

3 週間の短期海外語学研修が大学生の英語能力に及ぼす効果について

久野 寛之

抄録：本研究は、オーストラリア、ビクトリア州メルボルン市にあるラ・トロープ大学国際カレッジの外国人用英語研修プログラムで、現地の受入家庭にホームステイしながら、午前中は一般的な英語の授業を受け、午後は自分の専門分野に係る講義や活動に参加するという形で春期休暇3週間を過ごした大学1～3年生22名のうち、協力者8名に事前・事後テストを受けてもらい、その結果に基づいて、3週間のオーストラリアでの英語研修が大学生の英語力にどのような変化をもたらすかに関する予備研究を行った。この8名のグループと同時期に事前・事後テストを受けられる統制群を作ることができなかつたため、別の時期に全く同じ事前・事後テストを受けた大学1年生を、両テストをはさむ期間中に受けた処遇の種類によってグループに分け、対照群として比較することで、その8名の英語能力に何らかの有意な変化があったとみなすことができるかどうかを検証した。事前・事後テストの得点をクラスタ分析し、実験群の8名と同じクラスタに分類された学生24名を適合対照群として、実験群の8名と*t*検定及び分散分析によって比較したところ、日本の大学に在籍し、授業外で約10カ月自主的にEラーニング学習を行いながら、講義形式の教養英語教育を通年で受講したことによって対照群に起こった英語能力の変化と、3週間の海外研修中に実験群に起こった英語能力の変化がほとんど変わらない可能性のあることがわかった。

はじめに

2010年3月13日（金）～4月10日（金）までの3週間にわたり、北海道文教大学人間科学部の22名の学生（健康栄養学科5名、理学療法学科6名、作業療法3名、看護学科8名）がオーストラリアでの短期語学研修に参加した。この研修は、メルボルン市にあるラ・トロープ大学で開催され、2001年以来10年間にわたって外国語学部の学生が夏季と春季の年2度にわたって利用してきたプログラムである。諸般の事情から、期間は5週間から3週間に短縮されたが、参加学生の専攻に合わせた特別の研修内容を盛り込み、ESP（English for Specific Purposes）的な要素を加えた特別仕様のプログラムとして計画された画期的なものとなった。国際的な視野と世界で通用する英語力を兼ね備えた管理栄養士、理学療法士、作業療法士、看護師を育てたいという本学人間科学部の強い願いによって実現に漕ぎつけたもので、それだけでも意義は大きい。しかし、研修の意義はそれだけではなかった。海外研修という教育事業が人間科学部の目指す国際的な人づくりに貢献しうるものであることを実証するためのパイロット・プログラムとして今回の研修を役立てたいという人間科学部の願いが込められていた。本研究は、その願いを現実のものとするための努力の一つとして、3週間のような短期であっても、海外での研修が有効な教育効果を上げうるものであることを、英語力の変化という側面から実証しようと試みるものである。

1. 先行研究

2004年の82,945人をピークに、海外へ出る日本人留学生の総数は毎年減少してはいる¹が、今やどの大学も何らかの形で海外留学または語学研修のプログラムを持っていると言ってよいほど海外への門戸は開かれている。こうした機会の増加と反対に、期間は短期化している。アメリカの大学では、留学と言えば“junior year abroad”、すなわち、3年生の1年を海外で過ごすことを意味していたが、今や“semester abroad”（半期を海外で）となり、半期はさらに夏期休暇期間へと短期化しつつある。この傾向は、経済情勢、教養科目としての外国語学習の必修化、海外での長期研修が難しい専門教育課程の増加などによるものだという（Lafford & Collentine, 2006, p.115）。こうした事情は、日本においても同じであろう。

海外留学あるいは海外での語学研修の期間が短期化するにつれて、短期で効果の上がるプログラムが求められるのは当然である。また、通常の授業で学ぶ以外に海外で課外の英語学習をすれば何らかのプラス効果があって然るべきだが、たった3週間から5週間の短期で本当に有意な効果があるのかという疑問は、費用の大きさを考えると至極当然のものであり、その疑問に説得力のある回答ができるようになることは、短期海外研修を勧める教育機関側の責務と考える。

それにもかかわらず、短期海外研修が実際にどのような教育効果を上げているのかについて、日本で報告されてきた本格的な実証的研究の数はまだまだ乏しい。しかし、そのような事情は、最近まで欧米にも存在していたようである。Lafford & Collentine (2006)によると、アメリカでは、2,782人の大学4年生の語学力を調査し、単なる観光や夏休みの間だけの短期滞在でも外国語能力にプラスの影響があるとしたCarroll (1967)の研究がほぼ無条件に受け入れられ、外国での語学研修を促す根拠となってきた。しかし、90年代に入ると、60年代から80年代にかけて発表された海外での語学研修の効果に関する研究の問題点が明らかにされ始めた。サンプルのサイズや質が十分に制御されていない、処遇期間が短すぎる、対照する統制群を欠いている、語学力の測定テストの得点を拡大解釈しすぎているなど、実証研究としての問題点がMaera (1994)、Freed (1995)、Coleman (1996)らによって指摘され、90年代には、それらの不備を克服した研究が行われるようになってきた（Lafford & Collentine (2006), p.103）。

欧米の研究では、海外での学習が、語彙の習得を早めたり、発話の流暢さを高める効果のあることなどをはじめ、聞く・話す能力の向上へ貢献することが幅広く報告されている。（Lafford, 1995；Lafford & Collentine, 2006）。

一方、海外での留学や語学研修があまり効果をもたらさない技能分野としては、作文能力と文法能力が指摘されている（木村, 2006, p.3）。Lafford & Collentine (2006)は、90年代以降仏語、西語、露語、日本語学習者に関して行われたほとんどの研究が海外留学・語学研修のもたらすプラス効果を支持する結果を報告しているにもかかわらず、文法能力に対するプラス効果を報告するものがないのは「驚きの結果」だと指摘している（p.103）。ただ、注目すべきは、その「驚きの結果」が、初・中級レベルの学生が1 Semester（約4カ月）の短期海外研修に参加した場合のデータに限られているということだ。Isabelli & Nishida (2005)は、滞在期間が9ヶ月の上級学習者では文法能力に向上が見られたと報告しており（Lafford & Collentine, 2006, p.115）、文法能力の変化は、渡航時の言語能力や滞在期間によって影響を受けるという可能性を示している。

さて、欧米での研究は、共通項の多い印欧語族間での言語学習が中心で、短期でもプラスの効果が

あるという報告が主流を占めても不思議はない。しかし、英語と日本語のように、学習がきわめて困難とされている言語間²の場合は、どうなのか。本研究と同じく、休暇を利用した3～5週間程度の短期語学研修が日本人の英語学習に与える影響に関する研究を見てみる。

小林(1999)は、約3週間の夏季短期語学研修に参加した31名の学部生(2年生15名、3年生16名。男性は1名を除く30名が全て女性)に事前・事後テストとしてTOEFLのILPテストを行い、*t*検定の結果、総合点、語彙、聴解、読解の各セクションで有意差が出たことを報告し、「語彙・文法力の向上が最も顕著であり、続いて聴解と長文読解力の向上はわずかであった。」(p.97)「語彙・文法問題は短期間の学習で得点の向上が可能であることがわかる。これはTOEFLの指導に係っている教育者間での常識と一致する。」(p.94)としている。また、木村(2010)は、3週間の海外短期語学研修に参加した大学生に文法・作文・読解・聴解の問題75問と独立のライティング課題を与えたところ、作文・聴解・総合得点の3部門と、ライティングの流暢さに有意差があったという。1セメスターよりもはるかに短い語学研修で、作文能力や語彙・文法能力に有意な変化が見られたというのは、先に述べたLafford & Collentine(2006)の知見と矛盾するもので、今後の研究課題であると言える。

