

機関番号：27401

研究種目：基盤研究（C）

研究期間：2008～2010

課題番号：20520510

研究課題名（和文） マトリックスモデルによる語彙習得分析研究：シミュレーションと学習者データの検証

研究課題名（英文） Study on vocabulary acquisition analysis based on Matrix model: Investigation into simulation data and learner data

研究代表者

吉井 誠 (YOSHII MAKOTO)

熊本県立大学・文学部・教授

研究者番号：70240231

研究成果の概要（和文）：

この研究は第二言語語彙習得の予測は可能か検証したものである。マトリックスモデルという一つの理論を通して実証研究を行っている。モデルによる予測値と実際の学習者の実測値を比較しモデルの妥当性を検証している。またモデルに基づきシミュレーションを行い、語彙知識変化のパターンや変化の過程を観察している。シミュレーションによる仮想データと学習者のデータを比較し、シミュレーション研究の可能性を追求している。

研究成果の概要（英文）：

This study examines whether Matrix model enables us to make a reasonable prediction about the second language vocabulary acquisition. This model uses the two sets of data and calculates the transitional probability of vocabulary knowledge changes. Based on the probability, we can make prediction about the future vocabulary acquisition. The study compared the prediction and the learners' actual vocabulary scores in order to determine the validity of the model. The study also used simulation to compare the change processes in terms of frequency of changes between simulated data and learners' actual data. The results showed the model can make a reasonably good prediction and simulations showed that we can also make a reasonably good prediction about the frequency of changes.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
2008年度	1,100,000	330,000	1,430,000
2009年度	500,000	150,000	650,000
2010年度	900,000	270,000	1,170,000
総計	2,500,000	750,000	3,250,000

研究分野：人文学

科研費の分科・細目：言語学・外国語教育

キーワード：第二言語習得、語彙習得、シミュレーション

1. 研究開始当初の背景

(1) 第二言語習得ならびに語彙習得の分野で習得のモデル化の必要性が指摘されていた。このような必要性から Meara (1990) はマトリックスモデルを提唱し、モデルに基づく語彙習得の予測方法を紹介した。モデルを利用した研究の数はまだ少なく、日本人学習者を対象としてモデルの妥当性の検証が必要であった。また、これまでの語彙習得研究では短期的なデータ収集に基づくものが多く、長期的な視点の研究が必要であった。

(2) CALL などコンピュータを利用した語学教育の研究は進んできたが、コンピュータ、特に Simulation を利用した言語習得のプロセスを探る研究は少なく、これからの進展が期待されていた。

2. 研究の目的

(1) 日本人学習者を対象としてマトリックスモデルに基づく予測の妥当性について検証する。また、モデルの前提となっている推移確率の普遍性について検証している。

(2) マトリックスモデルに基づいたコンピュータ Simulation を行い、語彙習得の長期的プロセス、変化のパターンなどについて検証する。Simulation のデータと実際の被験者のデータを比較し、Simulation を用いた語彙習得プロセス研究の可能性を探る。

3. 研究の方法

(1) 様々な学習環境の中の日本人学習者の単語知識の変化を自己診断テストを通して長期的に観察する。被験者の語彙習得状況と、マトリックスモデルの予測による語彙習得状況を比較し、予測の妥当性について検証する。各テスト間の推移確率を比較し、推移確率の普遍性について検証する。

(2) 被験者の初期の自己診断テストの点数とテスト間の推移確率より Simulation を行い仮想のデータを産出し、語彙知識変化の長期的なプロセスの観察を試みる。実際の被験者のデータと Simulation のデータを比較し Simulation 研究の妥当性について検証する。

4. 研究成果

最初にマトリックスモデルによる予測の妥当性に関する調査の結果について、次に推移確率の普遍性の検証について、そして最後に、Simulation 研究の成果について報告する。

