

訳語ルビふり英文の可読性評価

大黒慶久

(株) リコー・研究開発本部・情報通信研究所

〒222 神奈川県横浜市港北区新横浜3-2-3

概要

英文原稿中の単語をOCR装置により読み取って、その和訳語をルビ状に添えて該原稿を印刷する、コピーマシン様の装置の英文読解支援効果を調査した。使用者は英単語と和訳語とを視覚的（本文とルビとして）に容易に対応つけられる。本装置では、英文中の多品詞語に対して品詞連接確率を利用して品詞推定を行い、その結果に基づいて和訳語を出力する。

本稿では、官能試験によって訳語がルビふりされた英文の英文読解支援効果を明らかにした。官能試験結果ではTOEICのreading問題を40問を使用し、結果を統計的検定の方法を用いて客観的に評価した。

Evaluating the validity of printing Japanese words alongside English text

Yoshihisa OHGURO

R&D Center, RICOH COMPANY,LTD

3-2-3, Shin'yokohama, Kohoku-ku, Yokohama, Kanagawa, 222, Japan.

ooguro@ipe.rdc.ricoh.co.jp

ABSTRACT

This paper describes an experimental result of validity of printing Japanese words alongside English text. A Copier-like machine produces such a text using OCR technology and estimating parts-of-speech using trigram. Full sentence translation is not possible, but English-Japanese word-to-word translation is easily achieved. The experiments, using 40 TOEIC reading questions, shows that printing Japanese words alongside English text are useful aids for Japanese reading practical English.

1. はじめに

近年、機械翻訳システムは実用化の段階に入っており、製品化も相次いでいる。しかしながら、翻訳精度が十分でないために、依然、人手による前編集あるいは後編集が必須である。ゆえに機械翻訳システムを効果的に使用するには、ある程度の英語力が要求される。

一方、コンピュータ機器が安価になり広く普及するに伴って、機械翻訳以外でも電子辞書を搭載した装置が多数製品化されている。

本稿では、英文原稿中の単語をOCR装置により読み取って、その和訳語をルビ状に添えて印刷するコピーマシン様の装置の英文読解支援効果を評価する。本装置によって、使用者は原稿を複写する感覚で和訳語ルビの付いた原稿を得ることができる。機械翻訳とは異なり、一文単位の翻訳ではないが、英単語とその和訳語を視覚的（本文とルビ）に容易に対応づけられる。英文原稿を読み進んでいく内、未知の単語に出会った場合にルビ状に添えられた和訳語を参照すればよい。すなわち、本装置は使用者の辞書引き作業を代替することを目的とする。

機械翻訳システムの評価には様々な観点があるが⁽¹⁾、本稿では、訳語がルビふりされた英文の実用局面における読解支援効果を官能試験によって明らかにした。官能試験結果は統計的検定の方法を用いて客観的に評価した。

2. 訳語ルビふり性能

2.1 品詞推定

本装置では、3単語の品詞並びの確率（trigram: 連接確率）と各単語がとる品詞の確率（語彙確率）とを用いて品詞を推定する⁽²⁾⁽³⁾。

品詞列 $T=t_1, t_2, t_3, \dots, t_n$ の出現する確率 $P(T)$ を連接確率 $p(t_i | t_{i-2}, t_{i-1})$ を用いて次式で定義する。

$$P(T) = \prod_{i=1}^n p(t_i | t_{i-2}, t_{i-1}) \quad (2.1)$$

品詞列 T の元になる単語列 $W=w_1, w_2, w_3, \dots, w_n$ の評価値 $E(W)$ を語彙確率 $p(t_i | w_i)$ を用いて次式で定義する。

$$E(W) = \prod_{i=1}^n \{ p(t_i | t_{i-2}, t_{i-1}) \times p(t_i | w_i) \} \quad (2.2)$$

