

# Generation effectとTesting effectを援用した 大学英語教育における補講の効果：項目反応理論による分析

## Effect of English remedial courses based on Generation effect and Testing effect in college education: an analysis based on Item Response Theory

横山 悟

Satoru YOKOYAMA

学習心理学、認知心理学、認知神経科学などの分野において、人間の記憶・学習への促進効果が見られるとされる、Generation effectとTesting effectという現象がある。本論文では、実際の大学英語教育の場において、特に英語能力が低い学習者層に対し、この二つの現象を援用した訓練を半年間週一回の補講において導入し、その効果が見られるか否かを検証した。方法として、補講を行った学習者群に対し、Generation effect/Testing effectに基づいて、通常講義で用いているテキスト内に提示された例文の和文英訳問題を繰り返し解かせる形式とした。本研究の解析では、講義開始前に行った簡易 placement test のスコアと定期試験のスコアを項目反応理論により解析した後、補講を行った群(71名)と補講対象にならなかった群の中で placement test が低かった学習者群(71名)とでスコアの群間比較を行った。その結果、講義前の placement test から講義後の定期試験のスコアにおいて、補講受講群・非受講統制群ともに有意にスコア上での改善は見られたものの、群間でのスコアの改善度に有意な差は見られなかった。よって、講義自体の英語力向上効果はあったものの、半年間の補講による英語力の改善効果は見られなかった。本結果となった理由はいくつか考えられるが、講義では e-learning を用いた適切なフィードバックが学習効果を促進させた可能性がある一方、補講では紙面による課題であったため、個別の学生へのフィードバックが機能し切れなかった可能性がある。この結果に基づき、補講の効果を上げる方法を改めて検討していく必要がある。

### 1. 研究の背景

現在の日本では、少子高齢化が進み、将来的に労働人口が減少していくことが予測されている。一方、労働条件としては高度な機械化が進んできており、今まで人間が行ってきた単純作業は、機械にとって代わられていくことも予測されているところである。よって、将来の日本にとっては、今までの状況では単純作業に従事すると

想定されていた層の若者を教育することで、機械化できる作業は機械に任せ、機械では行うことができないレベルの仕事に従事できる若者を育てていく必要があると考えるべきである。そのためには、人間の学習メカニズムを解明することで、それらの知見に基づいた、効果的かつ効果的な学習方法を開発していく必要がある。以上の背景より、人間の学習メカニズムの研究において、学習心理学・認知心理学・認知神経科学などの研究分野での知見を実際の教育現場に援用し、実用に耐えうる効果的・効果的な学習方法を確立することを目的とした研究が進められるべきと考える。

本論文では、大学英語教育の場において、英語力が低い学生の英語力を引き上げるための効果的な学習方法を

連絡先：横山 悟 syokoyama@cis.ac.jp

千葉科学大学薬学部薬学科

Department of Pharmacy, Faculty of Pharmacy, Chiba Institute of Science

(2015年9月27日受付, 2015年12月7日受理)

構築するため、Generation effect・Testing effectと呼ばれる人間の学習において効果があるとされる現象を援用した学習法を導入し、その効果を検証することを目的とする。

## 2. 先行研究

### 2. 1. Generation effect

Generation effectとは、学習したい内容をgenerate(産出)することで、より効率的に学習が進む、という効果のことである。特に、聞く・読むなどの受動的な学習よりも、しゃべる・書くといった能動的な学習の方が、記憶への定着効率が高い、ということが分かっている(e.g., Kinjo and Snodgrass 2000; Slamecka and Graf 1978; Smith and Healy 1998)。このGeneration effectは、上述した通り、聞く・読むといった受動的(passive)な学習と対比され、しゃべる・書くといった能動的(active)な学習として、近年日本でも着目されつつあるActive learningの基礎的知見として考えられているものである(e.g., Rosner 2012)。

このGeneration effectは認知神経科学においてもその効果が支持されている。Rosner et al. (2013)では、functional magnetic resonance imagingを用いて、Generation effectの効果を脳活動として可視化した。結果として脳内で記憶や学習に関わる脳部位に広く関与が見られ、左右両半球の下前頭回や中前頭回、海馬傍回などに活動が見られたことを報告している。