(1) 様々な学習環境の中の日本人学習者を対象に実証研究を行いマトリックスモデルの検証を行ってきた (吉井、2009a; 2010; 2011; Yoshii, 2011)。その中から具体的な実証例を2件報告する。一つはケーススタディ (吉井、2009a) であり、もう一つは学習者グループに対して行った研究 (吉井、2011)

である。

① 吉井 (2009a) では一人の学習者を対象にケーススタディを行った。英国在住の日本人を対象に 300 語の単語について「知っている (State 2 = S2)」、「知らない (State 0 = S0)」、「どちらとも言えない (State 1 = S1)」の 3 つの語彙知識に基づく自己診断テストを 1 週間に 1 回ずつ 10 週間実施した。被験者は一年間の英国滞在中であり、特に語学学校、語学学習などは行っておらず、英語圏の自然な生活の中でどのように単語の知識が推移するかに着目した。

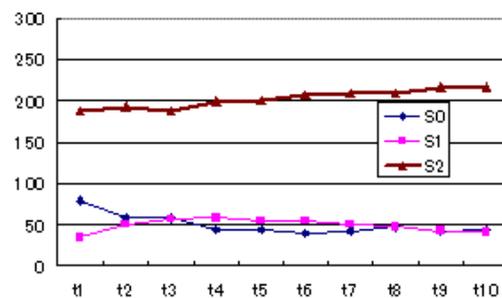


図1 被験者の実測値

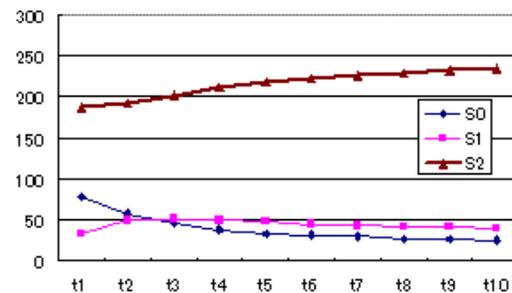


図2 マトリックスモデルによる予測値

図1と図2が示しているように、それぞれの単語知識 (S0, S1, S2) の実測値と予測値は、その変化がほぼ同じであることがわかる。予測値と実測値の相関関係を調べたところ、ピアソンの係数で $r = .97$ ときわめて強い相関関係がみられ、かなり正確に予測が可能であることを示していた。

② 吉井 (2011) では日本人大学生 22 名を被験者に吉井 (2010) と同様 10 週間にわたり 200 語の単語を意図的に学習したが、この研究では、集中学習後の語彙知識の変化、すなわち、語彙知識の長期的な保持について学習後約 2 カ月をかけて観察した。

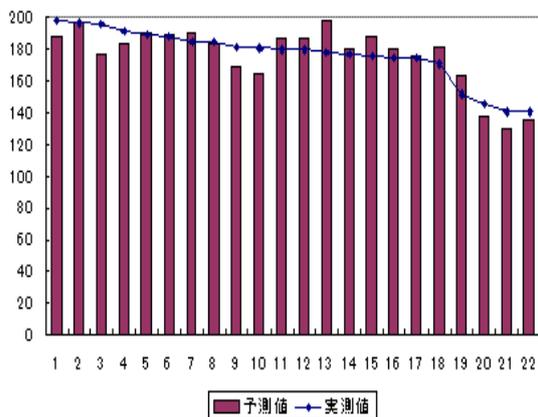


図3 学習者の実測値と予測

図3のグラフは、最終テストにおける22名の学習者の語彙テストの結果(Yesと認識された単語の数、実測値)が線で表わされ、マトリックスモデルによるそれぞれの学習者の予測値が棒グラフで表わされている。分かりやすいように被験者の語彙テスト得点が高い者から低い者へという順番に示している。

図3が示しているように、実測値よりも高い予測値(棒が線を越えている場合)もあれば、逆に、実測値よりも低い予測値(棒が線よりも下の位置に存在)もあるが、全体的に見ると二つの値は近似している。予測値と実測値の相関関係を調べたところ、ピアソンの係数で $r = .85$ と強い相関関係がみられ、かなり正確に予測が可能であることを示した。

