

科学研究費助成事業（科学研究費補助金）研究成果報告書

平成 25 年 5 月 23 日現在

機関番号：13903

研究種目：基盤研究(C)

研究期間：2010～2012

課題番号：22520561

研究課題名（和文）科学技術英語コーパスの分析に基づくコンピュータ適応型アカデミック表現テストの開発

研究課題名（英文）Development of Computerized Adaptive Tests for Academic Expressions Based on the Analysis of Science & Technology Corpora

研究代表者

小山 由紀江 (KOYAMA YUKIE)

名古屋工業大学・工学研究科・教授

研究者番号：20293251

研究成果の概要（和文）：本研究の目的は理工系学生の英語語彙・語句に関する能力を測定するために、科学技術コーパスデータの分析に基づいたコンピュータ適応型テスト(CAT)を作成することである。この目的ため、科学技術コーパスを分析し、重要語彙・語句を抽出しこれを試験項目に使用した。テスト項目の分析と受験者の能力推定には Latent Rank Theory (LRT) を採用し、予備テストを実施し 230 項目の項目バンクを作成した。これらを基に開発した理工系英語 CAT を moodle 上で実施した結果、科学技術英語のテストとして一定の妥当性があることが解った。

研究成果の概要（英文）：The purpose of this study is to develop a Computerized Adaptive Test based on the analysis of scientific corpora, which is appropriate for measuring English knowledge of engineering students. First, important vocabulary and phrases were extracted from the corpora in order to write test items. Secondly, Latent Rank Theory was adopted for the testing theory to analyze the preliminary test data and to measure each student's knowledge. After conducting preliminary tests of vocabulary and phrases, an item bank including 230 items was completed. Using this bank, the CAT was conducted in the moodle system. The results showed the CAT had potential as a method of measuring students' knowledge of English vocabulary relating to science & technology.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
2010 年度	1,800,000	540,000	2,340,000
2011 年度	800,000	240,000	1,040,000
2012 年度	600,000	180,000	780,000
総計	3,200,000	960,000	4,160,000

研究分野：人文学

科研費の分科・細目：言語学・外国語教育

キーワード：CAT、科学技術英語、コーパス、言語テスト、LRT、

1. 研究開始当初の背景

(1) コーパス言語学

ICT の発達に伴い、デジタル化された言語材料を集積して大規模コーパスを構築し、ま

たそれをコンピュータで分析することが比較的容易に可能となった。コーパス言語学は、このように大量の言語の実態を明らかにする研究として隆盛し、その成果は学習語彙リ

ストの作成等の形で言語教育にも反映されている。

(2) CAT

1990年代後半からアメリカを中心にTOEFL等の受験者の多い標準テストがCAT形式で実施され始めた(Weise, 1973)。これはItem Response Theory (IRT)に基づいて分析された大量の問題で項目バンクを作り、コンピュータが適切な難易度の問題を選んで個々の受験者の解答に応じた問題提示を行うものである。項目選定とテストの実施の両方の段階でコンピュータを使用することにより、受験者の能力を正確に効率的に測定するテスト方法として可能性が示されてきた(渡邊・野口, 1999)。CATは現在ではGRE, GMAT, COMPASS/ESLなどの著名な教育関連のテストや看護師、会計士などの専門職資格試験にも適用され、時間的な効率性とテストの信頼性が確立されつつある。

(3) 現状

日本国内においても約1700万語のPERCなど科学技術コーパスが公開されているが教育への応用は課題であり、科学技術コーパスに基づいた信頼性・妥当性のある英語テストはまだ開発されていない。一方、小山は平成10-18年に渡る三つの科学研究費の研究で、科学技術論文を中心とする約600万語のコーパスを構築し、この分析結果に基づいて工学系ESP語彙リスト(石川・小山, 2007)の作成やe-learning教材の開発を行ってきた。また単語のレベルに留まらない語句を含む特徴表現を抽出することも試みている(田中・小山, 2009)。CAT形式のテストに関しては、杉森他(2005)の研究において、日本人英語学習者を対象とした語彙、文法テストの開発研究を行ってきており、豊富な項目バンクとアダプティブエンジンを利用することで信頼性の高いCATテストを開発することに成功している。これらの研究を発展させ、本研究では信頼性と妥当性の高いCAT形式を用い、理工学部学生のための語彙・語句のテストを作成する。

2. 研究の目的

本研究の目的は、理工系大学の学部生、大学院生の科学技術分野一般における英語表現力(語彙・語句に関する知識)を測定するため、1)科学技術コーパスの分析に基づいて、2)語レベルのみならず、特徴的な語句や表現をも抽出し、3)これをターゲットとしたCAT型のテストを作成することである。

