

Internal Use Only (非公開)

TR-SLT-0083

TOEIC スコアを利用した翻訳能力評価実験

Experiments for Automatic Evaluation of Speech Translation System by TOEIC Score

中塚 裕之

竹澤 寿幸

Hiroyuki Nakatsuka

Toshiyuki Takezawa

2004年9月10日

概要

複数のデータセットに対して日英方向の翻訳評価法が英日方向に応用できることを確かめる。特に、MAD3とMAD4という特徴の異なる2つのデータを比べることで、話者への「簡潔に話せ」といった教示が機械翻訳の性能を高めることを確認する。翻訳一対比較法の自動化手法は、翻訳自動評価法（DPベース自動評価法、BLEU）を用いて、システムの翻訳能力をTest of English for International Communication (TOEIC) のスコアとして換算できる評価法であり、あまりコストをかけずに、翻訳一対比較法と非常に近い評価結果を得られるという優れた長所を持つ。本稿では、はじめにTOEIC_MAD3/4_EJ/JEデータについて説明し、その解析結果を報告する。次に、TOEIC_MAD3/4_EJ/JEデータとリファレンス（複数の正解訳）を用いて翻訳一対比較法の自動化手法により日英・英日双方向の翻訳システムに対する評価実験の結果を報告する。さらに、今回用いたMAD3とMAD4の差について検討、考察する。

(株) 国際電気通信基礎技術研究所

音声言語コミュニケーション研究所

〒619-0288 「けいはんな学研都市」 光台二丁目 2 番地 2 TEL : 0774-95-1301

Advanced Telecommunications Research Institute International

Spoken Language Translation Research Laboratories

2-2-2 Hikaridai "Keihanna Science City" 619-0288, Japan

Telephone: +81-774-95-1301

Fax : +81-774-95-1308

©2004 (株) 国際電気通信基礎技術研究所

©2004 Advanced Telecommunications Research Institute International

目次

1	はじめに	1
2	翻訳一対比較法の自動化手法	1
2.1	DP ベース自動評価法	1
2.2	BLEU	1
3	TOEIC_MAD データ	2
3.1	MAD について	2
3.2	TOEIC_MAD3/4 データについて	2
3.2.1	TOEIC_MAD3/4_JE データ	2
3.2.2	TOEIC_MAD3/4_EJ データ	2
4	TOEIC_MAD3/4 データの解析	3
4.1	word PerPlexity と平均単語数による文の難しさの解析	3
4.2	類似度による聞き取りと訳の正しさの解析	4
5	翻訳システムの評価実験	4
6	考察	6
7	まとめと今後の課題	7
	謝辞	7
	参考文献	7
A	グラフ	9

1 はじめに

音声翻訳システムの研究・開発において、翻訳評価法は必要不可欠である。低コストでの客観的な評価結果を与えるような翻訳評価法は優れた評価法であり、そのような翻訳評価法を考案することは非常に重要である。

日英方向の翻訳システムの評価法に翻訳一対比較法がある。翻訳一対比較法は、音声翻訳システムの翻訳能力を Test of English for International Communication (TOEIC) のスコアに換算して表すことができる [2]。この手法は TOEIC スコアという明瞭な評価結果を示すことができるが、その評価に非常にコストがかかるという問題があった。

翻訳一対比較法のコストの問題を解決する手法として、翻訳一対比較法の自動化手法が提案された。翻訳一対比較法の自動化手法は、翻訳自動評価法 (DP ベース自動評価法、BLEU) を用いて、音声翻訳システムの翻訳能力を TOEIC のスコアに換算する。この手法によって、あまりコストをかけずに翻訳一対比較法の評価結果と非常に近い評価結果を得ることができるようになった。

この翻訳評価法は、日英方向の音声翻訳システムの評価に対して考案されたものであるが、松吉らによってこの日英方向の翻訳一対比較法の自動化手法が英日方向の翻訳評価に応用できる見込みがあることが示された。

本論文では、松吉ら [1] の実験の追試を行なうとともに、話者に「簡潔に話せ」などの教示を与えることが翻訳に与える影響を調査した結果を報告する。

本論文は以下のように構成される。まず 2 章で翻訳一対比較法の自動化手法について、その概要を述べる。次に 3 章でテストセットと被験者の翻訳結果の集合である TOEIC_MAD データについて説明する。4 章ではその TOEIC_MAD データ自体の解析結果およびリファレンスと合わせた解析結果を示す。5 章で実際に翻訳システムの評価実験について述べ、結果を示す。6 章は前章までの結果の考察を行なう。7 章はまとめである。