田浦他(2009)は、春季休暇を利用してニュージーランドでの3週間の短期語学研修に参加した学部・男女混合の大学生20名(2年生16名、3年生1名、院生1名)に、TOEFLの聴解テストの問題50問を用いて事前事後のテストを行った結果、有意な得点変化が見られたと報告している。さらに、研修1日目と最終日に行った英語インタビューの発話データから文間及び文中のポーズの長さの変化を測定することで、発話能力の変化を客観的に測ろうと試みたが、有意な変化は見られなかったという。

夏季休暇中約1カ月の短期語学研修を経験した短大1年生のデータを使って、「ホームステイの英語力への効果」を調べた上斗他(1989)、沼本他(1990)、沼本他(1991)は、語学研修を目的としたホームステイ前後において、参加者の聴解力と会話表現の知識が有意に向上したこと、向上は成績上位群にも、下位群にも見られたこと、また、研修直後の事後テストで現れた聴解力の差が、研修終了2か月後のテストでも再び観察されたことを報告している。ただ、語法・読解力・作文力では有意な向上は認められなかったという。

野中(2001, 2002, 2005)は、3カ月と6カ月の海外語学研修の効果をPre-TOEFLで測定した結果、3カ月では「ゆるやかな伸び」(野中, 2001)が、6カ月では「明らかな効果」(野中, 2002)が総合得点に見られ、特に聴解能力においては変化が著しかったのに対し、3週間の海外語学研修に参加した51名の短大1年生の能力変化をPre-TOEFLで測定した結果は、聴解・読解・総合得点のいずれにおいても有意差がなく、むしろ、文法部門では、有意な下降が観察されたと報告している。ただ、得点の上位群と下位群をグループに分けて統計的な検定を行った結果、上位群ではどの部門でも有意差が検出されなかったが、下位群では、読解と総合スコアでプラスの有意差があったという(野中, 2005)。

渡航時の英語能力と能力変化の度合いの関係については、異なる研究報告もある。沼本他(1990)では、成績上位群の聴解力の方が下位群よりも有意に向上したとされている。ところが、同じ聴解力の分野で、吉田、小寺(2009)は、野中(2005)を支持する結果を報告している。吉田・小寺(2009)は、2週間オーストラリアに滞在した高専生15名に事前・事後テストを行い、その結果に基づいて、語彙、表現、聴解のすべての項目において、相対的に能力の低い学生の方が高い学生よりも「伸びた」ということをCASEC³でのテストスコアを用いた研究で示唆している。統計学的な検定の報告が含

まれていなかったため、論文に掲載されていた得点（吉田・小寺、2009、pp.115-117）を用いて筆者自身が *t* 検定を行った結果、CASEC テストの 4 部門（語彙の知識、表現力の知識、聴解での大意把握、具体的情報の聞き取り能力）のうち、成績上位群ではいずれにおいても有意差はなかったが、下位群では、「聴解での大意把握」と総合得点において、有意差が検出された（両側検定：聴解 $t(6) = 3.547$ 、 $p < .05$ ；総合得点 $t(6) = 4.257$ 、 $p < .01$ ）。

言語習得面での効果のほか、ホームステイ効果の心理的効果も研究されている。

沼本他（1991）は、ホームステイをすることによって、帰国後、英語学習意欲が増進し、ホームステイをした国の人々に対するイメージや価値観への好感度の向上が見られること（渡辺、1979；樋口他、1982）や、「外国人に対してより積極的に接近したり、受け入れたりしようとする態度が認められた」（北川、1989）ことを報告している（pp.227-228）。一方で、箕浦（1995）は、海外の異文化でのホームステイ体験が、異文化への好感度を高め、自己や自文化への認識を深め、その意味での人間的成長へとつながる機会となるかどうかは滞在期間に影響されると報告している（東 2007、p.3）。異文化間メンタルヘルスを医学的に研究している秋山（1998）は、「異文化接触による対応は接触者が旅行者・一時的滞在者（留学等）・定住者（移民等）」かどうかによって異なり、3 週間程度の短期滞在の場合には、「不安や緊張感…期待はずれによる体調不良やホームシック」といった要因が否定的な影響を及ぼしやすくなることを指摘している（田浦他、2009、pp.18-19）。

情意要因については、もうひとつ重要な側面が知られている。情意要因が留学や語学研修の効果としての言語習得に影響するということである。沼本他（1991）は、渡航前の言語能力と渡航後のホームステイの効果の関連についてふれ、聴解能力が高いほど、「外国人への接近および受容的態度」の度合いが高くなる傾向があることを報告している（p.229）。また、日本人に関する研究ではないが、スペインへの 5 週間の語学研修に参加したアメリカ人を研究した Talburt & Stewart（1999）によると、不安や緊張感、個人的な期待外れといった要因以外にも、人種や性差別といった、自分にとって文化的に重要な問題を受入家庭との日々の生活の中で感じると、そのことが外国語能力の習得にも影響を与えるという（Lafford & Collentine、2006、p.109）。海外での短期語学研修に参加する学生からホストファミリーの「当り外れ」はしばしば耳にする。受入家庭という環境要因が研修の主目的である言語習得に及ぼす偶然的な影響をできる限り少なくして、短期語学研修の学習効果を実証するためには、研究に用いる参加者の数を十分に多く確保することが求められるようだ。

まとめると、Lafford & Collentine（2006）の主張するように、海外での留学や語学研修が外国語習得にもたらすプラスの効果は、流暢さやコミュニケーション方略などの向上に見られる口頭発話能力の向上や語彙の増加をはじめ、さまざまな技能や知識の領域に及んでいるが、どの領域においても、その効果について主張の異なる報告があり、その相違は、滞在期間、学習環境（ホームステイ）、渡航前の学習者の言語能力という 3 つの要因によってもたらされているようだ。したがって、実証研究としての精度を確保するためには、単にどの技能にどのような変化が起こるかを調査するだけでも、研修参加前の英語能力はもとより、言語学習や習得に影響するとされているその他の内的要因（性別、性格、認知スタイル、学習スタイルなど）と、現地で受ける英語教育やホストファミリーとのコミュニケーションの質と内容といった環境要因を、変数としてできる限り上手く制御しなければならない。しかし、それは容易なことではなく、海外語学研修が言語習得に及ぼす効果の研究にとっての大きな課題である。

今回の我々の研究では、渡航前の学習者の言語能力についてのデータはあるが、実験群のサンプルのサイズが8と少なく、後でも述べるように、同時期に別の統制群からデータを取ることができなかった。また、何らかの技能が、一時的に観察される変化としてではなく、確かに習得され、真の意味での効果として学習者に定着したと言えるためには、最低でも1セメスター4カ月（約16週）の処遇期間が必要だというLafford & Collentine（2006）らの示唆が正しければ、海外語学研修の効果の検証に3週間は短すぎたことになる。また、今回は、学習者の心理的な変化を調査したデータもない。このような理由から、今回の研究はあくまで一つの予備調査であると位置づけ、たとえ3週間という短い期間ではあっても、今後の海外研修の継続、発展と、その効果についての研究継続の必要性を示唆する有意な学習効果、すなわち、言語能力の変化があったかどうかを以下検証していくことにする。

3. 研究の目的

3週間のオーストラリア滞在の前後で、参加した学生の英語力に何らかの有意な変化はみられるのかどうかを、(1) 聴解能力A（短い単文の英語を聞き、その意味を理解する力）、(2) 聴解能力B（単文レベルの短いやりとりを聞いて、その意味を理解する力）、(3) 語彙・文法、と(4) 読解能力（10～15行くらいの文章を読み、その内容を理解する力）の技能分野のそれぞれについて検証することが本研究の目的である。

4. 研究方法

3週間のオーストラリア滞を含め、異なる処遇を与えた4つのグループ間の事前・事後テストの得点を統計的に比較することによって、目的とする検証を行った。以下に述べる全ての統計的比較には、SPSS社のPASW Statistics Ver.17.0を使用した。