3. 研究の方法

(1) 概観

まず、科学技術コーパスを統計的に分析し重要な表現を抽出すると共に、これらの表現を含む文章をコーパスから検索し、これを元に

文脈性を持たせた問題を作成する。次に、これらの問題を使って予備テストを実施し、適切なIRT等の言語テスト理論による項目分析を行い、CATに採用する問題の精査を行う。最終段階として、moodle上に置いたCATモジュールによって、この科学技術英語CATを実施し、結果を分析して妥当性の検証を行う。以下方法について詳述する。

(2) 重要語句抽出とテスト理論

まず重要語句抽出の方法について理論的に研究し、抽出した対象語彙・語句の評価を行う。また同時にテスト項目作成の基盤である科学技術論文コーパスの拡充を行う。これは科学技術英語論文テキストの追加のみならず、特徴表現抽出と抽出結果の比較評価のため、英語国の高校の科学系教科書をコーパス化し英語難易度の異なるテキストを分析することも含んでいる。

さらに、作成された項目の難易度を推定する言語テスト理論の研究を行う。これについては、これまでCATの理論として主流であったIRTと新しい言語テスト理論であるLRTとを比較研究し英語教育の場により適切な理論を考察するためである。

(3) 項目バンクの作成

科学技術論文コーパスから抽出された重要語彙・語句を基に、それらを含む文章によるテスト項目を作成する。重要語彙・語句の抽出法は語彙に関しては「工学系ESP語彙リスト(EGSTリスト)」(石川・小山, 2007)を踏襲し、語句の抽出に関しては安定的な結果が得られるn-gramを使った抽出法を採用した。いずれも複数の分野に共通的に出現する語彙・語句を抽出する方法である。

語彙項目のテスト問題の形式は、単語―単語の日英あるいは英日の訳、英語による英単語の定義、文中の空所補充の4種類である。解答様式は全て多肢選択であるが、それぞれ正答以外の選択肢も「EGST語彙リスト」の同程度の頻度レベルから選んだものである。

語句は、Hyland (2008)によって適切な長さとしてされている4-gramと、より複雑な語句も対象とすることを視野に入れて5-gramとの両方を抽出することにし、応用物理、生物工学、電気工学、機械工学の4分野のコーパス約200万語から、AntConc3.2.4 (Anthony, 2011)を用いて4分野に共通して出現する語句を抽出した。問題形式は全て文中の空所補充で解答は多肢選択形式であるが、正答以外の選択肢も抽出されたリストにあるものを使用している。

問題文は、重要語彙・語句を含む文章をコーパスで検索し、検索された文の長さや他の使用語彙の難易度を調整し、さらに英語の母語話者と筆者が協議の上作成したものを使用している。

以上の方法で作成したテスト項目全てに

ついて予備テストを行い、結果を分析して各項目の難易度を推定する。本研究では、予備テストの結果を分析した際、学力の評価には連続尺度に基づくIRTよりも、順序尺度に基づくLRTの方がより適切な解像度を持つこと判断し、本研究にはLRTを採用することにした。従って、LRTによる分析によって、項目バンク全体として難易度が適切に分布するように調整を行った。また、予備テストは問題数が230問と多いため受験者に負担がかかることと、またその他の実施上の制約があることから、語彙(1回)と語句(2回)の計3回に分けて実施した。

(4) CATの実施

本研究のCATはmoodle 2.3.3に実装されたLRT-CATモジュール(秋山・木村・荘島, 2011)を使って実施する。対象は工業系大学1年生80名であり、大学のPCルームを使って全端末にインストールされているmoodle上で実施された。CATを実施する場合重要なのは、1) 最初のアイテムの設定、2) 次に出すアイテムの選択方法、3) 終了条件の決定の3点であるが、最初のアイテムはランダムに選択、次に出すアイテムの選択アルゴリズムは木村(2012)の提案した「暫定RMPとIRPの差分ベクトルの積和平均による項目選択ルール」(p76)を用いる。また終了条件はLRTの指標の1つRMP(Rank Membership Profile)の数値と最大受験問題数の両方を設定することができるが、今回は2クラスで実施するためそれぞれ「RMP<0.01あるいは最大問題数100問」、「RMP<0.03最大問題数50問」と異なる設定によって、CATを実施した。