2 翻訳一対比較法の自動化手法

翻訳一対比較法の自動化手法について概要を述べる。

翻訳一対比較法の自動化手法は、これまでに提案されている翻訳自動評価法を用いて TOEIC スコアが異なる複数の被験者の翻訳結果と翻訳システムの翻訳結果にスコアをつけ、回帰分析によりシステム TOEIC 換算点を求める評価手法である。

翻訳一対比較法の自動化手法では、翻訳自動評価法として DP ベース自動評価法と BLEU を使用している。以下、これら 2 つの翻訳自動評価法について概説する。

2.1 DP ベース自動評価法

DP ベース自動評価法では、以下に定義するスコア S_{DP} により翻訳の評価を行う。

$$S_{DP} = \frac{1}{N_{total}} \sum_{j=1}^{N_{total}} \max_i \left(\frac{T_{ij} - S_{ij} - I_{ij} - D_{ij}}{T_{ij}}, 0 \right) \quad (1)$$

ただし N_{total} はテストセットに含まれる発話数、 T_{ij} はテストセットの発話 j に対するリファレンス i の総語数、 S_{ij} はテストセットの発話 j に対するリファレンス i と評価対象の翻訳を DP マッチングにより比較したときの置換語数、 I_{ij} は同様に比較したときの挿入語数、 D_{ij} は同様に比較したときの脱落語数である。ここでは置換は「脱落+挿入」とは考えない。

2.2 BLEU

BLEU では、以下に定義するスコア S_{BLEU} により翻訳の評価を行う。

$$S_{BLEU} = \exp \left\{ \sum_{n=1}^N w_n \log(p_n) - \max \left(\frac{L_{ref}^*}{L_{sys}} - 1, 0 \right) \right\} \quad (2)$$

ただし、 N は n -gram マッチ率を計算する際に扱う最大 n -gram 長を表し、 w_n はその逆数を表す。また、 L_{sys} は評価対象の翻訳に含まれる単語数を表し、 L_{ref}^* は L_{sys} と最も単語数が近いリファレンスに含まれる単語数を表す。 p_n は修正 n -gram 適合率であり、 M_i^{n-gram} をテストセットの発話 i を翻訳した文の n -gram のうちリファレンスとマッチするものの数、 G_i^{n-gram} をテストセットの発話 i を翻訳した文の n -gram の数としたとき、次式により計算される。

$$p_n = \frac{\sum_i M_i^{n-gram}}{\sum_i G_i^{n-gram}} \quad (3)$$

3 TOEIC_MAD データ

3.1 MAD について

MAD3 と MAD4 は、対話内容は同じであるが、それを実際に話す際に「なるべく簡潔に話す」「私はなどの主語を補う」といった話し方に関する詳細な教示の有無が異なる。MAD3 は話し方に関する詳細な教示なし、MAD4 は話し方に関する詳細な教示ありである [5]。

3.2 TOEIC_MAD3/4 データについて

MAD3/4 と TOEIC を利用して日英・英日の翻訳能力を評価するために、JE データと EJ データを作成する。JE データとは日英方向の翻訳能力を評価するためのデータであり、EJ データとは英日方向の翻訳能力を評価するためのデータである。

3.2.1 TOEIC_MAD3/4_JE データ

TOEIC_MAD3/4_JE データの作成過程について詳説する。TOEIC_MAD3/4_JE データは次のようにして作成された。

- MAD3/4 における日本語の native speaker による日本語の発話から MAD3 は 509、MAD4 は 502 の発話を選択した
(以下その発話を日本語テストセットと呼ぶ)
- TOEIC スコアが異なる日本語 native の被験者を MAD3 では 32 名、MAD4 では 35 名集め、次の課題をさせた。
 1. 日本語テストセットの発話を聞いて、その英訳文を作る。(以下、その英訳文を「英訳文」と呼ぶ)

