指摘されている^{53), 54)}。

6. 読みやすさと内容理解の人による評価

英語には、リーダビリティの異なる健康医学テキストを対象として、テキストの特徴だけでなく人による検証で評価した研究が多く存在する。一

方、日本語には健康医学テキストを題材にはしていないが、一般的なリーダビリティ研究や言い換え研究の中で、人による検証が数は少ないが行われている。

これらの研究で用いられた評価方法には、まず、読みやすさの客観的な指標として、読みのスピード^{26), 55)}がある。主観的なものには、テキスト全体を対象とした研究で、読解力テストに似たクローズ法^{15), 26), 56)}と多肢選択問題²²⁾が内容理解の測定のために使われている。また、複数のテキストを比較して難易度を判定するものに、専門家によるもの⁵⁷⁾や中学生によるもの²⁵⁾がある。さらに、言い換え研究では、書き手による改善のための評価も行われている^{31), 32)}。

語彙の単位において、医学・医療用語を対象とした研究では、多肢選択問題^{34), 58)}、TOFHLAと呼ばれる人のヘルスリテラシーを測定するための英語やスペイン語のテスト^{59), 60), 61), 62)}と難易度判定^{15), 63)}が実施されている。

ほかに、テキストの内容理解を測定するには国語や英語の読解力テストや心理学で用いられる手法が応用可能である。しかし、客観性と経済性を兼ね備えた決定的なものは存在せず、目的に応じて最適な方法を適用するのが現実的であるとされている⁶⁴⁾。

III. 本研究の枠組み

本章では、第II章のレビューで概観したリーダビリティ研究およびその関連領域における成果から、一般市民向けの健康医学情報を伝える日本語テキストの改善と評価に役立つ成果をまとめ、本研究の枠組みとして目的と方法、検討すべき改善方法と評価方法について述べる。

A. リーダビリティに影響する要素、改善とその評価

先行研究の成果を、①構文的要素、②語彙的要素、③テキスト構造の別にリーダビリティに影響する具体的な要素と指標について整理し、続いて改善に対する評価についてまとめる。

構文的要素については、健康医学テキストに特

化した具体的な要素や指標の検討は行われていない。すでに英語には教育学分野で開発された伝統的なリーダビリティフォーミュラが存在したため、それを応用した研究の蓄積があることがわかった。構文的要素は日本語のリーダビリティ基礎研究や、自然言語処理分野で開発された測定ツールでも扱われているが、その指標として英語と同様に使われている最も簡素なものは文の長さである。日本語独自のものに文字種、すなわち漢字の割合があるが、ここでは意味としての難しさではなく、読めないに進まないとされる表面的な読みにくさを指している。自然言語処理の複雑な測定ツールでは、一文の述語数や文節数を測定するものもある。

語彙的要素については、一般的リーダビリティ研究において、その難しさの代替として英語では音節数、日本語では文字種が用いられる。個々の語彙を扱ったものでは、英語にも日本語にも教育語彙と特定コーパスとの照合があった。特定コーパスでは、特に英語のConsumer Health Vocabulary¹⁶⁾が専門家の用語と一般市民の用いる用語が含まれる語彙集として実用化され公開されている。日本語にはこれに相当する医学・医療用語の語彙集は現存しないため、評価や改善には個別の対応が必要となる。

テキスト構造については、英語のリーダビリティ研究や、日本語でも関連分野の諸領域で基礎的な研究があり、命題の密度を初めとした指標が各種検討されているが、具体的なテキストの改善方法についてまで検討した研究成果は見られない。

一方、リーダビリティ改善の評価に用いることのできる成果には、テキスト分析と人による評価がある。そのほとんどがテキストの要素を比較して改善を確認しており、人による評価の研究は圧倒的に少ない⁸⁾。テキストの分析では、構文的な要素であれば既存の測定ツールが適用できるが、医学・医療用語、テキスト構造について測定できる汎用的なツールは存在しない。また、数少ない人による評価研究でも、短文の難易度判定や医学・医療用語の意味を問うものが多く、ましま

たテキストの内容理解まで踏み込んだ成果は見当たらない。

B. 本研究の目的と方法

先行研究の成果からは日本語の健康医学テキストのリーダビリティに関して個々の要素の改善の手がかりが示されているが、まとまったテキストに対するテキスト構造も含めた複数要素の取り扱いや、具体的な改善手法は提示されていない。また、人による評価によって内容理解まで踏み込んだ研究はまだ少なく、評価方法も確立していない。すなわち、テキストのリーダビリティに影響するすべての要素を改善すれば、本当に読みやすく内容理解がしやすくなるのかは自明でなく、さらにそのためには具体的にどのような手続きと手法で改善し、どのような方法で評価をすれば良いのかも明らかににはなっていない。

本研究では、健康医学情報を伝える日本語テキストのリーダビリティの改善と評価の全体像を把握し、一連の流れの中で問題点を明らかにするために、実際に一般市民向けの日本語の疾病説明テキストの事例を取りあげ、リーダビリティにかか

わる要素を組み合わせで改善し、人の内容理解に踏み込んだ評価を行う実験を試みた。その結果から、包括的な改善方法と評価方法を記述的に確認し、併せて今後の研究課題を明らかにすることが本研究の目的である。

C. 本研究で用いるテキストの改善と評価方法

実験計画にあたり、具体的な改善方法と評価方法を、先行研究をふまえ以下のとおり検討し、選定した。

1. 改善方法

第Ⅱ章のレビューで扱われていた、リーダビリティに影響を与えるとされていた主な要素を①構文、②語彙、③テキスト構造の3種類に分け、先行研究の成果とともに第2表に整理した。実験では、これら3種類の観点それぞれから、文の長さや文字種の一部、医学・医療用語、パラグラフとトピックセンテンスと重点先行主義に焦点をあてて、日本語の健康医学テキストのリーダビリティの改善を試みることにした。これらの要素を選択した理由と具体的な改善方法は以下のとおり

第2表 先行研究と本研究で取り上げる要素と改善方法

要素の種類	先行研究	本研究	
	具体的な要素と指標	改善で取り上げる要素	改善方法
①構文	文の長さ 一文の述語数、文節数 文字種（日本語）	文の長さ 文字種	文の長さを短くし構文を単純にする 医学・医療用語に読み仮名を付記する
②語彙	音節数（英語） 文字種（日本語） 教育語彙との照合 特定コーパスとの照合	医学・医療用語	解説をカッコで補記する やさしい用語や表現に置き換える 複数の表現を統一する 不要な医学・医療用語を削除する
③テキスト構造	命題の密度 一命題あたりの異なる概念数 事象の連鎖を形成する構成要素 重点先行主義 概観から細部へ パラグラフとトピックセンテンス	パラグラフとトピックセンテンス 重点先行主義	トピックセンテンスを置いた パラグラフ形成 パラグラフ、文とも重点を先行する (先頭から内容の切れ目で段落変えをする) (段落の切れ目に空白行を入れる)

である。

①の構文的要素として文の長さや文字種を一部変更する。文の長さは英語でも日本語でも、構文的要素として複雑さを表す指標としてコンセンサスを得ている。また、複文を単文にかえることで、実際に簡易な改善ができる。文字種について、本研究では医学・医療用語に限定して、読み仮名を補記する手法を試すことにした。漢字の含有率は、日本語のリーダビリティ研究では多くとりあげられていたが、これは主に教育レベルに応じた漢字の難しさを問題としているものであった。そして、逆に説明文は漢字が多い方がわかりやすい²⁵⁾、漢字交じり文の方が読みやすい²⁶⁾という研究結果を考慮すると、単純に漢字を仮名表記に変える方法は必ずしも適切ではない。また、患者向けの説明テキストでは単独の漢字よりも、むしろ単語としての医学・医療用語の方が読みやすさの障害となることが予測される。なお、医学・医療用語の読み仮名付与は、意味を問題とする②の語彙で扱う医学・医療用語の改善とは異なる。

②の語彙としての医学・医療用語の改善には、解説の補記、置き換え、表現の統一、不要な用語の削除など、いくつかの手法を用いることとした。なお、日本語では専門用語と一般市民用語が結び付けられた網羅的な語彙集は現在のところ公開されていない。「病院の言葉」プロジェクト³⁴⁾で専門家と一般市民の用いる医学・医療用語の研究用コーパスもそれぞれ作成されているが、公開されていない。そこで、解説や置き換え作業では、個別に医学書等を参照することにした。

③のテキスト構造の改善として、前節で解説したトピックセンテンスを置いた、意味のまとまりのあるパラグラフをつくり、パラグラフ中や、文中で重点先行とした内容にするために、パラグラフ、文、語の順番の入れ替えを手作業で行うことにした。これらの操作を選択したのは、「論文の書き方」の具体的な解説書が存在し、手順を把握しやすかったためである。また、健康医学テキストの書き手である医師になじみのある医学論文でもとられている手法を採用することで、執筆にあ

たって同じ手法をとってもらえる可能性もある。

なお、テキスト構造の確認のためには評価も兼ねて内容分析を行い、段落ごとの内容を確認するラベル付けを実施した。また、このような改善・評価作業は労力がかかるため、テキストの先頭から内容が変わるところで段落を変え、段落の存在を際立たせるために空白行を挿入するという簡易な方法も考案した。

2. 評価方法

テキスト改善の評価にはテキスト分析と人による評価の両方を用いた。テキスト分析には、既存の日本語のリーダビリティ測定ツールと、前述の内容を表すラベル付けなどによるテキスト構造分析の2種類である。人による評価では、Webテストを実施した。読みやすさの客観的な指標として、テキストの読みの所要時間を適用することとした。これには後述のWebテストサイトのページの滞在時間を測定する機能を用いる。内容を理解したかどうかの確認のためには、テキストを読んだすぐ後に実施する多肢選択問題を用意した。多肢選択問題は適切な選択肢を作るのが難しいが、客観的かつ経済的に採点ができるためである。また、補足として読みにくい点・わかりにくい点の具体的な指摘を複数選択肢と自由記述で求めることとした。

IV. 一般市民向け疾病説明テキストの読みやすさと内容理解のしやすさの改善実験

A. 実験の実施手順

実験実施までの手順は以下のとおりである。

- 手順1: 医師の執筆した一般市民向けテキストの入手
- 手順2: 手作業による構文的要素と語彙的要素の改善
- 手順3: 手作業によるテキスト構造の分析と改善
- 手順4: 内容理解テストとしての多肢選択問題の作成
- 手順5: 別の医師による改善テキストと内容理解テストのチェックと手直し

手順6: (テキスト分析1) 既存の日本語リーダビリティ測定ツールによるテキストの構文的特徴の確認

手順7: (テキスト分析2) 内容分析によるテキスト構造の確認

手順8: Web テストの構築と実施

テキストの改善は、内容が変わらないように手順2と3で留意して手作業で実施し、手順5でオリジナルテキストの著者とは別の医師に、改善の趣旨を伝え医学的な説明として問題がないことを確認した。この医師にはまた、内容理解を確認する選択肢問題が医学的な見地から適切かのチェックもしてもらった。

B. テキストの改善とテキスト分析

1. テキストの改善

オリジナルテキストとして、医師が一般市民向けに執筆し Web 公開している「慢性化膿性中耳炎」⁶⁵⁾ を入手した。疾病に関する説明テキストを選択したのは、疾病に関する情報が、健康医学情報の中でも特に求められている主題⁶⁶⁾ とされているからである。同テキストは、段落の区切りははっきりしないが、空白の文字を段落とすると4段落、10の文、474文字から成る。以下に引用しているとおりの冒頭は解剖学的説明から始まる。

中耳は鼓室と乳突洞・乳突蜂巣からなる骨

の中の空洞で、耳管という管を介して鼻腔と交通しています。風邪や蓄膿症と呼ばれる副鼻腔炎などによって鼻腔から耳管に細菌が進入し、中耳で化膿性の炎症が生じます。

テキストの改善は第2表で提示した要素と改善方法に基づいて実施した。それぞれの要素について複数の方法が考えられたため、簡易な方法による改善テキストAと、労力を惜しまず改善する改善テキストBの2種類を用意して比較することとした。改善テキスト別の具体的な改善方法は第3表のとおりである。

改善テキストAでは、①構文的要素の改善として、複文から成る長い文章を二つに分け文の長さを短くした。また、医学・医療用語に読み仮名と、併せて②語彙の改善として医学・医療用語の解説を、“耳漏（じろう、耳だれのこと）”のようにかっこで補記した。この解説部分についても医師のチェックを受けている。③テキスト構造の改善のために、テキストを先頭から内容的に区切れるところで段落を変える改訂をした。その結果、改善テキストAはオリジナルテキストより2段落多い6段落となった。また、段落の切れ目には空白行も挿入した。ただし、パラグラフ、文、語の入れ替えは行わず、パラグラフとトピックセンテンスがあるか、重点先行となっているかなど、構造の精査はしていない。

第3表 テキストの改善方法

種類	改善テキスト A	改善テキスト B
①構文	文の長さを短くする	
	医学・医療用語へ読み仮名をかっこで補記する	
②語彙	医学・医療用語へ解説をかっこで補記する	医学・医療用語について やさしい用語や表現に置き換える 表現を統一する 不要な医学・医療用語を削除する
③テキスト構造	(先頭から内容の切れ目で段落変えをする)	トピックセンテンスを置いたパラグラフにする パラグラフ、文とも重点を先行する
	(段落の切れ目に空白行を入れる)	