4-1. 研究対象

4-1-1. 短期語学研修参加者

北海道文教大学人間科学部の1～3年生21名（専攻科、性別は「表1」参照）が、3月13日に日本を出発、オーストラリア、ビクトリア州メルボルン市に到着し、ラ・トロープ大学国際カレッジ（以下「カレッジ」）によってあてがわれたホストファミリーと週末を過ごした後、3月15日から、カレッジで開講されるELICOS（English Language Intensive Courses for Overseas Students）という集中講座で3週間英語を学んだ後、4月3日にメルボルンを出発し、帰国の途についた。人間科学部健康栄養学科の教員1名が引率者として同行した。

4-1-1-1 参加学生のプロフィール（表1）

学年 学科	参加者総数	内訳（男子／女子）		
		1年生	2年生	3年生
健康栄養学科	5	0／3	0／1	0／1
理学療法学科	6	0／0	0／3	1／2
作業療法学科	3	0／0	2／1	0／0
看護学科	8	0／5	1／2	0／0

4-1-1-2 短期語学研修の概要

初日の3月15日は、クラス分けテストがあり、日本人アドバイザーによるオリエンテーション、IDカード発行手続きとキャンパスツアーで一日が終わったが、それ以降は、毎日ホストファミリー宅からカ

レッジに通い、午前中はカレッジで4技能全般にわたる英語の授業を受け、午後は、学科別に分かれ、特別のプログラムに参加した。また、週末には団体での観光を楽しんだ。

(1) ELICOS の英語クラス (9:00 ~ 12:00)

クラス分けテストの結果、能力差があまりなかったことと、プログラム運営上の理由から、研修参加者22名が全員同じクラスで午前の授業を受けることになった。授業は全て英語で行われ、クラス内においては母語を使用しないように言われていたが、同じ大学の学生同士なので、教師からの指示などについて日本語で確認し合う場面も観察された。オーラルコミュニケーションを行うことに主体がおかれているようで、引率者の報告では、内容的にはきわめて簡単な話題で授業が進められていた。

(2) 昼休み (12:00 ~ 13:30)

長めの昼食時間が設定され、いろいろな自主活動プログラムに参加できるようになっていた。自主活動プログラムは、クラスの垣根を越えて、同じ趣味や関心を持つ学生が集まり、何か具体的な活動を通して英語を使い、使うことを通してより効果的に英語を学ぶことを目的としたもの。活動テーマには、会話、討論、各種のスポーツ、歌、料理、調査活動・調査報告、音楽鑑賞、オーストラリア映画の鑑賞、文法ゲーム、聞き取り練習、発音練習、就活技術、IELTS テストの準備、ロールプレイ、演劇、オーストラリア語などがあつた。引率教員の報告では、本学の学生は、あまり積極的にこの活動プログラムには参加していなかった。

(3) 各学科別に用意された午後の特別プログラム

学生たちは、英語で行われる専門別のプログラムに参加した。ラ・トローブ大学側から学生による日本語通訳が提供されたため、英語だけに集中できたわけではなかったが、学生通訳にも上手下手があり、常に通訳に頼りっ放しというわけではなかったようである。

- 講義・ワークショップ：「オーストラリアの保健制度」、「オーストラリアの教育」、「創傷の処置と管理」、「健康に配慮した学校給食」、「スポーツ栄養学」、「小児作業療法と理学療法」、「作業療法」
- 実習室見学：1時間半ずつ2つの実習室で、創傷管理、インドウェリングカテーテルの設置他の臨床技術を学ぶための看護実習を見学 [看護学科以外の学生も参加可]
- 病院・養護施設・学校訪問：Box Hill 病院、慈恵病院 (Mercy Hospital)、Villa Maria 高齢者養護施設、Warrigal 理学療法センター、Banksia 高等学校の食物工学の授業
- 実務家との交流：韓国人理学療法士の治療現場を見学、Isik 大学看護師訪問、ラ・トローブ大学看護学科学生と交流

(4) 観光旅行

メルボルン・クリケット競技場、LTU 野生動物特別保護区、メルボルン動物園、グレート・オーシャン・ロード (Great Ocean Road)、フィリップ島 (Phillip Island)、パフフィン・ビリー (Puffing Billy) 鉄道、メルボルン博物館

4-1-2. 比較のための学生群

4-1-2-1 学生群

今回の研究のために構成した学生群は合わせて4つあり、それぞれA群、B群、C群、D群と呼ぶことにする。短期海外語学研修に参加したグループA群。このいわゆる実験群としてのA群が受けた事前テストと事後テストとを同時期に受ける対照群を作ることができなかったために、同じテストを別の時期に受けていたグループを疑似対照群として用いることにした。(各グループの具体的なテ

スト実施時期は後述する。)

A 群： 短期語学研修参加者

前項 4-1-1 で述べた学生 8 名を A 群と呼ぶ。

B 群：

看護学科 1 年生 24 名。筆者の教える一般教養英語科目の「英語コミュニケーション I」、「同 II」を 2009 年度前・後期に通年で履修した。また、授業外で、E ラーニングの英語プログラム「ニュートン TOEIC®TEST 対策 A コース」を 4 月後半から翌年 3 月まで約 11 カ月使用し、4 技能にわたって基礎英語力を訓練した。この E ラーニング・プログラムは、TOEIC Bridge® の出題形式に即して学習を 5 つのパートに分け、各パート毎に 200 ～ 300 個の学習段階を設定して、学習者が、その段階を通して総計 2 万 2 千問を超える数の問題を解きながら学習を進めていくように作られている。この E ラーニングの課題達成率は、前期・後期の成績の一部とされ、30% の重みで成績評価に反映された。B 群と次に述べる C 群は、A 群が受けた事前テストと同じテストを受けていたので、得点合計ではなく、事前テスト問題 100 問に対する解答パターンをクラスタ分析し、A 群の 8 名のいずれかと同一クラスタに属する学生 24 名を集めて作ったのが B 群である。(クラスタ分析の結果の概要は後述する。)

C 群：

看護学科 1 年生 57 名。A 群と同じ事前テストでの解答パターンをクラスタ分析した結果、A 群の 8 名のいずれとも異なるクラスタに分類されたという点だけが B 群と異なる。

D 群：

短期大学経済学科の 1 年生 47 名。B 群、C 群と同じ教員(筆者)が教える一般教養英語科目「時事英語」を履修し、授業外で、無料で購読できるメールマガジン(週 4 回配信のテキストベースのものと、週 1 回配信の音声ファイル付のもの)を 1 年間購読した。各学生が配信された記事をどの程度読んでいるか、音声ファイルをどの程度聞いているかはデータ化されなかった。本研究にかかわる期間中は、授業でもテストでも、これらのインターネット教材の音声データは取り扱われず、英字新聞の見出しとリード(見出しに続く第 1 段落)の読解が指導の中心を占めた。B 群、C 群が 4 月下旬と 8 月上旬に繰り返して受けた事前テストの聴解・部門(Part 1, Part 2, Part 3)を 4 月下旬と 9 月上旬に繰り返して受けた。同じ事前・事後テストを受けた時期は異なるが、処遇期間が類似しているため、B 群・C 群の疑似対照群とした。この D 群は、「時事英語」と併行してもう一科目、4 技能を総合的に教える「総合英語」を履修していた。

4-2 各群の比較に用いたテスト

4-2-1 事前・事後テストの性格と内容

(1) テストの性格

事前・事後テストには、ニュートン社の E ラーニング・プログラム「TOEIC®TEST 対策 A コース」(以下「A コース」)の模擬テストとしてニュートン社から供与されていた 10 種類の模擬テストのうち 2 つ(以下「No.1」と「No.5」)を、ニュートン社の許可を得て使用した。2009 年度における同社の「TOEIC®TEST 対策 A コース」では、コンピュータが一定のアルゴリズムに従って学習用の問題群の中から 100 個の問題を選択し、それを TOEIC Bridge® の出題形式に再構成したものを、「模擬テスト」として、受講者がいつでもオンライン上で受験できるようになっていた。そのような仕組