単語列 W に対して評価値 $E(W)$ が最大になる品詞列 T が品詞推定結果である。本装置では、最大値探索アルゴリズムにはビームサーチを用いている。また実際の演算では、対数確率値を整数化し、(2.2)式を

整数の加算式に変換することによって、高速化を図っている。

使用した品詞種類と連接確率の学習データおよび trigram 数を表2-1に示す。

第2-1表 連接確率データ

品詞種数 [種]	学習データ量 [語]	trigram 数 [種]
43	約100万	13149

語彙確率は、辞書に記載されている品詞順位の逆数に比例すると仮定した。

連続する複数語から構成される熟語 ("a lot of" 等) についても辞書に登録しておき、同様に扱う。評価値算出の際には、特に正規化は行わない。このため、確率値の乗算回数が少なくなるので熟語候補を含む解の評価値が高くなる傾向になるが、熟語候補を含む解を優先するということは、仮名漢字変換における最小文節数優先のヒューリスティックに等しく、都合がよい。

2.2 訳語ルビふりの条件

使用した辞書の構成を、第2-2表に示す。表中、"単-単語"とは、複数単語から構成される熟語に対して、1単語の語を表す。

第2-2表 使用した辞書の構成

単語種類	見出し語数
単-単語	40172
動詞熟語	1612
その他の熟語	14470
計	56254

ルビふり対象である英文文書は、OCR 正解率 100% を仮定して、タイプ入力する。訳語ルビふり精度はOCRの精度に大きく影響されるが、本稿では訳語ルビふり英文の可読性に注目して検討するため、OCRの影響は極力排した。

また、単-単語に関しては、品詞推定結果が機能語の品詞（代名詞・冠詞・法助動詞・be動詞・接続詞・前置詞）である場合には訳語を出力しない。

3. 可読性評価

3.1 官能試験の方法

品詞推定が正解しても、その和訳語が正解である保証はない。これは機械翻訳システムにおいても同じく問題になることである。言語処理の精度を評価するならば、評価者によって訳語選択結果の可否を判定すべきであるが、訳語選択が正解であるとの基

準は評価者によって異なり、比較が難しい。一方、実用面からいえば、訳語選択が誤っていても正解が容易に推測できる場合には英文読解に役に立つ

本稿では、後者の観点から、訳語ルビふり英文が英文読解に実際に役に立つか否かを官能試験によって明らかにする。

官能試験では被験者にTOEICのReadingの読解問題を解いてもらい、訳語ルビふりの有無により正解数を比較する。TOEICは「英語による国際コミュニケーション能力を測定するためのテスト」であり、英語運用能力の絶対的評価基準である。これは、以下のことを意味する²⁾。

1. 同じ受験者が一定の期間においてTOEICを受験した場合、能力を時系列的に比較することができる。

2. 同一集団内の各受験者の位置づけや、他集団との比較が可能である。

よって、訳語ルビを添えることが英文読解に効果があるならば、訳語ルビの無い問題を解いた場合よりもスコアは上がるであろうし、効果がなければ、スコアは不変であろう。

具体的には以下の手順で官能試験を実施した。

<<有効性確認試験（有効な能力層の検出）>>

1. TOEICのReadingの読解問題（以下TOEIC問題）を被験者集団に制限時間内に解答してもらう。過去にTOEICを受験したことがある被験者にはそのときの総合スコアを申告してもらう。

2. 1.の試験結果と過去のTOEICスコアを考慮して、被験者を運用能力別に層化する（TOEIC運営委員会のガイドラインで公表している段階）。各層を二つのグループに分ける。

3. 各層の一方のグループは訳語無し、TOEIC問題を解答してもらう（対照群）。残りのグループは訳語付きのTOEIC問題を解答してもらう（調査群）。なお、訳語付き試験を受験するグループには、訳語はコンピュータで自動的に付けたものであり、必ずしも正解でないことを告げる。