4. 研究成果

(1) 重要語句分析の結果

本研究は科学技術英語のコーパス分析に基づく重要語彙・語句をテスト項目のターゲットとすることが特徴の1つであるが、以下、その手順と結果について説明する。

重要語句の抽出は既述のように工学4分野に共通する単語連鎖を分析することによって抽出したものだが、その結果、4-gramは658、5-gramは168が抽出された。それらの単語連鎖には“a function of the”、“as well as the”など定冠詞や不定冠詞で終わるものなど、意味的なまとまりのないものが含まれている。これらを手作業により除外した結果、頻度順で150位までの中で残った4-gramは71、5-gramは80の単語連鎖であった。それらの4-gramのうちアルファベット順にaから20番目まで抜粋したものが以下の表である。(5-gramに関してはスペースの関係で割愛する。)

科学技術英語 4-gram

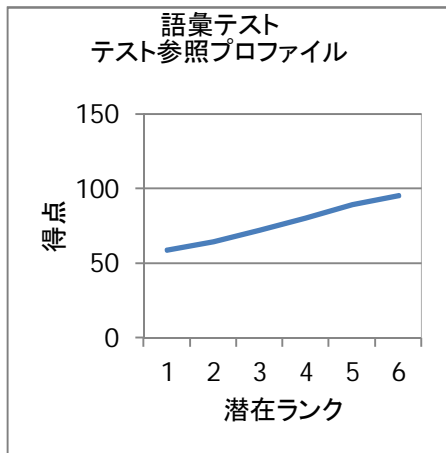
a large amount of
 a large number of
 a wide range of
 are plotted in fig
 are shown in fig.
 as a function of
 as a result of
 as the ratio of
 at a rate of
 at the beginning of
 at the center of
 at the end of
 at the same time
 be considered to be
 be taken into account
 by the fact that
 can be applied to
 can be attributed to
 can be considered as
 can be described by

(2) 予備テストの結果

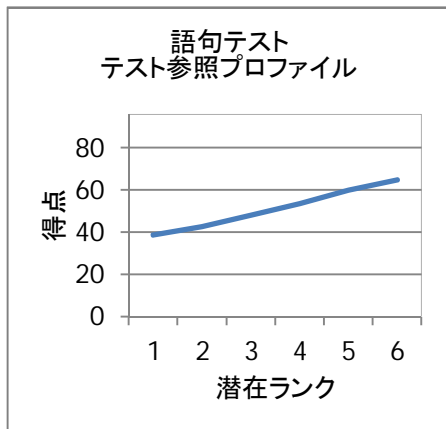
予備テストは前記のように、語彙と語句(単語連鎖)を分けて行った。語彙テストは受験者が大学1年生から大学院生まで134人、語句のテストは2回に分けて実施し、受験者は1年生74人である。両テストの結果の基本統計量は以下の表の通りである。語句のテスト用に作成した問題は80問であるが、語彙テストと受験者が異なるために語彙テストから16問のアンカーアイテムを採用した。そのために項目数は96間になっている。(以下、統計は全てExametrika by Shojima, 2008による。)

基本統計量	予備テストA (語彙)	予備テストB (単語連鎖)
受験者数(N)	134	74
項目数(n)	150	96
最小値(Min)	41	28
最大値(Max)	138	84
中央値(Median)	76	50.5
平均値(Mean)	77.60	51.64
分散(Var.)	357.3	160.72
標準偏差(SD)	18.90	12.67

テスト項目に関しては、言語テスト理論の研究により、順序尺度に基づく LRT を使って項目分析を行った。語彙テストの分析結果は以下のテスト参照プロファイルの図が示すように、ランク（能力）による傾きが小さく、識別力が余り高くないことが解った。