みで生成された10個のテストサンプルを紙ベースで使えるようにしたテストの中から、事前テストには「No.1」を、事後テストには「No.5」を使用した。このことからわかる通り、「Aコース」の「模擬テスト」は、「Aコース」の受講者が普段から学習用素材として接している問題をそのままテスト問題として使っているため、TOEIC Bridge®という習熟度テストの「模擬テスト」でありながら、同時に、「Aコース」を受講している学生にとっては到達度テストでもあった。したがって、「Aコース」の受講者であるB群・C群の学生がこのテストの或る問題に「正解した」ということは、その問題と同質のいかなる問題にも正しく応答できる一般的な能力を有していたことを意味するのでは必ずしもなく、単にその問題の「正解」を正しく記憶していたに過ぎない可能性もあるということである。

(2) テストによって測定できる英語能力の水準の目安

ニュートン社によると「Aコース」は、TOEICスコア450～570点、実用英語技能検定（英検）で言えば準2級～2級のレベル⁴を超える英語力の獲得を目指すために開発された。そのように低めにレベルが設定されているため、「Aコース」では、学習用項目も模擬テストも、ともに、TOEICテストのジュニア版であるTOEIC Bridge®（テスト時間、問題数ともTOEIC®の半分で、問題の難易度も低い）に準拠している。TOEIC Bridge®は、20点～180点の範囲で英語能力が点数化されるが、(財)国際ビジネスコミュニケーション協会が公開している換算データ⁵によると、TOEIC Bridge®の90点～160点はTOEIC®の230点～570点に対応し、「Aコース」の目標であるTOEIC®スコア450～570点は、TOEIC Bridge®の150～160点に相当する。「Aコース」の模擬テストでは、各問が1点ずつ均等の重みを持ち、100点満点で得点が算出され、その後ユーザがその得点を180点満点に換算することによって、目安としてTOEIC Bridge®の点数を推定するようになっている。したがって、「TOEIC®スコア450～570点超プログラム」を謳うニュートン社の「Aコース」では、全内容を完全に消化すれば、100点満点の「Aコース」模擬試験の点数にして83点～89点を超える得点が期待できるということになる。以上のことから、「Aコース」模擬テストの得点の持つ意味は凡そ次のように表すことができる（表2）。

表2

「Aコース」模擬テスト	50	56	61	67	72	78	83	89
TOEIC Bridge®	90	100	110	120	130	140	150	160
TOEIC	230	260	280	310	345	395	470	570

ただ、「Aコース」模擬テストの得点は素点合計で、TOEIC®やTOEIC Bridge®のように項目応答理論による素点の換算値⁶ではなく、「Aコース」模擬テストと実際のTOEIC Bridge®との間の得点比較データもないので、あくまで、客観的根拠を欠く目安に過ぎない。

(3) テストの構成・内容とテストが測定する能力

上で述べた通り、「Aコース」の模擬テストは、TOEIC Bridge®の出題形式に準じ、5つのパートから構成されている。Part 1からPart 3までは聴解の部門で、Part 4とPart 5は読解部門である。作文能力と発話能力をテストするセクションはない。時間は、聴解が24分、読解が36分の、計60分。

Part 1は、画像（Pictures）描写問題15問。1枚の絵について4つの短い説明文が1度だけ読まれるのを聞いて、絵を正しく描写しているものを選ぶ問題で15問。実際のTOEIC Bridge®では写真（Photographs）が使われるが、「Aコース」の模試では絵が使われている。

Part 2は、応答問題、20問。1つの短い質問とそれに対する3つの短い答えをそれぞれ1度だけ

聞いて、質問に対して最もふさわしい答えを選ぶ問題。選ぶ応答は印刷されておらず、聞こえてくる音声のみを手がかりに解答する。

Part 3 は会話問題、15 問。2 人の人物が、言って、返して、また言う、3 行程度の短い英語でのやりとりや、3 文程度から成る独白的説明文などが 1 度だけ読まれる。その会話や説明文などを聞いて、問題用紙に印刷された質問と解答を読み、4 つの答えの中から最も適当なものを選び解答用紙にマークする。

Part 4 は、文法・語彙問題、30 問。空所を補完して不完全な文章を完成させるために、4 つの答えの中から最も適当なものを選ぶ。

Part 5 は、読解問題、20 問。いろいろな文章と、それに関する数問の質問を読み、4 つの答えの中から最も適当なものを選ぶ。

重要なことは、聴解部門の 3 つのパートのうち Part 3 は、英語で書かれた 4 つの選択肢の意味を瞬時に読んで理解しなければ正答できないという意味で、聴解問題でありながら、聴解と読解の二つの技能を同時にテストしているということである。各部門の得点が表す意味を理解する意味で忘れてはならない。

4-3 研究手順

4-3-1 事前・事後テストの実施時期

ニュートン社からテスト No.1 とテスト No.5 の使用許可を得、下表 3 のスケジュールで事前テストと事後テストを実施した。

表 3

検証 1		事前テスト (「A コース」 No.1) 実施時期		事後テスト (「A コース」 No.5) 実施時期
	A 群	2010 年 2 月		2010 年 4 月
B 群 C 群	2009 年 4 月		2010 年 2 月	
検証 2		事前テスト (「A コース」 No. 1 Part 1, 2 & 3) 実施時期	事後テスト (「A コース」 No. 1 Part 1, 2 & 3) 実施時期	
	B 群 C 群	2009 年 4 月	2009 年 8 月	
	D 群	2010 年 4 月	2010 年 9 月	

「A コース」模擬テスト「No.1」(N=99)と「No.5」(N=95)(以後「テスト No.1」、「テスト No.5」)について、総得点と Part 1～Part 5 の 6 項目での信頼度を調べた結果、No.1 は $\alpha = .752$ 、No.5 は、 $\alpha = .765$ (ただし、Part 1 を除くと、 $\alpha = .786$) であった。また、Part 1～Part 5 の 5 項目での信頼度を調べた結果、No.1 は $\alpha = .635$ 、No.5 は、 $\alpha = .688$ (ただし、Part 1 を除くと、 $\alpha = .713$) となった。このことから、聴解能力の変化を見るための比較分析から Part 1 の得点は除外すべきであることが示唆されたが、本研究では、念のために Part 1 の分析も含めて報告する。

4-3-2 比較の手順・方法

短期海外研修に参加したグループ (A 群) を実験群とし、同時期に同じ事前・事後テストを実施する統制群を作って実証実験を行うことができなかった。また、処遇期間が 3 週間と短かったため、事前・事後テストに同じテストを繰り返し使うことができず、また、同じ能力を測定可能であることがわかっ

ているテストの二つの異なる版を使うこともできなかった。そのため、A群と同じ事前テストを受けたグループ（B群とC群）の中から、クラスタ分析によりA群と同じクラスタに分類されたケースで構成されるグループ（B群）を抽出して適合対照群とし、基本的に、この二つのグループ（A群とB群）の間のさまざまな差について統計的な検定や分析を行うことによって、A群における短期海外研修の効果について推論的な判断を行うこととした。

