

令和 6 年 6 月 25 日現在

機関番号：32657

研究種目：基盤研究(C)（一般）

研究期間：2021～2023

課題番号：21K00716

研究課題名（和文）基本語彙の受容的・発表的知識の多次元の変容の検証

研究課題名（英文）Investigation of Multidimensional Transformation of Receptive and Productive Knowledge of Basic Vocabulary

研究代表者

佐藤 研仁（Sato, Kento）

東京電機大学・工学部・講師

研究者番号：60835924

交付決定額（研究期間全体）：（直接経費） 3,200,000円

研究成果の概要（和文）：語彙知識を多面的に調査した先行研究はあるものの、語彙知識にまつわる各知識要素が互いどのように関わっているのか、そしてそれぞれの知識要素はどのような順序で習得されるのかについてはあまり一貫した研究結果が得られていない。本研究では、日本人英語学習者を対象とした語彙知識を多面的に測定するテストを開発し、語彙知識の発達プロセス解明を図ることを目的としていた。2種類のテスト（受容再生・高頻度の意味、受容再生・低頻度の意味）を行い、項目応答理論を用いて分析を行った。その結果、「日本人英語学習者が必ずしも英語圏における最頻の意味を最初に獲得しているわけではない」ということがわかった。

研究成果の学術的意義や社会的意義

本研究は、日本人英語学習者の語彙知識を多面的に調査し、その発達プロセスを明らかにすることを目的としている。特に、英語の単語の意味をどのように習得するかを調べるために、2種類のテストを実施した。その結果、英語圏で頻繁に使われる語彙の意味を日本人学習者が最初に習得するとは限らないことが分かった。つまり、日本人は英語の単語を学ぶ際に、必ずしも英語圏の人々と同じ順序で意味を覚えるわけではないということがわかった。

研究成果の概要（英文）：Although there have been previous studies that investigated vocabulary knowledge from multiple aspects, there have been few consistent findings on how the various knowledge components involved in vocabulary knowledge relate to each other and in what order each knowledge component is acquired. The purpose of this study was to untangle the developmental process of vocabulary knowledge by developing a test that measures vocabulary knowledge from multiple aspects for Japanese learners of English.

Two types of tests (receptive-recall, high-frequency meaning, and receptive-recall, low-frequency meaning) were administered and analyzed using item response theory. The results showed that Japanese learners of English do not necessarily acquire the most frequently used meaning in English-speaking countries first.

研究分野：語彙習得

キーワード：語彙テスト 語彙知識 多義語 項目応答理論

1. 研究開始当初の背景

語彙知識は、広さと深さという2つのメタファーで区別されてきた (Anderson & Freebody, 1981)。語彙知識の広さは「知っている単語の数」であり、語彙知識の深さは「単語をどのくらい深く知っているか」である。従来は、最初は語彙知識を広くし、ある一定のサイズに至った後で、深さの面を充実させる方法が奨励された。語彙能力構成要因の中で最も重要なのは語彙知識の広さであり、コミュニケーション能力を予測するための指標として最も有効と考えられてきた。

しかし、語彙知識の広さを測定するテストは、実際の伝達能力とは関連がない表面的な知識を測定しているとの批判もある。語彙知識は意味を「知っている」か「知らない」かの二分法で分けられるものではなく、Paribakht and Wesche (1993)の例で示すと、図1のように単語を全く知らない状態(A)から、受容語彙として中核的な意味が定着し(B)、やがて周辺的な意味、発音、用法などの語彙知識の深さを増し(C、D)、それから発信語彙として獲得(E)される、というように獲得の度合にはいくつかの階層性があると考えられる。

