

外国語教育メディア学会（LET）関西支部 メソドロジー研究部会

2015年度 第7号報告論集

(2015年9月)

和英辞書の記述改善に向けて

—日本語オノマトペの分析とその翻訳—

仁科 恒徳 1-14

留学環境での言語学習を研究対象とする際に注意すべきいくつかのこと

—効率的な調査をおこなうために—

三上 仁志 15-24

適応型テストへの応用

—ベイス EAP 推定法とフィッシャー情報量の概説—

住 政二郎 25-39

和英辞典の記述改善に向けて —日本語オノマトペの分析とその翻訳—

仁科 恭徳
神戸学院大学

Toward Improving the Compilation of Japanese-English Dictionaries: Investigating Japanese Onomatopoeias and their Authentic Translation

NISHINA, Yasunori
Kobe Gakuin University

Abstract

This article presents one of the methodologies for Japanese-English dictionary compilation especially in the case of treating untranslatable linguistic items, namely Japanese onomatopoeias. The study is primarily based on the corpus research (viz. KOTONOHA) of Japanese onomatopoeias and the informant's translation of the data extracted from the corpus. The method exemplified in this study will hopefully contribute to both the future bilingual lexicography and the material design of Japanese language education targeted for foreigners.

Keywords: Translation Studies, Lexicography, Japanese Onomatopoeia

1. はじめに

言語的祖先が全く異なる日英語間では、双翻訳方向において訳し難い言語項目が多々見つかる。人は国ではなく言葉に住むという言葉を借りれば(飯島 2004, p.243), 言葉にはその国の文化や慣習、歴史、思想が反映されていることから、翻訳者は起点言語 (source language) と目標言語 (target language) の双方に関して、これらを熟知していなければ正確な翻訳を産出することは難しいと言える。¹

しかしながら、二つの言語背景にある(国)の文化や慣習を理解し熟練した翻訳家であっても、翻訳に戸惑う場面に遭遇することは少なからずある。翻訳学の大家 Peter Newmark も、何事も翻訳自体は可能であるが、時には目標言語を通して起点言語で表現・描写された内容と同程度のインパクトを与えることは難しいと示唆する。また、翻訳前後に生じるニュアンスの差異がどの程度であれば質的に許容範囲であるのか、主観的な翻訳においてこれを数値化することは難しい。結局は、起点言語と目標言語の双方の文化や慣習、言語的知識を習得した経験豊富なバイリンガルにその判断を委ねるのが一つの得

策であろう。

“... everything without exception is translatable; the translator cannot afford the luxury of saying that something cannot be translated ... whilst translation is always possible, it may for various reasons not have the same impact as the original” (Newmark, 2003, p.6)

そこで、本稿では、Newmark(2003)の「翻訳不可能なものはない」という信念に従い、飯島(2006)で翻訳不可能(もしくは翻訳し難い)とされた日本語オノマトペを対象に、KOTONOHA コーパス(現代日本語書き言葉均衡コーパス)から現代日本語において頻度の高いオノマトペと動詞の共起表現を抽出し、両言語に精通したバイリンガルにそれらを日英翻訳してもらうことで、量・質の双方的アプローチから和英辞典に掲載すべき例文算出までのプロセスの一端を示す。本稿で示すプロセスを参考に、現行の辞書(特に海外で出版される日本語学習者向け日英・日本語辞書)や語学テキスト(特に日英翻訳・日本語学習テキスト)の記述がより信頼性あるものへと改善されることを願い、その一方法論を示すことが本稿の目標である。

2. The tip of my tongue

上級レベルの日本人英語学習者(もしくは英米人日本語学習者)であっても、適切な英語(もしくは日本語)が発話の途中で思い浮かばないthe tip of my tongue現象に遭遇する機会は多く、日本語オノマトペ表現の翻訳もその一つであろう。² 特に論者のアメリカ人への日本語教育並びに日英・英日翻訳教育の経験上、[副詞的オノマトペ+動詞]のような表現が英語には存在せず、その解釈と翻訳作業に苦しむ学生が多くいた。³ 彼らの多くが、HarperCollins 社から出版されている Collins Pocket Japanese Dictionaryなどの紙面版英日・日本語辞典や、ローマ字検索が可能な Bravolol Limited 社などから公開されているフリーの辞書アプリをスマートフォンにインストールして授業に挑んでいた。

無論、スピーディーな検索を重視するこのような辞書には、多数存在する日本語オノマトペが全て収録されているわけではなく、収録されている項目でもその意味が詳細に記述されているわけではない。日本語オノマトペに限定すると、専門辞典は別にして、一般的な大辞典クラスにおいても収録・記述されていないケースが多い。⁴ 日本人的感覚の獲得と日本語らしい日本語の産出を目指す中・上級日本語学習者にとって日本語オノマトペの正確な使い方の習得や、英語で表現できることは重要であり、そのような参考資料の存在は必要不可欠である。言いたい時に言えない、調べても掲載されていないことほど、学習者にとって辛いことはない。

日本語オノマトペが辞書に詳細に掲載されていない理由の一つは、基本的に逐語的に英訳することが難しいというオノマトペ本来の性質による。よって、翻訳時には共起語とのセットフレーズを翻訳ユニットと捉えるか、長々とパラフレーズするかのどちらかが常である。特に、Baker(1992)はパラフレーズ原則なるものを提唱し、そのうち日本語オノマトペに関してはその意味を分解することで翻訳の糸口を見出す Translation by paraphrasing using an unrelated word を用いるのが有効な方策の一つであると

論者は考える。

If the concept expressed by the source item is not lexicalized at all in the target language, the paraphrase strategy can still be used in some contexts. Instead of a related word, the paraphrase may be based on modifying a superordinate or simply on unpacking the meaning of the source item, particularly if the item in question is semantically complex. (p.38)

我々日本人にとって日本語オノマトペは意味的に複雑というわけではないが、それを説明しろと言われると一筋縄ではいかない。意味として定義する場合、その輪郭がはつきりせず、類似したオノマトペ間の違いを可視化することで初めてその輪郭や記述的な実態が浮かびあがる。次節では、日本語オノマトペを正確に英訳する上で必要な分析を試行し、日本語オノマトペの輪郭を求める。

3. 日本語オノマトペ再考

3.1 意味の輪郭化

一般的には、擬声語と擬音語の両者を総称してオノマトペと呼ぶ。松岡他(2000, p.382-384)では、様態副詞の例として擬声語と擬態語を挙げ、「何かの音をまねたもの」を擬声語、「ある動作をイメージ的に表したもの」を擬態語と定義し、特に「とんとん」や「そろそろ」など同じ語を繰り返すものを畳語と呼んでいる。オノマトペの多くは畳語だが、「バタンと」や「ねちやっと」など畳語に該当しないものも存在する(松岡他, 2000, 同上)。

日本語には類似したオノマトペが散見されるが、そのようなシノニムの意味も互いに微妙に異なる。Fukuda (2003, p.26)は「こんこん」と「ごほんごほん」、「ぱたん」と「ばたん」の違いを以下のように記述している。

表1

こんこん, ごほんごほん, ぱたん, ばたんの記述的違い (Fukuda, 2003, p.26 より)

こんこん	the sound of a dry, repetitive cough, in which the sound is relatively small with little reverberation
ごほんごほん	the resounding, repetitive sound of a wet cough with greater reverberation
ぱたん	indicates that a thin, hard object has collided with something else, one time
ばたん	indicates that a hard, heavy object of considerable size has collided with another object, one time

母語直感がある我々日本人には、この違いは納得できよう。このような記述的違いを可視化するには、語義を確認するためのテストを実施すればよい。例えば、「重い戸をバタンと閉めた」とは言えて、「軽い戸をバタンと閉めた」とは言えず、同様に「重い戸をパタンと閉めた」とは言えないが「軽い戸をパタンと閉めた」と言うとしつくりくる。このような共起語との相性テストを母語話者に調査することは、オノマトペの意味や定義の輪郭を知る上で極めて重要である。

特に, Fukuda(2003, p.29-31)によれば, 日本語オノマトペの文法機能は, (1) As Adverbs Modifying Regular Verbs, (2) Combined with する Suru or やる Yaru, (3) As Adjectivelike Words Combined with the Copular だ Da, (4) As Nominals Followed by the Particle の No in Adjective Phrases, (5) As Adverbial Phrases Followed by に Ni の 5 種類に分別できる。この各文法機能との相性をテストすることで各オノマトペの表面的な違いが可視化できる。例えば、「がんがん」は、「がんがん飲む」や「がんがんする」など(1)と(2)の文法機能は備わっているが、「がんがんだ」や「がんがんの頭」, 「がんがんになる」とは言えず,(3),(4),(5)いずれの文法機能も備わっていない。一方、「すべすべ」は、「(肌が)すべすべする」, 「(肌が)すべすべだ」, 「すべすべのお肌」, 「すべすべになる」とは言えるが(1)の文法機能にあてはまる用例は見当たらない。このように各オノマトペに備わっている文法機能の組み合わせは異なり, 輪郭を知る上での有効な手段となる。

つまり, オノマトペ間の中核的語義の違いを可視するには, まずこのような文法機能の差異や共起語との相性を調査することが有効なアプローチの一つと言える。飯島(2006)の言う「キラ・ピカ」系オノマトペも, 共起語調査からなんらかの違いが顕著となる。表2は「ビカビカ」と「ギラギラ」も加えた上で、論者が日本人インフォーマントに実施した簡単な共起語テストの結果の平均値である。⁵

表2

インフォーマント調査によるキラ・ピカ系オノマトペの共起結合度

	輝く	光る	する	の頭	に磨く	になる
ピカピカ	○	○	○	○	○	○
キラキラ	○	○	○	×	×	×
ビカビカ	×	○	×	×	△	△
ギラギラ	△	△	○	×	×	×

表2から, 意味が似通う4つの「キラ・ピカ」系オノマトペと各共起語との結合度には違いが見られ, この違いこそがオノマトペ間の中核的語義の違いを示す格好の材料となる。⁶ 例えば, 「ピカピカ」は物体の光沢を肯定的に表現したい時に, 「ギラギラ」はその輝きに嫌悪感を抱く際に用いられる傾向が強くこれを記述に活かせばよいのである。

3.2 日本語オノマトペを訳語に含む英語動詞

日本語オノマトペは単体で翻訳することが難しいため, 共起表現全体をセットフレーズとして翻訳することも多い。例えば, 「ゴクゴク飲む」のように様態副詞オノマトペは動詞と頻繁に共起し, 英語では「ゴクゴク」の様態を語義に含む動詞*gulp* 1語で表すことができる(Nishina, 2013, p.207-208; Takaie, 2002)。

また, 語義に同一の様態副詞オノマトペ(と動詞)が含まれるいくつかの英語動詞においても, その語義概念はいくばくか異なる。例えば, 飯島(2004, p.49-50)によれば, 自動詞 *glitter, shine, sparkle* はどれ

も「キラキラ」もしくは「ピカピカ」の様態副詞オノマトペが訳語に含まれているものの(飯島の言う「キラ・ピカ」系オノマトペ), 中心的語義概念はそれぞれ異なることから星の輝きを描写する際にはいずれも用いることができない。*glitter* は「派手に際立つ」という「けばけばしさ」に重点が置かれ,*shine* は金具の艶や磨きあげた靴などの描写にも使われるなど「広い適応力」がある。また,*sparkle* の中核的語義概念は「強烈なエネルギー」でシャンパンなどの泡立ちの勢いを描写する際にも用いられる。そして、飯島は,*twinkle*(原意は「またたく光を輻射状に放散する」)のみが小刻みに震える星の独特の輝きの描写にふさわしいと指摘する。このように、日本語のみで考えると類似したオノマトペの間にも違いが見られるが、同一の日本語オノマトペを語義に含む英語動詞間においても、その様態に違いが見られる。