4. 各個人毎、グループ毎に訳語の有無による有意な差があるかを検定する。

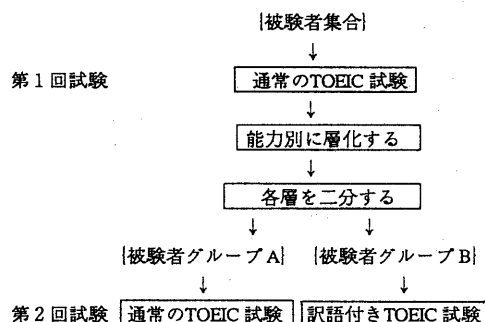
3.2 試験問題

官能試験に用いたTOEICのReading問題は、第1回、第2回ともに40問であり、これを制限時間40分で解いてもらった。第2回目の訳付き問題における品詞正解率を第3-1表に示す。

第3-1表 第2回試験における品詞正解率
品詞正解率=品詞正解数/訳語印刷箇所数

文数	単語数	品詞正解率 [%]
347	2537	97.9 (1011/1033)

注) 文数は署名などの一単語から成る文を含む



第2-1図 官能試験の進め方

3.3 結果

被験者は、当社中央研究所に勤務する者の内、協力可能であった者54名であり、特別な条件は指定しなかった。

第3-2表 第1回試験の結果（受験者全員54名）

E：標本平均 s：標本標準偏差
γ：標本相関係数

	TOEICスコア	正解数
E	469.64	28.17
s	138.37	5.36
γ	0.54	

第3-2表 官能試験結果

E：標本平均 s：標本標準偏差
1st：第1回正解数 2nd：第2回正解数

	グループA			グループB		
	TOEIC	1st	2nd	TOEIC	1st	2nd
E	472.41	28.15	24.37	466.64	28.19	25.96
s	139.45	5.16	5.96	140.01	5.66	4.92

第3-3表 TOEICスコアと正解数との相関係数

	グループA (27名)	グループB (25名)
第1回	0.52	0.54
第2回	0.64	0.73

第3-4表

第1回と第2回試験間の正解数の相関係数(γ)

	グループA (27名)	グループB (27名)
γ	0.71	0.59

4. 考察

4.1 受験者の特徴

第1-1図に受験者全員のTOEICのスコアのヒストグラムを示す。

第3-1表に受験者54名におけるTOEICスコアと第1回試験の正解数との、平均値および標準偏差、さらにTOEICスコアと第1回試験正解数との相関係数を示す。

TOEIC運営委員会の公表している定期受験におけるTOEICスコアを第4-1表、団体一括受験におけるTOEICスコアを第4-2表に示す。

第4-1表 定期受験におけるTOEICスコア*
受験者総数 384,306名

	Listening	Reading	Total
最高点	495	490	985
最低点	5	5	10
平均点	286.5	257.8	544.3
標準偏差	88.2	88.9	177.1

*TOEIC運営委員会による

第4-2表 団体一括受験におけるTOEICスコア*
受験者総数 1,494,900名

	Listening	Reading	Total
平均点	220.0	192.0	412.0

第3-1表と第4-1表とを比較すると、本官能試験の受験者集団は、定期受験におけるTOEIC受験者集団よりも平均点が70点ほど低いことがわかる。しかし、この結果から、本官能試験の受験者集団は一般的な日本人の集団に比べて英語運用能力に関して低いとは結論できない。なぜなら、定期試験におけるTOEIC受験者は自発的に受験した人々であり、英語に自信のある人々が多いことが予想されるからである。一方、団体一括受験の受験者は会社から強制的に受けさせられた人々が多く、こちらの方が平均的な日本人の英語運用能力を示している考えた方がよい。第3-1表と第4-2表とを比較すると、本官能試験の受験者集団は団体一括受験者集団よりも平均点が60点ほど高い。したがって、本官能試験の受験者集団の英語運用能力は、一般的な日本人ビジネスマンの集団より高いが、英語に自信のある集団よりは低いといえる。