改善テキスト B は、①文の長さは A と共通の方法で短くした。②語彙は医学・医療用語を一般的な語や表現に置き換え、表現の統一、不要な用語の削除を行い、加えて自然な文となるよう推敲した。表現を統一したのは「中耳で化膿性の炎症」と「中耳炎」などで、削除したのは、病気の一般的理解に必ずしも必要ではない「鼓室」「乳突洞」「乳突蜂巣」などの解剖学用語である。③テキスト構造の改善は、内容を精査しながらトピックセンテンスを置いたパラグラフを形成し、重点を先行する構成となるよう、パラグラフと文の入れ替え、さらに表現の変更も行った。これらの変更については、医学的な正確さを期すために医師と時間をかけて議論をしてテキストを整えた。最も大きく変更したのは、その病気がどのようなものであるかという基本的な情報を最初の段落に移動させたところである。これは、疾病の概要が、一般市民が一般的に一番知りたい情報であるといわれているためである⁶⁶⁾。文頭は「慢性化膿性中耳炎とは…」のように、題目を提示する主題主語を用いた表現に変更している。この冒頭の 2 段落は医師と調整した結果、以下の通り 60 文字を超す長い 1 文となった。構文的な観点からはリーダビリティが低くなってしまう恐れがあったが、内容の正確さを優先させ、このまま用いることにした。

慢性化膿性中耳炎とは、耳の奥にある中耳の骨の空洞にばい菌が住み着いて、炎症が慢性化し、鼓膜に穴があいた中耳炎のことです。

中耳炎は、風邪や蓄膿症と呼ばれる副鼻腔炎などによって、鼻の奥のばい菌が耳管という管を通じて中耳に入り、炎症がおこった状態です。

2. テキスト分析

a. リーダビリティ測定ツールによる構文的特徴の比較

オリジナルテキストと 2 種類の改善テキスト、あわせて 3 種類のテキストを、主に構文的要素の改善状況の評価のために、既存の日本語のリーダビリティ測定ツールで分析した (第 4 表)。『日本語テキストの難易度を測る』²⁷⁾、および『日本語リーダビリティ測定 Ver.0.5.0-UD』²⁸⁾ の判定結果の数値は学年レベルを意味し、数字が大きいほど難しいことを示している。『チュウ太の道具箱』²⁹⁾ の「語彙チェッカー」では、星が多いほど難しい語彙が多いことを示している。

3 種類のテキストの判定結果を比較すると、ツールによって差が出ていないものもあるが、おおむねオリジナルテキストより、改善テキスト A、および Bの方がやさしくなっていることを示

第 4 表 リーダビリティ測定ツールの判定結果と構文的特徴

	オリジナル テキスト	改善 テキスト A	改善 テキスト B
日本語テキストの難易度を測る ¹	9 学年 = 中学 3 年 ふつう	8 学年 = 中学 2 年 やさしめ	8 学年 = 中学 2 年 やさしめ
日本語リーダビリティ測定 Ver.0.5.0-UD ²	9.2 学年	7.44 学年	7.73 学年
チュウ太の道具箱 ³	★★★★★ 難しい	★★★★★ 難しい	★★★★ 少し難しい
一段落の長さ (文の数) ²	250	200	143
一文の長さ (文字数) ²	47.40	54.58	48.50
総文字数 ³	474	655	485
漢字の割合 ³	45.4%	35.7%	36.5%

¹ 日本語テキストの難易度を測る。 <http://kotoba.nuee.nagoya-u.ac.jp/sc/readability/>

² 日本語リーダビリティ測定 Ver.0.5.0-UD。 <http://readability.nagaokaut.ac.jp/readability>

³ チュウ太の道具箱。 <http://language.tiu.ac.jp/tools.html>

している。しかし改善テキスト A と B では、『日本語リーダビリティ測定 Ver.0.5.0-UD』²⁸⁾では改善テキスト A の方が、『チュウ太の道具箱』²⁹⁾では改善テキスト B の方がやさしいという判定が出ていて逆の結果を示している。

また、個別の構文的要素について測定した結果を見ると、一段落の長さはオリジナルテキスト、改善テキスト A, B の順で短くなっている。一文の長さや総文字数は、改善テキスト A ではオリジナルテキストより大きい数字となっている。また、改善テキスト B は、冒頭の 2 段落が長いいため、オリジナルテキストとほとんど変わらない。漢字の割合は、改善テキスト A, B と同程度で、オリジナルテキストより低くなっている。

b. 内容分析によるテキスト構造の比較

改善テキストがテキスト構造の観点から、オリジナルテキストより理解しやすいかを、テキストに書かれている内容をもとに、段落、文、語の三つのレベルで分析した。分析の観点は、トピックセンテンスを含む一つの内容から成るパラグラフになっているか、内容の重点が先行しているかである。第 1～3 図は各テキストの分析作業シートで、各段落に内容ラベルを付し、一段落に収められた内容、複数の種類の内容を表す文や段落同士の連節関係、内容の提示順、後に続く内容を表す主題主語の有無などについて点検した結果をコメントとして、「○良い点」「×問題となる点」「△どちらともいえない点」として書き出している。

オリジナルテキスト（第 1 図）では特に前半の 2 段落で、先行研究でもとりあげられていた、理解を難しくしていると思われる以下の 4 点の特徴が見られる。

- 1) 慢性化膿性中耳炎の定義が冒頭の段落にない
- 2) 接続関係がわかりにくい 2～3 文が一段落中に収められていて、第 1 文がトピックセンテンスになっていない。
- 3) 第 1, 2 段落の間の接続関係がわかりにくい。
- 4) 各段落の最後の文は、文の末尾の“～中耳で化膿性の炎症が生じます”“～が慢性

化膿性中耳炎です”まで読まないと内容がわからない。

1) と 4) の問題は、重点先行となっていない点にある。接続関係がわかりにくい原因は、2) においては段落中の文と文の関係にあり、意味のまとまりであるパラグラフが形成されていないこと、トピックセンテンスが前置されていないことに起因する。3) においては、段落間の関係を指摘していて、今回の改善手法の範囲外としたパラグラフ間の連節関係が明確でないことが問題となっている。実は、第 1 段落から第 2 段落の第 1 文までの三つの文は、「耳管」を中心に説明を進めている結束性が見られるが、段落のまとまりとは食い違いがある。その結果、第 2 段落は唐突に“特に小さなお子さんは、”から始まり、各段落にはそれぞれ二つの内容が含まれる結果になっている。第 3, 4 段落目は、それぞれ“症状は”“治療は”と題目を提示する主題主語を置いていて、前の段落とは主題の転換がはかられていることが明示されており、わかりやすい構造と言える。

改善テキスト A（第 2 図）では、長い段落を適宜分割した結果として、第 1 段落は 1 文となったため、一つの内容となり、指摘した文同士の接続関係やトピックセンテンスの問題は解消されたように見える。また、第 5, 6 段落も分割されて初期治療と根本治療が別段落になっている。しかし、第 2 段落には、「中耳炎の機序」と「小児が中耳炎になりやすい理由」の二つの内容が収まっている。また、やはり第 1～3 段落の間の接続関係はわかりにくい。さらに、文や段落の順序を変更していないため、1) と 4) の問題もそのまま残っている。

改善テキスト B（第 3 図）では、段落の順番も入れ替えを行ったため、慢性化膿性中耳炎の説明が最初の段落となっている。それぞれ一段落に一つずつの内容となり、各段落の冒頭はそれぞれ“慢性化膿性中耳炎とは”のようにトピックセンテンスにふさわしい題目を提示する主題主語による書き出しとなっているため、その文の後半を予想しながら読み進めることができる、内容を理解しやすい構造となっている。ただし、第 2 段落の

●オリジナルテキスト

慢性化膿性中耳炎

【第1段落：中耳の解剖学的説明 / 中耳炎の機序】
 中耳は鼓室と乳突洞・乳突蜂巣からなる骨の中の空洞で、耳管という管を介して鼻腔と交通しています。風邪や蓄膿症と呼ばれる副鼻腔炎などによって鼻腔から耳管に細菌が進入し、中耳で化膿性の炎症が生じます。

【第2段落：小児が中耳炎になりやすい理由 / 慢性化膿性中耳炎の機序と状態】
 特に小さなお子さんは、この耳管の働きが悪いため中耳炎になりやすいのです。急性中耳炎が抗生物質などの治療によって治癒すれば問題はないのですが、炎症が慢性化すると鼓膜に穴があいて、乳突洞・乳突蜂巣という骨の空洞に細菌が住み着いてしまいます。このような慢性化した状態が慢性化膿性中耳炎です。

【第3段落：症状】
 症状は難聴と耳漏で、耳鳴や耳閉塞感を伴うこともあります。炎症が内耳に波及すると難聴のなかでも感音難聴が進行し、めまいが生じるようになります。

【第4段落：治療】
 治療は抗生物質の内服や点耳によって細菌の除去と化膿性炎症を抑えることが第一ですが、根本的には手術で細菌が住み着いた骨の空洞を清掃して、さらに鼓膜を形成する必要があります（鼓膜形成術）。音を伝える耳小骨が障害されている場合には耳小骨を形成することになります（鼓室形成術）。感音難聴が軽度であれば手術によって聴力も改善します。

×慢性化膿性中耳炎の説明が最初の段落にない

×接続関係がわかりにくい2文が1段落に収められていて、第1文がトピックセンテンスになっていない。

×第2文は最後まで読まないと、「中耳炎」の説明だとわからない。

×第1段落との接続関係がわかりにくい。

×接続関係のわかりにくい2文が1段落に収められていて、第1文がトピックセンテンスになっていない。

×段落の最後まで読まないと、「慢性化膿性中耳炎」の説明だとわからない。

×「急性中耳炎」の説明から記述が続けられていて、「慢性化膿性中耳炎」の定義がわかりにくい。「炎症の慢性化」だけでなく、「鼓膜に穴があくこと」、「骨の空洞に最近が住み着くこと」が要件であることが把握しづらい

○1段落に一つの内容

○段落の冒頭に、「症状は」と主題主語があるトピックセンテンスが置かれている。

○1段落に一つの内容

○段落の冒頭に、「治療は」と主題主語があるトピックセンテンスが置かれている。

△初期治療と根本治療が一段落に収められているが、「第一」「根本的には」の語で区別はつく

第1図 オリジナルテキストのテキスト構造分析シート

“小児が中耳炎になりやすい理由”が唐突な点に関しては、恣意的に内容を変更することはできないため、前後の段落間の接続関係をわかりやすくすることはできなかった。

C. 内容理解テストの作成

1. 設問と選択肢の作成

テキストが読み手にとって内容を理解しやすいものになっているかをテストで確認するために、

多肢選択肢問題を5問作成した（第4図）。設問は、当該テキストに含まれていて、一般市民が病気にについて知りたい内容と言われている主要な情報として、「どんな病気なのか」「どうして起こるのか」「症状はどんなものか」「治療はどうするのか」を問うように作成している。設問と選択肢は医師のチェックを受け、医学的にも正しいことを確認した。

●改善テキスト A

慢性化膿性中耳炎（まんせい かもうせい ちゅうじえん）

【第 1 段落：中耳の解剖学的説明】
中耳は鼓室（こしつ）と乳突洞（にゅうとつどう）・乳突蜂巣（にゅうとつほうそう）からなる骨の中の空洞で、耳管（じかん）という管を介して鼻腔（びくう）と交通しています。

【第 2 段落：中耳炎の機序/小児が中耳炎になりやすい理由】
風邪や蓄膿症（ちくのうしょう）と呼ばれる副鼻腔炎（ふくびくうえん）などによって鼻腔から耳管に細菌が進入し、中耳で化膿性（かのうせい）の炎症が生じます。特に小さなお子さんは、この耳管の働きが悪いため中耳炎になりやすいのです。

【第 3 段落：慢性化膿性中耳炎の機序と状態】
急性中耳炎が抗生物質などの治療によって治癒すれば問題はありません。炎症が慢性化すると鼓膜（こまく）に穴があいて、乳突洞・乳突蜂巣という骨の空洞に細菌が住み着いてしまいます。このような慢性化した状態が慢性化膿性中耳炎です。

【第 4 段落：慢性化膿性中耳炎の症状】
症状は難聴と耳漏（じろう、耳だれのこと）で、耳鳴（じめい、耳鳴りのこと）や耳閉塞感（じへいそくかん、耳がつまった感じ）を伴うこともあります。炎症が内耳に波及すると難聴のなかでも感音難聴（かんおん なんちょう、内耳または神経の障害から音を感じられなくなること）が進行し、めまいが生じるようになります。

【第 5 段落：慢性化膿性中耳炎の初期治療】
治療は抗生物質の内服や点耳（てんじ、薬液を外耳から目薬のようににたらすこと）によって細菌の除去と化膿性炎症を抑えることが第一です。

【第 6 段落：慢性化膿性中耳炎の根本治療】
根本的には手術で細菌が住み着いた骨の空洞を清掃して、さらに鼓膜を形成する必要があります（鼓膜形成術）。音を伝える耳小骨（じしょうこつ）が障害されている場合には耳小骨を形成することになります（鼓室形成術）。感音難聴が軽度であれば手術によって聴力も改善します。