5. 結果

5-1 検証1

5-1-1 疑似対照群をグループ化するためのクラスタ分析

Eラーニングを受講しながら1年間英語の4技能を学んだ看護学科の学生グループ（N=91）を、事前テストでの解答パターンがA群と等質な適合対照群（B群）と非・適合対照群（C群）に分けるために、実験群の8名と看護学科学生91名を合わせた99名からなるグループの100問における正解（1）・不正解（0）の2値データに階層クラスタ分析（クラスタ化はWard法、測定方法は平方ユークリッド距離）を行った。8名が全て異なるクラスタに属することが考えられたので、範囲を8クラスタから始め、24クラスタまで分析した結果、17クラスタ目で、B群に含まれる等質ケース数が下限（N=27）に達し、A群の8名に最も近づくと同時に、A群内のばらつき（所属クラスタ数）が20クラスタ目で6から7に転じたので、クラスタ数19でグループ分けを終了することにした。その結果、得点順（昇順）に並べた各クラスタの特性は次のようになった（表4）。A群の8名が成績上位群と下位群に分かれているのが見て取れる。

表4

クラスタ	度数	平均値	標準偏差	最大値	最小値	クラスタ	度数	平均値	標準偏差	最大値	最小値
19	2	30.00	.000	30	30	7	8	53.75	7.421	67	42
15	4	37.75	5.679	44	32	17	4	54.50	4.203	59	50
9	5	42.20	4.087	47	38	13	5	54.60	4.159	60	50
8	4	42.50	3.697	47	39	18	2	55.00	2.828	57	53
5*	5(1)	45.20	6.140	53	38	11	3	58.00	4.359	61	53
3*	6(1)	45.50	3.391	52	42	16	10	58.50	4.950	67	53
14	3	46.33	5.508	50	40	2**	7(2)	59.29	6.525	70	52
12	10	49.80	3.910	56	43	4**	4(1)	62.75	6.652	68	53
10	4	50.50	5.745	58	46	6**	10(2)	66.10	4.701	73	59
1*	3(1)	53.33	5.033	58	48	合計	99	52.60	9.479	73	30

※印はA群の8ケースが所属するクラスタ、()内の数字はA群のケース数を示している。

5-1-2 検証1：A群とB群・C群との比較

5-1-2-1 記述統計

表5 全パートの総得点

群	事前テスト					事後テスト				
	N	Mean	SD	最高値	最小値	N	Mean	SD	最高値	最小値
A群	8	56.13	12.620	73	38	8	58.38	11.710	75	46
B群	24	55.96	9.299	72	40	24	60.92	8.772	74	48
C群	63	50.24	8.492	67	30	63	55.59	8.761	83	37
全体	95	52.18	9.385	73	30	95	57.17	9.223	83	37

表6 Part 1 (絵と英語の説明を一致させる問題 15問)

群	事前テスト					事後テスト				
	N	Mean	SD	最高値	最低値	N	Mean	SD	最高値	最低値
A群	8	13.88	0.991	15	12	8	13.75	0.886	15	13
B群	24	14.17	0.963	15	11	24	14.42	0.881	15	11
C群	63	12.79	2.041	15	4	63	14.25	0.897	15	12
全体	95	13.23	1.853	15	4	95	14.25	0.899	15	11

表7 Part 2 (英語の質問と応答を聞いて、適切なものを選ぶ問題 20問)

群	事前テスト					事後テスト				
	N	Mean	SD	最高値	最低値	N	Mean	SD	最高値	最低値
A群	8	11.75	3.576	17	5	8	12.13	2.232	15	8
B群	24	10.67	2.078	14	6	24	13.67	2.582	18	9
C群	63	9.22	2.814	15	2	63	12.22	2.871	20	5
全体	95	9.80	2.823	17	2	95	12.58	2.800	20	5

表8 Part 3 (短い会話、独白を聞いて、内容と一致する記述を選ぶ問題 15問)

群	事前テスト					事後テスト				
	N	Mean	SD	最高値	最低値	N	Mean	SD	最高値	最低値
A群	8	6.25	1.581	8	4	8	4.63	1.506	7	2
B群	24	5.96	2.156	11	3	24	4.79	1.719	8	2
C群	63	5.44	2.054	12	1	63	4.65	1.806	12	2
全体	95	5.64	2.047	12	1	95	4.68	1.746	12	2

表9 Part 4 (語彙・文法を問う選択式空所補充問題 30問)

群	事前テスト					事後テスト				
	N	Mean	SD	最高値	最低値	N	Mean	SD	最高値	最低値
A群	8	15.63	5.731	23	9	8	16.50	4.781	26	11
B群	24	16.42	4.074	23	10	24	17.71	3.962	24	12
C群	63	15.44	3.950	25	7	63	15.54	3.830	24	9
全体	95	15.71	4.120	25	7	95	16.17	4.012	26	9

表10 Part 5 (最大200語程度の英語を読んで、内容と一致する記述を選ぶ問題 20問)

群	事前テスト					事後テスト				
	N	Mean	SD	最高値	最低値	N	Mean	SD	最高値	最低値
A群	8	8.63	4.307	14	2	8	11.38	4.779	17	6
B群	24	8.75	3.791	16	3	24	10.33	2.944	17	6
C群	63	7.33	2.984	14	2	63	8.92	2.991	15	2
全体	95	7.80	3.347	16	2	95	9.48	3.228	17	2

A群、B群、C群の中から事前・事後テストのどちらか一方しか受けなかったケースを除き、5つのパートの合計100問についての各グループの処遇前、処遇後の「平均値」(Mean)、「標準偏差」(SD)、「最高値」、「最低値」は、上記表5～表10の通りであった。どのグループにも平均点の向上が見られるが、A群のPart 1の得点に0.13の減少、全グループのPart 3の得点に減少が見られる。テストの説明の際に述べたように、No.1とNo.5は同じ学習用問題プールから抽出されて作られた模擬テストだが、2つの模擬テスト間の同質性に関する統計的なデータがない。したがって、No.5のPart 3の問題は、たまたまNo.1の問題よりも難度の高いものが多く抽出されてテストとして作られてしまったのかもしれない。よって、各グループのNo.1とNo.5の得点平均の差の比較によって能力の変化を検証することはできないと考え、3グループ・2水準の群差(事前事後テスト間の処遇の差異)を要因、各テストの総合得点及び各パートの得点を従属変数として、一元配置の分散分析を行い、二つのテスト

の各項目において、各群間の差がどのように変化するかを調べ、群間の差の変化によって、言語能力の変化の有無を推論することにした。

5-1-2-1 A 群、B 群、C 群間の分散分析の結果

5-1-2-1-1 総得点の分散分析：

事前テストでは、3 群間に 5% 水準で有意な主効果がみられた ($F(2, 92) = 4.279, p < .05$)。そこで、サンプルサイズが等しくない場合に用いる Tukey-Kramer の多重比較を行った結果、A 群と C 群の間には 5% 水準でも有意差が見られなかったが、B 群と C 群の間の差は 1% 水準でも有意となった。A 群と B 群との間には有意差はなかった。B 群は A 群の等質グループとして構成したので当然の結果と言えるが、A 群と B 群が等質なら、A 群は等質グループの B 群が有意差を示した C 群に対して有意差を示してよいはずだが、有意な差はなかった。サンプルサイズの問題かもしれない。

一方、事後テストでも 3 群間に 5% 水準で有意な主効果がみられた ($F(2, 92) = 3.110, p < .05$)。そこで、多重比較 (Tukey-Kramer) を行った結果、総得点においては、3 つのグループの関係は事前テストの場合と全く同じで、B 群と C 群の間だけに 5% 水準で有意差が見られた。

5-1-2-1-2 各パート (Part 1 ~ Part 5) の分散分析：

総得点と同様に、事前、事後テストのそれぞれにおいて、3 グループの平均の差を見るために、各パートについて一元配置の分散分析を行った。

事前テストでは、読解の要素を含む Part 3 以外の、純粹に聴解能力だけを問う 2 つの部門 Part 1 と Part 2 で主効果があった (Part 1: $F(2, 92) = 5.841, p < .01$; Part 2: $F(2, 92) = 4.702, p < .05$)。一方、語彙・文法の部門 (Part 4) と読解・部門 (Part 5) では主効果がなかった。そこで、Part 1 と Part 2 について多重比較 (Tukey-Kramer) を行った結果、Part 1 では、B 群と C 群との間に 1% 水準で有意差があった。(ただし、Part 1 の得点では等分散性が検定されず、分散分析が使えないことがわかったため、平均値同等性の耐久検定 (Welch/Brown-Forsythe) を行ったところ、平均の差は有意であることが判明したので、分析を続行した。) Part 2 では、A 群と C 群、B 群と C 群の間に 5% 水準で有意差が見られた。