第1回試験はTOEIC 200問(Reading:100問、listening:100問)中の読解問題40問であるから、その正解数と受験者が過去に受験したTOEICスコアとは高い相関があることが予想される。TOEICスコアおよび第1回試験正解数の両者ともに、実際には二項分布するが、被験者数が多いので正規分布で近似して差し支えない。第3-1表によると両者間の相関係数は0.54である。

第2-1図にTOEICスコアと第1回試験正解数との散布図を示す。図中の点線は回帰直線である。

4.2 グループ分け

官能試験受験者を英語運用能力に関して等しい二つのグループに分けるために、第1回試験の正解数とTOEICスコアとの、平均値および標準偏差が等しくなるようにグループ分けした。第3-2表にその結果を示す。

第1-2図(a)にグループA、第1-2図(b)にグループB、のTOEICスコアのヒストグラムを示す。

等しく二つに分けられていることを確認するために、①正解数、②TOEICスコア、の平均値および分散に関して統計的に検定する。

①正解数

$$F_0 = \text{グループBの分散} / \text{グループAの分散} \\ = (5.66)^2 / (5.16)^2 \\ = 1.88 \cdots < F_{26,26;0.025} = 2.19 \cdots$$

有意差無しであり、分散に差があるとはいえない。分散が等しいとみなして、平均値の差を検定する。

$$t_0 = -0.025 \quad |t_0| < t(52, 0.025) = 2.00 \cdots$$

危険率5%で検定の結果、正解数の平均値の差は有意であるとはいえない。

②TOEICスコア

$$F_0 = \text{グループBの分散} / \text{グループAの分散} \\ = (140.01)^2 / (139.45)^2 \\ = 1.00 \cdots < F_{24,26;0.025} = 2.22 \cdots$$

有意差無しであり、分散に差があるとはいえない。分散が等しいとみなして、平均値の差を検定する。

$$t_0 = 0.149 \quad |t_0| < t(50, 0.025) = 2.00 \cdots$$

危険率5%で検定の結果、正解数の平均値の差は有意であるとはいえない。

①②の結果より、両グループの

- ・正解数
- ・TOEICスコア

には有意な差があるとはいえないと考えた方が妥

当である。ゆえに、両グループの英語運用能力に差はないものとする。

第2-2図(a)にはグループA、第2-2図(b)にはグループB、のTOEICスコアと第1回試験正解数との散布図を示す。図中の点線は回帰直線であり、第2-1図の受験者全員に対する散布図の回帰直線と比較すると、両グループともほぼ同じ傾向である。

4.3 第1回と第2回との個人別正解数の差

TOEICスコアは英語運用能力を測る絶対的尺度であり、英語運用能力が変化しない限りスコアは不変である(誤差±25)⁽²⁾。

第3-1図には、TOEIC総合スコアと、第1回試験と第2回試験との正解数の差(向上数)の関係を示す。A,B両グループともに、正解向上数が負である受験者が多く、第1回試験と比較して第2回試験は難しかったであろうことが予想される。

グループAは、第1回、第2回ともに通常の試験だったので、受験者の英語運用能力に変化が無い場合には、正解数の差とTOEICスコアとは無相関(=0)であることが予想される。一方、グループBは、第1回試験は通常試験、第2回試験は訳語ルビ付き試験であるから、語彙の乏しい受験者にとっては有利にはたらく、語彙の豊富な受験者にとっては無益にはたらくことが予想される。この仮説に基づけば、TOEICスコアが高い者ほど、第1回、第2回の正解数の差が小さくなる、負相関(<0)になると思われる。