×慢性化膿性中耳炎の説明が最初の段落にない
○1 段落に一つの内容

×第 1 段落との接続関係がわかりにくい。
×接続関係がわかりにくい 2 文が 1 段落に収められていて、第 1 文がトピックセンテンスになっていない。

×第 2 段落との接続関係がわかりにくい。
×段落の最後まで読まないと、「慢性化膿性中耳炎」の説明だとわからない。
×「急性中耳炎」の説明から記述が続けられていて、「慢性化膿性中耳炎」の定義がわかりにくい。「炎症の慢性化」だけでなく、「鼓膜に穴があくこと」、「骨の空洞に最近が住み着くこと」が要件であることが把握しづらい

○1 段落に一つの内容
○段落の冒頭に、「症状は」と主題主語があるトピックセンテンスが置かれている。

○1 段落に一つの内容
○段落の冒頭に、「治療は」と主題主語があるトピックセンテンスが置かれている。

○1 段落に一つの内容
○段落の冒頭に、「根本的には」と主題主語があるトピックセンテンスが置かれている

第 2 図 改善テキスト A のテキスト構造分析シート

2. 出題の意図

内容理解テストでは、特にテキストの②語彙、③テキスト構造の違いによって、以下のような結果を想定して出題している。

設問 1: “慢性化膿性中耳炎とは～”を冒頭に明示して、中耳炎との違いを明確にし（③テキスト構造）、難解な解剖学用語を削除した

（②語彙）改善テキスト B が最も正答を選びやすい。

設問 2: 第 2 段落に“中耳炎は～”と明示し（③テキスト構造），“鼻腔”を、正解と同じ“鼻の奥”と置き換えた（②語彙）改善テキスト B が最も正答を選びやすい。

設問 3: “耳漏”を“耳だれ”，“耳閉塞感”を

<p>●改善テキスト B</p> <p>慢性化膿性中耳炎</p> <p>【第 1 段落：慢性化膿性中耳炎とは】 慢性化膿性中耳炎とは、耳の奥にある中耳の骨の空洞にばい菌が住み着いて、炎症が慢性化し、鼓膜に穴があいた中耳炎のことです。</p> <p>【第 2 段落：中耳炎の機序と状態】 中耳炎は、風邪や蓄膿症と呼ばれる副鼻腔炎などによって、鼻の奥のばい菌が耳管という管を通じて中耳に入り、炎症がおこった状態です。</p> <p>【第 3 段落：小児が中耳炎になりやすい理由】 特に小さなお子さんは、この耳管の働きが悪いため中耳炎になりやすいのです。</p> <p>【第 4 段落：中耳炎と慢性化】 中耳炎が抗生物質などの治療によって治癒すれば、慢性化にいたらず問題はありません。</p> <p>【第 5 段落：慢性化膿性中耳炎の症状】 慢性化膿性中耳炎の主な症状は難聴と耳だれで、耳鳴りや耳がふさがった感じを伴うこともあります。炎症がさらに奥の内耳にまで達すると、難聴がひどくなり、めまいが生じるようになります。</p> <p>【第 6 段落：慢性化膿性中耳炎の初期治療】 治療の第一は、抗生物質を薬として飲んだり、抗生物質の入った点耳液を目薬のように耳の中にたらしたりしてばい菌を取り除き、炎症を抑えることです。</p> <p>【第 7 段落：慢性化膿性中耳炎の根本治療】 根本的には、手術でばい菌が住み着いた骨の空洞をきれいにし、さらに鼓膜を作りなおす必要があります。また、音を伝える耳小骨に壊れている箇所がある場合には、耳小骨を作りなおす手術もします。難聴が軽ければ手術によって聴力もよくなります。</p>	<p>○慢性化膿性中耳炎の説明が最初の段落にある。 ○1 段落に一つの内容 ○段落の冒頭に、「慢性化膿性中耳炎とは」と主題主語がある</p> <p>○第 1 段落との接続関係は明確 ○1 段落に一つの内容 ○段落の冒頭に、「中耳炎とは」と主題主語がある</p> <p>×第 2 段落との接続関係がわかりにくい。 ○1 段落に一つの内容</p> <p>×第 3 段落との接続関係がわかりにくい。 ○1 段落に一つの内容</p> <p>○1 段落に一つの内容 △段落の前半に「症状は」と主題主語がある。</p> <p>○1 段落に一つの内容 ○段落の冒頭に、「治療の第一は」と主題主語がある</p> <p>○1 段落に一つの内容 ○段落の冒頭に、「根本的には」と主題主語があり、トピックセンテンスとなっている。</p>
--	---

第 3 図 改善テキスト B のテキスト構造分析シート

“耳がつまった感じ”と解説，または言い換えをした改善テキスト A，B で正答を選びやすい (②語彙)。

設問 4: 最初に目指すことと根本治療の違いは，オリジナルおよび改善テキスト A でも特定できる。しかし，抗生物質が薬で，点耳は耳に薬液をたらすこと，と言い換えた改善テキスト B が最も正答を選びやすい。改善テキスト A も“点耳”を“薬液を外耳から

目薬のようにたらすこと”と説明を加えているので，オリジナルテキストより正答を得やすい (②語彙)。

設問 5: 抗生物質が薬で，点耳が耳に薬液をたらすこととし，「鼓膜を形成する」を「作り直す」と言い換えた改善テキスト B は正答を選びやすい。点耳に説明を加えた改善テキスト A もオリジナルテキストより正答しやすい (②語彙)。

設問 1.慢性化膿性中耳炎とはどんな病気ですか？

- 1) 鼻から耳に通じる管ができて起こる病気
- 2) 子どもが耳管の働きを良くする治療をしないとなる病気
- 3) 耳の奥にばい菌がすみつき鼓膜に穴があいている病気 [*正解]
- 4) 中耳炎の薬を飲み忘れると必ずなる病気

設問 2.中耳炎はどうやってひき起こされますか？

- 1) 耳の外からばい菌が入って
- 2) 風邪をひいて耳の中に空気が入るようになって
- 3) 鼓膜の働きが悪くなって
- 4) 鼻の奥のばい菌が中耳に入って [*正解]

設問 3.慢性化膿性中耳炎の症状としてあげられていたものはどれですか？

- 1) 耳垢がつまる
- 2) 耳だれが出る [正解]
- 3) 耳に強い痛みを感じる
- 4) ふらついて転ぶ

設問 4.慢性化膿性中耳炎の治療で、最初に目指すことは何ですか？

- 1) 薬でばい菌をなくし炎症を抑えること [*正解]
- 2) 抗生物質を点滴すること
- 3) 耳垢を清掃してきれいにすること
- 4) 手術によって聴力をよくすること

設問 5.慢性化膿性中耳炎の治療として間違っている、もしくは行わないものはどれですか？

- 1) 薬を飲む
- 2) 耳に薬をさす
- 3) 鼓膜を作りなおす
- 4) 耳小骨で穴をふさぐ [*正解]

第4図 内容理解を問う多肢選択問題と正解

D. Web テストの構築、実施と分析

1. Web テストの構築

改善したテキストの方が実際に読みやすいか、内容が理解しやすいかを人によって検証するために、Web 調査サイト Qualtrics.com に以下の 6 ページから成るテストページを作成した。第 2 ページのテキストだけはテスト実施期間に割り当てテキストを変更するために入替を行ったので、それ以外のページでは共通の 3 種類のテストを実施したことになる。

- 1) フェースシート
- 2) テキスト（オリジナルテキスト、改善テキスト A、改善テキスト B のいずれか）
- 3) 内容理解を問う 5 つの選択肢設問

- 4) 「読みにくい点・わかりにくい点」について選択肢設問、および自由記述
- 5) 中耳炎の罹患経験と、中耳炎について見聞、または読んだ経験を問う設問
- 6) 謝礼の抽選に参加するために電子メールアドレスを任意で入力

読みやすさの測定は、第 2 ページのテキストを読む所要時間と、第 4 ページの読みにくい点の選択肢設問で、また、内容理解のしやすさについては第 3 ページの内容理解に関する選択肢設問で分析する。それらの指標ではとらえきれない詳細なコメントや具体的指摘は、第 4 ページの「読みにくい点・わかりにくい点」で収集して分析を補強する。

テキストの特徴以外に、リーダビリティに影響を与える可能性のある読み手である人の特性を分析するために、第1ページのフェースシートには年齢と性別の質問項目を設定し、第5ページに中耳炎の罹患経験と、中耳炎に関して見たり聞いたり、読んだりした経験を問う設問を設けた。第2ページには、割り当てられた一つのテキストを読む所要時間を計測するため、このページの滞在時間を計測する機能を設定した。ここでは、このテストの趣旨に関わらず、テキストを読み飛ばして次ページの内容理解テストに進んでしまうことも予想された。そこで、自然に読んでいる時間なるべく正確に計測できるよう、前のページに注意書きをして、チェックしないとテキストのページに進めない設計にした。注意書きとしては、まじめに取り組んでもらうために、患者やその家族が病気の情報を得たいと思って読んでいる状況を想定するように促す文章を掲げた。また、第3ページの内容の理解に関する設問では、テキストを何度も注意深く読めば正答できる可能性も高くなることから、前のテキストページには戻れない設計にした。第4ページでは、具体的に読みにくいまたはわかりにくい点を自由記述で指摘してもらうため、第2ページと同じテキストを、行番号を添えて再掲した。ここから第3ページの内容理解テストに戻って回答を訂正することもできない。第6ページの謝礼の抽選への参加はあくまでも任意であること、預かった電子メールアドレスは回答と別に処理され、抽選と当選連絡用にのみ使用されることを明示した。

2. Web テストの実施

テストは2009年11月24日～12月20日に大学生を対象に実施した。実験の参加者として最も望ましいのは、一般市民を代表する人である。たとえば、教育レベルでいえば日本では高校を卒業した人であるが、実験では技術的に制約があるため、参加募集が容易な大学生とした。便宜的サンプリングによって参加者募集を行ったが、参加者は私立大学一校の文系の4学部および理系の1学部を加えた計5学部の学部生に限定した。ただ

し、この学部には、医学部や関連領域の学部は含まれていない。これは募集を容易にするためと、医学を体系的に学んだ経験がない、一定程度の基礎的な読解力が備わった同じ年齢層の参加者をそろえ、人の能力の違いによるテスト結果への影響をできるだけ排除するためである。学年は1年～4年生である。参加募集では、協力を承諾してくれた教員に、それぞれの担当授業の履修学生にテストサイトへ導く中間ページのURLを記載した依頼文書や電子メールを配布してもらった。一部、紙の参加用紙を集める方式を選択した教員の授業では、氏名と電子メールアドレスを用紙に記入してもらい、調査担当者から参加希望者に直接電子メールで中間ページのURLを送付し依頼した。いずれの方法においても文書や電子メールの本文中で、実験の目的が何であるか、参加は任意で授業評価とは独立していること、実験結果から個人は特定できないこと、個人情報の保護を遵守することを説明している。

参加者は3種類のテキストのうちいずれかのテキスト1種類についてのみ、1回だけテストに回答するようにした。テキストの割り当ては、テキスト中間ページのリンク先を3種類のテストへ順次手動で切り替えることで行った。つまり、実験に参加した時期によって割り当てられたテキストが異なる。これは、利用したQualtrics.comに複数テストを同時にオープンしランダム化割り当てをする機能がなかったためである。実際には、テスト開始時はオリジナルテキストをオープンして、当該URLを中間ページのリンク先とした。目標としていた25名以上が集まったタイミングで改善テキストAに、次に改善テキストBの順にオープンするテストを切り替え、リンク先URLを変更した。改善テキストBによるテスト実施中にそれぞれ30件程度の回答が集まる見込みとなったため、改善テキストBの回答が30件近くになった時点でオリジナルテキスト、改善テキストAの順で再度切り替えを行い、回答件数がほぼ同じとなるよう調整した。最終的に、オリジナルテキスト、改善テキストA、改善テキストBについてそれぞれ30、32、29の合計91件の有

効回答を得ることができた。

実験終了後、謝礼の抽選を研究協力者の立会いのもとで行った。10名の当選者には入力された電子メールアドレスあてに当選通知をし、それぞれ希望の送付先に3,000円相当の図書カードを郵送した。

3. 回答データの集計と分析

回答データを必要に応じて加工し、単純集計とクロス集計とを試みた。クロス集計項目は、テキストの種類および参加者の特性としてフェースシートにある年齢、性別、そして以下の手続きで独自に追加した推定学問分野と、別途設問を設けた中耳炎についての罹患および見聞した経験の各項目である。ここで「推定学問分野」は、「人文社会」「学際領域」「科学技術」のいずれかとし、協力してくれた教員が文書やメールを配布したタイミングから、調査者が参加者の所属学部から推定した。

テキストを読む所要時間として設定した第2ページの滞在時間は、91件の回答のうち2件に欠損値があったため、これを除いた合計89件の計測値で分析を行った。なお、標準偏差に基づいて、各テキストに2件ずつの外れ値が識別された。しかし、外れ値となった原因は、参加者がWebテストを一時放置したものか、内容理解テストに備えて読みなおしをしたものと推測できるが、はっきりした原因は決められないため、除外はしていない。