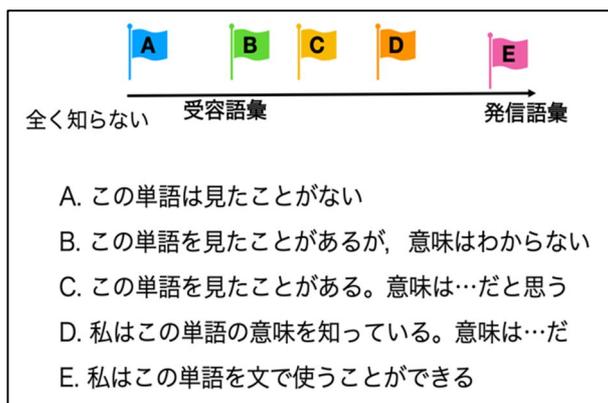


図1 受容語彙から発信語彙への段階的発達

また、どの段階までを「知っている」と線引きするのかによっても、推定される語彙サイズは異なる。さらに、受容語彙と発信語彙の関係性や習得プロセスは明らかになっておらず、語彙知識が受容語彙として定着した状態(D)から発信語彙として使える状態(E)へと段階を踏むとは限らない。

Laufer and Goldstein (2004)は、語彙習得で重要なのは語形と意味の間の結びつきで、結びつきの強さには4段階が存在すると仮定した。調査の結果、受容認識、発表認識、受容再生、発表再生の順に結びつきが強まることがわかった。Lauferらの調査は2つの問題点がある。まず、測定する語彙知識の側面である。語彙サイズの推定には、受容面と発表面の両面からが考えられるが、受容面(L2 L1)の方が、発表面(L1 L2)よりも易しいのは当然である。

(1) 発表再生 (L1 L2 翻訳)	e. _____	強調
(2) 受容再生 (L2 L1 翻訳)	_____	emphasis
(3) 発表認識 (L1 L2 選択)	強調	a. emphasis b. alarm c. damage d. castle e. 正解なし
(4) 受容認識 (L2 L1 選択)	emphasis	a. 認識 b. 理解 c. 強調 d. 協力 e. 正解なし

図2 Laufer and Goldstein (2004)のテスト形式

次に、問題の形式である。選択肢から解答を選ばせる認識課題は、翻訳や再生課題よりも、25%の期待値の分だけ有利である。Lauferらの結果では、例えば2000語レベルでは、期待値分を差し引くと、発表認識と受容再生の得点差はわずかであった。

発表認識と受容再生のテスト形式を、大幅に改善する必要があった。

語彙知識を多面的に調査したこれまでの研究 (e.g., Webb, 2005)から、語彙知識の諸側面が同時進行的に徐々に獲得されていくことが立証された。またその一方で語彙の知識の深化には階層性があることも示された (e.g., Chen & Truscott, 2010)。

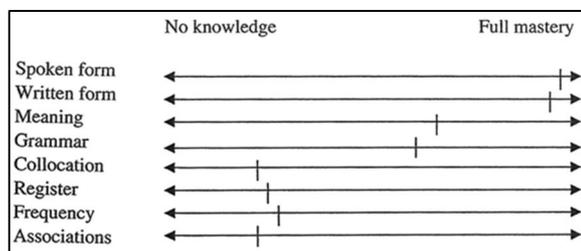


図3 語彙知識の階層性 (Schmitt, 2010)

しかし、先行研究の結果では各知識要素の習得順序や、それらの相互関係についてはあまり一貫性が見られない。以下の原因が考えられる。1)参加者の習熟度が異なる(標本依存の可能性)。2)正答率等をもとに習得順序を検討している(項目依存の可能性)。

このような点を克服するために各得点を比較できるような統一された1つの指標が必要である。項目応答理論 (IRT) では、能力値 ( ) という同一の尺度を用い、異なる項目から構成されるテストの比較が可能である。IRTを用いることで、語彙知識 ( ) の多面的側面を一つの共通した指標で測定し、習得状況の比較がより高い精度で行えると考えられる。

## 2. 研究の目的

上述の学術的背景をうけ、本研究では主として下記の2点について調査を行うこととした。当初の計画では、学習者の英語の基本語彙の習得を多面的に測定するテストを開発し、受容的知識から発表的知識への習得プロセスにおいてどのような変容をたどるのかを明らかにすることを最終目的としていたが、日本人英語学習者の多義語知識の測定方法の確立が急務であると判断されたため、この点については調査を断念せざるを得なかった。