この違いの要因を翻訳的観点から考察しよう。例えば、ある日本語オノマトペを A としよう。A には B や C, D という日本語動詞と共に起すると仮定する。この場合, [A+B], [A+C], [A+D] が翻訳ユニットとなり、これらに対応する英語動詞の違いは日本語の共起動詞の違いに起因することになる。文中の中核的な意味を構成する上で、基本的に adjunct(付加詞) や manner adverb (様態副詞) は義務的な項ではなく⁷、動詞こそが項の数や形式を決定づけることは理論言語学(特に語彙意味論)の分野では主流であり、翻訳的観点から言えば、日本語動詞の違いがこれら英語動詞間の違いに起因すると予測することは自然な摂理である。この考え方は、意味と形式が対応しているという前提にたてばもっともあるが、実際はそう簡単に一筋縄には解決できないのがオノマトペの扱いである。

例えば、次に、一つの英語動詞に固執してその訳語を概観してみよう。例えば、前述の英語動詞 *twinkle* の日本語訳には、様々な日本語オノマトペが出現する。表3 は *twinkle* に対応する日本語オノマトペのまとめで、「きらきら」「ちかちか」「ぴかぴか」のような疊語や「きらり」「ピカッ」となどヴァラエティに富んでいる。飯島自身も「こと星に関する限り全てが光り方の様態に応じて微妙に使い分けられる」(飯島, 2004, p.47-48)と指摘している。前述のように、副詞オノマトペに逐語的に対応する英訳は基本的に存在せず、あくまで「(副詞+)動詞」のみ英語動詞(表3 では *twinkle*)に対応する。

表3

星の輝きを表す副詞・動詞 (飯島, 2004, p.47-48 から一部抜粋)

	日本語	英語	
副詞+動詞	きらきら・ちかちか・ ちらちら・ぴかぴか(4) きらり(と)・きらッ (と)・ちかッ(と)・ぴか り(と)・ぴかッ(と)・ぴ ッかり(と)(6)	光る・輝く (10)	動詞 (自・他) <i>twinkle</i>
動詞=オノマトペ+す	きらきらする・ちかちかする・ちら		

る	ちらする・ぴかぴかする(4)		
動詞(=オノマトペ+接尾辞めく)	きらめく		

飯島(2006)も認めるように、表3に挙げているような各日本語オノマトペに対応する語が英語には存在せず、*twinkle* がこれら全ての日本語オノマトペと対応且つ内包していることから、その様態「ふるえる光を放射状に放つ」の幅は意味的にかなり広い。⁸ 飯島曰く「オノマトペ副詞とは、視・聴覚印象そのものをそのまま音声に写したもの」(同上, p.51-52)であり、「同じく光る様を極微の差異まで見極めて、あくまでオノマトペによって忠実に表現しようとする」(同上, p.51-52)のであるが、この例からも実際は日英間において 1 対 1 の対応関係になっておらず、複数の日本語動詞とオノマトペのセットが *twinkle* に対応している。

逐語的関係が保たれない以上、視点を変え、量的観点から翻訳傾向を見極めることは、このような問題に対処する有効な手法の一つである。例えば、Nishina(2013)では、日本語オノマトペの実態を、「笑う」という動詞と共に起する「くすぐす」、「うふふ」、「ふふふ」、「げらげら」、「からから」などのオノマトペに限定し、日英翻訳コーパス(通称、パラレルコーパス)を用いて計量的観点から調査した。表4は、Nishina(2013)の「くすぐす+笑う」の英語翻訳のまとめである。

表4

「くすぐす+笑う」に対応する英語動詞

	<i>giggle</i>	<i>snicker</i>	<i>laugh</i>	<i>titter</i>	<i>snigger</i>	<i>chuckle</i>	Total
平仮名	11	2	4	1	2	0	20 (67%)
片仮名	3	4	0	1	1	1	10 (33%)
Total	14 (47%)	6 (20%)	4 (13%)	2 (7%)	3 (10%)	1 (3%)	30 (100%)

表4から、*giggle* が最も好まれて翻訳時に使用されていることが分かる。確かに *snicker* や *snigger*, *titter*, *chuckle* なども散見されるが *giggle* が約半数を占めており、このような調査からある程度の翻訳傾向を捉えることも可能である。⁹ これら日本語オノマトペ(もしくは動詞との共起表現)の翻訳実態を概観すると、日英・英日の双翻訳方向において 1 対多の関係というよりは、互いが複雑に絡み合っていることが分かる。

3.3 翻訳できない日本語オノマトペ

飯島(2006 p.54-59)は、現代日本語オノマトペ一覧と称し、日常的に使われている日本語オノマトペ800語中、純粹なオノマトペとは捉えられないものや特殊なオノマトペを排除した上で、五十音図各行第一段の音から始まる 165 語をまとめている。各日本語オノマトペには対応する英訳が付されているが、その数は少なく「あたふた」=*rattle* や「かさかさ」=*rustle* など僅か 45 種(27%)のみに与えられている。

いた。これらは全て、副詞として捉えることができるが、英訳では自・他動詞、名詞、形容詞など様々な品詞に変化する。

表5

飯島(2006)に掲載された翻訳可能なオノマトペ⁹

がちっ *clank*, かちやかちや *clank*, がちやがちや *clank*, がちやん *bump*, かちん *clank*, がつがつ *gobble*, がぶがぶ *quaff*, *guzzle*, がみがみ *snap*, がやがや *buzz*, あたふた *rattle*, があがあ *squawk*, がくつ *drop*, かさかさ *rustle*, がさがさ *rustle*, かたかた *rattle*, がたがた *chatter*, がたん *bump*, かちっ *snap*, ざぶざぶ *splash*, たじたじ *flinch*, だらだら *slobber*, ぱくぱく *gobble*, ぱくり *snap*, がらん *clang*, がりがり *crunch*, かん *clang*, さあつ *swoop*, ざくざく *crunch*, ぱらぱら *patter*, ぱりぱり *crunch*, ぱりぱり *crisp*, わあわあ *bawl*, わいわい *jangle*, わななく *quiver*, ばしつ *flick*, はたはた *flutter*, ばたばた *flap*, ばたばた *flap*, ばたん *bump*, ぱちぱち *crackle*, ぱちやぱちや *splash*, ぱちん *snap*, ぱつ *flash*, ぱつたり *flop*(口語), はつはつ *chortle*

また、飯島(2006)は表5以外の残り73%は対応する翻訳語がないとする。この理由を二つ挙げ、「英語がそもそも、日本語のその語義の表す音や状態の特殊性を差異化していない(つまり感じとれない、気づいていない)」場合と、「それらが差異化されていても、その表現がオノマトペによらず一般の動詞の利用によって間に合わせてしまう」場合である。つまり、これら日本語オノマトペの翻訳時にはlost in translationが生じている可能性が高い。確かに、飯島が挙げるよう、「がばがばの大きな革手袋」や「しーんと静まりかえる」、「こんがり焼けた」などは明確にこの雰囲気を別言語で伝えるのは難しい。逐語的に対応する語が英語には存在しないのであれば、何かしらの翻訳基準を設け長々とパラフレーズしない限り、誤って他の語を訳語に割り当てる意味的にlost in translationが生じるか、翻訳 자체が不可能となる。同時に、このような日本語オノマトペの微妙なニュアンスは、言わずもがな日本語母国語話者でなければ理解できない点も多い。

4. 分析

上記の点をまとめると、日本語オノマトペの英訳は動詞とのセットフレーズを対象としても逐語的には訳せず、英訳が難しいとされる日本語オノマトペが存在することになる。この二つの問題点をなんとか解消し、これから和英辞典や日本語学習者用テキストの記述に貢献すべく、その方法論を探りたい。特に、前者に関しては、逐語翻訳が可能であれば、翻訳経験があまり無い辞書執筆者であっても対応できようが、オノマトペに関しては難しいといえよう。そこで、翻訳が難解とされるオノマトペに関して、日英間の言語と文化に精通し長年翻訳の仕事にも従事したバイリンガルに翻訳調査を実施する。¹⁰ 分析手順は、まず日本語コーパスから翻訳不可とされる副詞的オノマトペとその共起動詞を量的に抽出し、t-score順で高いものを選定する。そして選定した例文を翻訳調査することにした。

4.1 表現の選定

インフォーマント調査の実施にあたり、はじめに日本語オノマトペ表現の抽出を行う。逐語翻訳が難解と思われる日本語オノマトペは、紙幅の都合と調査協力者への労力・時間を勘案し、飯島(2006)で翻訳不可とされている45種をKOTONOHA コーパス(正式名、現代日本語書き言葉均衡コーパス)付属のコーパス検索アプリケーション「中納言」を用いて検索し(検索日時は2015年2月23日),一定の処理を加えた上でコーパス・コンコーダンサーで分析し、翻訳不可とされる日本語オノマトペの頻度リストを表6に完成させた。尚、KOTONOHA コーパスは、国立国語研究所が10年の期間を経て完成させた大規模日本語コーパスで、ユーザーの用途に応じて大納言、中納言、小納言の3種が用意されている。

次に、このリストの中で、平仮名・片仮名両表記の合計頻度が100を超えたtappuri, sappari, garari, pappiの計4種のオノマトペを調査対象とした。そして、これらの共起動詞も頻度が高いものを抽出することで、セットフレーズを含む例文を最終的な翻訳調査の対象とした。詳細な検索手順は平仮名表記と片仮名表記のオノマトペを書字形出現形で指定し、品詞も副詞とした。続いて、後方共起にオノマトペの直後を助詞、その次に動詞が出現する、いわゆるオノマトペ定型パターンで抽出した。図1は、中納言を用いた「ハラハラ」の抽出設定画面、図2はこの設定で抽出した計9例のコンコーダンスラインを示す。そして、この方法を各オノマトペの平仮名・片仮名表記に関して検索を繰り返し、作成したものが表6である。

図1. 中納言を用いた「ハラハラ」の検索方法

9件の結果が見つかりました。		□テーブルの幅を固定	
サンプルID	前文版	キー	後文版
OY04_04071	でなかった。」でももう(来週)にはいろいろハラハラみたいだから、(早くも)ドキドキ!	ハラハラ	[(は)(終わっ)ちゃうのかしら?](~;.)視聴率は、こちらのサイトより。 http://artv.info/ar0810.html
PS89_00017	がピクニックを楽しんでいた。ワインを飲んでババッキ(食べべ、チーズをかじる)。!	ハラハラ	[(と)いうよりババッキ(に)残ちる泡を眺めるながら、私(も)ちょっと懐かしく(だけ)。(お)
LB0_00001	【秋】内野さん(俳優) 我が家の娘は必ずアスリートに陥る。【出来得】であります。!	ハラハラ	[(と)](数ある)川の奥との(そして)のない(山)駆けっこであります。【都の市定ゴミ袋】が何
OC08_06759	なりません。もう真っ黒黒黒・!・!に(します)。そういうと、(皮)は!	ハラハラ	[(と)はがれます]。【(い)で水につけて(弱く)・!・!と(書いて)いる料理
OY04_00657	介して、打ち解けていく・!・!とい(う)ラブストーリー。ドロドロもしねければ、!	ハラハラ	[(も)](しません)ので、【(か)な(隠)のラブストーリー】を(楽しめた)方がオススメ
PB41_00056	ストレートとい(う)か勇気のある人(人)家敷は(投げき)れはし(い)かと	ハラハラ	[(と)](する)ほどでした。【(わ)たしの】町では、【(ま)】と(使)用(人)の食事を
PB10_00098	と、白い衣の女販たち三千人が身投げた(壁は)つづの花びらが外	ハラハラ	[(と)](数ある)さまで(似て)います。【(わ)しくも】美しい故郷に(ち)なみ、(後)
OY15_13258	ているヤツの物どが、毎日攻めてますし、に(う)して(いる)間に(も)	ハラハラ	[(と)](攻め込み)で(います)。【(下)】(奪)取(離)離(に)見えます。【(わ)たし】何(の)落ち葉(を)送(送)
OY13_02395	このさるの(家)構成は(=)・!・!・!(?)最後(まで)目が(離せ)ない内側(で)	ハラハラ	[(も)](した)さるの(は)娘女・!・!・娘娘(な)母親(に)肩(て)られて(来た)二十回