内容理解テストに対しては、1問あたり20点、5問全問正解で100点とした総スコアと、各設問の正答率で分析を行った。一方、「読みにくい点・わかりにくい点」については、各選択肢が選ばれた総回数と、各選択肢を選択した参加者の割合をテキスト別に集計した。また全回答の94.5%にあたる86件に自由記述の回答があった。ひとりで複数の指摘をしているため、切片化を行い、合計199件について、選択肢設問で整理したカテゴリ別に具体例を中心に分析した。

なお、統計学的有意性の検定では、所要時間と理解度テストの総スコアについては分布の同

等性検定としてKruskal Wallis検定またはMann-WhitneyのU検定を、その他の回答については独立性検定としてカイ二乗検定またはFisher検定を適用している。検定結果については、以下特にことわりのない箇所では、有意水準5%で帰無仮説を棄却した場合($p < 0.05$)に有意とした。

E. Webテストの結果

参加者の特性、テキストを読む所要時間、内容理解テストのスコア、「読みにくい・わかりにくい点」の順で、集計と検定した結果を以下に述べる。分析対象は所要時間のみ欠損値2件を除く89件、その他は91件の回答である。

1. 参加者の特性

参加者の年齢、性別、推定学問分野についてテキスト別の構成割合は第5表のとおりである。年齢は大学学部生に限定したため、いずれのテキストもほとんどが「15～19歳」または「20～29歳」のいずれかに属している。1名のみ改善テキストAに40歳代の参加者がいるが、テキスト別の有意差はない($p = 0.537$)。性別については、オリジナルテキストと改善テキストAでは女性の割合がそれぞれ6割を超えているのに対し、改善テキストBでは約5割をきっているという特徴があるが、これも有意差はない($p = 0.379$)。唯一、推定所属学部の学問分野に1%水準で有意差($p = 0.000$)が見られ、テキスト別で構成が異なる。合計では学際領域が全体で56.0%と3分野では一番大きな割合を占める。特にオリジナルテキストで83.3%と高いが、逆に改善テキストBでは割合が34.5%と低く、かわって人文社会が62.1%を占めている。

読みやすさや内容理解に影響を与える可能性のある、自分および家族等の中耳炎の罹患経験と、中耳炎に関して見たり聞いたりあるいは読んだりした経験をテキスト別に第6表に示した。本人または家族等の罹患経験が、改善テキストBではほかの2種類のテキストよりやや割合が高い。また、見たり聞いたり、あるいは読んだりした経験は、オリジナルテキストがほかの2種類のテキ

第5表 テキスト別の参加者の年齢・性別・推定学問分野別構成

	オリジナル テキスト		改善 テキスト A		改善 テキスト B		合計	
年齢	n	%	n	%	n	%	n	%
14 歳未満	0	0.0%	0	0.0%	0	0.0%	0	0.0%
15 ～ 19 歳	6	20.0%	4	12.5%	7	24.1%	17	18.7%
20 ～ 29 歳	24	80.0%	27	84.4%	22	75.9%	73	80.2%
30 ～ 39 歳	0	0.0%	0	0.0%	0	0.0%	0	0.0%
40 歳以上	0	0.0%	1	3.1%	0	0.0%	1	1.1%
合計	30	100.0%	32	100.0%	29	100.0%	91	100.0%
性別	n	%	n	%	n	%	n	%
男性	12	40.0%	11	34.4%	15	51.7%	38	41.8%
女性	18	60.0%	21	65.6%	14	48.3%	53	58.2%
合計	30	100.0%	32	100.0%	29	100.0%	91	100.0%
推定分野	n	%	n	%	n	%	n	%
人文社会	5	16.7%	8	25.0%	18	62.1%	31	34.1%
学際領域	25	83.3%	16	50.0%	10	34.5%	51	56.0%
科学技術	0	0.0%	8	25.0%	1	3.4%	9	9.9%
合計	30	100.0%	32	100.0%	29	100.0%	91	100.0%

年齢 $p = 0.537$
 性別 $p = 0.379$
 推定分野 $p = 0.000^{**}$

* $p < 0.05$ ** $p < 0.01$

第6表 テキスト別中耳炎に関する経験

	オリジナル テキスト	改善 テキスト A	改善 テキスト B	p
自分がなったことがある	36.7%	40.6%	48.3%	0.656
家族や親しい人がなったことがある	33.3%	21.9%	37.9%	0.694
見たたり、聞いたり、読んだりしたことがある	46.7%	65.6%	65.5%	0.225

ストよりやや少ないが、いずれもテキスト間で有意な差はない（それぞれ $p = 0.656$, $p = 0.694$, $p = 0.225$ ）。

2. テキストを読む所要時間

欠損値 2 件を除く 89 件の回答に記録された第 2 ページの滞在時間を、各テキストを読む所要時間として集計した（第 7 表）。テキスト別所要時間は、改善テキスト A が平均 127.35 秒、中央値 101.49 秒と最も長く、次いでオリジナルテキスト（それぞれ 92.73 秒、76.86 秒）、改善テキスト B（73.57 秒、60.11 秒）の順で、テキスト間に有意差がある（ $p = 0.015$ ）。テキスト 2 種ずつの検定

では、改善テキスト A と B の間に有意差が確認された（ $p = 0.003$ ）。

テキスト以外に読みの所要時間に有意差があったクロス集計項目は、「本人の罹患経験」である。「経験なし」の平均 115.69 秒、中央値 87.39 秒は、「経験あり」の 72.33 秒および 56.43 秒に対し有意に長い（ $p = 0.004$ ）。各テキスト内で経験の有無に分け所要時間を比較すると、特に改善テキスト A および B では、「経験なし」は「経験あり」の約 2 倍かかっている（それぞれ平均の比較で 157.80 秒と 74.75 秒、98.57 秒と 46.77 秒）。また、経験のある者だけまたはない者だけで読みの所要時間を比較すると、テキスト間の有意差は

第7表 読みの所要時間(実測)

欠損値除く n = 89		オリジナル テキスト	改善 テキスト A	改善 テキスト B
	件数	30	30	29
	平均(秒)	92.73	127.35	73.57
	標準偏差	90.06	114.43	84.64
	中央値	76.86	101.49	60.11
	最小値	6.25	2.53	14.38
	最大値	459.53	588.91	487.82

テキスト間 $p = 0.015^*$

経験なし	経験あり
53	36
115.69	72.33
107.77	78.44
87.39	56.43
6.74	2.53
588.9	459.5

経験の有無 $p = 0.004^{**}$

経験なし n = 53		オリジナル テキスト	改善 テキスト A	改善 テキスト B
	件数	19	19	15
	平均(秒)	87.10	157.80	98.57
	標準偏差	64.55	129.10	111.90
	中央値	86.00	130.68	60.64
	最小値	7.25	6.74	36.82
	最大値	250.40	588.91	487.82

テキスト間 $p = 0.018^*$

経験あり n = 36		オリジナル テキスト	改善 テキスト A	改善 テキスト B
	件数	11	11	14
	平均(秒)	102.44	74.75	46.77
	標準偏差	125.94	56.30	21.24
	中央値	68.91	55.01	45.13
	最小値	6.25	2.53	14.38
	最大値	459.53	215.27	74.54

テキスト間 $p = 0.226$ * $p < 0.05$ ** $p < 0.01$

「経験なし」のグループ(n=53)にのみ見られ($p = 0.018$),「経験あり」のグループ(n=36)では有意差はない($p = 0.226$)。なお,その他の参加者の特性である年齢,性別,推定学問分野,中耳炎についての見聞経験の各クロス集計項目に対して,所要時間に有意差は認められなかった。

次に,各テキストの文字数が異なるため,3種類のテキストの平均538文字あたりの秒数を算出した(第8表)。オリジナルテキストと改善テキストAの差は,それぞれ平均で105.24秒と104.60秒,中央値で87.23秒と83.36秒と,小さくなる。改善テキストBは平均81.60秒,中央値66.68秒と比較的短い,有意差はない($p = 0.259$)。

第8表 読みの所要時間(538文字あたり)

	オリジナル テキスト	改善 テキスト A	改善 テキスト B
平均(秒)	105.24	104.60	81.60
標準偏差	102.22	93.99	93.89
中央値	87.23	83.36	66.68
最小値	7.09	2.08	15.95
最大値	521.57	483.72	541.13

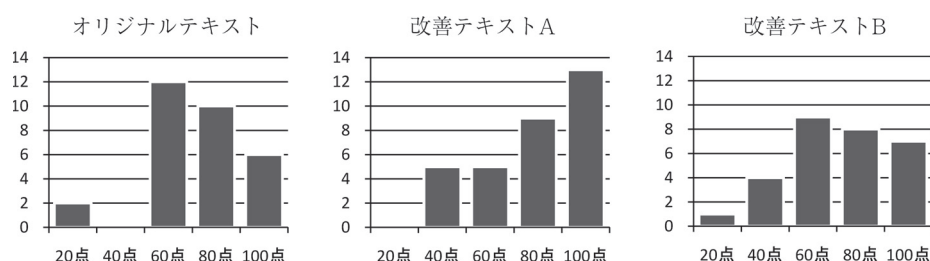
テキスト間 $p = 0.259$

3. 内容理解テストのスコア

第9表は内容理解テストの総スコアを集計したものである。テキスト別平均スコアは,改善テキストAが79.4点と最も高く,オリジナルテキス

第9表 内容理解テスト総スコア

	オリジナル テキスト	改善 テキスト A	改善 テキスト B
平均スコア	72.0	79.4	71.0
標準偏差	20.7	21.2	22.4
中央値	80.0	80.0	80.0
最小値	20.0	40.0	20.0
最大値	100.0	100.0	100.0

テキスト間 $p=0.196$ 

第5図 内容理解テスト総スコアの分布

トは72.0点、改善テキストBはこれに近い71.0点であるが、有意な差はない ($p=0.196$)。また、総スコアの分布をみると(第5図)、オリジナルテキストと改善テキストBは同じ60点が最頻値の山なりの分布となっているが、改善テキストBの方の山が低く、ばらつきが見られる。改善テキストAは100点満点が最も多く、右肩あがりの分布となっている。なお、テキスト以外のクロス集計項目には、内容理解テストの総スコアに有意差のある項目はなかった。

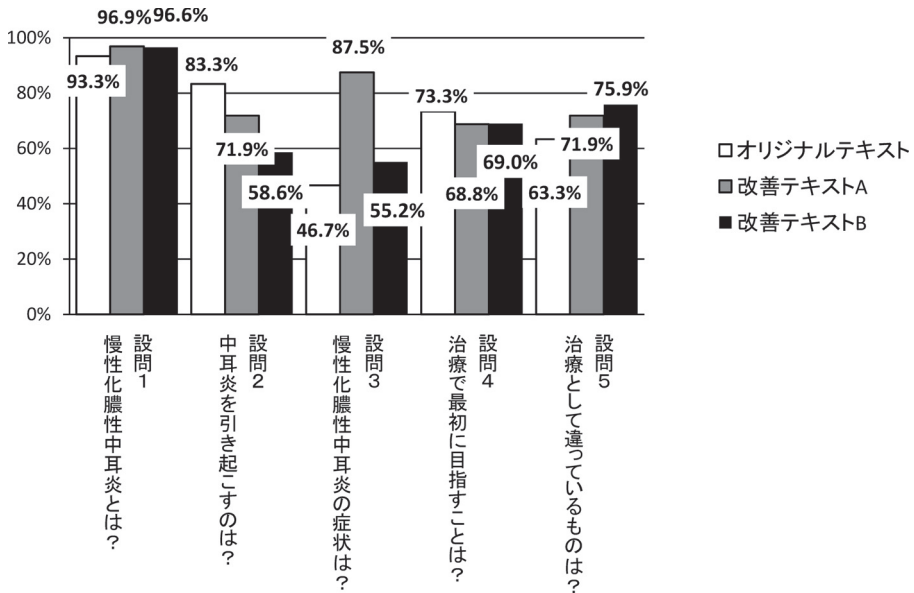
第6図は各設問のテキスト別の正答率である。設問によって正答率の高いテキストは異なり一定の傾向は見られないため、以下に各設問の結果を記述する。なお、記載のないクロス集計項目には、正答率に有意差はない。

「設問1: 慢性化膿性中耳炎とはどんな病気ですか?」はいずれも正答率は90%以上と高く、テキスト間で差はない。「本人の罹患経験」で正答率に有意差 ($p=0.028$) があるが、「経験がある」方が正答率が低く(89.5%)、ない方は100%の正答率である。「設問2: 中耳炎はどうやって引き起こされますか?」は、オリジナルテキスト

が最も正答率は高く83.3%で、改善テキストA(71.9%)、改善テキストB(58.6%)と順に低くなり、オリジナルテキストと改善テキストBの間に有意差がある ($p=0.035$)。「設問3: 慢性化膿性中耳炎の症状としてあげられていたものはどれですか?」は、オリジナルテキスト(46.7%)よりも改善テキストA、Bの正答率が高い。特に改善テキストAの正答率が87.5%と突出して高く、オリジナルテキストに対して有意差がある ($p=0.001$)。「設問4: 慢性化膿性中耳炎の治療で最初に目指すことは何ですか」では、オリジナルテキストが73.3%と若干正答率が高いが、改善テキストA(68.8%)やB(69.0%)と有意差はない。「設問5: 慢性化膿性中耳炎の治療として間違っている、もしくは行わないものはどれですか」では、オリジナルテキスト(63.3%)よりも改善テキストA(71.9%)、改善テキストB(75.9%)の正答率が高いが、有意差はない。