第3-1図によると、グループAは相関係数0.292と無相関に近い。また、グループBも相関係数0.220と無相関に近く、予想に反する。しかし、グループAとグループBとを比較すると、グループBの方が正解数の差が小さい(平均値:グループAの差=-3.778、グループBの差=-2.680)。つまりグループAと比べてグループBの方が正解数の下がり方が小さく、第2回試験においては、訳語ルビ付き試験を受けたグループBの方が正解数が多い。

第3-2図には、第1回試験の正解数と、第1回試験と第2回試験との正解数の差(向上数)の関係を示す。図によると、第2回目に通常試験を受けたグループAよりも、訳語ルビ付き試験を受けたグループBの方が、ちたばりの度合いが小さい。先に述べたように訳語ルビは、第1回試験の正解数が少ない受験者(=語彙の乏しい者)には有利にはたらく、第1回試験において正解数の多い受験者(=語彙の豊富な者)には無益にはたらくことが予想できる。したがって、第1回試験の正解数と正解向上数の関係は、

グループAが無相関(=0)、グループBが負相関(<0)になるはずである。試験結果から両グループの相関係数を算出すると、グループAが-0.188、グループBが-0.563であり、予想と一致する。

TOEICスコアと正解向上数との関係は両グループとも無相関であり、予想に反したが、第1回正解数と正解向上数との関係は予想どおりである。これは、TOEICスコアが

Listening (45分間)		Reading (75分間)	
写真描写	20	文法・語彙	40
応答	30	誤文訂正	20
会話	30	読解問題	40
説明文	20		
計	100問	計	100問

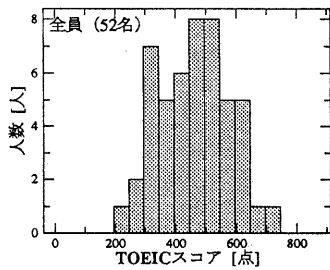
という複数種の問題から算出されており、読解力だけの指標ではないからだと思われる。一方、第1回試験正解数は読解力のみを表現する指標であり、加えて第1回試験と第2回試験とは同形式の試験であるから、その両者の正解数は直接に関係するであろうと推測できる。

第3-2図において、破線はグループBに対する回帰直線、一点鎖線はグループAに対する回帰直線である。二つの回帰直線は、各々のグループの散布図の重心を表現している。二つの回帰直線は(33.0, -4.5)において交わる。幾何学的には、この交点より左ではグループB(訳付き)の方がグループA(訳無し)よりも向上数が多く、この交点より右ではグループA(訳無し)の方がグループB(訳付き)よりも向上数が多いと考えられる。したがって、訳語ルビふりが有効に働くのは、第1回試験正解数が33問以下の集団であるといえる。第1回試験の正解数が33問である受験者は、第2-1図によるとTOEIC総合スコアが700点付近の受験者である。TOEIC運営委員会の資料によると、TOEICにおいて総合スコアが700点以下の受験者は総受験者の80%程度を占める。

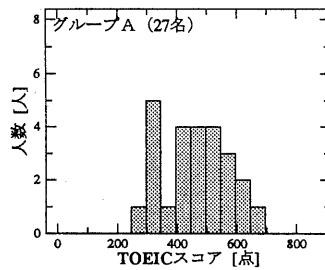
4.4 第2回試験正解数のグループにおける比較

第3-2表によると、両グループとも第2回試験正解数の平均は第1回試験正解数よりも低く、第2回試験の方が難しかったことがわかる。

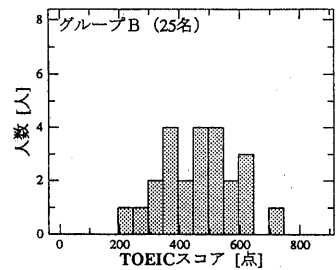
通常試験を受けたグループAの相関係数は0.643、訳付き試験を受けたグループBの相関係数は0.728であり、いずれも高い相関がある。第1回試験においては、相関係数が各々0.518、0.538であり、TOEICスコアと読解問題正解数との相関は低かった。先述し