一方事後テストでは、主効果が検出されたのは Part 5 のみ ($F(2, 92) = 3.318, p < .05$) で、Part 5 以外のどの部門でも主効果が検出されなかった。Part 5 の得点について多重比較 (Tukey-Kramer) を行ったが、A 群と C 群の間に 10% 水準で有意傾向 ($p = .101$) がみられただけで、有意なグループ間の差は見られなかった。

以上の結果をまとめると、事前・事後テストの総合点の比較では、はじめの 3 グループ間の差、すなわち、A 群と B 群、A 群と C 群には有意差がないが、B 群の得点は C 群の得点を有意に上回っているという関係が処遇後も変わらずにそのまま存在していた。ところが、各パートの比較をしてみると、事前テストで検出された聴解・部門 (Part 1、Part 2) での差が、事後テストでは消えていた。このことから、次の 2 つの解釈が可能になる。

- (1) 事前・事後のテストにおいて、A 群と B 群には変化がなかったが、C 群だけは、A 群、B 群よりも劣っていた聴解能力で、両グループに追いついた。
- (2) 事前・事後のテストにおいて、A 群と B 群はともに変化したが、いずれのグループよりも劣っていた C 群が、その変化を上回る変化を遂げた。すなわち、3 つのグループすべてが英語能力において向上した。

そこで、上記の (1)、(2) のどちらの解釈が妥当なのかを判断するために、次の検証 2 を行った。

5-2 検証 2：B・C 群と D 群との比較

5-2-1 記述統計

B・C 群と D 群は、表 3 に示した通り、それぞれ 4 月と 8 月、4 月と 9 月に全く同じテスト (No.1 の Part 1、Part 2、Part 3) を 2 度、繰り返し受けた。それぞれのグループの中から事前・事後テストのどちらか一方しか受けなかったケースを除き、両テストを受けた 131 名の、聴解部門 (Part 1、Part 2、Part 3) の合計 50 問について各グループの処遇前、処遇後の「平均値」(Mean)、「標準偏差」(SD)、「最高値」、「最低値」を記すと、下記表 11～表 14 の通りであった。E ラーニングで授業外の時間に聴解練習をするよう指示されていたか、いなかったかで異なるこの 2 つのグループの間の平均差を *t* 検定で調べてみる。

表 11 聴解・部門の合計点 (50 点)

群	事前テスト					事後テスト				
	N	Mean	SD	最高値	最低値	N	Mean	SD	最高値	最低値
B 群 & C 群	87	28.39	4.840	36	11	87	31.37	4.454	41	18
D 群	44	28.89	3.919	40	21	44	28.43	3.500	34	21
全体	131	28.56	4.542	40	11	131	30.38	4.372	41	18

表 12 Part 1 (絵と英語の説明を一致させる問題 15 問)

群	事前テスト					事後テスト				
	N	Mean	SD	最高値	最低値	N	Mean	SD	最高値	最低値
B 群 & C 群	87	13.17	1.912	15	4	87	14.78	0.769	15	9
D 群	44	13.70	1.268	15	10	44	13.48	1.320	15	10
全体	131	13.35	1.736	15	4	131	14.34	1.162	15	9

表 13 Part 2 (英語の質問と応答を聞いて、適切なものを選ぶ問題 20 問)

群	事前テスト					事後テスト				
	N	Mean	SD	最高値	最低値	N	Mean	SD	最高値	最低値
B 群 & C 群	87	9.63	2.711	15	2	87	10.63	2.520	16	5
D 群	44	9.91	2.568	17	5	44	9.89	2.244	13	5
全体	131	9.73	2.658	17	2	131	10.38	2.448	16	5

表 14 Part 3 (短い会話、独白を聞いて、内容と一致する記述を選ぶ問題 15 問)

群	事前テスト					事後テスト				
	N	Mean	SD	最高値	最低値	N	Mean	SD	最高値	最低値
B 群 & C 群	87	5.59	2.083	12	1	87	5.95	2.312	10	1
D 群	44	5.27	1.921	9	1	44	5.07	1.634	9	1
全体	131	5.48	2.028	12	1	131	5.66	2.144	10	1

5-2-2 事前・事後テストにおける各グループ間の差の検定

事前テストの得点を上記の各項目において *t* 検定を行ったところ、どの項目においても、両グループの有意差は検出されなかった。一方、事後テストにおいては、B・C 群の平均点がわずかずつであるが、表 11～14 の 4 つの項目のいずれにおいても、D 群の得点を上回った。*t* 検定をかけた結果、Part 3 では 5% 水準で有意差 (両側検定: $t(115.154)=2.535, p<.05$) が、Part 1 と、Part 1・Part 2・Part 3 の合計点とにおいて 1% 水準で有意差が検出された (Part1 両側検定: $t(58.159)=6.054, p<.01$; Parts 1, 2 & 3 合計 両側検定: $t(129)=3.815, p<.01$)。このことは、グループ B 群・C 群に何らかの変化が起こり、D 群との間に事前テストでは見られなかった差が事後テストにおいては現れたことを物語っている。では、どのような変化が起こっていたのか、それを調べるために、次に、対応のあるサンプルの *t* 検定を行い、各グループ内で事前・事後テストの平均を比較した。

5-2-3 各グループの事前・事後テストにおけるグループ内の差の検定

B群・C群全体と、D群、この二つのグループの事前・事後テストでの得点平均の差を t 検定にかけてみた結果、D群では、どの項目においても有意差は見られなかった。これによって、事前テスト (No.1) については、3カ月程度の期間を置いて繰り返して受けても、単純な<繰り返しによる慣れの効果>が学習効果と誤って検出される可能性がきわめて低いことが確認された。一方、B群・C群全体では、Part 1の平均と、Part 1・Part 2・Part 3の合計得点の平均に1%水準で事前・事後の有意差が見られた (Part 1 両側検定: $t(86)=-7.024, p<.01$; Part 1, Part 2, Part 3の合計点両側検定: $t(86)=-4.105, p<.01$)、Part 2の平均に5%水準で有意差が検出された (Part 2 両側検定: $t(86)=-2.544, p<.05$)。また、B群、C群それぞれの事前・事後の得点差を t 検定で比較してみると、B群ではPart 1に有意差 (両側検定: $t(22)=-2.405, p<.05$) が、C群では、Part 1, Part 2, Part 3の各平均とその合計得点のすべてに有意差が検出された (Part 1 両側検定: $t(59)=-6.874, p<.01$; Part 2 両側検定: $t(59)=-3.300, p<.01$; Part 3 両側検定: $t(59)=-2.149, p<.05$; Part 1, Part 2, Part 3 両側検定: $t(59)=-5.179, p<.01$)。

この検証2では、事前・事後テストに、両グループが使用していたEラーニングの到達度テストの性格を持つテスト (No.1) を用いたため、テスト上で検出された得点平均の有意差は、同じテストを繰り返し受験することによる慣れの効果ではないと言うことはできても、約4カ月の授業外の自主学習によって習得され、内在化した「英語能力」の変化だと結論づけることは難しい。むしろ、日頃の練習によるテスト問題やテスト形式への慣れによって、到達度テストにおける得点能力が向上したことを意味するにすぎないのかもしれない。