### 2. 2. Testing effect

学習したい内容を繰り返しテストすると、テストされた内容が効率的に学習できる、という現象のことをTesting effectと呼ぶ(Abbott 1909; Gates 1917; Carpenter and DeLosh 2006; Rowland et al. 2014; Allen et al. 1969; Toppino and Cohen 2009; Glover 1989; Roediger and Karpicke 2006; Carpenter and Pashler 2007; Kang 2010; McDaniel and Masson 1985; Carpenter 2009)。近年では、テストを行うこと自体ではなく、テストを通じて「思いだす」という経験を繰り返すことが長期記憶への定着に効果をもたらすことが分かっている。よって、「思いだす」ということに着目したRetrieval practice(思いだし練習、想起訓練)という訓練法として確立されつつあり、注目を浴びてきている(Karpicke et al. 2011; Karpicke and Blunt 2011; Karpicke and Smith 2012)。

このTesting effectも認知神経科学的方法論によって効果が検証されている。Generation effectと同様、functional magnetic resonance imagingを用いた研究結果が報告されており、結果として、未知の言語の単語を学習した一週間後、Testing effect/Retrieval practiceに

基づいて学習した単語は、記憶に関わる脳部位の活動が一週間前と同じように活動したが、通常の学習によって学習した単語は記憶に関わる脳部位での活動に低下が見られた(Keresztes et al. 2013)。よって、学習した内容が長期に脳内に記憶として定着する、という、先行研究でのTesting effectによる行動データによる知見を支持する結果となった。

これらGeneration effect及びTesting effectを実際の英語教育の現場に応用した例は、今のところ数が少ない。横山(2015)では、大学初年次の必修英語教育の場において上記2種のeffectに基づいた学習法を導入し、半年の講義において高い英語学習効果が見られたことを報告している。よって、補講という状況においても効果が得られる可能性は高いと思われる。

### 2. 3. 本研究の目的

本論文では、Generation effectとTesting effectを用いて、実際の大学英語教育における英語力の低い学習者に対し、英語力の引き上げに効果があるか否かを検証する。特に大学入学時に英語力が低い学習者は、少なくとも中学・高校の六年間を通じて英語を学習してきたにも拘らず、英語力が低いままの状態に留まっているということになる。六年間学習を続けてきた学習者に、半期の補講での訓練によって英語力の引き上げ効果を生じさせることは、困難であると予想されるが、そのような状況で効果が得られる学習法を確立できるならば、今後の大学における英語教育に意義のあることとなると期待される。

## 3. 本研究の方法

本研究では、本論文著者が所属する大学の初年次英語教育の講義受講者で、講義開始前に行ったplacement testのスコアが20点未満であった補講受講群のうち、データの欠損のない71名を補講受講者群として解析した。統制群としては、同様のplacement testで20点以上の学生で、補講対象にはならなかったが、その中で点数の低い群、つまり、補講非受講者のうちで点数が下位の学生71名とした。

補講における学習介入方法は、Generation effect/Testing effectに基づいた方法を採用した。具体的には、定期試験の範囲内のうち、使用するテキスト内における基本語順及び受動態に関する例文を対象とし、和文英訳を繰り返し行わせる訓練形式の課題を行わせた。頻度は週に一回、主となる英語の講義とは別途、補講対象者のみ行った。修了者は随時補講を卒業できる形式とし、最小受講回数は4回、最大で10回、中央値は6回であった。よって、補講受講者は和文英訳課題を非受講者と比べ、別途行ったことになる。この訓練以外の介入は行っていない。

最後に期末試験を行い、そのスコアにおいて、補講効果があったのか否かを検証した。期末試験では、講義の各回において学習内容として含まれている例文部分の和文英訳問題、文法問題を出題した。これらは補講対象者・非受講者共に共通のテキストで、このテキストの中から、上述した補講での訓練用課題を作成した。これに加えて、定期試験にはテキスト外からの長文読解問題も出題した。また、placement test で使用した問題も全て出題し、直接的な英語力の伸びも測定した。placement test で出題した問題は、解答例の配布や解説は行っていない。