第1-1図

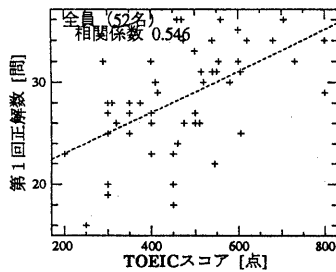


第1-2(a)図

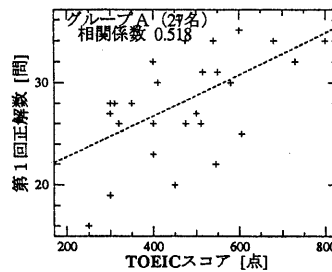


第1-2(b)図

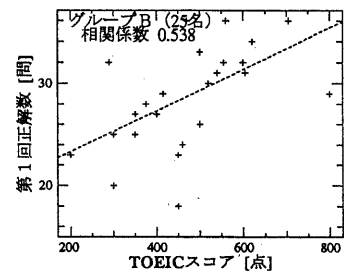
被験者におけるTOEICスコアの分布



第2-1図

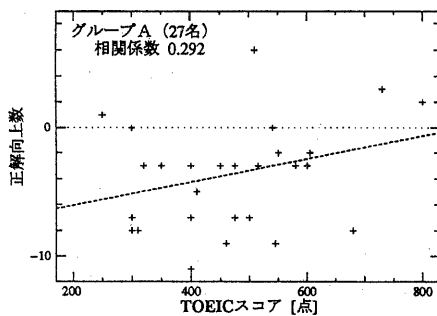


第2-2(a)図

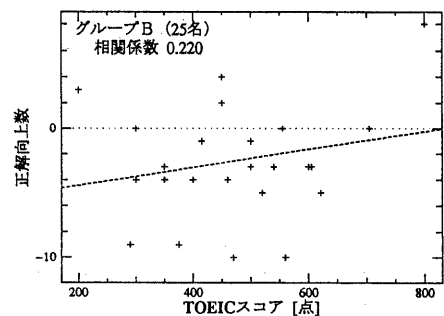


第2-2(b)図

被験者のTOEICスコアと第1回試験との関係

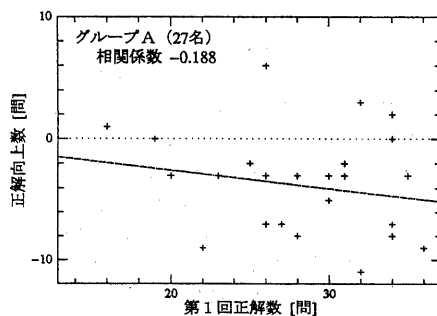


第3-1(a)図

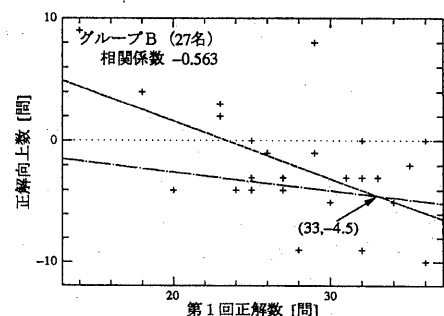


第3-1(b)図

個人別の正解数の向上数 (第2回正解数-第1回正解数)



第3-2(a)図



第3-2(b)図

個人別の正解数の向上数 (第2回正解数-第1回正解数)

たように、TOEICスコアは読解問題を含む複数種の問題から算出されるから、TOEICスコアと読解問題正解数とは高い相関があるはずである。第2回試験では両グループともに高い相関があるから、第1回試験は一般的な水準よりも易しかったといえる。

4.5 問題毎の比較

(1) 第1回試験

第3-5表は、第1回試験において被験者の選んだ選択肢と正答率を問題毎にまとめた表である。第1回試験は両グループともに通常の読解問題を解いた。両グループの総合的な正答率を比較すると、グループAが70.28%、グループBが70.46%とほぼ等しく、同じ英文読解力を持つグループへ分割されていることが確認できる。