### (1) 日本人英語学習者の多義語の意味の獲得順序の特定

英語基本語彙は多義語であり、複数の意味を持っている。例えば、*line* の場合は多くの意味があるが、意味頻度順に「綱」「線」「家系」などで使用される。日本人英語学習者が、多義語の意味のうち、どの意味をどの順番で獲得するのかを調査することを目的の一つとした。

### (2) 多義語テストの項目の選定

最終的に項目応答理論を用いて等化をする際、共通項目を必要とする。そのため、本調査では、共通項目の妥当性検証を行うことを第二の目的とした。これには目標語の妥当性検証、文脈の妥当性検証が含まれる。

## 3. 研究の方法

上記の目的を達成するために以下のリサーチクエスチョンを立てた。

- (1) 項目応答理論による適合度に当てはまらない項目は、どの程度あるか。(構造的側面)
- (2) 受験者の幅広い潜在特性(latent trait)を測定できる程度、項目困難度に十分な幅があるか。(構造的側面)
- (3) 各テストは、想定する能力値の範囲で、より精度の高い推定を行えるか。(一般化可能性の側面)

\*括弧内はMessick(1995)による妥当性の側面

## 実験

日本人英語学習者370名を対象に調査を行った。目標語としてMaby(2016)の20語を使用した(*line, order, board, cover, body, course, face, hand, air, arm, cut, heart, position, character, class branch, form, point, case, head*)。これらの語は3つ以上の高頻度の意味を持ち、かつ新JACET8000の2000位以内の語である。問題の項目はMaby(2016)の意味頻度(高・中・低)に基づき、1語につき3つの意味について項目を作成した。文脈の作成には、Maby(2016)のほか、Oxford Advanced Learner's Dictionary、Longman Dictionary of Contemporary English、Cambridge Learner's Dictionaryに記載されている例文を参考にした。テストの実施は、授業内で行われた。LMS(Learning Management System: 学習管理システム)を使用し、オンライン上でテストを実施した。

テスト項目は60項目で、それぞれの受験者に対し、ランダム順で出題された(図4)。

次の文の赤文字で示した単語の意味のうち、文脈に合う意味を日本語で教えてください。

例: *Their new house has four **stories** ( )* →階

*I have a house in the **heart** of the city.*

回答欄が空欄のままでは、クイズを終了できません。

分からない場合には **x** (全角または半角のエックス) を回答欄に入力してください。

図4 出題画面サンプル

採点は、英語に精通する日本語母語話者3名で行い、全回答の中から正解・不正解の振り分けを行った。その際に、正解であれば1点、不正解であれば0点とした。部分的正解はないものとした。採点の基準には、Maby(2016)、上記の英英辞典のほか、リーダーズ英和辞典、ジーニアス英和大辞典を参考にした。