本研究では、研究のためのデータではない講義内におけるテストのデータを使用しているが、ヘルシンキ宣言に則り、倫理的な側面を配慮したデータの収集を行った。個人情報の保護のため、データ解析の前段階で、個人情報と本研究に使用したテスト等でのデータとを、被験者IDによる管理とすることで連結不可能な匿名化状態としたうえで、データの解析を行った。

### 3. 解析方法及び結果

解析では、まず placement test と定期試験が異なるテストであるため、比較可能な形にするために、項目反応理論による解析を行った。

項目反応理論とは、出題されたテスト内の問題に対する各受験者の応答データに基づき、テストが測定している項目に対する受験者の能力値、及びテスト内に含まれる問題項目の難易度につき、確率的に推定する、という理論である。この理論による解析により、異なるテスト間の難易度を客観的に評価でき、かつ異なるテスト間

での受験者の能力値を直接比較することができるようになる。

よって理論上は、項目反応理論に基づいて解析したスコアは客観性を持って他のテストと比較が可能な形にはなるが、より精度を高めるために、テスト間で等化も行った。等化とは、異なるテストを困難度に基づいて比較可能な同じ評価軸上に置くことである。本研究のデータの生値と、項目反応理論適用後の値は、表1、及び表2に示す。place は講義前の placement test、place 2 は定期試験での placement test と同一問題のスコアを指す。項目反応理論による解析後のスコアは、おおむね3から-3の値を取るとされる。

次に等化後のデータを用いて、以下の三つの仮説を検証した。

1. 定期試験での和文英訳問題のスコアの差が、placement 時の両群の差よりも縮まった
2. 定期試験での文法問題のスコアの差が、placement 時の両群の差よりも縮まった
3. 定期試験での長文読解問題のスコアの差が、placement 時の両群の差よりも縮まった

もし定期試験でのスコアの両群の差が、placement 時のスコアの差よりも縮まっていれば、両群の差は補講受講の有無であるため、補講の効果があった、ということになる。

上記の3つの仮説を検証するため、IRT による解析後の値を用いた ANOVA による解析を行った。被験者内

表1. データの平均値及び標準偏差 (生値)

	受講者	非受講者
place	6.1 (4.28)	17.5 (3.19)
place 2	33.6 (17.74)	35.7 (18.58)
英訳	66.5 (20.50)	72.9 (14.16)
文法	55.7 (22.33)	59.43 (18.50)
読解	49.0 (10.44)	51.6 (13.37)

表2. データの平均値及び標準偏差 (IRT 解析後)

	受講者	非受講者
place	-2.312 (0.270)	-1.717 (0.326)
place 2	-1.573 (0.952)	-1.121 (1.020)
英訳	-0.725 (1.589)	-0.173 (1.296)
文法	-0.942 (1.350)	-0.394 (1.354)
読解	-0.979 (1.008)	-0.516 (1.097)

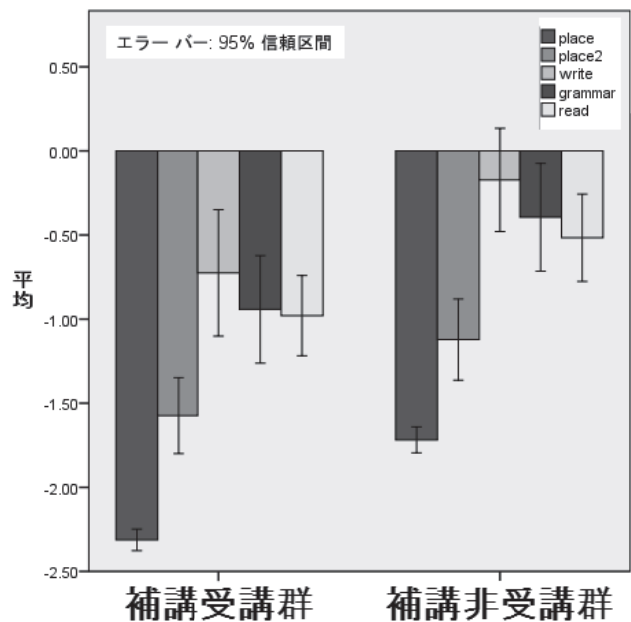


図1. IRT 解析後の結果のグラフ

要因は試験種別の placement、定期和文英訳、定期文法、定期長文読解の4水準、被験者間要因は補講受講と非受講の2水準とした。placement test時と比べ、定期試験時のスコアで差が縮まったならば、交互作用が有意となると考えられる。