4. 「読みにくい点・わかりにくい点」

「読みにくい点・わかりにくい点」については、まず、各選択肢が選ばれた回数について比較



第6図 各設問のテキスト別正答率

する。選択肢設問のうち、具体的な読みにくい点・わかりにくい点13項目について、オリジナルテキスト、および改善テキストAではのべ141回選択されている。これに対し改善テキストBでは、のべ71回と約半数である。以下では、選択肢設問、自由記述の回答の順で、テキスト別とクロス集計項目別に比較する。なお、記載のないクロス集計項目については、「読みにくい点・わかりにくい点」の選択に関する有意差は認められなかった。

選択肢設問について、第10表にテキスト別、推定学問分野別、本人の罹患経験別、見聞経験別に、各項目を選択した参加者の割合を示した。表中の p 値はカイ二乗検定をした結果で、「*」は5%水準で、「**」は1%水準で有意であることを意味している。太枠と網がけは、有意差のある項目で割合が高いグループの数値である。

漢字についての「漢字が多い」「読めない漢字がある」の2項目は、オリジナルテキストについて改善テキストよりも指摘が多く、特に「漢字が多い」は73.3%とどの項目のどのグループよりも高い割合である。「文が長い」「一段落が長

い」という構文的な特徴については、テキスト間で有意差は見られない。「かっこの読みでつかえる」「かっこの説明書きでつかえる」は、改善テキストAが有意に高い。また、推定学問分野別で有意差が出ていて、科学技術分野の割合が高い($p=0.039$ と $p=0.044$)。

医学・医療用語に関する項目のうち「医学用語が多い」「医学用語がわかりにくい」の2項目は、改善テキストAがそれぞれ68.8%と59.4%、オリジナルテキストがそれぞれ56.7%と56.7%と、半数以上の参加者が選択している。これに対し改善テキストBは、24.1%と20.7%で、ほかの2テキストと有意差がある。「医学用語が読みにくい」は改善テキストAが68.8%、オリジナルテキストが43.3%、改善テキストBが20.7%と順に少なくなり、それぞれ二つのテキスト間にも有意差がある。「医学用語以外の表現が難しい」は、オリジナルテキストが13.3%で、改善テキストA、Bそれぞれ6.3%、6.9%の倍の割合の参加者が選択しているが、有意差はない。医学・医療用語に関する項目では、クロス集計項目で有意ではないが、中耳炎に関して見聞または読んだ経験の有無で差

第 10 表 「読みにくい点・わかりにくい点」を選択した者の割合と p 値

10-1 テキスト別および推定分野別

	テキスト別				推定分野			
	オリジナル テキスト	改善 テキスト A	改善 テキスト B	p	人文社会	学際領域	科学技術	p
件数	30	32	29		31	51	9	
漢字が多い	73.3%	37.5%	27.6%	0.001 **	38.7%	54.9%	22.2%	0.114
読めない漢字がある	40.0%	9.4%	0.0%	0.000 **	12.9%	19.6%	11.1%	0.657
文が長い	30.0%	25.0%	31.0%	0.854	25.8%	31.4%	22.2%	0.783
一段落が長い	10.0%	6.3%	3.4%	0.595	6.5%	7.8%	0.0%	0.682
かっこの読みでつかえる	10.0%	31.3%	3.4%	0.007 **	12.9%	11.8%	44.4%	0.039 *
かっこの説明書きでつかえる	6.7%	34.4%	0.0%	0.000 **	3.2%	17.6%	33.3%	0.044 *
医学用語が多い	56.7%	68.8%	24.1%	0.002 **	45.2%	51.0%	66.7%	0.522
医学用語が読みにくい	43.3%	68.8%	20.7%	0.001 **	35.5%	49.0%	55.6%	0.392
医学用語がわかりにくい	56.7%	59.4%	20.7%	0.004 **	32.3%	56.9%	33.3%	0.069
医学用語以外の表現が難しい	13.3%	6.3%	6.9%	0.560	9.7%	7.8%	11.1%	0.929
文章がかたい	56.7%	25.0%	31.0%	0.025 *	38.7%	39.2%	22.2%	0.612
耳のしくみがよくわからない	60.0%	53.1%	55.2%	0.856	58.1%	56.9%	44.4%	0.757
手術の方法がわからない	13.3%	15.6%	20.7%	0.739	16.1%	15.7%	22.2%	0.886

* $p < 0.05$ ** $p < 0.01$

10-2 本人の罹患経験・見聞経験別

	本人が中耳炎に罹患した経験			中耳炎に関して見聞した経験		
	なし	あり	p	なし	あり	p
件数	53	38		50	41	
漢字が多い	43.4%	50.0%	0.670	43.2%	48.1%	0.674
読めない漢字がある	13.2%	21.1%	0.394	16.2%	16.7%	1.000
文が長い	26.4%	31.6%	0.642	35.1%	24.1%	0.345
一段落が長い	5.7%	7.9%	0.691	10.8%	3.7%	0.219
かっこの読みでつかえる	17.0%	13.2%	0.771	13.5%	16.7%	0.774
かっこの説明書きでつかえる	17.0%	10.5%	0.546	10.8%	16.7%	0.549
医学用語が多い	47.2%	55.3%	0.526	51.4%	50.0%	1.000
医学用語が読みにくい	39.6%	52.6%	0.286	32.4%	53.7%	0.055
医学用語がわかりにくい	47.2%	44.7%	0.835	40.5%	50.0%	0.400
医学用語以外の表現が難しい	9.4%	7.9%	1.000	16.2%	3.7%	0.058
文章がかたい	37.7%	36.8%	1.000	43.2%	33.3%	0.382
耳のしくみがよくわからない	56.6%	55.3%	1.000	59.5%	53.7%	0.699
手術の方法がわからない	11.3%	23.7%	0.154	16.2%	16.7%	1.000

* $p < 0.05$ ** $p < 0.01$

が見られる。「医学用語以外の表現が難しい」は経験がない方が、ある者より多く選択している（16.2% と 3.7%, $p = 0.058$ ）。逆に「医学用語が読みにくい」は、見聞経験がある者が、ない者より

多く選択している（53.7% と 32.4%, $p = 0.055$ ）。

「文章がかたい」はオリジナルテキストで半数以上が選択していて、他のテキストよりも多く選択している。

「耳のしくみがわからない」はいずれのテキストでも半数以上が選択していて、テキスト間に有意差はない。「手術の方法がわからない」もテキスト間に有意差なく選択されているが、選択した者の割合はそれぞれ1～2割である。

自由記述の「読みにくい点・わかりにくい点」に関する回答は、選択肢設問の順にそって、具体例を記載する。

漢字について指摘した具体例は、オリジナルテキストと改善テキストAにみられる「乳突洞」「耳閉塞感」「感音難聴」といった用語で、医学用語の例としても取り上げられている。医学用語に関する例には、改善テキストBで置き換えをした「耳だれ」なども含め、すべてのテキストのあらゆる用語が例示されていた。

「耳のしくみがわからない」という指摘が、自由記述でも3種類のテキストのすべてで見られた。該当箇所は、オリジナルテキストと改善テキストAでは、解剖学的な説明をしている第1段落であり、改善テキストBでは置き換えをした「耳の奥の骨の空洞」である。耳の構造については「図がないとイメージしにくい」「具体的なイメージがわからない」という指摘もあった。解剖学的な記述以外にも、手術方法や「蓄膿症」など解説が補記されていない疾病名も例としてあげられている。

構文的な特徴に関しては、オリジナルテキストの治療についてふれた最終の第4段落、改善テキストBの根本治療についての第7段落が“一段落が長い”。また、改善テキストBの慢性化膿性中耳炎と中耳炎の定義をそれぞれ1文で記述した第1, 2段落について、“文が長い”と指摘されている。改善テキストAに対しては、自由記述でも“かっこの読みや説明でつかえ、かえって読みづらい”などの指摘が多数なされていた。

選択肢にはなかったテキスト構造についても、自由記述では指摘があった。オリジナルテキストに対しては、“文章の順番がおかしい。第2段落よりも、3, 4段落を先に持ってくるべき”として、病気の説明よりも、症状や治療といった具体的な診療に関する内容を優先するよう提案してい

る。“各段落を、たとえば慢性化膿性中耳炎とは、中耳炎とは、とすべき”といった意見も見られた。また、改善テキストAについては、“慢性化膿性中耳炎が何か直接説明している文にたどりつくのに時間がかかった”との指摘があった。しかし、テキスト構造を改善した改善テキストBについても、何人かの被験者が“慢性化膿性中耳炎と中耳炎の違いがわからない”と指摘している。

V. 一般市民向け疾病説明テキストのリーダビリティの改善とその評価

実験の結果、構文および語彙の要素を機械的に改善し、テキスト構造を簡易な方法で改善したテキストAは、オリジナルテキストと比較して読みの所要時間は長かったが、内容理解テストの平均スコアは高かった。一方、ほぼ同じ手法で構文を改善した上で、手間のかかる方法で語彙の難しさを改善し、加えて綿密にテキスト構造を改善したテキストBでは、オリジナルテキストより読みの所要時間は短縮されたものの、内容理解テストの平均スコアは変わらなかった。

以下では、実験結果を踏まえ、一般市民向け日本語の疾病説明テキストの「読みやすさ」と「内容理解のしやすさ」の具体的な改善方法の評価を改善項目別に試みる。続いて、読み手としての人の特性の影響についても検討し、「読みやすさ」と「内容理解のしやすさ」の観点からみたリーダビリティの評価方法についても考察する。

A. 文と段落の長さ

改善テキストA, Bとも同じ操作を施した、構文的特徴について検討する。文の長さが短くなるようにテキストを修正したものの、本実験ではこの改善の効果を確認できなかった。改善テキストA, Bとも他の改善を施した結果、文を短くすることができなかったためである。改善テキストAでは、読み仮名と医学・医療用語の意味をかっこで補記したため全体的に文字数が増えた。改善テキストBでは、疾病の定義に正確を期すためやむなく残した冒頭の2段落の長い文が残った。そのため、それぞれ平均54.58文字、48.50文字とオ

リジナルテキスト 47.40 文字より一文が長くなってしまった。参加者の「文が長い」という指摘もそれぞれ 30.0%, 25.0%, 31.0% と同程度で変わらない。複数の要素の改善ではそれぞれが対立する可能性があることが具体的な課題として明らかになった。

テキスト構造の改善作業の結果、構文的特徴として一段落の長さが短くなった。これは、構文的複雑性を減じ、読みやすい印象を与えたと考えられる。一段落の文の数は、オリジナル 2.50 に対し改善テキスト A では 2.00, B では 1.43 と実際に短くなっている。これに対応して参加者からの「一段落が長い」という指摘も、オリジナルテキスト 10.0% に対し、改善テキスト A で 6.3%, 改善テキスト B では 3.4% と有意ではないが少なくなっている。

ただし、読みの所要時間は、改善テキスト A では読み仮名と解説で文字数が多くなっていることと、かっこでつかえてしまうという別の要因で長くなっているため、段落が短くなった効用があったとは言えない。改善テキスト B は所要時間が短くなっているが、ほかの構文的要因、語彙的要因、テキスト構造の改善と切り離して評価する判断材料がない。また、構文的特徴の改善が内容理解に貢献することを想定した内容理解テストの設問は今回設定していない。

構文的改善は、その効用は明確ではないが、少なくとも読みやすい印象を与えていると考えられる。文の長さを短くする操作は比較的容易なことから、まず取り入れることのできる改善点の一つと言えるだろう。ただ、前述のとおり、対立する要素があったときにどちらを優先すべきか判断する結果を今回は得ていない。また、改善手法の課題としては、今回の実験では書き換え規則が明確ではなく、汎用性に欠けていた。今後、文章論の成果⁴⁹⁾などを踏まえて規則性を持たせる必要があるだろう。

B. 医学・医療用語の扱い

本実験では、読みやすさと内容理解の双方の障壁となると予測された医学・医療用語について、

改善テキスト A と B で異なる方法で改善を行った。それぞれの手法について読みやすさと内容の理解のしやすさの観点からその効果を考察し、両方の手法に共通の課題についてもふれる。

1. 読み仮名と解説の補記

改善テキスト A では構文的読みやすさの観点から医学・医療用語の漢字に読み仮名を、内容理解の観点から解説をかっこ内に補った。その結果、構文的特徴である文字種の割合に変化が生じ、漢字が多いという印象も減じている。実際にオリジナルテキストでは 45.4% を占めた漢字が 35.7% へと減少し、これに伴い読みにくい点・わかりにくい点として「漢字が多い」「読めない漢字がある」を選択した割合が有意に少なくなっている（それぞれ 73.3% 対 37.5%, 40.0% 対 9.4%）。漢字表記については、読めない文字があるとつかえる⁶⁷⁾という報告がある反面、漢字交じり文の方が読みやすいという研究結果もある²⁶⁾。また内容理解のためには既知の漢字が役立つ。したがって、漢字が多いことが必ずしも読みやすさや内容理解のしやすさを阻害するとは限らないが、上記の補記の結果、少なくとも心理的負担が減ったという効果があったと言える。「文章がかたい」という指摘が、オリジナルテキスト 56.7% に対して 25.0% と大幅に少なくなっているのも、漢字の割合が減ったことと、解説が付されたための印象と推測できる。