この「英訳文」を発話ごとに集めて整理したものが TOEIC_MAD3/4_JE データである。

また被験者は、TOEIC 試験を受けてから半年以内の TOEIC スコア 300 点台から 900 点台まで、各 100 点台ごとに 5 名ずつとなっている。ただし、MAD3 においては 300 点台のみ 2 名となっている。

3.2.2 TOEIC_MAD3/4_EJ データ

TOEIC_MAD3/4_EJ データの作成過程について詳説する。TOEIC_MAD3/4_EJ データは次のようにして作成された。

- MAD3/4 における英語の native speaker による英語の発話から MAD3 は 504、MAD4 は 502 の発話を選択した
(以下その発話を英語テストセットと呼ぶ)
- TOEIC スコアが異なる日本語 native の被験者を MAD3 では 20 名、MAD4 では 21 名集め、次の課題をさせた。

1. 英語テストセットの発話を聞いて、それを書き取る。
(以下、被験者が聞き取った英語をテキストにしたものを「聞き取った英語の発話」と呼ぶ)
2. 「聞き取った英語の発話」の和訳文を作る。
(以下、その和訳文を「聞き取った英語の発話の和訳文」と呼ぶ)
3. テキストとして示された英語テストセットの発話(以下、「正解の英語の発話」と呼ぶ)を見て、それを和訳する。(以下、その和訳文を「正解の英語の発話の和訳文」と呼ぶ)

この「正解の英語の発話」、「聞き取った英語の発話」、「聞き取った英語の発話の和訳文」および「正解の英語の発話の和訳文」を発話ごとに集めて整理したものが TOEIC_MAD3/4_EJ データである。

また被験者は、TOEIC 試験を受けてから半年以内の TOEIC スコア 300 点台から 900 点台まで、各 100 点台ごとに 3 名ずつとなっている。ただし、MAD3 では 400 点台のみ 2 名となっている。

以下、表および図の中では、「聞き取った英語の発話」を LE、「聞き取った英語の発話の和訳文」を LJ、「正解の英語の和訳文」を TJ と略記することにする。

上記の整理としては、音としては意味がとれるが、綴りが誤っているという単語のスペルミスの修正、および、何か単語が存在することを示すために被験者が独自に使用した特殊記号の削除を人手で行っている。

なお、JE データと EJ データで作成方法が異なるが、これは被験者が日本語 native であるため、JE データに関して、被験者は発話された日本語を 100 % 聞き取ることができであろうという予測により、発話の書き取り等のタスクを省略している。よって、JE データにおける「英訳文」は EJ データにおける「正解の英語の和訳文」に相当すると考えられる。

4 TOEIC_MAD3/4 データの解析

まず前章のようにして作成した TOEIC_MAD3/4_JE/EJ データの「英訳文」や「聞き取った英語の発話」、「聞き取った英語の発話の和訳文」、「正解の英語の発話の和訳文」がどのような特徴を持っているかということを明らかにするためにデータの解析を行った。

さらに翻訳一対比較法の自動化手法が適用できるかどうかを明らかにするために、TOEIC_MAD3/4_JE/EJ データとリファレンス(複数の正解翻訳)を合わせて解析した。松吉らにより翻訳一対比較法の自動化手法が英日方向の翻訳評価に応用できる見込みがあることが示されているが、本実験はその追加実験の意味も持つ。なお、そのとき同時に自動評価に必要なリファレンス数が 10 程度で十分であることも示されているが、今回はリファレンスとして、各テストセットの発話 1 発話に対して 15 の翻訳を人手で作成したものを用いた。

被験者の翻訳結果対リファレンスの類似度と TOEIC のスコアとの間に正の相関があれば、翻訳一対比較法の自動化手法を英日方向の翻訳評価に応用することができる。

以下、各 TOEIC_MAD3/4 データにおいて調査した結果を示す。なお類似度の計算には、DP マッチングと最大 n -gram 長を 4 とした BLEU を用いた。また word PerPlexity の計算には 3-gram で行なった。

4.1 word PerPlexity と平均単語数による文の難しさの解析

テストセットの発話の難しさを word PerPlexity と平均単語数を用いて調査した。その結果を表 4.1 に示す。

MAD3 と MAD4 を比較すると、JE/EJ 両方について、PP も平均単語数も MAD4 の方が低い。これは MAD3 より MAD4 の発話の方が簡単であることを表しており、教示の効果が発話に現れていることを示している。