また、問題毎に正答率を比較すると、グループAの方が正答率が高い問題が15問、グループBの方が正答率が高い問題が15問、正答率が同じ問題が10問であり、極端な偏りはなかった。この正答率の勝敗を符合検定する。

《符合検定》

グループAの方が正答率の高い問題を"+"

グループBの方が正答率が高い問題を"-"

両グループの正答率が等しい問題を"0"

符合	+	-	0
問題数	15	15	10

符合検定表より、 $n=30$ ("+"-"の合計)のときの危険率5%で有意な問題数(少ない方の問題数)は9(<15)であるから、符合検定の結果、両グループの正答率には有意な差はないとする。

符号検定の結果からも、均等な英語実力グループに分割されているといえる。

(2) 第2回試験

第2回試験はグループAは通常の読解問題、グループBは訳付きの読解問題を解いた。両グループの40問に対する総合的な正答率は、グループAが60.93%、グループBが64.91%であり、グループBの方が3.75%正答率が高い。

また、問題毎に正答率を比較すると、グループAの方が正答率が高い問題が12問、グループBの方が正答率が高い問題が25問、正答率が同じ問題が3問であり、訳付き試験を受けたグループBの方が正答率の高い問題が多かった。この正答率の勝敗を符合検定する。

《符合検定》

符合	+	-	0
問題数	12	25	3

符合検定表より、 $n=37$ ("+"-"の合計)のときの危険率5%で有意な問題数(少ない方の問題数)は12(=12)であるから、符合検定の結果、両グループの正答率には有意な差があるとする。

両グループともに通常試験を受けた第1回では、総合正答率および問題毎の正答率はほぼ同じであったのに対し、一方が訳付き試験を受けた第2回では、総合正答率および問題毎の正答率に差が生じていることから、訳がついたことにより正答率が向上したものと考える。

4.6 統計的検定⁽⁵⁾

両グループの正解の平均値および分散に関して統計的に検定する(t検定)。4.1節でも述べたように、正解数の分布は離散分布の二項分布にしたがうが、データ数が多いので連続分布の正規分布で近似しても構わない。まず、TOEICスコア800以下の集団(被験者全員)を対象にし、次にTOEICスコア730(TOEFL500点に相当)以下の集団を対象にする。

(1) TOEICスコア 800 以下

第4-3表 TOEIC読解問題40問の正解数
TOEICスコア 800 以下 (54名)

	グループ A (27名)	グループ B (27名)
第1回	①訳語なし 平均 = 28.148 標準偏差 = 5.157	③訳語なし 平均 = 28.185 標準偏差 = 5.657
第2回	②訳語なし 平均 = 24.370 標準偏差 = 5.962	④訳語あり 平均 = 25.963 標準偏差 = 4.918

○グループ間に関して統計的検定

(a) 第1回

$$F_0 = \frac{\text{グループ B の分散}}{\text{グループ A の分散}} \\ = \frac{(5.657)^2}{(5.157)^2} \\ = 1.203 \dots < F_{26,26;0.025} = 2.19 \dots$$

有意差無しであり、分散に差があるとはいえない。分散が等しいとみなして、平均値の差を検定する。

$$t_0 = -0.025 \quad |t_0| < t(52, 0.025) = 2.00 \dots$$

危険率5%で検定の結果、正解数の平均値の差は有意であるとはいえない。

(b) 第2回

$$F_0 = \text{グループ A の分散} / \text{グループ B の分散} \\ = (5.962)^2 / (4.918)^2 \\ = 1.470 \dots < F_{26,26;0.025} = 2.19 \dots$$