## 4. 研究成果

受験者370名のうち、11名のデータを分析から除外した。実験のテストの結果、信頼度係数(クロンバックの )は、.86であった。

表 1 基本統計量

	最低点	最高点	平均	95%信頼区間	標準偏差	
受容再生	1	44	22.91	[22.1、 23.71]	7.76	.86

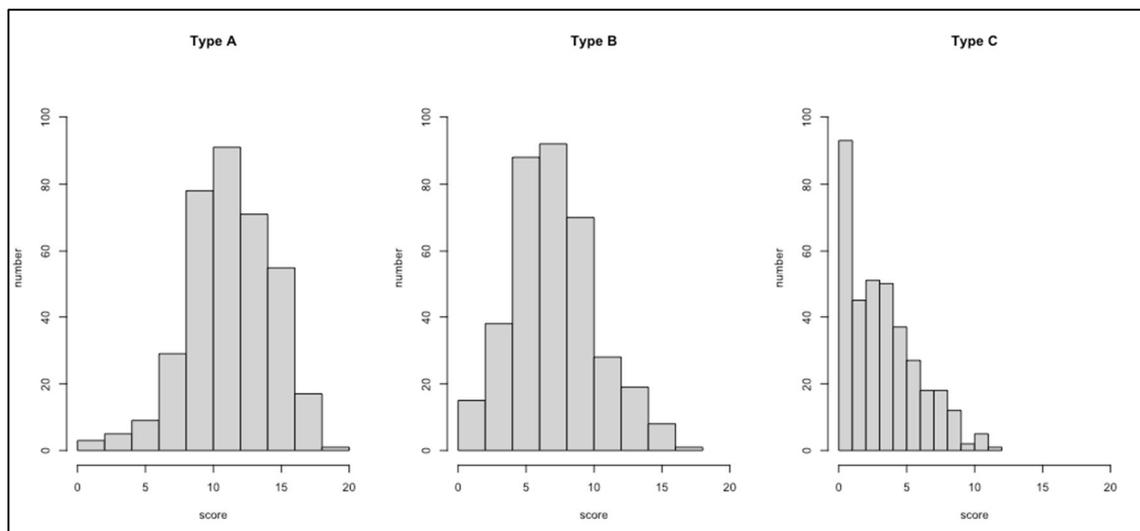


図 5 意味頻度別得点分布：Type A = 高頻度、Type B = 中頻度、Type C = 低頻度

図 5 は意味頻度別に得点の分布を表している。意味頻度が低い項目になるにつれ、得点が左によっていく、つまり得点が下がっていることがわかる。

項目応答理論モデルでは、観測データがモデルに当てはまっているかを確認する必要がある。項目応答理論分析には、R Studio の *irt* (Partchev, 2017) パッケージを使用し、2 パラメータ・ロジスティックモデル (2PLM) を用いた。