③このほかにも吉井(2010)では意図的学習という環境の下でモデルの予測の妥当性について検証した。被験者は日本人大学生26名であり、通常の授業の中でカリキュラムの一環として毎週20単語を10週間にわたり学習し、目標単語200語の語彙知識の変化を「知っている」「知らない」の2つの語彙知識状態を仮定した自己評価テストを計5回実施した。

語彙知識の保持に関しては、Yoshii(2011)のStudy1において46名の日本人大学生からデータを収集し、目標単語200語の語彙知識の変化を2つの語彙知識の状態を仮定した自己評価テストを使用して観察している。

どの研究においても観察最終時の被験者の語彙テストの結果と、マトリックスモデルで予想される語彙テストの予測とはかなり高い相関関係が示され、モデルの妥当性と信頼性を確認する結果となった。

(2)次に、推移確率は一定しているというマトリックスモデルの前提の検証報告を行う。ここでは吉井(2011)を例として挙げ報告する。推移確率とはあるテストと次のテスト

の間に語彙知識が変化する確率のことであり、その変化は安定しており一定のものであるということが前提となっている。吉井(2011)では被験者22名一人一人の推移確率を算出し、各テスト間の確率が安定しているかどうか検証している。表1では例として一人の被験者の推移確率が記載されている。

表1 被験者Aの各テスト間の推移確率

推移パターン	T1-T2	T2-T3	T3-T4	T4-T5
s0s0	.73	.83	.81	.69
s0s1	.27	.17	.19	.31
s1s0	.14	.13	.10	.06
s1s1	.86	.87	.90	.94

この実験では2つの語彙レベルを想定し、それに基づき表にあるように4つの推移パターンが存在した。S0(知らない)という状態がS0にとどまる場合、S0がS1(知っている)に移行する場合、逆にS1であったものがS0に戻ってしまう場合、そしてS1がS1にとどまる場合である。そして、これをTest1とTest2の間(T1-T2)、Test2とTest3の間(T2-T3)、Test3とTest4の間(T3-T4)、Test4とTest5の間(T4-T5)とそれぞれのテスト間で推移確率を算出した。表1のような形式で被験者22名全員の各テスト間の推移確率を算出し、それぞれの推移パターンが安定しているのかどうかを、各テスト間の相関関係を用いて分析した結果を表2が示している。

表2 全員の各テスト間の推移確率の相関関係

	T1-T2 (1wk)	T2-T3 (4-5wks)	T3-T4 (1wk)	T4-T5 (1wk)
T1-T2(1wk)	1.00			
T2-T3(4-5wks)	.84	1.00		
T3-T4(1wk)	.93	.88	1.00	
T4-T5(1wk)	.86	.80	.94	1.00

表2が示しているように、各テスト間の推移確率は強い相関関係を示しており、推移確率が安定していたことを示している。テスト間の間隔が他のテスト間と比べかなり長かったT2-T3においても、他のテスト間の推移確率と高い相関関係を維持しており、テスト間の間隔の長さに関わらず安定していたことが判明した。これにより、マトリックスモデルの前提である推移確率は不変ではないにしても、かなり安定したものであることが分かった。

(3) 最後に Simulation 研究の成果を報告する。Simulation を行うために、Delphi 言語を用いて図 4 のような Simulation プログラムを作成した。

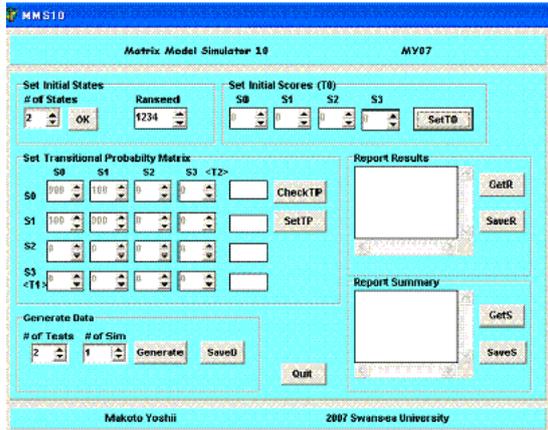


図 4 Simulation プログラ

このプログラムでは、語彙知識レベルを 1 つから 4 つまで設定でき、まずは語彙知識の初期状態の数値と推移確率を挿入する。そしてテスト回数と Simulation の回数を指定する。プログラムはこの初期設定に基づき仮想データを算出、研究ではこの仮想データと被験者の実際のデータを比較することで検証を行っている (吉井, 2009a; 2009b; 2010)。

①吉井(2009b)では推移確率の微妙な変化が語彙知識の変化にどのような影響を与えるかを検証している。Meara(1990)の追実験を行った形であるが、語彙知識の初期状態は同じで (知っている単語数 30 個、知らない単語数 70 個)、推移確率だけがわずかながら異なる 3 つのパターンを例に挙げてそれぞれのデータを産出した。実験の結果、推移確率が少しでも変化すると長期的な変化に大きな違いが出てくることが判明した。

②一人の被験者の長期的な語彙知識の変化に着目した吉井(2009a)では、被験者の語彙知識初期状態、推移確率をもとに Simulation を行い仮想データを産出した。Simulation では特に長期的な語彙知識の安定性について、テスト間で語彙知識が変化している単語、または変化せず安定している単語を被験者の実際のデータと比較している。

図 5 では、被験者の実際の測定値の安定した単語を実測値(安定)、変化した単語を実測値(変化)と示している。同様に、シミュレーションで安定した単語を Simulation(安定)、変化した単語を Simulation (変化)としている。Time 1 と Time 2 の間の比較がグラフ横軸の 1 番として示している。また、Time 2

と Time 3 の間の比較が 2 番で、残りも各テスト間の違いを比較したものである。最後の 9 番は Time 9 と Time 10 のテスト間の差を示している。

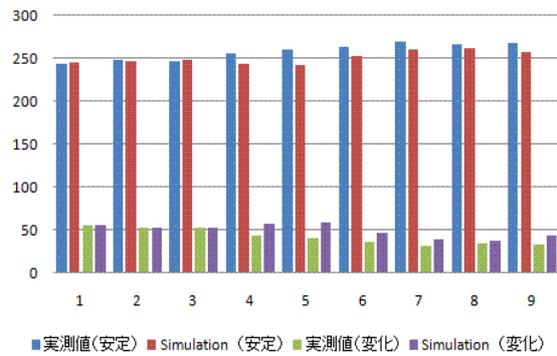


図 5 実測値とシミュレーションの比較：安定した単語数と変化した単語数

このグラフから、被験者のデータとシミュレーションのデータにはほとんど差がないことが分かる。安定している単語と変化している単語の割合を平均すると、被験者のデータとシミュレーションのデータともほぼ同じであった (安定=83%、変化=17%)。更に、安定性に関しては、被験者のデータとシミュレーションのデータとに非常に高い相関関係 ($r = .998, p = .000$) がみられ、シミュレーションによる安定性予測の可能性を支持するものとなった。

③吉井(2010)では日本人大学生 26 名を対象に意図的語彙学習の環境における語彙知識の変化について観察したが、シミュレーションのデータは実際の被験者のデータと比べ、どのくらい正確に語彙知識変化パターンを反映させられるかについて考察している。この実験においては語彙知識を 2 つのレベルとして行っており、語彙知識の変化について以下の 4 つのパターンが想定される。知らない単語 (State 0=S0) がそのまま知らない状態を保つ場合 (S0→S0)、知っている単語 (State 1=S1) がそのまま知っている状態を保つ場合 (S1→S1)、知らない単語が知っている状態に変化する場合 (S0→S1)、そして知っている単語が知らない状態に戻る場合 (S1→S0) の 4 つである。表 3 では被験者全体の各テスト間における変化パターンの割合を実際の被験者のデータとシミュレーションのデータとで比較している。

表 3 から分かるように、全体的な変化パターンの割合は、実際のデータとシミュレーションのデータとはきわめて近いものとなっている。S0 から S1 に変化する割合は 17-18%を示しているが、テストが 2 週間ごとに行われ、その間学習すべき単語が 40 語(200

表3 変化パターンの比較

変化のパターン	実際のデータ	シミュレーション
S0→S0 (0)	26% (SD=16) max=56, min=6	28% (SD=17) max=59, min=5
S1→S1 (1)	47% (SD=17) max=77, min=17	44% (SD=18) max=80, min=13
S0→S1 (+1)	18% (SD=3) max=24, min=11	17% (SD=4) max=23, min=9
S1→S0 (-1)	9% (SD=3) max=17, min=3	11% (SD=3) max=19, min=4

語の20%)であったことを考えると、妥当な割合といえよう。忘却のパターン、すなわち、S1からS0に戻る割合は約1割であった。学習者間で幅が広がったのが、残りの二つのパターンであった。すなわちS0にとどまるパターンがその1つで、割合は26-28%であった。標準偏差値はSD=16であり、これは平均値の半分以上の値を占め、広範囲に及ぶ。S1にとどまるパターンにおいても平均で5割近くを占め、記憶の定着の良さを物語る一方、標準偏差SD=17という高い数値に象徴されているように広範囲に及ぶことが分かる。このことから、被験者全体の変化のパターンを考えたとき、学習者の語彙変化を左右したのは、S0からS1に増えたり、S1からS0へ減ったりするパターンよりも、S1をS1に保持する、あるいはS0がS0にとどまるパターンであったと考えられる。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計5件)

- ① Makoto Yoshii, A new approach to vocabulary assessment: A matrix model for predicting future vocabulary performance, Malaysian Journal of ELT Research, 査読有、Vol. 7 (1)、2011、94-117
- ② 吉井 誠、マトリックスモデルの検証：長期的な語彙知識変化の安定性に関する一考察、熊本県立大学文学部紀要、査読有、第17巻第70号、2011、37-45
- ③ 吉井 誠、長期的語彙知識変化パターンの検証：被験者データとシミュレーションデータの比較、熊本県立大学文学部紀要、査読有、第16巻第69号、2010、33-46
- ④ 吉井 誠、Investigating into word knowledge change: Comparison of

simulation data and actual data、Language Issues: The international journal of the academic information and media center、査読有、2009a、1-22

- ⑤ 吉井 誠、マトリックスモデルを用いたシミュレーションスタディー：長期的な語彙知識変化の検証、熊本県立大学文学部紀要、査読有、第15巻68号、2009b、65-79

[学会発表] (計4件)

- ① Makoto Yoshii, Long-term vocabulary change: Can we predict? 英国スオンジー大学応用言語学語彙習得学会、2011年3月16日、英国ウェールズ州スオンジー大学
- ② 吉井 誠、長期的な語彙知識変化の安定性に関する一考察：マトリックスモデルの検証、第36回全国英語教育学会大阪研究大会、2010年8月7日、関西大学
- ③ Makoto Yoshii, An investigation into long-term vocabulary knowledge change: Comparison of actual data and simulation data, EuroCALL 2009, September 10, 2009, Gandia, Valencia, Spain
- ④ Makoto Yoshii, An attempt to measure long-term vocabulary knowledge change: Comparison of actual data and simulation data, British Association of Applied Linguistics 2008, September 11-13, 2008, Swansea, Wales, UK

6. 研究組織

(1) 研究代表者

吉井 誠 (YOSHII MAKOTO)
熊本県立大学・文学部・教授
研究者番号：70240231

(2) 研究分担者

なし