結果として、上記ANOVAの結果は、群間では有意差が見られたものの、どれも交互作用では有意にはならなかった。統計の結果は表3に示す。

表3. 2 × 2 ANOVAの結果

	F 値	p 値	効果量*
試験種別間の主効果	235.657	0.001	0.627
群間の主効果	249.067	0.001	0.640
交互作用	0.160	0.69	0.001

\*効果量は偏  $\eta^2$

post-hoc testとして、補講受講群・補講非受講群を含めた試験種別間の多重比較の結果は以下の表4の通りであった。

表4. 試験種別間の多重比較

	p 値
place – place 2	0.001*
place –和文英訳	0.001*
place –文法	0.001*
place –長文	0.001*
place 2 –和文英訳	0.001*
place 2 –文法	0.001*
place 2 –長文	0.001*
和文英訳 –文法	0.384
和文英訳 –長文	0.079
文法 –長文	1.000

\* (Bonferroni,  $p < 0.05$  にて有意)

#### 4. 考察

本研究では、認知心理学・認知神経科学等の分野において学習促進効果が見られると報告されている Generation effect と Testing effect を用いて、実際の大学英語教育における英語力の低い学習者に対し半年間週一コマの補講を行うことにより、英語力の引き上げに効果が見られるか否かを検証した。

結果として、2×2でのANOVAにおいて、テスト間及び群間での主効果が有意となった(表3)。一方交互作用は有意にならなかったため、本結果からは、群間での英語力の変化には差がなかったことになる。よって、補講

による直接的な英語力の押し上げ効果はあるとは言えない、という結論となる。

一方、データから見えてきたことも多くある。講義前のスコアと比べ、講義後のスコアは間違いなく上昇しており、講義自体の効果があつたことは確認ができる。この成果は、多重比較の結果からも明らかで、講義前の placement test のスコアと比べ、定期試験時の placement test スコア及びその他の定期試験スコアが統計的に有意に上昇していることから確認できる。逆に定期試験のスコアにおいては、特に和文英訳・文法・長文のスコア間では差がない。つまり、有意な差があつたのが講義前の placement test と講義後定期試験との間であつたことから、講義自体の効果は強く見られたと言える。

第二に、講義開始前の非受講統制群の placement test のスコア -1.717 と、講義後の定期試験内における補講受講群の placement test 同一問題のスコア -1.573 との比較すると、若干後者の方が上回っている。講義開始前に行った placement test は、問題用紙も全て回収し、かつ問題の解説も行っておらず、その上で項目反応理論に基づく解析、及び等化を行ったことで、英語力の直接比較が可能と考えられる。よって、半年の講義後の補講受講者が、講義前時点の補講非受講者の英語力までは追いついた、ということになる。つまり、少なくとも中学・高校の6年間で付いてしまった英語力の差を半年で埋めることができた、という事実は、ポジティブに解釈している部分であると考えられる。

本研究では、講義と別の形で行った補講の形式では、効果的な英語力の向上には繋がらないことが示唆された。一方、講義自体の効果は強く見られることも確認された。これらの結果より、講義とは全く別の形で補講を行うよりも、講義と繋がった、一貫性のある教授・学習の仕方でない場合、効果が上がらない可能性があることが明らかになった。また、講義では一部の教員により e-learning の導入を行っている。特に Generation effect、Testing effect を最大限に活用するには、課題遂行直後の適切なフィードバックが重要である一方(横山2015)、今回の補講では紙面での課題となつてしまい、フィードバックが適切に機能しなかった側面があつたと思われる。Testing effect ではフィードバックの有無により効果が変わることが報告されている(横山2015)。よって、週に講義を1コマ、補講を1コマという形式ではなく、一貫性を持たせた講義を週に2コマ、可能ならば e-learning を2コマ通して導入というカリキュラムの構築といった方向で考えなければならぬと思われる。