一方、医学・医療用語の読み仮名と解説をかっこ内に補記する手法は、明らかに円滑な読みを妨げてしまった。所要時間はオリジナルテキストの 92.73 秒の約 1.4 倍の 127.35 秒かかっている。これは単純に補記分の文字数が増えているためもあるが、平均文字数あたりの所要時間もオリジナルテキストと変わらず、情報量を無視したとしても読みやすいテキストになったとは言えない。「かっこの読みでつかえる」「かっこの説明書きでつかえる」という指摘が、他の二つのテキストより有意に多く、これが原因と特定できる。科学技術分野の参加者では、有意にこの 2 項目を選択するものが多かった。この理由は、同分野の参加者

9人のうち88.9%にあたる8人がかっこを多用した改善テキストAを読んだため、分野の違いではないと考えられる。また、用語自体は置き換えをしていないのに、“医学用語が多い”“医学用語が読みにくい”“医学用語がわかりにくい”という指摘がオリジナルテキストより多い。これも、かっこで医学・医療用語が際立ったためと、かっこで途切れるためであろう。

さらに、改善テキストAにも「最初の一文からとつきにくいと読みたくなる」という自由記述のコメントがあった。これは、冒頭の解剖学用語が並ぶ箇所を指していて、読みが付されていても医学・医療用語の存在自体が読みにくい印象を強めているものと推測される。改善テキストAで用いた手法は、一見読みやすい印象を与えているものの、実際の読みやすさにはつながっていないと言える。

内容理解についても、この手法による改善の効果は明確ではなかった。内容理解テストのうち設問3～5が医学・医療用語を対象とした改善によって正答率があがると実験前には予想していた。実際そのうち、設問3と5で改善テキストAの方がオリジナルテキストより正答率が高かった。特に、「設問3: 慢性化膿性中耳炎の症状としてあげられていたものはどれですか?」では、オリジナルテキストでは46.7%の正答率だったのに対し、改善テキストAの正答率が87.5%と突出して高い。しかし、これは「耳漏(じろう、耳だれのこと)」と正答の「耳だれ」が表記として印象に残り、選びやすかっただけで予測される。内容を理解して回答したかどうかは不明である。

「設問5: 慢性化膿性中耳炎の治療として間違っている、もしくは行わないものはどれですか?」では、正答率がオリジナルテキストで63.3%だったのに対し、改善テキストAは71.9%と有意差はないものの、予測どおりやや良くなっている。この設問に回答するには、「抗生物質の内服」「点耳」「鼓膜を形成する」の3種類の医学・医療用語および表現を理解して、それぞれに対応する「1)薬を飲む」「2)耳に薬をさす」「3)鼓膜を作りなおす」を誤った選択肢として排除し、

残りの「4)耳小骨で穴をふさぐ」を選択する必要がある。改善テキストAでは「点耳」に解説を補記しているが、それが正答の選択に直接影響したかどうかかわからない。

もう一問の「設問4: 慢性化膿性中耳炎の治療で、最初に目指すことは何ですか?」では、結果が予想と異なり、オリジナルテキストが73.3%の正答率で最も高かった。この設問の正答は「1)薬でばい菌をなくし炎症を抑えること」である。改善テキストAは68.8%とやや低かったが、その原因の一つには、オリジナルテキストでは“治療は抗生物質の内服や点耳によって細菌の除去と化膿性炎症を抑えることが第一ですが…”とあるが、かっこで補記された読みや意味のために文が長くなり、「最初に目指すこと」の特定がしにくかった可能性がある。もう一つには、炎症性疾患に関する基本的な知識があれば正答を選ぶことが容易な選択肢だったため、オリジナルテキストでも正答率が高かったと考えられる。

改善テキストAに適用した読み仮名と解説のかっこ内での補記は、ある程度機械的に作業できるという利点がある。漢字の割合が自動的に減って印象はやさしくなる効果があるが、実際に読みやすく内容理解しやすいテキストになっていなかった。かっこの問題に対して、今後、読み仮名の補記が必要な場合はかっこではなくルビを振る手法や、解説が必要な場合は脚注とするか、文中に自然な形で説明語句を補足するなどの手法も検討する余地があるだろう。

2. 置き換えと削除

改善テキストBでは自然な読みやすさを目指し、内容理解を促進する観点から、一般的な用語や表現に置き換える作業を行った。加えて、全体の説明に必ずしも必要ではない解剖学的用語を、改善テキストと選択肢のチェックを担当した医師の了解のもとに削除した。その結果、この手法においても漢字の割合が36.5%に減り、これに伴い「漢字が多い」という指摘も減り(27.6%)、「読めない漢字がある」を選択した回答はなかった。

改善テキストBは、読みの所要時間について

も有意差はないが、オリジナルテキストより短かった。また、「読みにくい点・わかりにくい点」の複数選択肢の選択数はオリジナルテキストと改善テキスト A の 1/2 に減っていた。ここから改善テキスト B はオリジナルテキストや改善テキスト A よりも「読みやすい」テキストになっていたと考えられる。しかし、前述のとおり、改善テキスト B は改善テキスト A には適用しなかったテキスト構造の改善も入念にしているため、医学・医療用語に対する改善方法とテキスト構造の改善方法のどちらが貢献したか、断定することはできない。

内容理解テストで改善テキスト B の方がオリジナルテキストより正答率が高かったのは、この手法においても設問 3 と 5 であるが、置き換えの手法と内容理解との関係は明確ではない。「設問 3: 慢性化膿性中耳炎の症状としてあげられていたものはどれですか?」について、46.7% だった正答率が 55.2% に上がっている。正解の“耳だれ”が“耳漏”に置き換わっていたためと推測できるが、“耳だれがどういう状態かわからない”という自由記述も見られ、改善テキスト B の改善がその症状を具体的に理解するのに貢献したという証拠は得られない。「設問 5: 慢性化膿性中耳炎の治療として間違っている、もしくは行わないものはどれですか?」では、有意差はないがオリジナルテキストの 63.3% より改善テキスト B の正答率が 75.9% と高かった。この設問に回答するために必要な「抗生物質の内服」「点耳」「鼓膜を形成する」の 3 種類の医学・医療用語や独特の表現は、それぞれ“抗生物質を薬として飲んだり”“抗生物質の入った点耳液を目薬のように耳の中にたらしたりして”“鼓膜を作りなおす”と一般的な用語や表現に置きかえている。このことが正答率の向上に貢献している可能性はある。

以上、二つの手法を医学・医療用語の改善に適用してきたが、少なくとも改善テキスト A の改善手法が読みやすさを損ね、改善テキスト B の方が読みやすかったことから、文章を寸断するかっこによる解説の補記は避けた方が良いと考えられる。しかし、文中に自然な形で説明語句を補

足する手法であれば、置き換えと同様の効果も期待できるかもしれない。ただし、説明の付与にしても、置き換えにしても、参照できる日本語の大規模な医学・医療用語の言い換えコーパスなどは存在しない。そのため、今回の実験においては、医学の専門教育を受けていない人間が作業にあたるため、個別に医師に確認する手順が必須であった。事例研究の域を超えて改善方法を実装するためには、これ以上の個別対応には限界がある。一般市民向けの日本語による医学・医療用語集の組織的な開発が期待される。

C. テキスト構造とその改善

テキスト構造の改善として、改善テキスト B に対してだけ、意味的単位と等しいパラグラフの形成、トピックセンテンスの前置、パラグラフ間およびパラグラフ内の重点主義を適用する入念な改善を実施した。改善テキスト A は簡易な「先頭から内容の切れ目で段落変えをする」という手法を適用したのみであり、これが大きな効果をもたらすとは予想していなかった。したがって、ここでは改善テキスト B についてのみ、考察する。

テキスト構造分析においては、オリジナルテキストや改善テキスト A で問題となっていたテキスト構造の問題点が、改善テキスト B ではほとんど改善されていることが確認された。しかし、恣意的に内容が変更される書き換えはできなかったため、前半の段落にパラグラフ間の接続関係がわかりにくいという問題が残っていた。

改善テキスト B で読みの所要時間の短縮が見られたことから、このテキスト構造の改善も読みやすさに貢献している可能性がある。テキスト構造の改善の単独の効果なのか、医学・医療用語の改善の効果なのか、あるいは両方の要因に交互作用が存在するのか、本実験では根拠となる結果を得ていない。しかしながら、自由記述で得た意見に、改善テキスト B の重点先行主義やトピックセンテンスの前置を支持するものが見られたことから、この改善手法が読みやすさおよび内容理解のしやすさとして受け入れられる可能性があることが示唆された。

一方、内容理解テストでは主にテキスト構造の改善によって正答率が向上することを意図して作成した設問1および2の結果が、予想と異なりオリジナルテキストと同等かむしろ低くなっていた。したがって、ここからテキスト構造の改善が内容の理解のしやすさに貢献したという明確な証拠はやはり得られない。

設問1および2のテストの正答率が予想より低かった原因の一つは、テキスト構造の改善を試みたことが、構文的なわかりやすさを損ねてしまったことにあって考えられる。「設問1: 慢性化膿性中耳炎とはどんな病気ですか?」の回答が書かれている箇所は、オリジナルテキスト、および改善テキストAでは第2段落にある。加えてトピックセンテンスにあたる“このような慢性化した状態が慢性化膿性中耳炎です”が第2段落の最後に位置している。改善テキストBでは同内容を第1段落に移し、“慢性化膿性中耳炎とは…”と冒頭で内容を明示する文で始める改善を行った。しかし、この1文は医学的な正確性を重視して内容を詰めこんだため、1文あたり平均48.5文字であるのに対し、60文字と長いものになってしまっている。同様に「設問2: 中耳炎はどうやって引き起こされますか?」の回答が書かれている、「中耳炎は…」で始まる第2段落も1文で63文字と長い。もう一つ考えられる要因は、今回のテキスト構造の改善では、トピックセンテンスを段落の冒頭に置き、重点先行主義でトップダウン式に理解を促進させた点である。しかし、先行研究には、学習レベルによって高い読解力をもつ人は、むしろボトムアップ式の方が、労力をかけた分、理解が促進されることがあるという結果も出ている^{42), 68)}。今回の参加者の所属大学から彼らの読解力は一定水準以上に高いことが予測され、トップダウン式のテキスト構造でなくとも、理解に差がでなかった可能性がある。

今回実施したテキスト構造の改善では、改善の効果は確認できなかったが、改善やテキスト分析の視点を学習することができた。また、テキスト構造の変更によって伝達内容が変わっていないかどうかを医師へ確認するのに時間と労力がかかる

こともわかった。より効率的な改善方法はないか、ほかの要素と切り離して読みやすさと内容理解のしやすさへの影響を測定することができるか、読み手としての人の特性である読解力の影響はあるか、検討すべき課題が多くある。

D. 読み手としての人の特性とその影響

本研究では参加者の年齢や教育程度をそろえる配慮をしてテキストの改善を目指したが、実際には読み手としての人の特性が、読みやすさの指標に影響を与えていた。本実験ではこれが内容の理解のしやすさにも影響したかどうかまでは確認できないが、以下、人の特性についての整理を試みる。

1. 経験と事前知識

「本人の罹患経験」がある参加者のテキストの読みの所要時間は、有意に短かった。これは、罹患経験があると内容に親しみがあがり、スムーズに関連主題のテキストを読むことができるためであると考えられる。特に改善テキストAおよびBで、罹患経験のある人の方が読みの所要時間が短い。罹患経験者に特に影響を与えるのが改善テキストのどの要素なのかは特定できない。また、「見たり聞いたり読んだりした経験」のある参加者からの「医学用語以外の表現が難しい」という指摘は少なかった。しかし、「医学用語が読みにくい」については、むしろ見聞経験がある方が高い割合を示している。これは、見聞経験は必ずしも事前知識につながっていないためであると考えられる。

同様に、「本人の罹患経験」があると内容理解テストのスコアや正答率が高くなるという有意差は見られなかった。むしろ、「本人の罹患経験」がある方が「設問1: 慢性化膿性中耳炎とはどんな病気ですか?」の正答率は低かった。その理由として、罹患経験は必ずしも疾病に関する網羅的な知識には結びつかないためであると考えられる。たとえば、改善テキストA、Bを読んだ参加者の中の12名が、“「耳だれ」という単語も、どういう状態になるのかわからない”など、「耳漏」に

代えた「耳だれ」が理解できていないことをコメントしていたが、そのうち9名には本人に罹患経験があった。また、「見たり聞いたり読んだりした経験」も、内容理解につながるとは言えない。たとえば今回の実験結果では、「耳のしくみがよくわからない」「手術の方法がわからない」など専門性が高いと思われる医学知識に関する選択肢が、テキストの違いはもちろん、罹患経験あるいは聞聞や読んだ経験と関連なく選ばれていた。

これらのことから、罹患経験は読みやすさに相対的な影響を与えるが、内容理解に影響する人の特性はむしろ事前知識の有無ではないかと推測される。

2. 知りたい内容とモチベーション

読みにくい点・わかりにくい点の選択肢による指摘と自由記述によれば、知りたい内容ではないために、読むモチベーションが失われ、その結果、十分な内容理解に結びつかないことが推測される。