図2. 抽出した「ハラハラ」のコンコーダンスライン

表6

翻訳不可能オノマトペの頻度リスト (H は平仮名, K は片仮名を示す)

#	Onomatopoeia	H	K	Total	#	Onomatopoeia	H	K	Total
1	tappuri	420	6	426	24	bassari	14	5	19
2	sappari	163	12	175	25	gakkari	13	5	18
3	garari	79	70	149	26	waku-waku	9	8	17
4	parii	17	86	103	27	nayo-nayo	7	0	7
5	darari	83	13	96	28	zakuu	4	3	7
6	akkerakan	72	7	79	29	patan	6	0	6
7	gasshiri	69	5	74	30	gabuu	0	5	5
8	gacchiri	54	16	70	31	katsu-katsu	2	2	4
9	gatsun	7	53	60	32	zaa-zaa	2	1	3
10	zakkuri	47	9	56	33	gappuri	2	0	2
11	zawa-zawa	26	12	38	34	barii	2	0	2
12	saku-saku	7	31	38	35	appu-appu	1	0	1
13	nami-nami	37	0	37	36	kachi-kachi	0	1	1
14	kara-kara	23	11	34	37	dabu-dabu	1	0	1
15	kara-kara	23	11	34	38	yakimoki	1	0	1
16	hara-hara	24	9	33	39	kachin-kachin	0	0	0
17	pasa-pasa	6	24	30	40	abekobe	0	0	0
18	zara-zara	12	14	26	41	gachi-gachi	0	0	0
19	haki-haki	16	8	24	42	zaba-zaba	0	0	0
20	gakkuri	18	5	23	43	yare-yare	0	0	0
21	gan-gan	4	18	22	44	wasshoi	0	0	0
22	patari	11	11	22	45	wahaha	0	0	0
23	gaburi	11	10	21					

次に、選定した4種のオノマトペのファイルを中納言からダウンロードし、Excel上で生テキスト部分のみを抽出してテキストファイル化した。そして、Mac専用コンコーダンサーCasualConcで読み込み、共起しやすい動詞とのセットフレーズをt-scoreを基準に抽出した。¹¹ 図3は、「たっぷり」と共起する動詞(設定はR1を助詞、R2を動詞)をt-scoreに基づいて視覚化したものである。動詞の文字の濃さは頻度情報を表し、濃いもの程、生起頻度が高いことを示す。「言う」などいくつかの動詞のフォントサイズは大きく文字も濃いことが分かる。t-scoreは他の共起関係を示す統計指標と比較して生起頻度の影響を受けやすい傾向にあることから、図3では文字が大きいものは濃く表示されている。このような分析を経て、頻度の高い共起関係である「たっぷり」+「言う」、「さっぱり」+「する」、「がらり」+「変わる」、「ぱりっ」+「した」を用いた例文を翻訳調査の対象とした。



図3. 「たっぷり」の t-score に基づく視覚化 (頻度情報も含めた場合)

無論、検索時には平仮名と片仮名の両表記において各オノマトペを検索する必要があるが、 CasualConc には両検索ワードを一つのキーワードとして指定し、共起動詞の一括抽出を可能にするモードが設定可能であることから、検索時にはこのモードを利用した。参考までに、このモードを利用して抽出した parii の共起動詞リスト(t-score 順)を表 7 に挙げる。

表7

parii の共起動詞リスト(t-score 順)

#	Word	T-Score	Freq.	#	Word	T-Score	Freq.
1	し	6.830366929	51	14	焼きもどす	0.975317517	1
2	さ	4.284484144	20	15	乾い	0.925952552	1
3	する	2.0042628	5	16	割る	0.975317517	1
4	仕上がり	1.675049056	3	17	いつ	0.901270069	1
5	焼い	1.632297742	3	18	頑張る	0.975317517	1
6	焼き	1.326947808	2	19	着込ん	0.975317517	1
7	剥がれる	0.975317517	1	20	なる	0.827222622	1
8	割れ	0.975317517	1	21	離れ	0.975317517	1
9	焼き上がり	0.925952552	1	22	出る	0.975317517	1
10	焼く	0.827222622	1	23	いう	0.876587587	1

11	焼け	0.950635035	1	24	砕け	0.975317517	1
12	凍り	0.975317517	1	25	焼き上げる	0.950635035	1
13	揚げ	0.925952552	1				

試しに、KOTONOHA から抽出したオノマトペの共起動詞と現行和英辞典に掲載されている共起動詞とを比較してみる。調査には、アマゾンの人気度ランキング(2015年2月25日付)の上位を占めた『スーパー・アンカーコ・英辞典第2版(SA2)』と『ウィズダム・英辞典第二版(WD2)』の2種を用いた。中高大の現場でも好評で現行の和英辞典をリードしている和英辞典である。表8は、「がらり」と「ぱりっ」の例文で用いられた共起動詞を掲載順に並べたものと、KOTONOHA から抽出した同オノマトペの共起動詞を t-score 順に並べたものである。尚、「たっぷり」と「さっぱり」は掲載されている例文の数が予想以上に豊富であるため表8では割愛している。

表8

現行の和英辞典の日本語オノマトペ記述比較

SA2	WD2	KOTONOHA
変わった (今まで親切だったおば の態度ががらりと変わった)	開ける (彼はノックもせずにがらりと戸 を開けた)	変わる, 変える, 開く, 異なる
がらり 変えた (彼は彼女が社長の一人娘 だとわかるとがらりと態度を変え た)	変わった (その後私たちに対する彼の態 度はがらりと変わった)	
***	している (ぱりっとしている)	する, 焼く, 剥がれる, 割れる, 凍る, 揚 げる, 乾く, なる, 離れる, 砕ける
ぱりっ ***	した (ベッドにぱりっとした木綿のシ ツを敷いた)	

上記の辞書の記述と KOTONOHA のデータを比較すると、飯島(2006)では翻訳が難しいとされている日本語オノマトペにおいても、現行の和英辞典には何かしらの翻訳例文を与えていていることが分かる。但し、「ぱりっ」に関しては SA2 では掲載されておらず、また KOTONOHA のデータから多様な動詞とも共起することが分かるため、このような共起関係とその例文を辞書に掲載する必要があるかもしれない。

4.2 インフォーマント調査結果

次に、前節までの分析を通して選定した計4種のオノマトペの[副詞的オノマトペ+動詞]の共起表現(t-scoreに基づき最上位のもの)の日本語例文とその翻訳をインフォーマントに作例して頂いた。日本語文も作例して頂いた理由は、コーパスからの用例をそのまま転用することには、辞書の記述上向い

ていないからである。インフォーマントの選出には、下記3つの条件を全て満たした方にお願いした。

[翻訳調査を依頼したインフォーマントの条件]

- ・英米において10年以上の滞在経験がある。
- ・日英・英日翻訳・通訳業務などに携わった経験がある。
- ・大学レベル以上で言語教育に携わっている現役大学教員である。

表9は、以上の条件を満たした自他共に認めるアメリカ系の英日バイリンガルの方による翻訳例である。

表9

バイリンガルの日本語オノマトペの英訳(論者により日本語を一部修正)

「たっぷり」+「言う」	説得するために理由を <u>たっぷり</u> 言う方がいいです。 In order to be persuasive, give a lot of reasons.
「さっぱり」+「する」	シャワーを浴びて <u>さっぱり</u> することにした。 He decided to shower and <u>freshen up</u> .
「がらり」+「変わる」	海外研修でがらりと変わる生徒がいます。 Some students <u>change drastically</u> after overseas study.
「ぱりっ」+「した」	あたらしいオーブンで温めたら <u>ぱりっと</u> した。 Warming it in the new oven <u>made it crispy</u> .

表9に示されるように、一見翻訳が難しいとされる[副詞オノマトペ+動詞]の共起表現においても全て翻訳可能であることが分かる。ただし、完全に逐語的に訳してあるのは「がらり」(drastically)のみであり、他に関しては動詞とのセットフレーズで訳している感がある。これが実際の翻訳文であり、次世代の和英辞典の編纂時には、このように経験豊富な翻訳のプロに作例して頂くことで信頼性ある情報を学習者に提供することができるのではなかろうか。今後の課題としては、調査協力して頂いたような翻訳経験豊富なインフォーマントの数を増やし、現行の和英辞典との英文の比較や、実際の辞書編纂者とインフォーマントに同じ日本語文を翻訳してもらい、その質の違いを調査する、などが考えられる。いずれにしても、インフォーマント調査においても量が求められる時代であろう。

5. おわりに

以上、本論では翻訳が難しいとされる日本語オノマトペの翻訳を試みた。翻訳できないものでも可能な限り別言語で簡潔に説明・記述しようとする努力の姿勢こそが、最終的には日本の文化を伝え日本と他国の交流、並びにグローバル化を促進する礎となると論者は信じている。そして、このような小さな取り組みが国益へつながり、日本に興味を持つ外国人の日本語学習の促進と、海外で活躍する日

本人のコミュニケーション力の向上にも繋がると信じている。本稿で記した知見が、今後の和英辞典の編纂や外国人向け日本語テキストの記述に貢献することを願ってやまない。

注

1. 「正確」な翻訳というものはおそらく存在しない。起点言語で書かれた命題を目標言語を通して如何に表現することができるか、もしくは可能な限り内容を近似させることができるかが、翻訳作業の課題である。
2. 論者が2014年に米国Hope Collegeで教鞭をとっていた際、上級アメリカ人日本語学習者の多くが日本語オノマトペの習得、並びにその翻訳に困難を感じていた。
3. 担当していた授業はApplied Linguisticsで、授業内では英日・日英翻訳も教えていた。
4. Fukuda (2003)の記述も参考にすると、日本語オノマトペを英語で表現するには言い換えや説明が必要であるため、記述スペースが限られた辞書には詳細に記述することが難しい。例えば、ぴちぴち(pichipichi)は、(1) The sound or feeling of a small object bouncing or flapping in a lively manner. Often used to describe live fish. (2) By extension, a spirited, active person--usually a young woman who is full of vim and vigor (Fukuda, 2003, p.141).と説明されている。
5. 表2は論者以外に2人の日本語インフォーマントに調査を実施した結果に基づいている。結合度が強い順に○、△となり、×は共起不可を表す。
6. インフォーマント間で判断のずれはほぼ認められなかった。
7. Goldberg and Ackaman (2001)は実際には義務的な付加詞(adjunct)が存在すると主張し、言語学的な議論を展開している。
8. 表3における副詞+動詞タイプのオノマトペの数は計10種であり、この数からも *twinkle* の意味の幅が分かる。
9. 「オノマトペ+笑う」に対応する他の翻訳に関しては Nishina(2013)を参照のこと。
10. インフォーマントとして調査協力して頂いた方にここで感謝の意を示す。
11. 中納言から直接ダウンロードしたファイルには、CasualConc を用いて分析する上で余計なノイズが含まれるため、この整形作業を施した。

参考文献

- Baker, M. (1992). *In other words: A coursebook on translation*. London: Routledge.
- Danielsson, P. (2007). In other words: Using paraphrases in translation. *Translating and the Computer*, 29, 1-11.
- Fukuda, H. (2003). *Jazz up your Japanese with onomatopoeia*. Kodansha.
- 飯島英一 (2006). 『日本の猫は副詞で鳴く、イギリスの猫は動詞で鳴く』 朱鳥社
- 松岡弘(監修) (2000). 『初級を教える人のための日本語文法ハンドブック』 スリーエーネットワーク
- Newmark, P. (2003). *A textbook of translation*. Pearson Education.
- Nishina, Y. (2013). Translation, paraphrasing and onomatopoeia. *Kyoto Working Papers in English and General*

Linguistics, 2, 203-216.