しかし、いずれの場合であっても、Eラーニングを4カ月間使用し、テストNo.1を事前・事後テストとして受けたB群・C群の事前・事後の得点平均差が統計的に有意で、その統制群Dの事前・事後の得点平均差が有意でなかったことによって、B群にもC群にも得点の有意な上昇をもたらすプラスの変化があったということが検証2において判明したので、前項検証1の最後に提示した2つの解釈のうち、第(2)の解釈、すなわち、「事前・事後のテストにおいて、A群とB群はともに変化した、いずれのグループよりも劣っていたC群が、その変化を上回る変化を遂げた。すなわち、3つのグループすべてが英語能力において向上した。」を採用すべきであることがわかった。

6. まとめ

6-1 まとめと考察

検証1において、それぞれ、3週間と10カ月という異なる処遇期間に海外語学研修とEラーニングという異なる処遇を受けたA群と、その対照群B群、C群が互いにグループとして示した関係をまとめると、下記表15のようにまとめることができる。

表 15

総合得点		Part 1		Part 2	
事前テスト	事後テスト	事前テスト	事後テスト	事前テスト	事後テスト
A ≐ B	A ≐ B	A ≐ B	A ≐ B	A ≐ B	A ≐ B
A ≐ C	A ≐ C	A ≐ C	A ≐ C	A > C	A ≐ C
B > C	B > C	B > C	B ≐ C	B > C	B ≐ C

A: 短期海外語学研修群 B: Aとの等質Eラーニング群 C: 非等質Eラーニング群
この表内の等号≐は有意差が見られなかったことを示している。□は事前と事後でグループ間の得点平均に変化があった場合を示している。

このデータを根拠として、検証1と検証2による結論を次のようにまとめることができる。

検証1によって、同一処遇群内（Part 1におけるB群とC群）だけでなく、異なる処遇群間（Part 2におけるA群とC群）にも事前・事後の有意差が検出された。したがって、もし異なる処遇群のA群とB群のいずれか一方に有意な変化が生じたとすれば、その変化が検出可能であることが分かった。一方、検証2によって、B群とC群においてプラスの変化が生じていたことを確認することができた。B群ではPart 1に、C群では、Part 1, Part 2, Part 3とその合計得点のすべてにプラスの有意差が検出された。おそらくその変化のために、Part 2で、事前テストではA群、B群のいずれよりも有意に低かったC群が、事後テストでは有意差を検出されなくなるほど得点が向上していたと考えられる。同様に、もし、B群にもC群にもプラスの変化が生じる一方で、本来等質であるはずのA群に何の変化もなかったならば、検証2でB群に有意なプラスの変化が検出されたPart 1で、A群と有意差がなかったB群の得点が、事後テストにおいてA群よりも有意に高くなってもおかしくはない。しかし、そのような有意差は検出されなかった。このことは、B群の変化と同様のプラスの変化がA群にもあったが、しかし、その変化の度合いは、C群の聴解能力にもたらされた効果ほどではなかった、と推論することによって説明がつく。

以上のことから、3週間の海外語学研修中に、ホストファミリーとの英語での生活をはじめ、英語による様々な学習や体験活動に参加することによって、参加学生の語彙・文法能力や読解力には有意な変化はなかったが、簡単な短い発話を英語で聞いて理解する能力はプラスに向上したと考えられ、また、その変化は、1年間授業外でEラーニング学習を行った学生群に見られた効果に匹敵するものであった可能性があるとして推論することができる。⁷

6-2 この研究の問題点と今後の課題

今回は、推論と入手可能な限りの客観データを組み合わせることで、3週間の海外研修の効果を示唆する結果を出すことができた。ただ、用いた方法は透明性の高い実証的検証とはほど遠く、導き出した結果は多くの見えざる変数によって汚染されている。今回はあくまで予備検証なので、次回までには、3週間の海外研修という処遇期間に対応して実験群を設定できるような工夫をし、能力測定手段として、到達度テスト的な要素が一切なく、信頼性の高い熟達度テストを確保する方法も確立しなければならない。また、研究の内容で言えば、今回は紙面の都合で触れることができなかったが、先行研究のいくつかで言及されている能力の上位・下位という要因の関与についての分析が必要で、今回のデータからわかることについては別の場で報告する。また、B群、C群の得点変化とEラーニングの学習時間との関係⁷についても触れる余裕がなかった。これについても場を改めて報告することとする。

謝辞： 最後に、引率の労を取り、現地で直接観察された貴重な情報をご提供くださった健康栄養学科准教授峯尾仁先生、ラ・トロープ大学から送られてきた全ての情報を日本語に翻訳され、その翻訳資料を快く使わせてくださった理学療法学科教授宮本重範先生、学生と一緒に事前・事後テストを受け、学生にこの研究への協力を率先して呼びかけ、データ収集に決定的な役割を果たしてくださった作業療法学科教授渡辺明日香先生に心からお礼を申し上げます。

注：

1. 文部科学省のホームページ「平成 22 年 12 月文部科学省 報道『日本人の海外留学者数』について」(p.2) の「日本から海外への留学者数の推移」(http://www.mext.go.jp/b_menu/houdou/22/12/_icsFiles/afeldfile/2010/12/22/1300642_1.pdf) より
2. ある言語の母語話者にとって相対的に学習困難度が高い又は低い言語は存在する。最もよく知られる基準は、アメリカの外務職員養成所 (Foreign Service Institute) が使用している ILR Scale (Interagency Language Roundtable Scale 連邦政府機関言語問題連絡会議基準) と呼ばれる基準で、発話能力と読解能力がレベル 3 (一般実務レベル) に達するまでに要する時間に基づいて、アメリカ人にとっての学習困難度をカテゴリー I ~IV の 4 つに分けている。日本語は、アラビア語、中国語、韓国語と並んで、最も学習困難なカテゴリー IV に属している。アメリカ人にとっての日本語の難しさがそのまま日本人にとっての英語の難しさとして解釈できるわけではないが、FSI の分類は、少なくとも、欧米の外国語学習者がお互いの言語を学び合うのと同じ速さで日本人が英語を学習することが困難であることを示唆しており、印欧語族の間での短期語学研修の効果についての研究報告を、そのまま単純に日本人の英語学習に適用することはできないということの意味している。
3. CASEC (Computerized Assessment System for English Communication) は、財団法人日本英語検定協会と教育測定研究所が開発し、後者が運営する英語コミュニケーション能力判定テスト。
4. 2001 年 4 月から 2010 年 3 月までのデータによると、準 2 級保持者の平均は 401 点、2 級保持者の平均は 533 点となっている。<http://www.toeic.or.jp/toeic/pdf/data/DAA2009.pdf> (2010 年 12 月 22 日アクセス)
5. 日本と韓国の受験者 15,569 名のデータに基づく「予測」として下記 URL に掲載されている。http://www.toeic.or.jp/toeic/pdf/data/Comparison_BridgeandTOEIC.pdf (2010 年 12 月 22 日アクセス)
6. 偶然による正解 (Chance Score) の確率や、異なるテスト毎の難易度差による測定結果の誤差を最小に抑えるために、素点による単純な得点算出法ではなく、項目応答理論にもとづく換算方式が採られている。http://www.toeic.or.jp/toeic/faq/faq_03_9.html (2010 年 12 月 22 日アクセス)
7. 実は、E ラーニング学習を行った学生群の問題修了率は、Part 1 から Part 5 まで、それぞれ、Part 1 (90%)、Part 2 (67.8%)、Part 3 (31.8%)、Part 4 (13.9%)、Part 5 (6.1%) で、また、平均学習時間は、Part1 (約 19 時間)、Part2 (約 10 時間)、Part3 (約 8 時間)、Part4 (約 2 時間)、Part5 (約 1 時間)、年間総平均学習時間は約 40 時間であった。つまり、聴解力だけでなく、速読力をも訓練する必要のある Part 3 の聴解問題や、文法・語彙の Part 4、読解力の Part 5 はほとんど学習されなかったことがわかっている。したがって、Part 3 から Part 5 で実験群 A と対照群の B 群・C 群との間に有意差が見られなかったのは、単に、E ラーニングという処遇が発生していなかったことによるのであって、決して、その処遇の効果がなかったということではないことをここで明確に指摘しておかなければならない。