まず、わかりにくい点について指摘が最も多かったのは、解剖学的な説明箇所である。たとえば「耳のしくみがわからない」という選択肢は、テキストによる差がなく、53.1%～60.0%が選択している。自由記述にも知りたい内容との食い違いとモチベーションへの影響をうかがわせる、以下のようなコメントが見られる。

- ・知りたい情報ではない部分が少々多い気がしました。
- ・病気の説明よりも、症状や治療といった具体的な診療に関する内容を優先した方がいい。
- ・乳突洞、乳突蜂巣、副鼻腔炎など、その後の段落でも同様に、医学用語や普段親しみのない用語が並ぶことで、文章に対して距離感を感じてしまい、思考にすんなりと入ってこないように感じられる。
- ・鼓室（こしつ）と乳突洞（にゅうとつどう）・乳突蜂巣（にゅうとつほうそう）からなる骨の中の空洞という説明を聞いても具体的にイメージできない。耳の奥の骨の中の空洞で十分。最初の一文からとっつきにくいと

読みたくなくなる

これらの結果から、医師が伝えたい内容をすべて、患者を含む一般市民に高いモチベーションをもって読んでもらえるよう伝えるには、テキストの改善だけでは限界があることがうかがえる。読み手のモチベーションについては、リーダビリティに影響を与える要素としてリーダビリティ研究でもたびたび取り上げられてきた⁸⁾。また、医師が伝えようとする内容と一般市民が必要とする内容の不一致について、一般市民のメンタルモデルによって複雑な医学概念を結び付けるのは困難であるという先行研究もあり^{69), 70)}、それらの研究結果と合致する結果と言える。

E. リーダビリティの評価方法

リーダビリティの評価のためには、テキストの改善状況の確認として、テキスト構造について手作業の内容分析が必要であるが、限定的に既存のリーダビリティ測定ツールを使うこともできる。今回の実験では、文や段落の長さ、漢字の割合などの構文的改善の程度は、参加者の回答の傾向と既存の公開されているリーダビリティ測定ツールによるテキスト分析の結果が合致していた。したがって、あらためて人による検証を行わなくても、個々の要素についてはテキスト分析の客観的な結果を確認すれば十分である。しかし、テキストの特徴だけを分析するリーダビリティ判定ツールのスコアだけで、そのまま読みやすさと内容理解のしやすさを測ることはできない。今回用いたツールでも、構文的要素から判定する『日本語リーダビリティ測定 Ver.0.5.0-UD』²⁸⁾のリーダビリティスコアと、語彙的要素から判定した『チュウ太の道具箱』²⁹⁾の結果に食い違いがあったことからそれがわかる。現在実用化されている日本語のリーダビリティ測定ツールは、あくまでも測定変数としている個別の要素の改善の目安として活用されるべきである。

人によるリーダビリティの改善評価では、読みの所要時間は、人にとって実際に読みやすくなったかの客観的な目安となる。しかし、外れ値が存在したように、実際の読みの時間を計測できない

ような想定外の行動をとる参加者も存在する可能性がある。評価指標としてより妥当な実験の運営方法にも工夫が必要である。

内容理解の指標として、本実験では選択問題の正答率を用いた。しかしながら、想定していたとおりの正答率の差を得ることができなかったり、予想と異なる正答率を得た設問に対し、その要因が構文的改善、語彙の改善、テキスト構造の改善のいずれにあるのか見極めることが困難であったりした。その原因としては、複数の改善方法の一つのテキストに同時に適用したことに加え、選択肢の作り方にも問題があったと言える。医学的正確さを優先したため、結果としてテキストの特徴にかかわらず正答しやすい選択肢となってしまった。

当初は、改善テキストの特徴によって正答率があがるような誤答の選択肢を用意していた。しかし、医師のチェックの結果、当該テキストを読んでもわからないが、医学的には正しいため、やむなく変更した選択肢がいくつかある。たとえば、設問1で出題している慢性化膿性中耳炎については、オリジナルテキストで冒頭から繰り返し説明されている「耳管」をとりあげ、「耳管の炎症が長引く病気」を誤答の選択肢としていた。しかし、「耳管にも炎症が起こっていることもあるので不正解とは言えない」との医師の指摘を受け取り下げた。代って作成した誤答は、「鼻から耳に通じる管ができて起こる病気」と、結果的に常識でも選択しないような記述に改めざるを得なかった。同様に「鼓膜に穴をあけたために炎症がおこる病気」を用意していたが、これも“外傷性の鼓膜穿孔から起こる中耳炎もある”との指摘で取り下げることにした。

また、選択肢方式自体の限界があった。設問3の「耳だれ」のように、文字列として認識しているため正答はできても、自由記述の回答を見ると、実際の用語の意味は理解していなかったと思われる設問もあった。選択肢型の設問は採点が効率的で正誤の判定が明確であるという利点はあるものの、内容理解をより正確に測るには、読解力テストなどで実施されているほかの方法を合わせ

てとることが望ましいだろう。

今回の Web テストのように「読みにくい点・わかりにくい点」として選択肢や自由記述を設けることは必須である。自由記述も分析には手間がかかるが、実際の改善の評価に欠かせない。所要時間や内容理解テストでは測定できない、具体的な指摘から分析を行うことができるからである。今回は、改善テキスト A の所要時間が有意に長かった理由が「かっこでつかえる」で、改善方法に問題があったことが、選択肢と自由記述の両方から特定できた。また、回答が予想できなかったため選択肢には設けなかったが、それに対する指摘を自由記述から得ることができた。

人によるリーダビリティの改善評価では、相対的にリーダビリティに影響を与える読み手の特性に注意が必要である。これらの特性に関するクロス集計の分析結果は、本論文では取り上げないが、同じ大学の学部生で同じ学問分野であっても、個人の基礎的な読解力の違いが読みやすさや内容理解のしやすさに影響を与えた可能性がある。これについては、一般的な読解力を測定するテストを合わせて実施することで、その影響を測ることができるだろう。

さらに、特に改善テキスト B は、読みやすい・わかりやすいという印象を参加者に与えていながら、内容理解テストの正答率からは内容理解しやすいテキストと判断できなかった。これは、読みやすい、わかりやすいテキストがかならずしも内容が理解しやすいテキストとは限らないという先行研究の結果と一致する⁷¹⁾。この不一致を踏まえた、内容理解のしやすさの妥当な評価方法については今後の課題である。

VI. 健康医学情報を伝える日本語テキストのリーダビリティの改善とその評価

構文的要素、語彙的要素、テキスト構造の3種類の要素から読みやすさと内容理解のしやすさの改善を試みた日本語による説明テキストの改善実験の結果から、健康医学情報を伝える日本語テキストのリーダビリティの改善とその評価の難しさが改めて明らかになった。改善の効果は限定的

だったが、医師との協働のもとでの実際の改善方法と、その効果の評価方法について、以下の点が示された。1)構文的改善は読みやすい印象を与えるが、他の要素の改善と対立が生じることがある、2)医学・医療用語の改善は自然な置き換えや文中での説明の補足が望ましい、3)テキスト構造は、今回の改善・評価方法では問題が多くその効果は明らかではない、4)事前知識やモチベーションなど、読み手としての人の特性を考慮した評価が必要である、5)内容理解を正確に測定する選択問題の作成に課題がある。

本研究における実験には以下の点で限界がある。今回は、オリジナルテキストに対して最大限の改善を加えるため、同時に3種類の要素を一度に盛り込んだ。実際に利用されるテキストに対する改善方法を検討できたというメリットがあった反面、両立しない要素が生じたり、読みやすさと内容理解のしやすさへの影響がどの要素によるものか区別ができなかったりする問題が生じた。また、参加者募集の技術的な制約から、同じ大学の学生を対象としながら、個人的能力を示す標準化されたデータを収集しなかったため、実際には基礎的な読解力の異なる学生が混在していて実験結果に影響を与えた可能性もある。さらに、内容理解テストの設問と選択肢の完璧な組み合わせを作成するのは難しく、予測したテキストの改善要素とは異なる要因によって正答率がかわってしまっていたということがあった。

今後の研究では、リーダビリティの改善に貢献した要素を特定できるよう、各要素を個別に改善したテキストを用意し、個人の読解力の違いを考慮するために読解力の標準テストを併用するなど、精緻な実験の実施が必要となる。また、内容理解テストについては、英語のリーダビリティ研究でも用いられ、日本語テキストでもその有用性が確認されていて、なおかつ比較的处理をしやすい穴埋め問題のクローズ法など、複数の方法を用いることで、より妥当性と信頼性の高いデータを集めて分析を行うことが考えられる。さらに、改善の手続きとして、医師とのやりとりの結果、医学的正確さを優先させるため、リーダビリティに

貢献する要素について妥協しなければならないという実際の難しさがあることもわかっている。この点については、医師とのやりとりを分析することで問題の糸口を探り、健康医学情報を伝える日本語テキストのリーダビリティの改善とその評価について研究を深めなければならない。

謝 辞

本研究のために、テキストの改訂に根気強くつきあい協力してくださった医療法人ふくいく会の鳥谷部郁子医師に御礼申し上げます。また、テストに参加してくれた学生みなさんと募集にご協力をいただいた教員の方々にも感謝します。調査結果の統計解析については国立保健医療科学院研究情報センター長の緒方裕光氏にご指導いただくことができました。御礼申し上げます。さらに、テストの実施から論文の作成まで、常に惜しみない支援と的確な助言をくださった慶應義塾大学文学部の倉田敬子教授に、心より御礼を申し上げます。最後に、本論文のまとめ方について建設的かつ仔細なご意見をくださった本誌の査読者にも感謝申し上げます。

注・引用文献

- 1) 正確には、米国の医学研究所 (Institute of Medicine) の定義は以下のとおりで、情報提供側の問題でもあることが明記されている。すなわち“ヘルスリテラシーは、健康に関する適切な意思決定をするために必要な、基本的な健康医学情報やサービス入手し、処理し、また理解する個人の持つ能力の度合いである。それは個人の能力だけでなく、健康情報や医療の提供者（医師、看護師、管理者、家庭看護師、メディアなど）の情報提供のスキルや好み、期待によっても左右されるもの” [p.2]. Nielsen-Bohlman, L. et al. Health Literacy: a Prescription to End Confusion. National Academies Press, 2004, p. 345.
- 2) Center for Health Care Strategies. Healthy People 2010: Understanding and Improving Health. 2nd ed. U.S. Government Printing Office, 2000, 2 vols.
- 3) ヘルスリテラシーに関する問題に、図書館情報学分野の研究が貢献する可能性については以下に詳しい。酒井由紀子. ヘルスリテラシー研究と図書館情報学分野の関与: 一般市民向け健康医学情報サービスの基盤として. Library and Information