語句（単語連鎖）のテストについても以下の図を見る限り傾きが小さく、識別力の高いテストとは言えない結果になった。

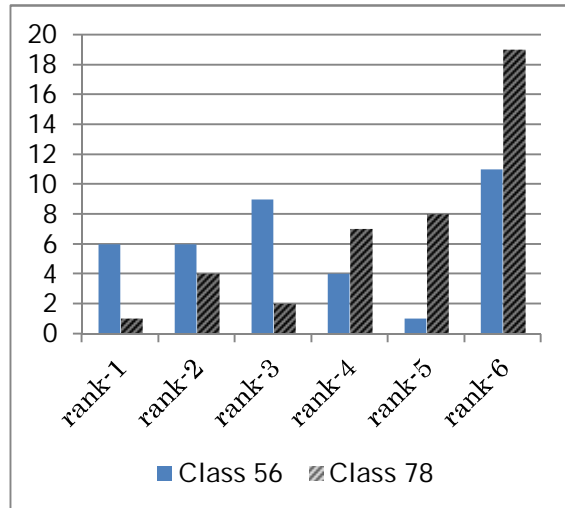


(3) CAT の結果

以上のように全体として識別力が高い項目バンクとは言えないが、ランク 1 とランク 6 に属する受験者の推定得点にはそれぞれ一定の差が出ているため、この項目バンクを使用して CAT の試行を行った。実施したのは 2 つのクラス（クラス 56 とクラス 78）で受験者数は各 37 名と 43 名である。CAT 実施には前述のように moodle 2.3.3 に実装された LRT-CAT モジュール(秋山・木村・荘島, 2011) を使った。各クラスの終了条件はやはり前述のように「RMP<0.01 あるいは最大問題数 100 問」と、「RMP<0.03 最大問題数 50 問」とそれぞれ異なる設定であった。

CAT 試行の結果を各ランクに属す受験者数のヒストグラムで示すと以下の通りになる。二つのクラスは一見してかなり異なる結果となり、クラス 78 は多くの受験者がラン

ク 6 に属していることが解る。それに対しクラス 78 ではランク 1 から 6 まで受験者が散らばっている。ちなみに今回実施の CAT、TOEIC のリーディングセクション、科学技術英語の期末試験リーディングセクションの得点の平均値は次の表にある通りで、やはりクラス 78 が全てクラス 56 より高いという結果である。



平均点	CAT	TOEICR	期末R
クラス56	3.6	211.7	72
クラス78	4.6	228.4	79.3

ランクが確定するまでの受験アイテム数（解答数）と所要時間は CAT の成否を決める重要なファクターであるため、これらの点について述べることにする。全体に最大受験項目数の 100 問あるいは 50 問を受ける受験者が多く、特にクラス 56 では 100 問を受験した者が半数以上に上った。これは CAT の終了条件の設定が不適切であったことが考えられる。全体としてはランク 6 に属する受験者は短時間で終了したことが解った。

最後に本研究で開発した CAT の精度について述べておく。CAT の能力測定の精度を確かめるには、アイテム・バンクの元の問題に対する解答行動によって得られた受験者の能力値（真値）と CAT を受験することによって得られた能力値を比較する必要がある。しかし、今回の CAT の受験者は予備テスト A（語彙テスト）によって構築されたアイテムは解答しておらず、予備テスト B（単語連鎖テスト）のみしか受験していないため、解答全体に基づいた推定ランクを比較することができない。従って、今回の CAT に関しては、同じ受験者が同時期に受けた科学技術英語（リーディング）の期末試験と、TOEIC（リーディングセクション）の得点と CAT によって推定されたランクとのスピアマン相関を求めることによってその精度を検証する

ことにした。2つのクラスは実施条件が異なるため別々に値を求めた。対象としたのは全てのテストを受けたクラス56の30名、クラス78の41名である。結果は以下の表が示す通りCATと科学技術英語の期末試験は両クラスとも0.6以上の比較的高い相関係数を得たが、TOEICとはクラス78において0.48という低い相関係数となった。科学技術英語の期末試験は学習内容が科学技術に関するものであることから、今回のCATが測定する能力とは類似度が高い。一方、TOEICは内容、受験時期ともにこのCATとの差異が大きく、相関も低い結果となったと考えられるだろう。

CATと他のテストとの相関

	TOEICR	期末R
クラス56	0.630748	0.634739
クラス78	0.481363	0.655065

(4) 今後の課題

以上、本研究のCATの開発及び実施結果について述べたが、本CATにはCATの適切な実施に必要なプロセスが2点欠けている。即ち、アイテムの精選とシミュレーションの十分な実施である。アイテムの作成と予備試験の実施上の時間的な制約から、作成したアイテムを精査し十分精選することができなかった。そのため、語彙の予備試験Aも語句の予備試験Bもテスト参照プロファイル(TRP)の傾きが緩やかな、識別度の高くないテストとなり、結果的にCATの精度も上がらなかった可能性が考えられる。LRTにおいてはアイテムの識別力をIRPのグラフの傾きで判断するが、今回は予備テストの結果から右がり、あるいはフラットなグラフも散見された。今後、不適切なアイテムを除去する必要があるが、その基準の明確化も課題の一つである。

しかし、テストとしての内容的妥当性を高めるために本研究で試みた科学技術英語コーパスの分析に基づくアイテム構築と、これをアイテム・バンクとして作成したCATは一定の精度を示しており、今回の手法はこのICTの時代に大きな可能性を秘めていると言える。今回のCAT実施のプロセス一つ一つに吟味を加え、改良版のCATを作成することが次のステップであり、この改良によって、科学技術英語の能力を効率的なかつ正確に評価することが近い将来実現するであろう。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計9件)