文献

- Carroll, J. B. (1967). Foreign language proficiency levels attained by language majors near graduation from college. *Foreign Language Annals*, 1, 131-51.
- Coleman, J. (1996). *Studying languages : a survey of British and European students*. London : CILT.
- Freed, B.R. (1995). *Second language acquisition in a study abroad context*. Amsterdam/Philadelphia: John Benjamins Publishing Co.
- Isabelli, C. A., and Nishida, C. (2005). Development of Spanish subjunctive in a nine-month study-abroad setting. In *Selected proceedings of the 6th conference on the acquisition of Spanish and Portuguese as first and second languages*, ed. D. Eddington, 78-91. Sommerville, MA: Cascakilla.
- Lafford, B.(1995). Getting into, through and out of a situation: A comparison of communicative strategies used by students studying Spanish abroad and 'at home.' In B. R Freed (Ed.) *Second Language Acquisition in a Study Abroad Context*. (pp. 97-122). Amsterdam/Philadelphia: John Benjamins Publishing Co.
- Lafford, Barbara. (2004). The effect of the context of learning on the use of communication strategies by learners of Spanish as a second language. *Studies in Second Language Acquisition*, 26, 201-225.
- Lafford, B., and Collentine, J. (2006). The effects of study abroad and classroom contexts on the acquisition of spanish as a second language: From Research to Application. In Salaberry, Rafael & Lafford, Barbara A. (eds.). (2006) *The Art of Teaching Spanish: Second Language Acquisition from Research to Praxis* (PP. 103-126). Washington, D.C.: Georgetown University Press.
- Milton, J., and Meara, P. (1995). How periods abroad affect vocabulary growth in a foreign language. *ITL Review of Applied Linguistics*, 107/108, 17-34.
- Talburt, S., and Stewart, M.A. (1999). What's the Subject of Study Abroad?: Race, Gender, and "Living Culture." *The Modern Language Journal*, 83 (2), 163-175.
- Tanaka K., and Ellis, R. (2003). Study abroad, language proficiency, and learner beliefs about language learning. *JALT Journal*, 25(1), 63-55.
- 吉田三郎、小寺光雄 (2009) 「短期海外語学研修が高専学生の英語力にもたらす効果」『福井工業高等専門学校研究紀要 人文・社会科学』43, 111-122.
- 山内優佳 (2010) 「留学前後における英語聴解方略使用の変化」第36回全国英語教育学会大阪研究大会自由研究発表・事例報告 (口頭発表・ポスター発表・賛助会員プレゼンテーション) プログラム・アブストラクト (2010-07-21 版)」2010年8月8日 (日) 関西大学千里山キャンパス第1学舎 (2011年1月5日アクセス : http://keles.web.infoseek.co.jp/jasele2010osaka/abstracts_web/abstracts/pa8.html)
- 小林敏彦 (1999) 「海外短期語学研修で英語力はどのくらい伸びるものか」『小樽商科大学 人文研究』97, 83-100.
- 沼本健二, 黒田ディアナ, 北川歳昭, 上斗晶代, 福森護 (1991) 「ホームステイの英語力への効果 (III)」『中国短期大学紀要』22, 227-242.
- 沼本健二, 上斗晶代 (1990) 「ホームステイの英語力への効果 (II)」『中国短期大学紀要』21, 135-142.
- 上斗晶代, 沼本健二 (1989) 「ホームステイの英語力への効果 (I)」『中国短期大学紀』20, 165-177.

- 大塚賢一 (2009) 「茨城工業高等専門学校における海外短期語学研修の言語学的教育効果検証」. 『茨城工業高等専門学校研究彙報』 44, 23-30.
- 田浦秀幸, 堀井耕太郎, 馬西卓徳, 岡田宏子, 清水大介, 柏本恵未, 戸成辰也 (2009) 「ニュージーランド短期英語研修の効果に関する一考察」『大阪府立大学人間社会科学部言語文化学科 言語文化学研究 (言語情報編)』 4, 1-22.
- 東正訓 (2007) 「高校生の海外留学プログラムの評価—地域連携型研究の一試みとして—」『追手門学院大学地域支援心理研究センター紀要』 4, 2-12.
- 樋口勝也 (1983) 「短大生の異文化体験」、文化と人間の会 (編) 『異文化との出会い』 (pp. 53-69.) 川島書店
- 北川歳昭 (1989) 「意識の国際化に関する研究 (I) —短大生のホームステイ経験がおよぼす心理的効果」『中国短期大学紀要』 20, 75-86.
- 箕浦康子 (1995) 「若者の異文化体験と国際理解」高木修 (編) 『社会心理学への招待』 有斐閣
- 木村啓子 (2006) 「英語圏滞在が学生の英語力に及ぼす影響：短期語学研修により英語力は向上するか」『尚美学園大学総合政策研究紀要』 12, 1-20.
- 木村啓子 (2010) 「大学生の海外短期語学研修の効果—英語の聴解力とライティング力 (fluency) に焦点を当てて」「留学前後における英語聴解方略使用の変化」第 36 回全国英語教育学会大阪研究大会自由研究発表・事例報告 (口頭発表・ポスター発表・賛助会員プレゼンテーション) プログラム・アブストラクト (2010-07-21 版)」2010 年 8 月 8 日 (日) 関西大学千里山キャンパス第 1 学舎 (2011 年 1 月 5 日アクセス：http://keles.web.infoseek.co.jp/jasele2010osaka/abstracts_web/abstracts/pa8.html)
- 木内亜紀 (1996) 「独立・相互依存的自己理解—文化的影響, およびパーソナリティ特性との関連」『心理学研究』 67, 308-313.
- 木内亜紀 (2000) 「相互独立・相互協調的自己観における社会・文化的適応と異文化間コミュニケーション—日本人学生, 日本人留学生, およびアメリカ人学生の比較研究」『日本教育心理学会発表論文集』 42, 467.
- 野中辰也 (2005) 「海外語学研修の効果測定」『新潟青陵大学短期大学部研究報告』 35, 7-12.
- 野中辰也, 田中ゆき子, 隅田朗彦 (2002) 「短期語学研修プログラムの効果測定 (2)」『新潟青陵女子短期大学研究報告』 32, 33-38.
- 野中辰也, 田中ゆき子, 隅田朗彦 (2001) 「短期語学研修プログラムの効果測定 (1)」『新潟青陵女子短期大学研究報告』 31, 71-78.

Effects of a Short-term (3-week) Study Abroad Program on the English Proficiency of Four-year College Students

KUNO Hiroyuki

Abstract: This is a preliminary study that investigated if there was any difference made in the English proficiency of the 22 college students who spent three weeks during their spring break participating in a study abroad program hosted by La Trobe University in Melbourne, Victoria, Australia, where they learned English in EFL classes in the morning five days a week, attended special classes and lectures delivered both on and off the campus by professionals in the fields of the students' specialties and other tourism-oriented activities in the afternoon, and enjoyed interactions with their host family at home in the evening. Eight of the 22 students volunteered to join this study by taking pre-tests before the departure and post-tests after their return. Circumstances prevented us from forming a control group to contrast with the treatment group of the comparable size at the time when the latter group took the pre-tests. So, using cluster analysis, their pre-test scores were used to form their matched control group of 24 students who also took the same tests at different times, receiving a different treatment, i.e., 10-month-long e-learning. The *t*-test and ANOVA results showed that these two groups may have experienced the same degree of gains in their abilities to listen to and comprehend short utterances and dialogs in English.