- Science. 2008, no. 59, p. 117-146.
- 4) Kars, M; Baker, L.; Wilson, F. L. The Medical Library Association Guide to Health Literacy. Neal-Schuman Publishers, 2008, 314p.
 - 5) 以下のサイトには健康医学情報の読みやすく内容理解しやすいテキストの書き方の手順, リーダビリティ測定ツールやガイドラインへのリンクが集められている。U.S. National Library of Medicine. How to Write Easy-to read Health Materials. <http://www.nlm.nih.gov/medlineplus/etr.html>, (accessed 2011-03-30).
 - 6) 野呂幾久子, 邑本俊亮. インフォームド・コンセント説明文書のわかりやすさと情緒的配慮の記述が患者アウトカムに与える影響: 大学生を対象とした調査. 日本保健医療行動科学会年報. 2009, vol. 24, p. 102-116.
 - 7) 英語の Web 上の健康医学情報の評価について下記のシステムティックレビューがあるが, とりあげられた 79 の研究のうち 11 の研究でリーダビリティが評価指標の一つとなっている。Eysenbach, G. et al. Empirical studies assessing the quality of health information for consumers on the World Wide Web: a systematic review. JAMA. 2002, vol. 287, no. 20, p. 2691-2700.
 - 8) Klare, G. R. "Readability". Handbook of Reading Research. Pearson, P., Barr, R., Kamil, M. L. eds. Longman, 1984, p. 681-744.
 - 9) Dale, E. デールの視聴覚教育. 西本三十二訳. 日本放送教育協会, 1957, 242p.
 - 10) Gray, W. S.; Leary, B. E.; Joint Committee on the Reading Interests and Habits of Adults. What Makes a Book Readable, with Special Reference to Adults of Limited Reading Ability. University of Chicago Press, 1935, 358p.
 - 11) Nielsen-Bohlman, L.; Panzer, A. M.; Kindig, D. A.; Institute of Medicine. Committee on Health Literacy. Health Literacy: a Prescription to End Confusion. National Academies Press, 2004, 345p.
 - 12) Ley, P.; Florio, T. The use of readability formulas in health care. Psychology, Health & Medicine. 1996, vol. 1, no. 1, p. 7-28.
 - 13) Friedman, D. B.; Hoffman-Goetz, L. A systematic review of readability and comprehension instruments used for print and web-based cancer information. Health Education & Behavior. 2006, vol. 33, no. 3, p. 352-373.
 - 14) Chall, J. S.; Dale, E. Readability Revisited: the New Dale-Chall Readability Formula. Brookline Books, 1995, 159p.
 - 15) Gemoets, D. et al. Assessing readability of consumer health information: an exploratory study. Medinfo. 2004, vol. 11, no. Pt 2, p. 869-873.
 - 16) Consumer Health Vocabulary Initiative. <http://www.consumerhealthvocab.org/>, (accessed 2010-12-19).
 - 17) MedlinePlus. <http://medlineplus.gov/>, (accessed 2010-07-13).
 - 18) Zeng, Q. et al. "A text corpora-based estimation of the familiarity of health terminology". Biological and Medical Data Analysis: 6th International Symposium, ISBMDA 2005. Aveiro, Portugal, 2005-11-10/11. Springer, 2005, p. 184-192.
 - 19) Elhadad, N. "Comprehending technical texts: predicting and defining unfamiliar terms". Biomedical and Health Informatics: Proceedings of the 2006 AMIA Annual Symposium. Washington, D.C., USA, 2006-11-11/15. American Medical Informatics Association, 2006, p. 239-243.
 - 20) UMLS®. <http://www.nlm.nih.gov/research/umls/>, (accessed 2010-07-13).
 - 21) Tse, T.; Soergel, D. "Exploring medical expressions used by consumers and the media: an emerging view of consumer health vocabularies". Biomedical and Health Informatics: From Foundations to Applications: Proceedings of the 2003 AMIA Annual Symposium. Washington, D.C., USA, 2003-11-8/12. American Medical Informatics Association, 2003, p. 674-678.
 - 22) 森岡健二. 読みやすさの基礎調査. 国立国語研究所年報. 1953, vol. 4, p. 114-131.
 - 23) 庵谷巖. 文章の難易度よりみた適書選択の基準について (試案). 読書科学. 1956, vol. 1, no. 1, p. 29-33.
 - 24) 阪本一郎. 教育基本語彙. 牧書店, 1958, 377p.
 - 25) 堀川直義. 文章のわかりやすさの研究. [朝日新聞調査研究室], 1957, 213p.
 - 26) 北尾倫彦. ひらがな文と漢字まじり文の読みやすさの比較研究. 教育心理学研究. 1960, vol. 7, no. 4, p. 195-199.
 - 27) 日本語テキストの難易度を測る. <http://kotoba.nuee.nagoya-u.ac.jp/sc/readability/>, (accessed 2010-07-13).
 - 28) 日本語リーダビリティ測定 Version 0.5.0-UD. <http://language.tiu.ac.jp/tools.html>, (accessed 2010-07-13).
 - 29) チュウ太の道具箱. <http://language.tiu.ac.jp/tools.html>, (accessed 2010-07-13).
 - 30) 乾健太郎. 自然言語処理と言い換え (特集 言葉を言い換える). 日本語学. 2007, vol. 26, no. 13, p. 50-59.
 - 31) 山本聡美, 幹健太郎, 幹裕子. 聾者向け読解支援のための文可読性基準のモデル化. 情報処理学会研究報告. 自然言語処理研究会報告. 2001, vol. 2001, no. 9, p. 131-136.
 - 32) 藤沢仁子, 神門典子, 相原健郎, 安達淳. 言い換え箇所と言い換え候補の提示による解説文リライト支援の書き手の評価実験 (言い換え・略語・要

- 約). 情報処理学会研究報告. 自然言語処理研究会報告. 2009, vol. 2009, no. 36, p. 59-66.
- 33) 独立行政法人医薬品医療機器総合機構. 患者用語集 (副作用用語集). 患者向医薬品ガイド. http://www.info.pmda.go.jp/guide_ippan/guide.html, (accessed 2010-07-13).
- 34) 国立国語研究所「病院の言葉」委員会編著. 病院の言葉を分かりやすく: 工夫の提案. 勁草書房, 2009, 234p.
- 35) 図書館情報学におけるジャンル分析の有用性について, 以下のレビューがある.
Anderson, J. The concept of genre in information studies. Annual Review of Information Science and Technology. 2008, vol. 42, no. 1, p. 339-367.
- 36) “談話”. 言語学大辞典. 亀井孝, 河野六郎, 千野栄一編著. 三省堂, 1988, p. 897-901.
- 37) Halliday, M. A. K.; Hasan, Ruqaiya. Cohesion in English. Longman, 1976, 374p.
- 38) Van Dijk, T. A. Text and Context: Explorations in the Semantics and Pragmatics of Discourse. Longman, 1977, 261p.
- 39) 片岡邦好. “結束性と一貫性”. 応用言語学辞典. 小池生夫編. 研究社, 2003, p. 227.
- 40) Bakhtin M. M. “ことばのジャンル”. ことば対話テキスト. 新谷敬三郎ほか訳編. 新時代社, 1988, p. 113-189.
- 41) Samuels, S. J. et al. Adults' use of text structure in the recall of a scientific journal article. Journal of Educational Research. 1988, vol. 81, no. 3, p. 171-174.
- 42) Meyer, B. J. F. Reading research and the composition teacher: the importance of plans. College Composition and Communication. 1982, vol. 33, no. 1, p. 37-49.
- 43) Kemper, S. Measuring the inference load of a text. Journal of Educational Psychology. 1983, vol. 75, no. 3, p. 391-401.
- 44) Dillon, A. Bringing genre into focus: Why information has shape. Bulletin of the American Society for Information Science and Technology. 2008, vol. 34, no. 5, p. 17-19.
- 45) Kintsch, W.; Vipond E. “Reading comprehension and readability in educational practice and psychological theory”. Nilsson, L. G. ed. Perspectives on Memory Research. Wiley, 1979, p. 329-365.
- 46) 木下是雄. 理科系の作文技術. 中央公論社, 1981, 244p.
- 47) 酒井聡樹. これからレポート・卒論を書く若者のために. 共立出版, 2007, 225p.
- 48) 酒井聡樹. これから論文を書く若者のために. 共立出版, 2002, 232p.
- 49) 永野賢. 文章論総説: 文法論的考察. 朝倉書店, 1986, 380p.
- 50) 佐久間まゆみ. 「文段」認定の一基準 (I): 提題表現の統括. 文藝言語研究. 言語篇. 1987, vol. 11, p. 89-135.
- 51) 佐久間まゆみ. 文段認定の一基準 (II): 接続表現の統括. 文藝言語研究. 言語篇. 1990, vol. 17, p. 35-66.
- 52) 佐久間まゆみ. 中心文の「段」統括機能. 日本女子大学紀要. 文学部. 1994, vol. 44, p. 93-109.
- 53) 市川孝. 段落・文段をめぐる. 言語生活. 1959, p. 42-49.
- 54) 文章構成法. 新版. 東海大学出版会, 1995, 213p.
- 55) 永田亮, 井口達也, 榊井文人, 河合敦夫. リーディングスピードに基づいた文章の読み易さについて. 電子情報通信学会技術研究報告. TL, 思考と言語. 2002, vol. 102, no. 491, p. 13-18.
- 56) クローズ法は内容理解テストの一手法で, 決まった数おきに語を削除して空所を設ける穴埋め問題のこと. 出題者が意図的に空所を選ぶ完成法とは異なる.
- 57) Rosembat, G. et al. “Text features and readability: expert evaluation of consumer health text”. MEDNET 2006: 11th World Congress on Internet in Medicine the Society of Internet in Medicine. Toronto, Canada, 2006. http://www.mednetcongress.org/fullpapers/MEDNET-192_RosembatGracielaA_e.pdf, (accessed 2010-07-13).
- 58) Chapman, W. W.; Aronsky, D.; Fiszman, M.; Haug, P. J. “Contribution of a speech recognition system to a computerized pneumonia guideline in the emergency department”. Converging Information, Technology, and Health Care: Proceedings of the AMIA Annual Symposium. Los Angeles, CA, USA, 2000-11-4/8. American Medical Informatics Association, 2000, p. 131-135.
- 59) Keselman, A.; Tse, T.; Crowell J.; Browne, A.; Ngo L.; Zeng Q. Relating consumer knowledge of health terms and health concepts. AMIA Annual Symposium. Proceedings. 2006, p. 980.
- 60) Zeng, Q.; Kim, E.; Crowell, J.; Tse, T. “A text corpora-based estimation of the familiarity of health terminology”. Biological and Medical Data Analysis: 6th International Symposium, ISBMDA 2005. Aveiro, Portugal, 2005-11-10/11. Springer, 2005, p. 184-192.
- 61) Keselman, A.; Massengale, L.; Ngo, L.; Browne, A.; Zeng, Q. “The Effect of user factors on consumer familiarity with health terms: using gender as a proxy for background knowledge about gender-specific illnesses”. Biological and Medical Data Analysis: 7th International Symposium, ISBMDA 2006. Thessaloniki, Greece, 2006-12-7/8. Springer, 2006, p. 472-481.
- 62) Keselman, A.; Tse T.; Crowell J.; Browne A.; Ngo L.; Zeng Q. Assessing consumer health vocabulary familiarity: an exploratory study. Journal of Medical Inter-

- net Research. 2007, vol. 9, no. 1, p. e5.
- 63) Leroy, G.; Miller, T.; Rosembat, G.; Browne, A. A balanced approach to health information evaluation: a vocabulary-based Naïve Bayes classifier and readability formulas. *Journal of the American Society of Information Science and Technology*. 2008, vol. 59, no. 9, p. 1409-1419.
 - 64) 高梨庸雄, 卯城祐司. 英語リーディング事典. 研究社出版, 2000, 435p.
 - 65) 慶應義塾大学耳鼻咽喉科学教室. 慢性化膿性中耳炎. KOMPAS 慶應義塾大学病院医療・健康情報サイト. 2009-02-01. <http://kompas.hosp.keio.ac.jp>, (accessed 2009-10-01).
 - 66) Sakai, Y.; Kunimoto, C.; Kurata, K. "How they 'change': health information consumers in Japan". Poster presented at iFusions: MLA09: Medical Library Association Annual Meeting, May 15-20, 2009, Hawaii Convention Center, Honolulu, HI. http://www.openaccessjapan.com/resources/pdf/05_sakai.pdf, (accessed 2010-07-13).
 - 67) 久保鈴子. 厚生労働科学研究研究費補助金 医薬品・医療機器等レギュラトリーサイエンス総合研究事業 患者及び国民に理解される副作用等医薬品情報内容の構築と医薬品適正使用への患者参加推進に関する研究 平成 16 年度総括・分担研究報告書. 2005, 160p.
 - 68) “身近な題材と意外性のある題材における読解過程: 不当なスキーマによる解釈のゆがみ (大学生対象)”. 学習者中心の英語読解指導. 津田塾大学言語文化研究所読解研究グループ編. 大修館書店, 1992, p. 54-71.
 - 69) Patel, V. L.; Arocha, J. F.; Kushniruk, A. W. Patients' and physicians' understanding of health and biomedical concepts: relationship to the design of EMR systems. *Journal of Biomedical Informatics*. 2002, vol. 35, no. 1, p. 8-16.
 - 70) Slaughter, L.; Soergel, D. How physicians' answers relate to health consumers' questions. *Proceedings of the American Society for Information Science and Technology*. 2003, vol. 40, no. 1, p. 28-39.
 - 71) Charrow, V. "Readability vs. comprehensibility: a case study in improving a real document". *Linguistic Complexity and Text Comprehension*. Davison, A., Green, G. M. eds. L. Erlbaum Associates, 1988, p. 85-114.

要 旨

【目的】 本研究の目的は、健康医学情報を伝える日本語テキストのリーダビリティの改善と評価の一連の手続きにおいて、包括的な改善方法と評価方法を記述的に確認し、今後の研究課題を明らかにすることにある。

【方法】 医師が執筆した、慢性化膿性中耳炎の一般市民向けの日本語説明テキスト 1 件を取り上げ、改善と評価の実験を行った。最初に、リーダビリティに影響するとされる構文、語彙、テキスト構造のすべての要素について、方法を変えて改善したテキストを 2 種類用意し、オリジナルテキストと合わせテキスト分析を行った。次に、大学生 91 名にこれら 3 種類のテキストのいずれかを割り当て、Web テストを行った。テストでは、読みやすさの指標としてテキストを読む所要時間を計測した。内容理解のしやすさの指標として、選択肢問題の回答と、「読みにくい点・わかりにくい点」の具体的な指摘を求めた。

【結果】 構文的要素を改善し、語彙的改善として医学・医療用語に解説をかって補記し、テキスト構造を簡易な方法で改善したテキスト A は、内容理解テストの平均スコアはオリジナルテキストより高かったが、読みの所要時間は長かった。構文的要素を改善し、医学・医療用語を一般的な用語や表現に置き換え、テキスト構造を入念に改善したテキスト B は、読みの所要時間は短縮されたが、内容理解テストの平均スコアはオリジナルテキストと変わらなかった。また、本人に罹患経験があると所要時間が有意に短かった。これらの結果から、1) 構文的改善は読みやすい印象を与えるが、他の要素の改善と対立が生じることがある、2) 医学・医療用語の改善は自然な置き換えや文中での説明の補足が望ましい、3) テキスト構造は、今回の改善・評価方法では問題が多くその効果は明らかではない、4) 読み手としての人の特性を考慮した評価が必要である、5) 内容理解を正確に測定する選択肢問題の作成に課題があることが確認された。