60 項目のうち、4 項目 (6.67%) が適合が悪い結果となった (図 6)。

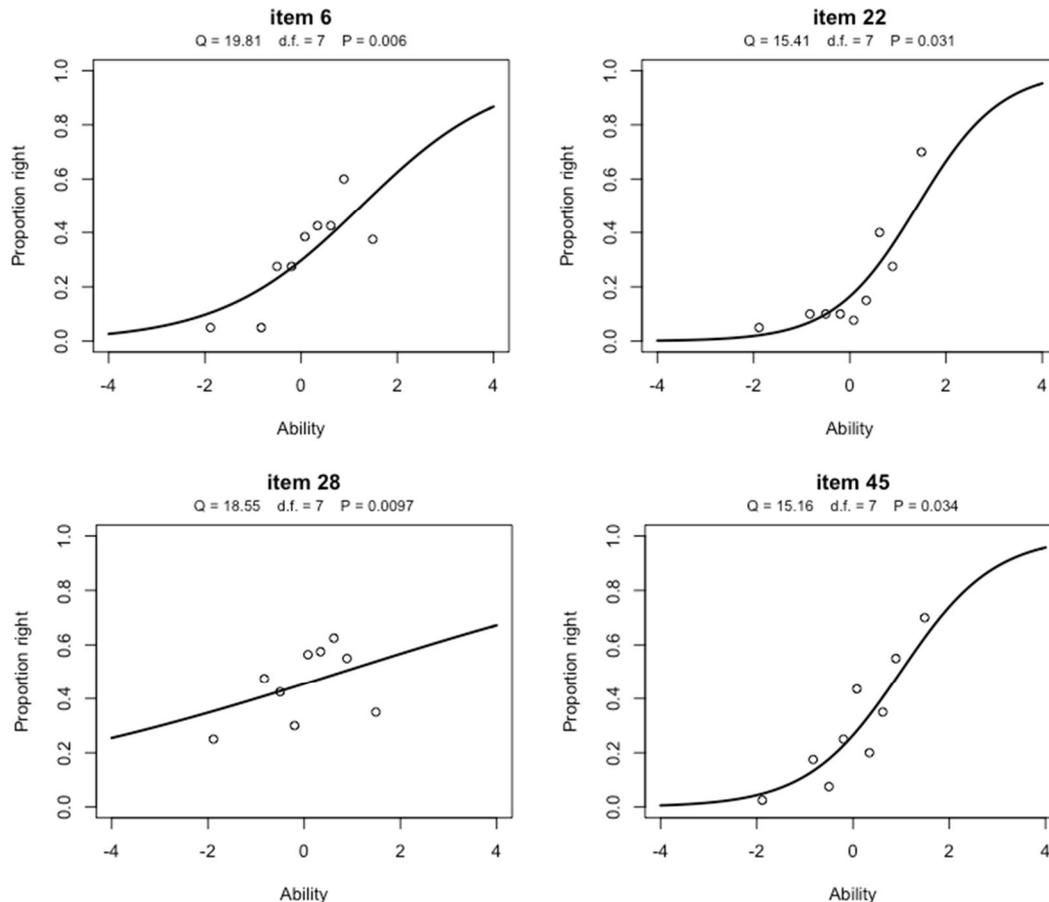


図 6 IRT に基づく項目特製曲線と実際の正答率 (適合が悪い項目)

つまり、RQ1 の項目応答理論による適合度に当てはまらない項目は、どの程度あるか、に対する答えは、4項目であった。特に適合度が悪かった項目は item 6 の *Please don't sit on the arm of the chair.* (意味頻度: 低) (正答例: 肘掛け、手すり、腕置き) と item 28 の *There is a plastic cover over the meal.* (意味頻度: 高) (正答例: 覆い、蓋、被せるもの) だった。これらの項目は、意味はわかるが、当てはまる日本語が思いつかなかったという可能性がある。今後、このような項目は再度検討していく必要がある。

RQ2 の受験者の幅広い潜在特性(latent trait)を測定できる程度に項目困難度に十分な幅があるか、に関しては図7に示す通り、全ての意味頻度を含め項目困難度を考慮すると、-4.0 から 4.0 の幅広い困難度をカバーしていると言える、という結果になった。しかし、今回の一度の調査だけでは集団依存性を完全に避けることは難しいため、今回よりも習熟度が高い集団、低い集団で今後実施していく必要がある。

RQ3 の各テストは、想定する能力値の範囲で(容易な項目が能力値の低い範囲で、難しい項目が能力値の高い範囲で)より精度の高い推定を行えるか、に関しては、図8の項目情報曲線を考慮すると、おおよそ想定した能力値の範囲で高い推定を行えていると考えられる。特に、意味頻度が中程度の項目群と意味頻度が低い項目群に関しては高い推定を行えている。

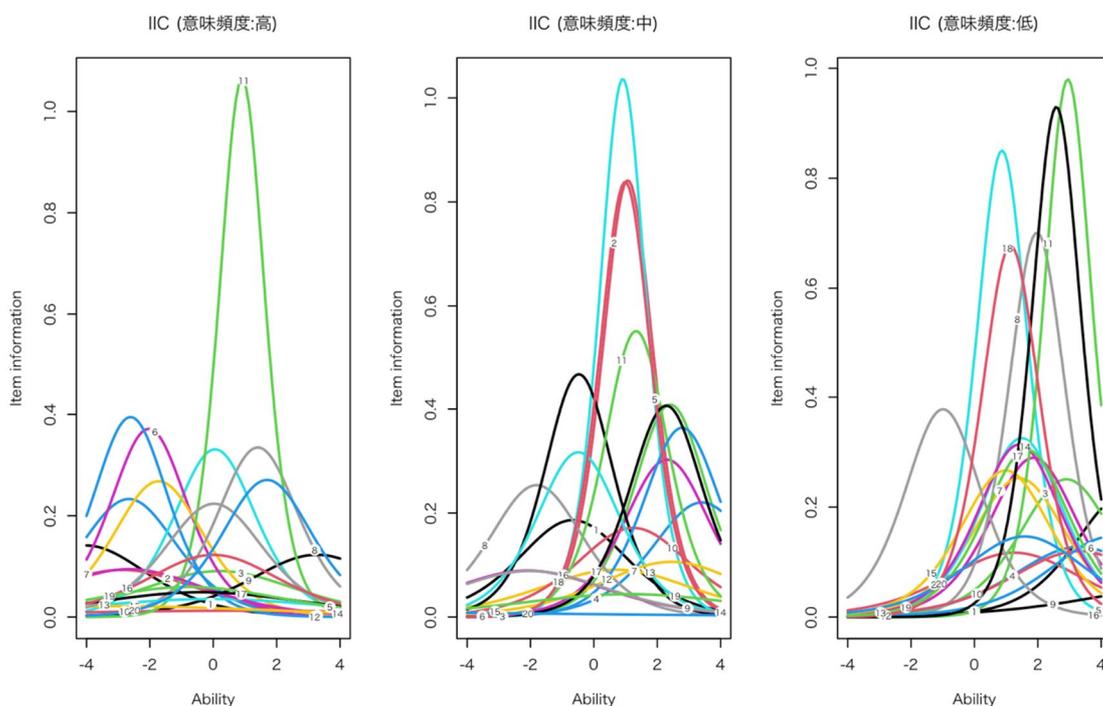
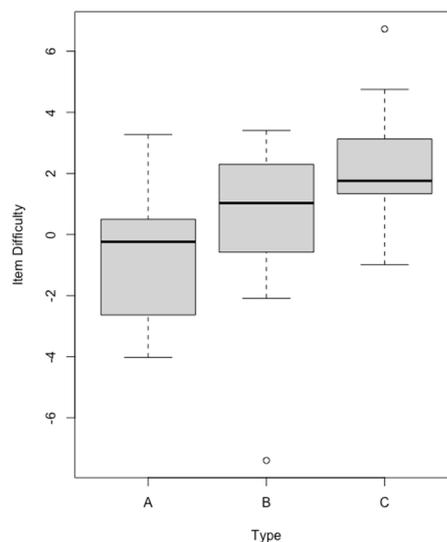


図8 意味頻度別 項目情報曲線

意味頻度が高い項目のテストにおいても(図8左)能力値の比較的高いところになっている( = 1 付近)項目がある原因は、再生テストが形式上、難易度が比較的高いためと考えられる(e.g., González-Fernández & Schmitt, 2020)。今後、認識形式の項目を加えていくことで、能力値のより低い範囲もカバーできるのではないかと考えられる。

今回の調査を通し、語彙知識を多面的に測定するテストを開発するために重要な知見が得られた。まず、基本語彙の持つ複数の意味のうち、意味頻度が高い順に意味を獲得している傾向がある一方で、必ずしも最も頻度が高い意味を最初に獲得しているとは限らないという点である。これは、学習者や語によって外れ値となる例が見られた。

また、日本語で意味を書かせるテストには限界があることが明らかになった。語彙知識の側面のうち、「意味がわかる」ということが必ずしも日本語で相応しい語が書けるということではないことから、受容知識を測定するテストを開発することは非常に難しいと考えられる。

5. 主な発表論文等

〔雑誌論文〕 計0件

〔学会発表〕 計1件（うち招待講演 0件 / うち国際学会 0件）

1. 発表者名 佐藤 研仁, 相澤 一美, 磯 達夫, Paul Nadasdy
2. 発表標題 基本語彙の習得プロセス解明に向けた多面的診断テストの開発 -経過報告-
3. 学会等名 大学英語教育学会 英語語彙研究会 年次大会
4. 発表年 2022年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究分担者	相澤 一美  (Aizawa Kazumi)  (00222448)	東京電機大学・工学部・教授    (32657)	
研究分担者	磯 達夫  (Iso Tatsuo)  (40438916)	東京電機大学・工学部・教授    (32657)	
研究分担者	ナダスティ ポール  (Nadasdy Paul)  (60807303)	東京電機大学・未来科学部・准教授    (32657)	

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------