Takaie, H. (2002). The use of the Japanese-English parallel corpus in English education. *English Corpus Studies*, 9, 93-104.

留学環境での言語学習を研究対象とする際に留意すべきいくつかのこと —効率的な調査をおこなうために—

三上 仁志

中部大学

概要

本稿は、留学環境での言語学習（もしくは言語習得）を研究／調査対象とする際に留意すべきいくつかのポイントを紹介するものである。本稿で紹介する情報は、特に留学環境での言語学習に研究上の関心を持ち始めた、もしくはこれから本領域で研究をおこなうと考えている読者に向けたものとなっている。本稿で取り扱うテーマは、大きく分けて以下の 3 つである：(a) 調査対象期間は、どの程度取るべきなのか；(b) 研究をおこなう上で必要となるデータには、どの様な種類のものがあるのか；(c) 実際にデータを収集し、データ分析をおこなう際に留意するべき事柄は何か。

Keywords: 留学と言語学習、調査方法、言語能力、サンプリング、一般化

1. はじめに

我々は、留学を通してどのくらい外国語の能力を高めることができるのか？より効果的な留学プログラムをデザインするためには、どうすればよいのか？外国語の能力を効率的に高めるために、留学参加者がすべきことは何か？そもそも、（国内での学習環境と比べて）留学環境ではどれくらい多く目標言語を読んだり、書いたり、聞いたり、話したりするのだろうか？これらの研究テーマは、すべて留学環境での言語学習（もしくは言語習得）に関わるものであり、これまで多くの研究者によって取り上げられてきたものもある(Allen, 2010; Davidson, 2010; DeKeyser, 2010; Dewey, 2004; Freed, Segalowitz, & Dewey, 2004; Golonka, 2006; Hernández, 2010; Isabelli-García, 2010; Llanes & Muñoz, 2013; Magnan & Back, 2007; Sasaki, 2011; Segalowitz & Freed, 2004; Taguchi, 2008; Trentman, 2013; Yashima & Zenuk-Nishide, 2008)。上記の研究テーマ例からも分かる通り、留学環境における言語学習は、多くの研究者や教育関係者の関心事項となりえる。

本稿は、この様な留学環境での言語学習を研究テーマとして扱う際に、調査者や研究者（以下、まとめて調査者とする）が留意すべきいくつかのポイントを紹介するものである。留学環境での言語学習に関して、研究知見をレビューした論文／書籍(Kinginger, 2011; 塩澤, 吉川, & 石川, 2010)や、研究方法の妥当性について議論した論文(Davidson, 2007; DeKeyser, 2007, 2010)は、これまで存在した。しかし、これから本領域で研究を

行おうと考えている人間が、自分でデータ収集やデータ分析をおこなう際に留意すべきポイントについて解説した論文は、著者が知る限り存在しない。この点について著者は、自分が調査を進める中で「避けるべき落とし穴などについての事前情報を知っていれば、無駄な調査をせずに済んだのに…」と思うことが幾度かあった。そこで本稿では、データの収集や分析を効率的におこなう上で有用と思われる情報を、自身の経験や、これまで本研究領域で重ねられてきた議論にもとづいて整理し、まとめてみることとした。本稿で取り扱うテーマは、大きく分けて以下の3つである：(a) 調査対象期間は、どの程度取るべきなのか；(b) 研究をおこなう上で必要となるデータには、どの様な種類のものがあるのか；(c) 実際にデータを収集し、データ分析をおこなう際に留意するべき事柄は何か。以下の第2章では、この内(a) 調査対象期間、および(b) 収集すべきデータについて解説してゆく。

2. 調査対象期間、言語能力と心的特性、留学プログラムのデザイン

留学環境での言語学習を研究テーマとする場合、その調査は、以下の理由から通時的なものとなる。留学の効果を検証するために頻繁に用いられる研究手法は、留学参加者の(a) 言語能力や(b) 心的特性（例：学習動機づけ、スピーキング不安）を留学の前後で測定し、各測定結果を比較することである(Allen & Herron, 2003; Allen, 2010; Davidson, 2010; DeKeyser, 2010; Hernández, 2010; Magnan & Back, 2007; Segalowitz & Freed, 2004)。また、留学環境での言語学習の特性を調べるために、留学を経験する学習者達と国内で学習を続ける学習者達の双方を調査対象とし、各群の一定期間内での(a) 言語能力や心的特性の変化の程度や、(b) 教室内外での目標言語との接触量や言語接触の質を比較することもある(Freed, Segalowitz, et al., 2004; Kohro, 2001; Llanes & Muñoz, 2013; Sasaki, 2011)。なお、言語接触（Language Contact）とは、目標言語を読んだり、聞いたり、書いたり、話したりする活動の全てを意味する(Freed, Dewey, Segalowitz, & Halter, 2004)。

この内、言語能力の変化を調査対象とする場合、統計的に意味のある変化が観察されるのは、中期（1セミスター=4～6カ月程度）以上の留学を調査対象とした場合が多い(Allen & Herron, 2003; Allen, 2010; Sasaki, 2011)。これと比べて、2カ月以下の短期留学を調査対象とする場合、心的特性に関しては統計的に意味のある変化が見られる場合もあるものの、言語能力については変化が観察されにくい傾向がある（上記の参考文献の内、特にAllen, 2010を参照のこと）。よって、留学環境での言語能力の変化に関心があり、かつ興味のある言語能力について取り扱った先行研究がない場合、まずは中期以上の留学を対象として（パイロット）調査をおこなうことが勧められる。

言語能力の変化を調査対象とする場合、どの様な言語能力に注目するのかについても考えておく必要がある。例えば、スピーキングテストの得点は、上記に挙げた中期（1セミスター=4～6カ月）程度の留学の前後で、統計的に意味のある変化が認められること

も多い(Davidson, 2007; Hernández, 2010; Magnan & Back, 2007)。一方で、ライティングやリーディングの能力については、一定の変化が見られた事例と見られなかつたものが混在している(Dewey, 2004; Kohro, 2001; Llanes & Muñoz, 2013; Sasaki, 2011)。また、例えばスピーキングの能力についても、発話の流暢性(Fluency)については変化が認められやすい一方(Freed, Segalowitz, et al., 2004; Segalowitz & Freed, 2004), その正確性(Accuracy)や複雑性(Complexity)については、短～中期の留学期間中での成長を疑問視する向きもある(DeKeyser, 2007, 2010)。リスニング能力についても、Taguchi (2008)では、リスニング課題の理解の正確性について、中期の留学期間中を通して統計的に意味のある変化は観察されなかつた。この研究では、婉曲的な発話表現(遠回しな断りの表現など)を理解する能力が測定の対象とされた。Taguchi (2008)は、この研究結果について以下のように考察している。婉曲的な発話表現に触れる機会は、より直接的な表現に触れる機会と比べて非常に限られている。また婉曲表現は、直接的な表現とは異なり典型的なスキーマや枠組みに沿って構成されるものではない(=言語形式と意味の対応が一対一では無い)。この様な(a) 出現機会が少なく、かつ(b) 非典型的な発話内容を理解する能力は、仮に留学環境であつても伸ばしにくいものであり、限られた留学期間の中では、統計的に有意味な成長は観察しにくい。

Taguchi (2008)のこの指摘は、心理測定に対する重要な示唆を含んでいる：言語能力や心的特性の測定に使用するテストや課題を選ぶ際には、調査対象者達が留学中に成長させる(ことが見込まれる)能力や特性を、十分敏感に測れるものを選ぶ必要がある。例えば、IELTS® Academic Module のテストスコアを使用して、調査対象者達の留学前後の言語能力を比較するとしよう。日本英語検定協会のホームページにあるように、IELTS® Academic Module は、Cognitive Academic Linguistic Proficiency (CALP) (Harsch, 2014)の様な、学術領域での英語運用能力の測定を目的としたテストである¹。そのため、調査対象者達が留学中に経験する学習活動の多くが学術的な内容であることが予測できる場合、このテスト得点を用いて留学中の言語能力の変化を検証することには、一定の妥当性があると言えるだろう。これとは逆に、CALP の成長を目的としない留学(例えば、ワーキングホリデー)に参加した調査対象者達の留学前後の言語能力について IELTS® Academic Module のテスト得点を用いて比較した場合には、ごく限られた変化しか観察できない可能性がある。しかし、仮に CALP の育成を目的としない留学であったのなら、IELTS® Academic Module のテスト得点を用いて留学の効果を評価することは、そもそも留学の目的と無関係な言語能力について調査をしていることになり、不適切である。

より適切な言語能力や心的特性を調査対象として選ぶ上で重要なことの1つは、調査対象者達が参加する留学の内容、特に教室内で経験する言語学習の内容を可能な限り詳しく把握しておくことである。留学環境での言語学習を調査対象とする場合、一定のサンプルサイズを得る目的から、1つの教育機関(高校や大学など)が実施する留学プログラム

に参加する学習者達を調査対象とすることが圧倒的に多い（この様なサンプリング方法には、メリットもデメリットもあるが、その話は後述する）。この様な留学プログラムの実施目的は、外国語能力の育成である場合が多い。そのため、留学参加者達は、留学中に受け入れ先の教育機関で一定数の（a）語学に関する授業（例えば、English as Second Language）や（b）目標言語で開講される一般科目を履修することが多い。これらの授業のシラバスや、授業への参加記録（例えば、授業への出席率、成績表や提出課題など）を参照することで、学習者達が参加した授業の内容や性質を把握することが出来る（もちろん、調査対象者達へのインタビュー調査などから情報を集めるという方法もある）。例えば、学習者達の参加した授業がインストラクション主体（＝授業の大部分において教師が話す授業）だったのか、それともコミュニケーション主体（＝授業中に教師・生徒間や生徒・生徒間で多くの会話が交わされる授業）だったのかは、研究対象とする言語能力や心的特性を選定する上で重要な情報となる。一般的に抱かれるイメージとは異なり、留学環境であっても、目標言語を使用する主な機会が教室内に限られていることは、例えばスピーキングに関しては珍しくない（DeKeyser, 2007; Freed, Segalowitz, et al., 2004; Trentman, 2013）。仮に教室外での発話機会が限られており、かつ教室内でもあまり目標言語を話す機会が無かった場合、中期の留学を経験した場合であってもスピーキング能力に変化が見られないことは、十分に想定できるだろう。これとは逆に、調査対象者全員について教室内の発話機会が限られていたことが分かっている場合、スピーキング能力の変化に関する個人差を説明する要因として教室外での言語接触の（a）量的な多寡や（b）質的な違いを仮定した調査をおこなうことは、ある程度妥当だといえるだろう（Allen, 2010; Freed, Segalowitz, et al., 2004; Hernández, 2010; Llanes & Muñoz, 2013）。この様に、留学参加者達が教室内で（a）どの様な学習活動をするのか、そして（b）それらの活動をどの程度おこなうのか、について把握をしておくことは、仮に調査者の主たる関心が教室外での言語接触の量や質にある場合でも重要である。