①Miyazaki, Y., Ikemoto, T. and Tanaka, S., Development of Web Application Supporting to Write Technical Documents in English: Using Corpus for Language Teaching, 査読有り, Proceedings of the ICTATLL 2010 Kyoto Conference, 2011, 149-158

②田中省作・宮崎佳典・池本孝徳・小山由紀江, 英作文支援のためのクラス n-gram モデルに基づいた文例の汎化, 審査無し, 統計数理研究所共同研究レポート, 254, 2011, 1-19

③小山由紀江・木村哲夫, Neural Test Theory を使った Can-do Statements の分析, 審査無し, 統計数理研究所共同研究レポート, 254, 2011, 59-77

④小山由紀江, Coh-Metrix を使った科学技術英語テキストの分析, 統計数理研究所研究レポート, 審査無し, 2012, 276, 81-96

⑤Miyazaki, Y., Tanaka, S. and Koyama, Y., Development and Improvement of a Corpus-based Web Application to Support Writing Technical Documents in English, ICCE2011, 審査有り, 2011, 263-270

⑥石川有香, 小山由紀江, 大学初年次学生の工学系準専門語彙の理解度, 中部地区英語教育学会紀要, 審査有り, 41, 2011, 257-264

⑦小山由紀江, 科学技術コーパスの分析に基づくアイテム・バンクの構築と潜在ランク理論によるコンピュータ・アダプティブ・テストの試行, 統計数理研究所研究レポート, 審査無し, 2013, 295, 29-49

⑧田中省作, 安東奈穂子, 富浦洋一, コーパス構築と著作権-Webを源とした質情報付き英語科学論文コーパス-, 英語コーパス研究, 審査有り, 19, 2012, 31-42

⑨杉森直樹, WordNet を用いた英語多義語リストの構築, 立命館言語文化研究, 審査無し, 24(4), 2013, 171-182

[学会発表] (計9件)

①小山由紀江, Neural Test Theory を使った Can-do Statements の分析, 言語研究と統計 2011, 2011年3月14日, 統計数理研究所 (東京)

②小山由紀江・田中省作, 構造を考慮した科学英語の特徴語抽出, 外国語教育メディア学会第50回研究大会, 2010年8月3日, 横浜市立横浜フロンティア高等学校 (神奈川)

③小山由紀江, Moodle を使ったコンピュータ適応型言語テストの実施, 大学英語教育学会, 宮城大学 (宮城)

④宮崎佳典・田中省作・小山由紀江, 技術文献コーパスを活用した英語技術文書作成支援 Web アプリケーション開発, LET2011, 2011年8月8日, 名古屋学院大学 (愛知)

⑤Koyama, Y. and Kimura, T., Linking Can-do Statements with Language Test

s Using Neural Test Theory, JACET,
2011年9月1日, 西南学院大学 (福岡)

⑥杉森直樹, 多義語を用いた日本人英語学習者の語彙知識の測定について, 第37回全国英語教育学会山形研究大会, 2011年8月21日, 山形大学 (山形)

⑦小山由紀江, 科学技術コーパスの分析に基づく アイテム・バンクの構築と潜在ランク理論 によるコンピュータ・アダプティブ・テストの試行, 言語研究と統計 2011, 2011年3月14日, 統計数理研究所 (東京都)

⑧田中省作, 富浦洋一, 宮崎佳典, 小林雄一郎, 徳見道夫, 機関リポジトリを活用した部局別英語表現リストの作成支援, 情報処理学会第75回全国大会, 2013年03月08日~2013年03月08日, 東北大学 (宮城)

⑨Sugimori, N., Exploring a New Battery of Multi-Facet Vocabulary Tests 2: Measuring the Polysemy Knowledge of L2 Learners, JACET CONVENTION 2012: The 51st International Convention, 2012年08月31日~2012年09月02日, 愛知県立大学 (愛知)

6. 研究組織

(1)研究代表者

小山 由紀江 (KOYAMA YUKIE)
名古屋工業大学・工学研究科・教授
研究者番号: 20293251

(2)研究分担者

杉森 直樹 (SUGIMORI NAOKI)
立命館大学・情報理工学部・教授
研究者番号: 40216338

田中 省作 (TANAKA SHOSAKU)
立命館大学・文学部・教授
研究者番号: 00325549

(3)連携研究者

()

研究者番号: