

# 英語コースウェアにおける繰り返し学習時間の MTS を用いた分析

## An Analysis of Iterative Learning Time in English Courseware by MTS

渡邊 博之<sup>†</sup>  
Hiroyuki Watanabe

川嶋 正士<sup>†</sup>  
Masashi Kawashima

### 1. まえがき

学習支援システム (LMS) の運用によって得られるデータには得点と学習時間がある。一般に、演習問題型コースウェアでは制限時間を設けるため、得点が分析の対象となる。これに対して、繰り返し学習型コースウェアでは得点 (合格点) を設けるため、学習時間が分析の対象となる。学習時間の分析には、得点の SP 表を学習時間に拡張する方法[1]や、学習時間に対する累積回答率の推移を数理モデルとして推定する方法[2]などがある。これらの分析法は、問題項目や学習者の回答状況の把握には有用である。しかし、任意の問題項目で再構成されたコースウェアを再び同じ学習者に提示する場合には、問題項目が既知となるために同一結果が得られず、コースウェアの効率よい作成や回答状況の推定へ利用するには限界がある。このような分析には、品質管理工学で用いられている MTS (Mahalanobis-Taguchi System) の利用が提案できる[3]。MTS はマハラノビス距離を求めることによって単位空間とデータとの差異を分析し、SN 比を求めることによってマハラノビス距離を長引かせる要因を分析する方法である。教育工学の分野への適用も進められており[4]、繰り返し学習における選択式と記述式の合格点に到達したときの最終回の学習時間 (合格時間) を単位空間として用いることによって、任意の問題項目で構成されるコースウェアの学習時間を推定している。また、今後の活用例として、問題形式の差異や年度間の差異の分析を提案している。

本論文では、繰り返し学習型コースウェアにおいて、英単語の混合式を対象に、混合式の合格時間を単位空間として学習時間を分析し、選択式の合格時間で混合式と選択式との差異を求めている。また、英文法では記述式を対象に、音声ありの合格時間を単位空間として学習時間を分析し、音声なしの合格時間で音声の有無による差異を求めている。分析の結果、英単語ではマハラノビス距離から、選択式に比べて混合式の学習時間は長い、問題形式によらず、一部の英単語と熟語の SN 比の利得が大きいことを明らかにした。また、英文法では音声の有無によらず、SN 比の利得から Be 動詞より一般動詞、現在形より過去形、YES/NO 疑問文より WH 疑問文の学習時間が長いことを明らかにした。さらに、学習時間では音声を何度も聞いているが、合格時間では音声を聞かずに回答していることを明らかにした。

### 2. MTS による分析法

#### 2.1 マハラノビス距離

学習者が問題項目に回答し、合格点に到達するまで繰り返し学習に要した全ての時間を学習時間とする。また、学習時間の中で、合格点に到達した最終回の学習時間を

<sup>†</sup>日本大学工学部

合格時間 (学習時間  $\geq$  合格時間) とする。合格時間は初回に比べてばらつきが小さく、定常状態で均質なデータのため、単位空間として用いることができる。MTS では、先ずマハラノビス距離を求める。このため、問題項目  $j(=1, 2, \dots, m)$  における合格時間の平均値  $\bar{t}_j$  と標準偏差  $\sigma_j$  を用いて、学習時間  $t_{ij}$  を次式で正規化する。

$$V_{ij} = \frac{t_{ij} - \bar{t}_j}{\sigma_j} \quad (1)$$

次に、 $r_{jk}$  を単位空間における問題項目の  $j$  列と  $k$  列のピアソンの積率相関係数として、相関行列の逆行列  $a_{jk}$  を求める。 $a_{jk}$  を用いると、学習者  $i(=1, 2, \dots, n)$  の学習時間のマハラノビス距離  $D_i^2$  は次式で求められる。

$$D_i^2 = \frac{1}{m} \sum_{j=1}^m \sum_{k=1}^m V_{ij} a_{jk} V_{ki} \quad (2)$$

ただし、 $m$  は全問題項目数、 $n$  は全学習者数である。単位空間の  $V_j$  の平均値は 0 であり、 $D_i^2$  の平均値は 1.0 である。 $D_i^2$  の値は、大きいほど単位空間から離れることを示している。

#### 2.2 直交表と SN 比

任意の問題項目で構成されたコースウェアを全学習者が回答したとする推定値を求めるため、2 水準系直交表の各行に学習時間を割り付ける。ただし、直交表の各行に対応する第 1 水準を「その問題項目をコースウェアの構成に加える」、第 2 水準を「その問題項目をコースウェアの構成から外す」とする。次に、直交表に割り付けた学習時間が得られたものとして、学習者  $i$  の学習時間のマハラノビス距離  $D_i^2$  を式(2)で求める。 $D_i^2$  を用いると、問題項目  $j$  の望小特性の SN 比  $\eta_j$  は次式で求められる。 $\eta_j$  の値は、0db に近づくほど単位空間に近づくことを示している。

$$\eta_j = -10 \log \left( \frac{1}{n} \sum_{i=1}^n D_i^2 \right) \quad (3)$$

### 3. 英単語の分析

#### 3.1 混合式コースウェア

英単語の学習において、約 25 問を 1 節とし、全 3 節 (1 は高校英語 II 単語、2 はセンター試験単語、3 は熟語) で、全 5 章 (A~E) から成る全問題項目数  $m$  が 15 (=3 節  $\times$  5 章) のコースウェアを作成した。A2 は annual, immediately などタイプミスしやすいセンター試験単語である。また、E2 は cash, fog など短時間の入力で、かつタイプミスしにくい英単語である。選択式では一つの英単語を一つの画面で構成している[4]。英単語の意味の正答を 4 つの選択肢から選ぶ形式であり、A1~E3 の全てが選択式コースウェアである。これに対して、混合式では 1 画面目は選択式、2 画面目はその英単語をタイプする記述式で構成される A1~E3 の全てが混合式コースウェアである。したがって、混合式は選択式の 2 倍の画面数となる。ただ

し、D2 と E1 における混合式の英単語数は 35 問あり、選択式に比べて 10 問多い。混合式と選択式の各学習集団は、あらかじめコースウェアとは内容の異なるマーク式プレイメントテストを行い、得点によって振り分けた。その結果、混合式コースウェアは高得点集団に提示し、選択式コースウェアは低得点集団に提示した。60 分の授業の後、CALL 教室で LMS を用いて[5]、一斉に学習者自身のペースで、一つの問題項目(節ごと)の正答率(=正答画面数/全画面数)が 90%以上を合格点として、合格点に到達するまで繰り返し学習を行った。全学習者数  $n$  は 49 名 ( $n \geq 3m$ ) である。

図 1 は問題項目 C3 (C 章の熟語) における混合式中の記述式問題 (2 画面目) を示している。フレーム右下の「判定」ボタンを押すと、正誤判定の結果と共に正答が表示される。正誤判定で「次へ」ボタンを押すと、次の英単語が選択式で提示される。正誤判定や画面移動の時間は学習時間に含まれない。学習時間は、問題項目ごとの合格点に到達するまでの実質的な回答時間となる。

### 3.2 学習時間と合格時間 (混合式)

図 2 は混合式の問題項目に対する学習時間 (実線) 及び合格時間 (破線) の平均値と標準偏差を示している。全体的に全ての章・節において、合格時間の標準偏差は、学習時間の標準偏差より小さい。この傾向は選択式と同様であり、合格時間を単位空間として用いることができる。また、例えば C3 の合格時間の平均値と標準偏差は 499sec と 117sec であるが、選択式では 110sec と 43sec である[4]。混合式は選択式の 2 倍の画面数であるため、合格時間の平均値も 2 倍以上長い。A2 は学習時間と合格時間との比が大きく、繰り返し回数が多い。E1 は学習時間と合格時間の差異は小さいが、問題項目数が多いため、合格時間は 770sec と長い。これに対して、E2 は差異が小さく、合格時間も 268sec と短い。A1, A2, A~E の 3 (熟語) の多くは、B~E の 1・2 と比較して学習時間がやや長い。なお、この傾向は選択式と同様である[4]。



図 1 問題項目 C3 (混合式中の記述式問題)

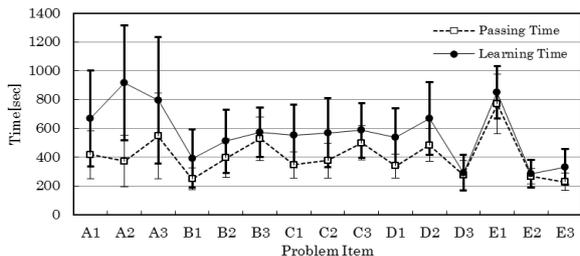


図 2 学習時間と合格時間 (混合式)

### 3.3 学習時間のマハラノビス距離 (混合式)

図 3 は、図 2 の合格時間を単位空間として、式(2)で求めた学習時間のマハラノビス距離  $D^2$  を示している。ウォード法によるクラスター分析では、18 以上が  $D^2$  の値が大きい集団である。図 3 でもピークが 12 と 22 に見えるように見えるが、全体的には 12 以上から徐々に減少しており、選択式と記述式との混合した分布となっている。図 3 の  $D^2$  の平均値は 12.1 であるが、選択式では 14.6 である[4]。混合式は選択式に比べて学習時間と合格時間との比が小さいため、繰り返し回数が少ない集団であることがわかる。

### 3.4 学習時間の SN 比 (混合式)

図 4 は、学習時間を直交表  $L_{16}(2^{15})$  に割り付け、式(2)を用いてマハラノビス距離を求め、式(3)を用いて SN 比を求めた結果 (要因効果図) を示している。●はその問題項目をコースウェアの構成に加えた値 (直交表の第 1 水準), ○は外した値 (直交表の第 2 水準) を示している。すなわち、左下がり的问题項目が図 3 のマハラノビス距離を長引かせる要因となる。特に、A2 は SN 比の幅 (利得) が大きいため、その英単語がマハラノビスを長引かせている。その他に A1, A~E の 3 (熟語) の多くもまた左下がりであるため、マハラノビス距離を長引かせている。なお、この傾向は選択式と同様である[4]。図 2 は問題項目間の相互相関を考慮してないが、図 4 は相互相関を考慮した結果を示している。しかし、概して図 2 で学習時間と合格時間との比の大きい問題項目は、図 4 で SN 比の利得が大きい。図 4 の SN 比の平均値  $\eta_0$  は -9.1db であるが、選択式では  $\eta_0 = -10.1db$  である。前者は後者に比べて SN 比がやや大きいため、繰り返し回数の少ない集団であることがわかる。この結果は 3.3 で述べた結果とも一致している。また、図 4 の SN 比  $\eta_{all}$  ( $=\eta_0 + \text{利得} = -9.1 - 1.7$ ) は -10.8db であり、この値は  $10 \log 12.1$  の値と一致している。図 4 において、本コースウェアからマハラノビス距離の長い問題項目 (A1, A2, A~E の 3) を外したコースウェアの SN 比は -8.9db (利得は +0.2db) となる。

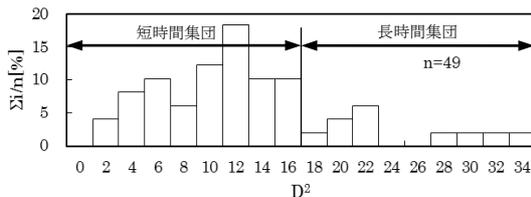


図 3 学習時間のマハラノビス距離 (混合式)

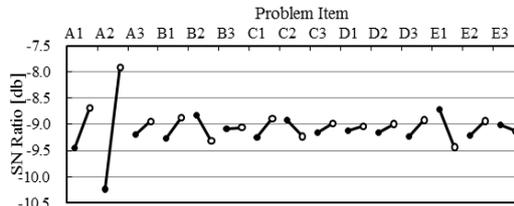


図 4 学習時間の SN 比 (混合式)

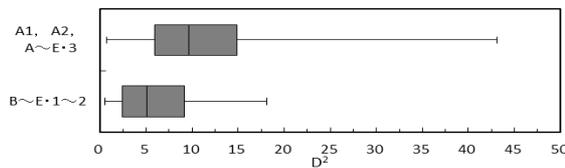
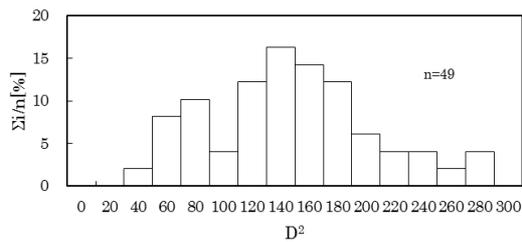
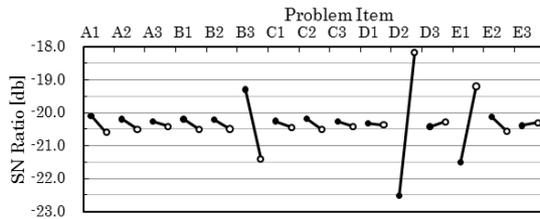


図 5 学習時間のマハラノビス距離 (混合式)



(a) マハラノビス距離



(b) SN比

図6 選択式と混合式との合格時間の差異

図5は問題項目が(A1, A2, A~Eの3)または(B~Eの1・2)で構成されるコースウェアを全学習者が回答したとするマハラノビス距離  $D^2$  の箱ひげ図を示している。 $D^2$ の平均値は前者が11.9, 後者が6.3であり, 前者が学習時間を長引かせる要因であり,  $D^2$ の最大値や標準偏差も大きい。

### 3.5 問題形式による差異

選択式と混合式とを比較するため, 選択式の合格時間を単位空間として, 混合式の合格時間のマハラノビス距離及びSN比を求める。

図6(a)は選択式の合格時間のマハラノビス距離  $D^2$  の平均値を1.0としたときの, 混合式の  $D^2$  を示している。 $D^2$ の平均値が137.9と大きいのは, 3.1で述べたように, 混合式の画面数は選択式の2倍のためである。

図6(b)は混合式の合格時間のSN比を示している。SN比の平均値  $\eta_0$  は-20.3dbであり, 0dbより極めて小さいため, 問題形式による合格時間の差異が大きい。しかし, 左上がりの問題項目が多いため, 選択式の学習者集団に比べて混合式は, SN比を小さくする要因(問題項目)が少ない。このことは3.1で述べたように, 高得点集団と低得点集団との差異に起因している。混合式と選択式の多くの問題項目数はほぼ同じであるため, SN比の幅(利得)は小さい。しかし, D2とE1は3.1で述べたように, 選択式に比べて混合式の英語数が多いため, SN比の利得が大きい。

## 4. 英文法の分析

### 4.1 記述式コースウェア

英文法の学習において, 10画面を1節とし, 全2節(1~2)で, 全6章(AはBe動詞現在形WH疑問文, Bは一般動詞現在形WH疑問文, CはBe動詞過去形YES/NO疑問文, Dは一般動詞過去形YES/NO疑問文, EはBe動詞過去形WH疑問文, Fは一般動詞過去形WH疑問文)から成る  $m=11$  のコースウェアを記述式で作成した。ただし, 音声なし(前年度)はテキストだけによる問題画面の提示である。音声あり(本年度)は音声なしと各節の問題画面数及び内容は同じであるが, 問題画面でネイティブ

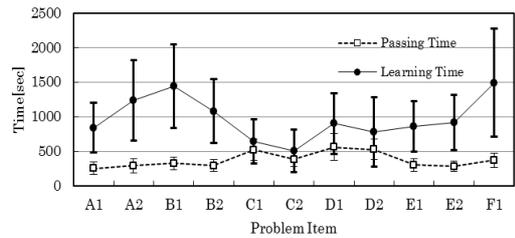


図7 学習時間と合格時間(音声あり)

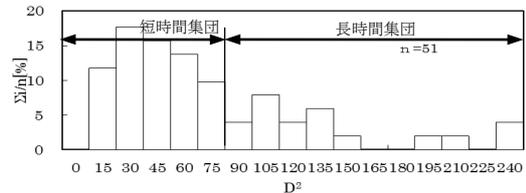


図8 学習時間のマハラノビス距離(音声あり)

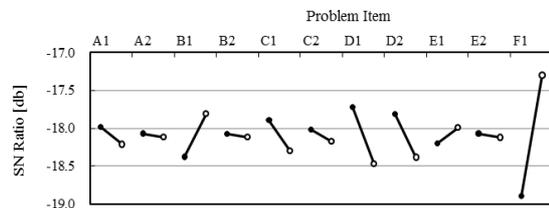


図9 学習時間のSN比(音声あり)

による音声を, 「再生」ボタンを押すことで何度も聞けるようにしている。60分の授業の後, CALL教室でLMSを用いて, 一斉に学習者自身のペースで, 節ごとの正答率が80%以上を合格点として繰り返し学習を行った。全学習者数  $n$  は51名 ( $n \geq 3m$ ) である。

### 4.2 学習時間と合格時間(音声あり)

図7は音声ありの問題項目に対する学習時間(実線)及び合格時間(破線)の平均値と標準偏差を示している。Cに比べてDは学習時間が長い。また, CとDを除く問題項目は, 学習時間と合格時間との比が2倍以上である。章ごとに平均した学習時間の問題項目を昇順に並べると, C, D, E, A, B, Fとなる。音声ありの合格時間の平均値は375secであるが, 音声なしは372secであり, 音声の有無によらずほぼ同じ値である。したがって, 音声なしの合格時間を単位空間とすることもできるが, 合格時間はそれぞれの学習時間から作成することに意味があるため, 後者の方法で単位空間を作成する。

### 4.3 学習時間のマハラノビス距離(音声あり)

図8は, 図7の合格時間を単位空間とした学習時間のマハラノビス距離  $D^2$  を示している。クラスター分析では75未満とそれ以上に2分化され, 図8からもピークが30と105になっている。音声ありの  $D^2$  の平均値は65.3であるが, 音声なしは10.1であり, 前者は後者の6倍である。音声の有無によらず合格時間はほぼ同じ値であるが, 学習時間の平均値はそれぞれ974secと533secのため, 学習時間が音声の有無による差異を大きくしている。

### 4.4 学習時間のSN比(音声あり)

図9は学習時間を  $L_{12}(2^{11})$  の直交表に割り付けて求めたSN比を示している。左下がりの問題項目はE1(Be動詞過去形WH疑問文), B1(一般動詞現在形WH疑問文), F1(一般動詞過去形WH疑問文)の順にSN比の利得が大きい。学習者の多くは疑問詞の追加や, WH疑問文のよう

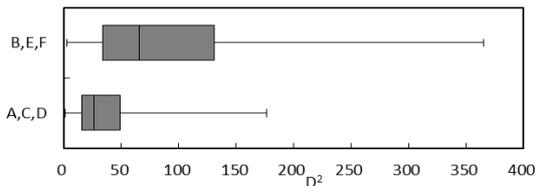
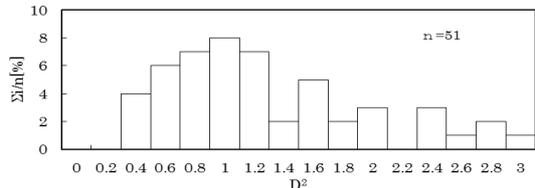
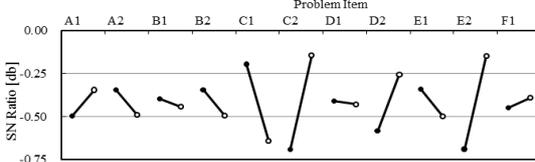


図10 学習時間のマハラノビス距離 (音声あり)



(a) マハラノビス距離



(b) SN比

図11 音声の有無による合格時間の差異

に考えなければならない問題項目を不得意とするため、SN比の利得が大きい。これに対して、C、DのようなYES/NO疑問文は単語の入れ替えだけのため、SN比の利得が小さい。なお、これらの問題項目は音声なしと同様である。図9のSN比の平均値 $\eta_0$ は-18.1db (10log65.3とも一致)であり、音声なしの-9.5dbに比べて小さい。4.2で述べたように、音声の有無によらず合格時間はほぼ同じ値であるが、音声ありの学習時間は合格時間の2倍以上のためである。学習時間が長いと、音声を何度も理解するまで聞いていることがわかる。また、図9のSN比 $\eta_{all}$ は-18.2dbであるが、本コースウェアからマハラノビス距離の長い問題項目(B, E, F)を外したコースウェアのSN比は-17.0db (利得は+1.1db)となる。

図10は問題項目が(B, E, F)または(A, C, D)で構成されるコースウェアを全学習者が回答したとするマハラノビス距離 $D^2$ の箱ひげ図を示している。 $D^2$ の平均値は前者が65.7で、後者が26.1である。前者の問題項目が合格時間を長引かせているのは、図9の考察で述べたとおりである。

#### 4.5 音声の有無による差異

図11(a)は音声なしの合格時間を単位空間とした音声ありの合格時間のマハラノビス距離 $D^2$ を示している。 $D^2$ の平均値は1.2であり、単位空間の平均値1.0に近いと、音声の有無による合格時間の差異はほとんどない。すなわち、合格時間では音声をほとんど聞いてないといえる。学習時間では音声を何度も聞いているが、合格時間では理解が深まるためと考えられる。

図11(b)は音声ありの合格時間のSN比を示している。SN比の平均値は0dbに近い。また、どの問題項目もSN比の利得が0.5db以下のため、音声なしの合格時間との差異は小さい。

## 5. まとめ

英語の繰り返し学習型コースウェアを対象に、学習時間に含まれる合格時間から単位空間を作成した。英単語では混合式の学習時間を分析し、混合式と選択式の合格時間の差異を求めた。また、英文法では記述式の学習時間を分析し、音声の有無による合格時間の差異を求めた。分析の結果、次のことが明らかとなった。

- (1) 英単語において、混合式は繰り返し学習の少ない集団であるため、学習時間のマハラノビス距離が短く、SN比が大きいことを明らかにした。また、SN比の利得からA2(タイプミスしやすいセンター試験単語)と、A~Eの3(熟語)は、繰り返し回数の多い問題項目であることを明らかにした。
- (2) 混合式は選択式に比べて画面数が多いため、合格時間のマハラノビス距離が長いことを明らかにした。しかし、混合式は高得点集団のため、選択式の低得点集団に比べてSN比を小さくする要因は少ないが、問題項目数が多い場合、選択式と混合式の合格時間のSN比の利得が大きくなることを明らかにした。
- (3) 英文法において、音声を導入すると、何度も聞くために学習時間のマハラノビス距離が長くなることを示した。また、SN比の利得からBe動詞より一般動詞、現在形より過去形、YES/NO疑問文よりWH疑問文の学習時間が長引くことを明らかにした。
- (4) 合格時間のマハラノビス距離及びSN比から、問題形式、問題画面数及び内容が同じなら、音声の有無による差異はほとんどないことを明らかにした。学習時間では音声を何度も聞いているが、合格時間では理解が深まるため、音声を聞かずに回答していることを明らかにした。

本分析法は単位空間を、特定の問題形式や特定年度の学習時間から作成するのではなく、分析対象の学習時間に含まれる合格時間から作成するため、単位空間の全学習者数を考える必要がない利点がある。また、得られた結果は英語の指導上、有用な情報を与えるものと考えられる。

## 参考文献

- [1] 渡邊博之, 加藤勝洋, “CAI コースウェアにおける学習時間分布の分析”, 電子情報通信学会論文誌 D-I, Vol.J83-D-I, No.7, pp.789-796 (July, 2000).
- [2] 植野真臣, 永岡慶三, “ガンマ分布によるeラーニング所要時間データのオンライン解析”, 日本教育工学会論文誌, Vol.29, no.2, pp.129-137 (2005).
- [3] 例えば, 立林和夫, 手島昌一, 長谷川良子, “入門 MT システム”, 日科技連出版社 (2010).
- [4] 渡邊博之, 川嶋正士, “繰り返し学習型コースウェアにおける学習時間のMTSによる分析”, 電子情報通信学会論文誌 D, Vol.J98-D, No.1, pp.163-166 (Jan. 2015).
- [5] Masashi Kawashima and Hiroyuki Watanabe, “Developing English Learning Courseware for an e-learning System Named WebCAI”, Thailand TESOL 35<sup>th</sup> International Conference, p.147 (Jan.29, 2015).