本章のまとめとして、ここまで紹介した情報を以下にまとめる。次章では、上記で触れた「一つの教育機関が実施する留学プログラムに参加する学習者達を調査対象とする場合の研究上のメリットとデメリット」について見て行く。

- (1) 調査対象が心的特性である場合、2カ月以下の短期留学を対象として（パイロット）調査をおこなっても統計上意味のある変化が観察される可能性はある。
- (2) 調査対象が言語能力であり、かつ該当する言語能力について取り扱った先行研究がない場合、さしあたり中期（1セミスター=4~6カ月程度）以上の留学を対象として（パイロット）調査をおこなうことが勧められる。
- (3) どの様な言語能力を、どの様なテストや課題を使用して測定するかは、留学プログラムの特性やその実施目的と併せて考え、決定する。

- (4) 調査対象者（達）が参加する留学のデザイン、特に教室内での学習内容や学習経験については、可能な限り詳細に把握しておく。教室外での言語接触についても、必要と判断すれば記録する。

3. サンプリングについて

前章で、1つの教育機関が実施する留学プログラムに参加する学習者達を調査対象とすることには、メリットとデメリットの双方があると書いた。本章では、その理由について解説してゆく。

メリットとして挙げられる事項には、以下のものがある。第一に、この様なサンプリングをおこなった場合、留学参加者達の渡航前の言語能力は、（入試などの選抜の結果として）ある程度均質となることが予想される。学習者達の言語能力が変化した要因を留学経験に求める場合、渡航前の言語能力が近い学習者達をサンプリングすることは、選択バイアスへの対処として有効である。選択バイアスがある状態とは、サンプリングが偏った結果として、調査対象者達の母集団に対する代表性がなくなっている状況のことである（Brown, 2006）。例えば、前章でも触れた中期以上の留学の前後で観察される言語能力の成長は、主に渡航前の言語能力が初級～中級の学習者達を調査対象とした研究の結果である。つまり、渡航前のテスト得点が既に上級に達している学習者については、これとは異なる研究結果が得られる可能性がある（例えば、既に上級に達している学習者であれば、2ヵ月以下の留学期間中に言語能力を著しく向上させることが可能かもしれない）。このように考えれば、渡航前のテスト得点が（a）初級と判定される学習者と（b）上級と判定される学習者を同じ“留学を経験した学習者”として一括りに扱うことで、データ分析の結果が曇ることを容易に想像できるだろう（ただし、何らかの方法を用いて留学経験者間の渡航前の言語能力の違いを調整できる場合は、この限りでは無い）。この様に、1つの教育機関が実施する留学プログラムに参加する学習者を調査対象とすることには、渡航前の諸条件が比較的均質な調査対象者達を集めやすいというメリットがある。

また、この様な留学プログラムであれば、渡航先で全ての学習者達が同じ授業を履修するケースも多い。この場合、留学参加者達が教室内で経験した授業内容は、全て同じであつたと見なすことができる（あくまでも「研究の上ではそう見なせる」ということであり、もちろん同じ授業を受講したことが同じ学習を経験したことと同義ではないことに留意されたい）。教室内での学習内容に均質性が認められる場合、その影響も考慮しつつ、教室外での言語接触（の量や質）が言語能力や心的特性の変化に影響した程度を検証することが容易となる。調査対象者達の留学中の教室内活動が把握しやすく、またその影響の統制が容易なことは、このサンプリング方法の第2のメリットといえる。

最後のメリットも、選択バイアスに関連するものである。単一の教育機関が実施する留学プログラムを調査対象とができる場合、調査者は、その教育機関もしくは留

学プログラムの関係者である可能性がある（例えば、留学プログラムの責任者である、もしくは留学プログラムの実行員と知り合いであるなど）。この様な立場にいる場合、調査者は、調査対象者達（や留学プログラムの実行員）と頻繁に連絡を取り合うことが出来る。著者の経験から言って、調査者が（a）調査対象者とある程度顔見知りである場合や（b）調査対象者と頻繁に連絡を取り合っている場合には、データの回収率を高めることが比較的容易となる。一般的に言って、データの回収率は、最低でも 60%を上回っているべきであるとされる。データの回収率が 60%以下である場合、收取されたデータにはかなりの程度バイアスが含まれていることが想定されるため、歪んだデータをもとに議論を進める危険性が高まることになる(Dörnyei & Taguchi, 2010; 木下, 2011)。また、調査の結果が議論や査読の対象になった際に、回収率の低さを致命的な欠陥として指摘されることもある。つまり、データの回収率が低い場合には、再調査の必要性が高まるのである。この点について、通時的な研究において回収率を高い水準に保つことは簡単ではない（複数回に渡ってデータを収集する必要があるためである）。よって、小まめな連絡という比較的簡単な対処法を用いることでデータの回収率を高めることが出来るのであれば、その様な方法は積極的に活用すべきであろう。

この様なサンプリング方法を用いることのデメリット（の 1 つ）は、統計的分析の結果が一般化しにくい、ということである。1 つの教育機関が実施する 1 つの留学プログラムの参加者数は、統計的分析結果の一般化という側面から見て十分でない場合が多い。この種のサンプリングを実施した場合の調査対象者数は、少ない場合で 10 人未満、多い場合でも 50 人以下である場合が圧倒的に多い(Allen, 2010; Taguchi, 2008)。この点について、先ずは 10 人未満のサンプルから得られたデータを使って母集団での傾向を統計的に推測することは、不適切であるとされる(Chernick, 2008)。統計的分析の結果が一般化可能であるか否かは、効果量、サンプルサイズ、危険率、検定力の 4 つの側面から考えるべきである（またサンプリングの適切さや、測定や分析（結果の解釈）の妥当性についても考慮する必要がある）(Borsboom, Cramer, Kievit, Scholten, & Franic, 2009; Brown, 2006; Kane, 2012; Plonsky & Brown, 2014; Plonsky, 2012)。この点について、サンプルサイズが小さい場合に一般化の基準を満たすことは容易ではない。また厳密にいえば、1 つの教育機関（大学とする）の出身者（日本人英語学習者とする）数十人の渡航前の性質に均質性が認められるからといって、彼／彼女らが母集団（日本人大学生英語学習者全員）に対して代表性を有していると考えることには飛躍がある(Brown, 2006)²。よって、統計的分析結果の一般化という側面から考えれば、複数の教育機関が実施する複数の留学プログラムに参加する学習者を調査対象とし、バイアスや測定など様々な側面に気を配りながら調査を実施する方が望ましいといえる。

しかし、複数の教育機関が主催する複数の留学プログラムに参加する学習者を調査対象とする場合には、調査対象者達の（a）渡航前の性質がある程度均質であり、かつ（b）

留学中の教室内での学習活動にも一定の類似性が認められる必要がある。このどちらの条件も満たす留学プログラムや留学参加者達を探し出し、かつ調査への協力を取り付けることは、言うまでもなく簡単ではない。また、調査対象者数が増えれば、彼／彼女らと小まめに連絡を取り合い、データの回収率を増やすことも難しくなるだろう。ここまで記述から分かる通り、留学環境での言語学習を調査対象とする際に選択し得る主要なサンプリング方法には、それぞれメリットとデメリットが存在する。研究の妥当性が非常に重要である一方、現実的には調査の実現可能性も無視できない問題である(Bachman & Palmer, 2010)。この様な理由から、サンプリング方法に関しては、上記で紹介した 2 つの方法の内、どちらの方が良いかを断定することは、現時点では残念ながら不可能である。サンプリング方法に関して重要なことは、最終的にどの様な方法を選ぶにせよ、その方法を採ることで生じ得るバイアスに考えを巡らし、それに最良と思われる対処を施すことである。また、データの分析結果が自身の研究意図にとって都合の良いものであった時にも、バイアスの存在や調査全体の妥当性について再考し、分析結果の一般化に慎重になる姿勢が重要である。最後に、本章の内容を以下にまとめる。

- (5) 1 つの教育機関が実施する留学プログラムに参加する学習者達を調査対象とするとのメリットの 1 つは、選択バイアスへの対処のしやすさである。
- (6) 1 つの教育機関が実施する留学プログラムに参加する学習者達を調査対象とするとのデメリットは、統計的分析結果の一般化のむずかしさと、サンプリングの性質上残ってしまう選択バイアスにある。
- (7) 複数の教育機関が主催する複数の留学プログラムに参加する学習者を調査対象とすることは、分析結果の一般化の面から見てより望ましいといえる。
- (8) 複数の教育機関が主催する複数の留学プログラムに参加する学習者を調査対象とする場合、(5) と (6) では余り考えずに済んだ選択バイアスへの意識的な対応が必要となり、調査の実行可能性が低下する。

4. おわりに

本稿で紹介した情報の内、(1) ~ (4) は、これから留学環境での言語学習に関する(パイロット)調査をおこなう調査者への Tips であり、(5) ~ (8) は、取集されたデータやその分析結果をより適切に取扱うための前提となる知識である。もちろん、本稿で紹介した情報は、これさえ知っていれば留学中の言語学習について必ず効率的な調査が出来る、というほど網羅的なものではない。これらの情報はしかし、出来るだけ費用や時間を無駄にせずに調査をおこなうための第一歩程度にはなると考える。本稿で紹介した情報が、これから留学環境での言語学習を研究テーマとして(パイロット)調査をおこなう調査者の役に少しでも立てば、これに勝る喜びは無い。

謝辞

本稿の着想は、中部大学の塩澤正教授、名古屋学院大学の中山麻美先生、名古屋大学大学院の平石順久先生と開いている勉強会から得たものである。また、本稿で提供した情報の一部は、筆者が学術誌に論文を投稿した際に査読員の先生方から頂いたコメントに基づくものである。以上の通り、本稿の執筆にあたって多くの先生方からアドバイスや助力を頂いた。ここに記して感謝したい。

注

1. <http://www.eiken.or.jp/ielts/test/>を参照のこと。
2. 無作為抽出が行われてはいないことにも留意すべきである(木下, 2011)。

参考文献

- Allen, H. W. (2010). Language learning motivation during short-term study abroad: An activity theory perspective. *Foreign Language Annals*, 43, 27—49.
- Allen, H. W., & Herron, C. (2003). A mixed-methodology investigation of the linguistic and affective outcomes of summer study abroad. *Foreign Language Annals*, 36, 370–385.
- Bachman, L., & Palmer, A. (2010). *Language assessment in practice: Developing language assessments and justifying their use in the real world*. Oxford: Oxford University Press.
- Borsboom, D., Cramer, A., Kievit, R., Scholten, Z., & Franic, S. (2009). The end of construct validity. In R. W. Lissitz (Ed.), *The concept of validity* (pp. 135–170). North Carolina: Information Age Publishers.
- Brown, J. D. (2006). Generalizability from second language research samples. *Shiken: JALT Testing & Evaluation SIG Newsletter*, 10, 24—27.
- Chernick, M. R. (2008). *Bootstrap methods: A guide for practitioners and researchers* (2nd ed.). New Jersey: Wiley.
- Davidson, D. E. (2007). Study abroad and outcomes measurements: the case of russian. *The Modern Language Journal*, 91, 276–280.
- Davidson, D. E. (2010). Study abroad: When, how long, and with what results? New data from the Russian front. *Foreign Language Annals*, 43, 6–26.
- DeKeyser, R. M. (2007). Study abroad as foreign language practice. In R. M. DeKeyser (Ed.), *Practice in a second language: Perspectives from applied linguistics and cognitive psychology* (pp. 208–226). Cambridge: Cambridge University.
- DeKeyser, R. M. (2010). Monitoring processes in Spanish as a second language during a study abroad program. *Foreign Language Annals*, 43, 80—92.

- Dewey, D. P. (2004). A comparison of reading development by learners of Japanese in intensive domestic immersion and study abroad contexts. *Studies in Second Language Acquisition*, 26, 303–327.
- Dörnyei, Z., & Taguchi, T. (2010). *Questionnaires in second language research: Construction, administration, and processing* (2nd ed.). New York: Routledge.
- Freed, B. F., Dewey, D. P., Segalowitz, N., & Halter, R. (2004). the Language contact profile. *Studies in Second Language Acquisition*, 26, 349–356.
- Freed, B. F., Segalowitz, N., & Dewey, D. P. (2004). Context of learning and second language fluency in French: Comparing regular classroom, study abroad, and intensive domestic immersion programs. *Studies in Second Language Acquisition*, 26, 275–301.
- Golonka, E. M. (2006). Predictors revised: Linguistic knowledge and metalinguistic awareness in second language gain in Russian. *The Modern Language Journal*, 90, 496–505.
- Harsch, C. (2014). General Language Proficiency Revisited: Current and Future Issues. *Language Assessment Quarterly*, 11, 152–169.
- Hernández, T. A. (2010). The relationship among motivation, interaction, and the development of second language oral proficiency in a study-abroad context. *The Modern Language Journal*, 94, 600–617.
- Isabelli-García, C. (2010). Acquisition of spanish gender agreement in two learning contexts: Study abroad and at home. *Foreign Language Annals*, 43, 289–303.
- Kane, M. T. (2012). Validating score interpretations and uses. *Language Testing*, 29, 3—17.
- Kinginger, C. (2011). Enhancing language learning in study abroad. *Annual Review of Applied Linguistics*, 31, 58–73.
- 木下徹. (2011). 「アンケート調査データと多変量解析」. In 藤村逸子 & 滝沢直宏 (Eds.), 『言語研究の技法—データの収集と分析』 東京:ひつじ書房, pp. 93–110.
- Kohro, Y. (2001). A pilot study of the linguistic impact of study abroad experiences on the writing of Japanese college ESL learners in America. *JACET Bulletin*, 33, 57–72.
- Llanes, A., & Muñoz, C. (2013). Age effects in a study abroad context: Children and adults studying abroad and at home. *Language Learning*, 63, 63–90.
- Magnan, S. S., & Back, M. (2007). Social interaction and linguistic gain during study abroad. *Foreign Language Annals*, 40, 43–61.
- Plonsky, L. (2012). Replication, meta-analysis, and generalizability. In G. Porte (Ed.), *A guide to replication in applied linguistics* (p. 116—132). New York: Cambridge University Press.
- Plonsky, L., & Brown, D. (2014). Domain definition and search techniques in meta-analyses of L2 research (Or why 18 meta-analyses of feedback have different results). *Second Language Research*, 31, 267—278.

- Sasaki, M. (2011). Effects of varying lengths of study-abroad experiences on Japanese EFL students' L2 writing ability and motivation: A longitudinal study. *TESOL Quarterly*, 45, 81–105.
- Segalowitz, N., & Freed, B. F. (2004). Context, contact, and cognition in oral fluency acquisition: Learning Spanish in at home and study abroad contexts. *Studies in Second Language Acquisition*, 26, 173–199.
- 塩澤正, 吉川寛, & 石川有香 (Eds.). (2010). 『英語教育大系第3巻 英語 教育と文化—異文化間コミュニケーション能力の養成』 東京:大修館。
- Taguchi, N. (2008). Cognition, language contact, and the development of pragmatic comprehension in a study-abroad context. *Language Learning*, 58, 33–71.
- Trentman, E. (2013). Arabic and English during study abroad in Cairo, Egypt: Issues of access and use. *The Modern Language Journal*, 97, 457–473.
- Yashima, T., & Zenuk-Nishide, L. (2008). The impact of learning contexts on proficiency, attitudes, and L2 communication: Creating an imagined international community. *System*, 36, 566–585.

適応型テストへの応用 —ベイズ EAP 推定法とフィッシャー情報量の概説—

住政二郎

関西学院大学

概要

本稿の目的は、ベイズ EAP 推定法とフィッシャー情報量について概説することである。著者は、項目反応理論に関連して、ラッシュモデルの導出(住, 2013), 各モデルの概説(住, 2014), PROX 法と同時最尤推定法(住, 2015)についてこれまでまとめてきた。項目反応理論は、テストの開発および結果の分析のみならず、適応型テスト(computer-adaptive testing)にも応用されている。適応型テストは、受験者能力を推定し、受験者能力にあった問題項目を出題することを特徴とする。このとき、受験者能力の推定にはベイス EAP(expected a posteriori) 推定法が、問題項目の選択にはフィッシャー情報量が利用されることが多い。しかし、適応型テストに関する論文では、紙面の都合から、各手法を使った事実や数式のみが記載されていることが多く、詳細な説明を得ることが難しい。また、項目反応理論に関する書籍では、数学的な知識が前提とされていることが多く、文系読者には理解が難しい。加えて、数式の記述形式に書籍間で微妙な違いがあり、初学者の理解を妨げている。本稿は著者が複数の関連書籍を通読し、その理解をまとめたものであるが、これから項目反応理論を学ぶ読者の助けになれば幸いである。

Keywords: 適応型テスト, ベイズ EAP 推定法, フィッシャー情報量, 項目反応理論

1. 適応型テストについて

適応型テストは、以下のアルゴリズムを基本とする(大友, 1996, p. 273; Thissen & Mislevy, 2000, p. 101)。

- a. How to start: どのように開始するか
- b. How to continue: どのように続けるか
- c. How to stop: どのように終了するか

上記 b のプロセスにおいて、適応型テストは、受験者能力を逐次的に推定し、受験者能力にあった問題項目を出題する。この時、受験者能力の推定にはベイス EAP 推定法が、問題項目の選択にはフィッシャー情報量が利用されることが多い。

2. ベイズ EAP 推定法について

ベイズ EAP (expected a posteriori) 推定法は、ベイズの定理を応用し、事後分布 (posteriori distribution) に基づき未知の受験者能力パラメータ (θ) を推定する手法である。同様の手法には、ベイズ MAP (maximum a posteriori) 推定法もある。これは単にベイズ最頻値 (Bayesian modal) 推定法と呼ばれることもある (村木, 2011, p. 83)。MAP 推定法とは、「被験者パラメーターの事後確率分布を最大化するようなパラメータ値を求める方法である」(村木, 2011, p. 82)。ただし、MAP 推定法は、最尤推定法 (maximum likelihood estimation method) と同様に、全問正答あるいは全問誤答の場合など、受験者の回答パターンによって推定値が得られない場合があり、適応型テストには適していない。MAP 推定法と最尤推定法の性質については、涌井・涌井 (2010, p. 190) と豊田 (2002, p. 35) が詳しい。また、MAP 推定法の問題については、植野・永岡 (2009, p. 56) と村木 (2011, p. 83) が詳しい。一方で、EAP 推定法は、少人数の受験者集団に適応可能で、また、回答パターンに関わらず推定が可能であることから、適応型テストに適していることが指摘されている (村木, 2011, p. 85; 植野・永岡, 2009, p. 56)。

2.1 ベイズの定理

涌井・涌井 (2012, p. 39) を参考に、ベイズの定理について概説する。

ここにジョーカーを抜いた 1 組 52 枚のトランプがあるとする。この 1 組 52 枚のトランプは、これから起こり得るすべての事象が含まれている標本空間である。この時、抜いた 1 枚のカードがハートで絵札である確率を求める。標本空間を U 、ハートを A 、そして絵札を B とすると、図 1 のように表現できる。

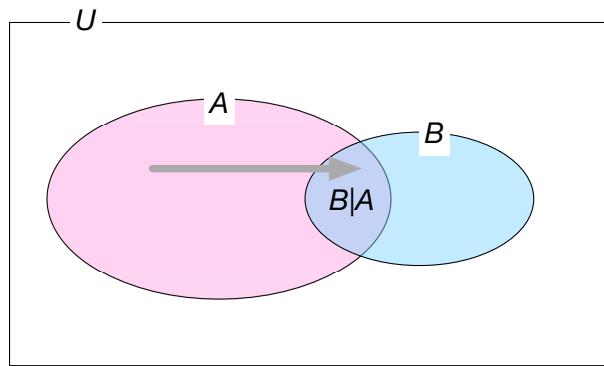


図 1. 条件付き確率 1

図 1 は、以下の式で表現できる。

$$P(B|A) = \frac{P(A \cap B)}{P(A)} \quad (1)$$

式(1)の左辺 $P(B|A)$ は、 A が起こった時 (A が起こるという条件下で) の B の確率、という意味である。これを条件付き確率という。図 1において、 A から B に向かって矢印が伸びている。注意すべき点は、 A で且つ B 、という意味ではない。これは、例えば、サイコロを 2 回振って、1 回目が 3 で、2 回目が 5 であるような場合を指す。これは同時確率と呼ばれ、式(1)では右辺の分子 $P(A \cap B)$ にあたる。

図 1 を改めて見てみると、式(1)の左辺 $P(B|A)$ の意味は、右辺の構造から、 $P(A)$ (ハート全体: 52 枚中 13 枚) に占める $P(A \cap B)$ (ハートで且つ絵札: 52 枚中 3 枚) の確率 (面積) と理解することができる。計算すると以下のようになる。

$$P(B|A) = \frac{P(A \cap B) : \text{ハートで且つ絵札}}{P(A) : \text{ハート全体}}$$

$$= \frac{\frac{3}{52}}{\frac{13}{52}} = \frac{3}{13}$$

今度は逆に、抜いた1枚のカードが絵札で、ハートである確率を求めてみる。図2において、図1とは逆にBからAに向かって矢印が伸びており、条件付き確率もAとBの順番が図1とは逆になっていることに注意する。

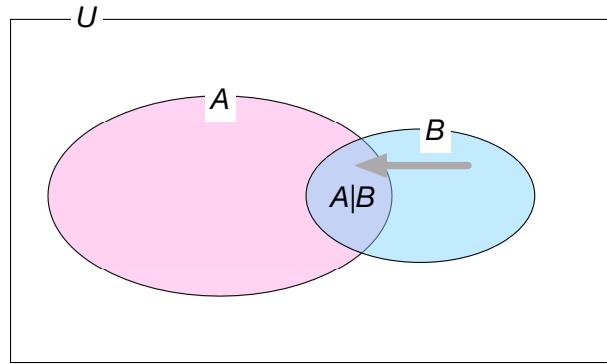


図2. 条件付き確率2

図2は、以下の式で表現でき、計算すると以下のようになる。

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

$$= \frac{\frac{3}{52}}{\frac{12}{52}} = \frac{3}{12} = \frac{1}{4}$$

図2と式(2)から、式(2)の左辺 $P(A|B)$ の意味は、右辺の構造から、 $P(B)$ (絵札: 52枚中12枚)に占める $P(A \cap B)$ (ハートで且つ絵札: 52枚中3枚)の確率(面積)と理解することができる。

ここで先の式(1)を整理する。

$$\text{式(1)} : P(B|A) = \frac{P(A \cap B)}{P(A)}$$

両辺に $P(A)$ をかけて、左辺と右辺を入れ換える。

$$P(A \cap B) = P(B|A)P(A) \quad (3)$$

一方、式(2)は、以下のように整理することができる。

$$\text{式(2)} : P(A|B) = \frac{P(A \cap B)}{P(B)}$$

両辺に $P(B)$ をかけて、左辺と右辺を入れ換える。

$$P(A \cap B) = P(A|B)P(B) \quad (4)$$

式(3)と(4)は乗法定理と呼ばれ、 $P(A \cap B)$ に関して結合することができる。

$$\text{式(3)} : P(A \cap B) = P(B|A)P(A)$$

$$\text{式(4)} : P(A \cap B) = P(A|B)P(B)$$

$$P(B|A)P(A) = P(A|B)P(B) \quad (5)$$

式(5)の両辺を $P(B)$ で割り、左辺と右辺を入れ換えて $P(A|B)$ に関して解くと、以下のようになる。式(6)をベイズの定理と呼ぶ。

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \quad (6)$$

2.2 ベイズの定理の発展

涌井・涌井(2010, p. 39–42)と涌井・涌井(2012, p. 72–79)を参考に、ベイズの定理を発展させる。

標本空間において、ある事象が観察された。その事象には、ある原因または仮説が影響している。このとき、標本空間を U 、原因または仮説を H 、事象を D とする(図3)。

D	$D \cap H$
U	

H

図 3. 原因・仮説 H と事象 D の関係性 1

図 3 の関係性を踏まえ, A を H に, B を D に書き換え, ベイズの定理を以下のように再定義する。

$$\begin{aligned} \text{ベイズの定理 式 (6)}: P(A|B) &= \frac{P(B|A)P(A)}{P(B)} \\ P(H|D) &= \frac{P(D|H)P(H)}{P(D)} \end{aligned} \quad (7)$$

さらに, 標本空間で観察された事象 D に対して, 複数の原因または仮説 $H_i (i = 1, 2, 3, \dots, N)$ が影響しているとする(図 4)。ただし, 原因または仮説 H_i は, それぞれ独立して重ならないことを仮定する。

D	$D \cap H_1$	$D \cap H_2$	$D \cap H_3$	$D \cap H_{\dots}$	$D \cap H_N$
	H_1	H_2	H_3	H_{\dots}	H_N

図 4. 原因・仮説 H と事象 D の関係性 2

図4において、 D の生起確立は、 H_i が独立して重ならないことを仮定すると、

$$P(D) = P(D \cap H_1) + P(D \cap H_2) + P(D \cap H_3) + \cdots + P(D \cap H_N) \quad (8)$$

となる。

ここで、式(3)の乗法定理 $P(A \cap B) = P(B|A)P(A)$ から、式(8)の $P(D \cap H_i)$ を $P(A \cap B)$ と考えと、 $P(D)$ は、以下のように整理することができる。

$$\text{式8 : } P(D) = P(D \cap H_1) + P(D \cap H_2) + P(D \cap H_3) + \cdots + P(D \cap H_N)$$

$$P(D \cap H_i) = P(D|H_i)(H_i) \text{ から}$$

$$P(D) = P(D|H_1)P(H_1) + P(D|H_2)P(H_2) + P(D|H_3)P(H_3) + \cdots + P(D|H_N)P(H_N) \quad (9)$$

式(9)を

$$\text{式(7) : } P(H|D) = \frac{P(D|H)P(H)}{P(D)}$$

に代入すると

$$P(H_i|D) = \frac{P(D|H_i)P(H_i)}{P(D|H_1)P(H_1) + P(D|H_2)P(H_2) + P(D|H_3)P(H_3) + \cdots + P(D|H_N)P(H_N)}$$

これを一般化すると

$$P(H_i|D) = \frac{P(D|H_i)P(H_i)}{\sum_{i=1}^N P(D|H_i)P(H_i)} \quad (10)$$

になる。

式(10)は、変数が、サイコロの目やトランプのカードのように中間値の存在しない離散変数の場合の一般解である。変数が、例えば体重・身長などのように無数の中間値が存在

する連続変数の場合は、分母で和の代わりに積分を用いて以下のようなになる。この点に関しては、豊田(2002, pp. 49–50)が詳しい。

$$P(H_i|D) = \frac{P(D|H_i)P(H_i)}{\int_{-\infty}^{+\infty} P(D|H_i)P(H_i) dx} \quad (11)$$

式(11)において、

$P(H_i|D)$ は、結果・データ D が原因・仮説 H_i から得られる確率を意味する。これを事後確率という。この確率に従う分布を事後分布という。

$P(D|H_i)$ は、原因・仮説 H_i の下で結果・データ D が得られる確率を意味する。これを尤度(ゆうど)という。標本空間 U で観察される結果・データ D は、ある原因・仮説 H_i に基づき生起する。しかし、事象の背景にある原因・仮説 H_i は、未知の値であるため、実際に観察された結果・データとの確率的な連鎖の値として尤度を使って表現する。

$P(H_i)$ は、結果・データ D を得る前の原因・仮説 H_i の確からしさを意味する。これを事前確率といい、この確率に従う分布を事前分布という。

式(11)において、右辺の分母は、分子の積分になっている。ということは、分子が分かれれば分母も分かる。この時、分母、つまり観察された事象の生起確立 $P(D) = \int_{-\infty}^{+\infty} P(D|H_i)P(H_i)$ を定数と見なすと、式(11)は以下のように省略することができる。

$$P(H_i|D) \propto P(D|H_i)P(H_i) \quad (i = 1, 2, 3, \dots, N) \quad (12)$$

上記式(12)の意味は、事後確率は、尤度と事前確率の積に比例する(\propto)、という意味である。

ベイズの定理は、標本空間 U において観察される事象=事後分布は、原因・仮説 H と結果・データ D との確率的な連鎖として捉えている。さらに、事象の生起に対して常に一定の影響を与える値を事前確率として取り込むことによって、人間の経験や常識を反映した結果=事後確率を計算することができるとされている。計算された事後確率は、次の計算の事前確率として利用され、新たな事後確率が計算される。このプロセスをベイズ更新と呼び、これを繰り返すことによって未知のパラメーターを推定しようとするのが、ベイズの定理を利用した推定の基礎である。

2.3 ベイズ EAP 推定法

ベイズ EAP 推定法は、EAP (expected a posteriori) の名称が示すとおり、「被験者パラメーター θ_i の事後確率分布の期待値を推定する」方法である (村木, 2011, p. 84)。期待値 (expected value) とは、確率変数の平均である。観察される事象 (x_1, x_2, \dots, x_n) が、確率分布に従って一定の頻度で生起する場合に、その平均を求めるのに等しい。では、なぜ期待値を利用するのだろうか。それは、「能力値の尤度が最も高くなるのは、その能力値のもとで期待得点が実際の得点に等しくなるとき」(静, 2007, p. 234, p. 239; 住, 2015, p. 97) であるからである。つまり、ベイズ EAP 推定法とは、(誤差を含む応答データの) 観測値と（潜在的な真の力の）期待値との差である残差を限りなく小さくし、2つの値を近づける作業ともいえる。

期待値は、記号で $E(X)$ と表す。以下、[Khan Academy \(2009\)](#) の教材を参考に概説する。

サイコロを 6 回降って、

2, 2, 3, 5, 5, 6

が得られたとする。これを標本空間 U とする。この変数の平均値は、

$$\frac{(2 + 2 + 3 + 5 + 5 + 6)}{6} = 3.8 \quad (13)$$

になる。期待値は、確率変数の平均であることから、式 (13) を以下のように変形する。括弧で添えられた数字は、各変数の出現回数を表している。また、% の表記を加えることで、各変数の出現頻度が分かる。

$$\begin{aligned} \frac{(2)2 + 1(3) + 2(5) + 1(6)}{6} &= \frac{1}{6}(2 \cdot 2 + 1 \cdot 3 + 2 \cdot 5 + 1 \cdot 6) \\ &= \frac{2}{6} \cdot 2 + \frac{1}{6} \cdot 3 + \frac{2}{6} \cdot 5 + \frac{1}{6} \cdot 6 \\ &= 0.33 \cdot 2 + 0.17 \cdot 3 + 0.33 \cdot 5 + 0.17 \cdot 6 \\ &= 33\% \cdot 2 + 17\% \cdot 3 + 33\% \cdot 5 + 17\% \cdot 6 \\ &= 0.66 + 0.51 + 1.65 + 1.02 = 3.8 \end{aligned} \quad (14)$$

式(14)の形式を念頭に置き、確率変数 X と確率 $P(X = x)$ ($x = 1, 2, 3, \dots, n$) が得られるとき、離散型確率変数の期待値 $E(X)$ は、以下のように一般化できる。

$$E(X) = \sum_{i=1}^n x_i \cdot p_i \quad (15)$$

連続確率変数の期待値 $E(X)$ は、以下のように一般化できる。

$$E(X) = \int_{-\infty}^{+\infty} x \cdot f(x) dx \quad (16)$$

連続確率変量のベイズの定理の発展公式は、式(11)より以下のとおり得られた。

$$\text{式(11)}: P(H_i|D) = \frac{P(D|H_i)P(H_i)}{\int_{-\infty}^{+\infty} P(D|H_i)P(H_i) dx}$$

式(11)において $P(D|H_i)$ は、尤度 (likelihood) であることから、記号 L に書き換える。

$$P(H_i|D) = \frac{L(D|H_i)P(H_i)}{\int_{-\infty}^{+\infty} L(D|H_i)P(H_i) dx} \quad (17)$$

連続確率変数の期待値 $E(X)$ の式(16)に、式(17)を代入し、未知のパラメータ θ の期待値を求める。この時、事前分布の $P(H)$ は、正規分布 $g(\theta)$ に従うこととする。

$$\begin{aligned} E(\theta_i|D) &= \int_{-\infty}^{+\infty} \theta_i \cdot f(\theta_i) d\theta_i \\ &= \int_{-\infty}^{+\infty} \theta_i \cdot \frac{L(D | \theta_i) g(\theta_i)}{\int_{-\infty}^{+\infty} L(D | \theta_i) g(\theta_i) d\theta_i} \\ &= \frac{\int_{-\infty}^{+\infty} \theta_i L(D | \theta_i) g(\theta_i) d\theta_i}{\int_{-\infty}^{+\infty} L(D | \theta_i) g(\theta_i) d\theta_i} \end{aligned} \quad (18)$$

式(18)において、項目応答理論の表現形式に合わせて D を項目困難度パラメーターに変更すると、村木(2011, p. 85)のベイズ EAP 推定法の式(5.19)と一致する。

3. フィッシャー情報量について

適応型テストは、受験者能力を逐次的に推定し、受験者能力にあった問題項目を出題することを特徴とする。この時、問題項目の選択には、フィッシャー情報量が利用されることが多い。では、なぜフィッシャー情報量が利用されることが多く、どのような基準で問題項目は選択されるのだろうか。

3.1 最尤推定の性質

N 個の観測値が、ある一定の母数分布に基づく確率変数 $P(X)$ に従って生起するとき、観測値が最も生起しやすい尤度関数を推定しようとするのが最尤推定 (maximum likelihood estimation) である。

豊田 (2002, pp. 64–65) は、最尤推定の特徴として以下の 3 つをあげている。

1. 最尤推定値の標本分布は、 n が大きくなるに従って、限りなく正規分布に近づく
2. 最尤推定値の標本分布の平均（期待値）は、 n が大きくなるに従って、限りなく真値 θ_i に近づく
3. 最尤推定値の標本分布の分散は、 n が大きくなるに従って、限りなく $\frac{1}{I(\theta)}$ に近づく。このとき、 $I(\theta)$ をフィッシャー情報量と呼ぶ

$$V[\hat{\theta}_i | \theta_i] = \frac{1}{I(\theta_i)}$$

$V[\hat{\theta}_i | \theta_i]$ は、 θ_i が与えられたときの $\hat{\theta}_i$ の分散という意味である。

上記 3 に関して、豊田 (2002, p.65) は、途中式を省略し、以下のようにフィッシャー情報量を定義している。

$$I(\theta_i) = E \left[\left(\frac{\partial}{\partial \theta} \log L(u_i | \theta) \right)^2_{\theta=\theta_i} \right] \quad (19)$$

以上のことから、式 (19) $I(\theta)$ の逆数は、最尤推定値の標本分布の分散 $\frac{1}{I(\theta)}$ になる。その平方根は、最尤推定値の標本分布の標準偏差 $\frac{1}{\sqrt{I(\theta)}}$ になる。標準偏差は、標本分布の母平均からの誤差を表す。 $I(\theta)$ と最尤推定値の標本分布の分散とが逆数の関係であることから、 $I(\theta)$ が最大となる問題項目を選択することは、標準偏差が最小となる問題項目を選択することになり、結果的に母平均に近づくことを意味する。これがフィッシャー情報量 $I(\theta)$ が問題項目の選択に利用される原理になる。

3.2 フィッシャー情報量の導出

確率変数 $X(x_i = x_1, x_2, x_3, \dots, x_N)$ が、未知のパラメータ θ を持つ関数に従う時、以下のようないちばん簡単なデータが得られたとする。

No.	x_i	確率
1	x_1	$P(X = x_1 ; \theta)$
2	x_2	$P(X = x_2 ; \theta)$
3	x_3	$P(X = x_3 ; \theta)$
\vdots	\vdots	\vdots
N	x_N	$P(X = x_N ; \theta)$

$x_1, x_2, x_3, \dots, x_N$ となる確率 $P(x_i ; \theta)$ は、以下の同時確率で求められる。

$$P(x_i ; \theta) = P(X = x_1 ; \theta) \cdot P(X = x_2 ; \theta) \cdot P(X = x_3 ; \theta) \cdot \dots \cdot P(X = x_N ; \theta) \quad (20)$$

このとき、式(20)をパラメータ θ の尤度関数と呼び、以下のようにあらわす。

$$L(\theta ; X = x_1, x_2, x_3, \dots, x_N) = P(X = x_1, x_2, x_3, \dots, x_N ; \theta) \quad (21)$$

そして、 $L(\theta ; X = x_1, x_2, x_3, \dots, x_N)$ を最大化する未知のパラメータ θ を推定することを最尤法という（石村・劉・石村, 2010）。尚、上記の最尤法の表現形式は、石村・劉・石村（2010, p. 62）を参照にした。

式(21)の $L(\theta ; x_i)$ において、 $P(x_i ; \theta)$ とは、 x_i と θ の順番が入れ代わっていることに注意する。 $L(\theta ; x_i)$ の意味は、標本母集団の大きさ N において、観測値 $x_1, x_2, x_3, \dots, x_N$ が得られたことを条件として θ の値を最大化するという意味である。文献によっては、条件確率を使って $L(\theta | x_i)$ と表現しているものもある。また、特に x と θ の順番を入れ換えていないものもある。

$L(\theta ; x_i)$ は、式(20)のとおり、 N 個の積から成る同時確率である。かけ算は計算量が多く複雑であるため対数変換をする。これを対数尤度関数という。

$$\log L(\theta ; x_1, x_2, x_3, \dots, x_N) \quad (22)$$

次に, $L(\theta ; x_1, x_2, x_3, \dots, x_N)$ の値を最大化する未知のパラメータ θ を求めるために, 対数誘導関数を微分する。微分することによって, 瞬間的な関数の傾きの値を得ることができる(静, 2007, pp. 53–74)。これをスコア関数という。

$$U(\theta ; x_i) = \frac{\partial}{\partial \theta} \log L(\theta ; x_i) \quad (23)$$

スコア関数は, 確率変数であるため期待値と分散を求めることができる。

$$E[U(\theta ; x_i)] \quad (24)$$

式(24)のスコア関数の期待値は, 0になることが知られている。

確率変数の分散は, 以下の要領で求められる。これを分散公式と呼ぶ(村木, 2011, p.14)。

確率変数の分散 = [確率変数の2乗の期待値] - [確率変数の期待値の2乗]

$$Var[U(\theta ; x_i)] = E[U(\theta ; x_i)^2] - (E[U(\theta ; x_i)])^2$$

スコア関数の期待値は0であるから, $(E[U(\theta ; x)])^2 = 0$ となる。

$$Var[U(\theta ; x_i)] = E[U(\theta ; x_i)^2] \quad (25)$$

$U(\theta ; x_i)$ はスコア関数であるから, 式(25)は以下のように整理できる。式(26)は, フィッシャー情報量 $I(\theta)$ と呼ばれ, 豊田(2002, p. 65) 式(18)とも一致する。尚, 豊田(2002)では, 尤度関数を表現する際に x_i と θ の順番を入れ換えておらず, また, 条件付き確率の形式を使っているため, 本文の式とは若干異なるものになっている。

$$\begin{aligned} Var[U(\theta ; x_i)] &= E[U(\theta ; x_i)^2] \\ &= E\left[\left(\frac{\partial}{\partial \theta} \log L(\theta ; x_i)\right)^2\right] \\ I(\theta) &= E\left[\left(\frac{\partial}{\partial \theta} \log L(\theta ; x_i)\right)^2\right] \end{aligned} \quad (26)$$

さらに、豊田(2002, p. 66)では、項目反応理論の各モデルのフィッシャー情報量を紹介している。項目 j を受験者能力 θ を持つ受験者 i が受験した際のフィッシャー情報量は、項目反応理論の各モデルにおいて以下のようなになる（豊田, 2002, pp. 68–69）。

1 パラメーター・ロジスティックモデル

$$I_j(\theta_i) = 1.7^2 a^2 p_j(\theta_i) q_j(\theta_i) \quad (27)$$

2 パラメーター・ロジスティックモデル

$$I_j(\theta_i) = 1.7^2 a_j^2 p_j(\theta_i) q_j(\theta_i) \quad (28)$$

3 パラメーター・ロジスティックモデル

$$I_j(\theta_i) = 1.7^2 \frac{a_j^2 (p_j(\theta_i) - c_j)^2 q_j(\theta_i)}{p_j(\theta_i) (1 - c_j)^2} \quad (29)$$

上記モデルの内、1パラメーター・ロジスティックモデルと2パラメーター・ロジスティックモデルにおいて、フィッシャー情報量は項目困難度と等しいとき、その値が最大となる。3パラメータ・ロジスティックモデルにおいても、項目困難度からは大きくずれることはない（豊田, 2002, p. 69）。

受験者に対して問題項目を選択する際に重要な点は、受験者の能力を十二分に発揮させることである。そのためには、受験者が解けるか、それとも解けないか五分五分の難易度の問題項目を出題することが必要になる。なぜならば、易しすぎる問題や難しすぎる問題では、受験者が問題項目に取り組む前に正誤が決定してしまい「意外性」（静, 2007, p. 272）がなく、受験者の能力に関して新たな情報を得ることができないからである。以上のことから、項目選択においては、フィッシャー情報量を使って、 $I(\theta) = b$ 、またはその近似値を持つ問題項目が選択されることになる。

4. まとめ

本稿では、ベイズEAP推定法とフィッシャー情報量について概説をした。項目反応理論は外国語教育の分野でも適応型テストとして幅広く活用されながらも、その原理を理解することは敷居の高いものであった。数式の細かな展開式は理解しなくとも、原理を理解すれば数式の意味が見えてくる。原理が分かると絵のような見えていた記号の集まりの数式

から意味が見えてくる。そして、項目反応理論の醍醐味をより深く理解することができる。このことは著者自身の経験とも重なる。本稿がこれから項目反応理論を学ぶ読者の一助となれば幸いである。そして本稿の誤りや不十分な点は、新たな読者に是非とも指摘して頂きたい。

参考文献

- 石村 貞夫・劉 晨・石村 光資郎 (2010). 『入門はじめての統計的推定と最尤法』 東京図書株式会社.
- Khan Academy. (2009, February 24). *Expected Value: E(x)*. [Video file]. Retrieved from https://www.youtube.com/watch?v=j_Kredt7vY
- 村木 英治 (2011). 『シリーズ〈行動計量の科学〉8 項目反応理論』 朝倉書店.
- 大友 賢二 (1996). 『項目応答理論入門』 大修館書店.
- 静 哲人 (2007). 『基礎から深く理解するラッシュモデリング：項目応答理論とは似て非なる測定のパラダイム』 関西大学出版.
- 住 政二郎 (2013). 「ラッシュモデルの導出」『メソドロジー研究部会 2012 年度報告論集第 3 号』 83–101. Retrieved from http://www.mizumot.com/method/2012-07_Sumi.pdf
- 住 政二郎 (2014). 「項目反応理論：1PLM, 2PLM, 3PLM」『メソドロジー研究部会 2013 年度報告論集第 4 号』 34–62. Retrieved from http://www.mizumot.com/method/04-04_Sumi.pdf
- 住 政二郎 (2015). 「PROX 法と同時最尤推定法の概説」『メソドロジー研究部会 2014 年度報告論集第 6 号』 96–116 Retrieved from http://www.mizumot.com/method/06-06_Sumi.pdf
- 豊田 秀樹 (2002). 『項目反応理論 [入門編]：テストと測定の科学』 朝倉書店.
- 植野 真臣・永岡 慶三 (2009). 『e テスティング』 培風館.
- 涌井 良幸・涌井 貞美 (2010). 『Excel でスッキリわかるベイズ統計入門』 日本実業出版社.
- 涌井 良幸・涌井 貞美 (2012). 『史上最強図解 これならわかる！ベイズ統計学』 株式会社ナツメ社.

外国語教育メディア学会（LET）関西支部 メソドロジー研究部会
2015 年度 第 7 号報告論集

発行 2015 年 9 月 30 日
編集・発行 外国語教育メディア学会（LET）関西支部 メソドロジー研究部会
代表 山西 博之
〒564-8680
大阪府吹田市山手町 3-5-35
関西大学 外国語学部
Tel : 06-6368-1121
E-mail : hiyamanishi@kansai-u.ac.jp