#### 謝辞

本研究の一部は科学技術研究費の補助を受けて行われた(挑戦的萌芽研究：課題番号26580104)。



## 参考文献

- 1) Abbott, E. E. (1909). On the analysis of the factors of recall in the learning process. *Psychological Review: Monographs Supplements*, 11, 159–177.
- 2) Allen, G. A., Mahler, W. A., & Estes, W. K. (1969). Effects of recall tests on long-term retention of paired associates. *Journal of Verbal Learning and Verbal Behavior*, 8, 463–470.
- 3) Carpenter, S. K. (2009). Cue strength as a moderator of the testing effect: The benefits of elaborative retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 35, 1563–1569.
- 4) Carpenter, S. K., & DeLosh, E. L. (2005). Application of the testing and spacing effects to name learning. *Applied Cognitive Psychology*, 19, 619–636.
- 5) Carpenter, S. K., & Pashler, H. (2007). Testing beyond words: Using tests to enhance visuospatial map learning. *Psychonomic Bulletin & Review*, 14, 474–478.
- 6) Gates, A. I. (1917). Recitation as a factor in memorizing. *Archives of Psychology*, 6, 1–104.
- 7) Glover, J. A. (1989). The “testing” phenomenon: Not gone but nearly forgotten. *Journal of Educational Psychology*, 81, 392–399.
- 8) Kang, S. H. (2010). Enhancing visuospatial learning: The benefit of retrieval practice. *Memory & Cognition*, 38, 1009–1017.
- 9) Karpicke, J. D., & Blunt, J. R. (2011). Retrieval practice produces more learning than elaborative studying with concept mapping. *Science*, 331, 772–775.
- 10) Karpicke, J. D., Smith, M. A. (2012). Separate mnemonic effects of retrieval practice and elaborative encoding. *Journal of Memory and Language*, 67, 17–29.
- 11) Keresztes, A., Kaiser, D., Kovács, G., & Racsmány, M. (2014). Testing Promotes Long-Term Learning via Stabilizing Activation Patterns in a Large Network of Brain Areas. *Cerebral Cortex*, 24, 3025–3035.
- 12) Kinjo, H., & Snodgrass, J. G. (2000). Does the generation effect occur for pictures. *American Journal of Psychology*, 113, 95–121.
- 13) McDaniel, M. A., & Masson, M. E. J. (1985). Altering memory representations through retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 11, 371–385.
- 14) 水本篤・竹内理 (2008), 研究論文における効果量の報告のために—基礎的概念と注意点— *英語教育研究*, 31, 57–66.
- 15) Roediger III, H. L., & Karpicke, J. D. (2006a). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science*, 17, 249–255.
- 16) Rosner, Z. A. (2012). *The Generation Effect and Memory*. UC Berkeley, Ph.D thesis.
- 17) Rosner, Z. A., Elman, J. A., & Shimamura, A. P. (2013). The Generation Effect: Activating Broad Neural Circuits During Memory Encoding. *Cortex*, 49, 1901–1909.
- 18) Rowland, C. A. (2014). The Effect of Testing Versus Restudy on Retention: A Meta-Analytic Review of the Testing Effect. *Psychological Bulletin*, 140, 1432–1463.
- 19) Slamecka, N. J., & Graf, P. (1978). The generation effect: Delineation of a phenomenon. *Journal of Experimental Psychology: Human Learning and Memory*, 4, 592–604.
- 20) Smith, R. W., & Healy, A. F. (1998). The time-course of the generation effect. *Memory & Cognition*, 26, 135–142.
- 21) Toppino, T. C., & Cohen, M. S. (2009). The testing effect and the retention interval: Questions and answers. *Experimental Psychology*, 56, 252–257.
- 22) 横山悟. (2015). 『実証的方法論により検証された最新の科学的知見「産出効果とテスト効果」に基づく効率的学習法』, 総説出版