有意差無しであり、分散に差があるとはいえない。
分散が等しいとみなして、平均値の差を検定する。

$$t_0 = 1.071 \quad |t_0| > t(52, 0.300) = 1.04 \dots$$

危険率 30% で検定の結果、正解数の平均値の差は有意であるといえる。

(2) TOEIC スコア 730 未満

第4.4表 TOEIC読解問題40問の正解数
TOEIC スコア 730 未満 (50名)

	グループ A (25名)	グループ B (25名)
第一回	①訳語なし 平均 = 27.760 標準偏差 = 5.158	③訳語なし 平均 = 27.880 標準偏差 = 5.711
第二回	②訳語なし 平均 = 23.480 標準偏差 = 5.229	④訳語あり 平均 = 25.240 標準偏差 = 4.304

○グループ間に関して統計的検定

(a) 第 1 回

$$F_0 = \text{グループ B の分散} / \text{グループ A の分散} \\ = (5.711)^2 / (5.158)^2 \\ = 1.226 \dots < F_{24,24;0.025} = 2.269 \dots$$

有意差無しであり、分散に差があるとはいえない。
分散が等しいとみなして、平均値の差を検定する。

$$t_0 = -0.078 \quad |t_0| < t(48, 0.025) = 2.011 \dots$$

危険率 5% で検定の結果、正解数の平均値の差は有意であるとはいえない。

(b) 第 2 回

$$F_0 = \text{グループ A の分散} / \text{グループ B の分散} \\ = (5.229)^2 / (4.304)^2 \\ = 1.47 \dots < F_{24,24;0.025} = 2.269 \dots$$

有意差無しであり、分散に差があるとはいえない。
分散が等しいとみなして、平均値の差を検定する。

$$t_0 = -1.299 \quad |t_0| \geq t(48, 0.100) = 1.299 \dots$$

危険率 20% で検定の結果、正解数の平均値の差は有意であるといえる。

(1)(2)の結果により、TOEIC総合スコアが800以下の集団を対象とした場合よりも730以下に絞った集団の方が危険率が下がる。これは、800以下の集団で有効というよりも、730以下の集団に有効であるという方が確からしさが増すことを意味する。

5. 結論

両グループともに通常試験を受けた第1回では

- ・総合的な正答率
- ・問題毎の正答率に対する符合検定
- ・平均値の差に関するt検定

のいずれにおいても両グループの正解数には明確な差はみられなかった。

一方、第2回試験では、

- ・総合的な正答率
- ・問題毎の正答率に対する符合検定
- ・平均値の差に関するt検定

のいずれにおいても訳付き試験を受けたグループ Bの方が正解数が多かった。

以上の結果より、「訳語ルビあり」は英文読解に有効に機能すると結論する。統計的検定によると、特にTOEICスコアが730以下の集団に有効らしいことがわかった。また回帰直線から推定すると、TOEICスコアが700点までの人に有効に機能し、そのような人はTOEIC受験者の80%に相当する。

6. 今後の展開

通常試験と訳付き試験とにおいて、受験者の解答傾向が大きく異なっている問題がある。受験者にインタビューすることによって、この原因を追求し、読解支援効果に影響する要素を明らかにする。

また、(紙の)辞書引きを許して問題を解いた場合との比較も行う。

謝辞

官能試験に快く協力して下さった中央研究所の皆さんに感謝します。

【参考文献】

- (1) 電子協編: "機械翻訳システムの実用化に関する調査研究", 平成4年度調査事業報告書, 93-計-6 (1993)
- (2) 木村恒夫: "TOEIC リーディングの徹底対策", 語研 (1987)
- (3) Church K.W.: "A Stochastic Parts Program and Noun Phrase Parser for Unrestricted Text", ACL Proceedings 2nd Applied Natural Language Processing (1988)
- (4) DeRose, S.: "Grammatical Category Disambiguation by Statistical Optimization", Computational Linguistics, Vol.14, No.1 (1988)
- (5) ガットマン他 (石井他訳): "工科系のための統計概論", 培風館 (1968)