「英訳文」、「聞き取った英語の発話の翻訳文」、「正解の英語の翻訳文」の難しさを word PerPlexity を用いて求め、それと TOEIC のスコアの関係性を調査した。その関係を図 1- 図 6 に示す。また、それぞれの相関係数を表 2 に示す。さらにその被験者の平均とリファレンスの平均をまとめたものを表 3 に、訳文の平均単語数について被験者の平均とリファレンスの平均をまとめたものを表 4 に示す。

表 1: MAD3/4 のテストセットの発話の word PerPlexity と 1 発話あたりの平均単語数

		PP	平均単語数			PP	平均単語数
MAD3	JE	31.71	12.0	MAD4	JE	23.28	11.4
	EJ	42.79	13.5		EJ	28.18	11.2

表 2: MAD3 と MAD4 の word PerPlexity 相関係数

	JE	EJ	
		LJ	TJ
MAD3	-0.540	-0.373	-0.588
MAD4	-0.697	-0.592	-0.654

表 2 より、MAD3_EJ_LJ の相関係数が極端に低い値となっているが、全体的にも TOEIC スコアとの相関はそれほど高くない。また、LJ と TJ を比較すると、TJ の方が相関が高くなっている。これは表 3 の word PerPlexity が LJ から TJ で高くなっていることと合わせると、被験者はリスニングでは聞き取れない部分は訳さないが、テキストを見せられると、不自然ではあっても訳そうとしたため、TJ では相関が強まったと推察できる。

表 3 より、MAD3 と MAD4 のリファレンスの値を比べると、MAD4 の方が低い値となっている。これはテストセットの発話の簡単さが、翻訳結果の簡単さにつながったためと考えられる。また、表 4 と合わせて LJ と TJ の差を見ると、MAD3 より MAD4 の方が差が小さい。これは LJ と TJ が余り変わらないことを示しており、テストセットの発話が簡単になったことにより、リスニングの段階で被験者は十分に聞き取れていることが分かる。

4.2 類似度による聞き取りと訳の正しさの解析

まず「聞き取った英語の発話」と「正解の英語の発話」の類似度を求め、それと TOEIC のスコアの間を調査した。この類似度の計算には DP マッチングを用いた。その関係を図 7、図 8 に示す。

次に「英訳文」、「聞き取った英語の発話の和訳文」、「正解の英語の和訳文」それぞれとリファレンスの類似度を求め、それと TOEIC のスコアの間を調査した。これらの類似度の計算には DP マッチングと BLEU を用いた。その関係を図 9- 図 20 に示す。

また、これらの相関係数を表 5、表 6 に示す。

表 5、表 6 より、JE の DP による相関係数が少し低いものの、全体的には相関係数は高めの値をとっており、それぞれの類似度と TOEIC スコアは相関しているようである。中でも LE は非常に高い値となっており、強い相関があることが見て取れる。

5 翻訳システムの評価実験

前章で、それぞれの訳文に対して、訳文とリファレンスの類似度と TOEIC のスコアが相関していることが分かった。そこで TOEIC_MAD データとリファレンスを用いて、翻訳一対比較法の自動化手法を日英・英日の両方向の翻訳システムの評価に応用する実験を行った。ここでは評価対象として、音声翻訳システム ATR-MATRIX の言語翻訳サブシステム TMDT の MAD3/4 version を用いた。またこの TMDT に対する入力、各言語の native speaker が発話した音声をタイプストがタイプしたものとした。タイプストはほとんどタイプミスをしなため、発話そのもの (EJ データであれば「正解の英語の発話」に相当) が入力されたと考えてよい。

DP マッチングと BLEU による「システムの翻訳」とリファレンスの類似度を表 7 に示す。これらと前章で求めた被験者の訳文のスコアからシステム TOEIC 換算点を求めた。

表 3: MAD3 と MAD4 の word PerPlexity

	MAD3					
	JE		EJ			
			LJ		TJ	
	PP	標準偏差	PP	標準偏差	PP	標準偏差
被験者	82.44	27.76	39.07	4.02	41.06	4.59
リファレンス	65.20	15.43	-		40.38	5.66

	MAD4					
	JE		EJ			
			LJ		TJ	
	PP	標準偏差	PP	標準偏差	PP	標準偏差
被験者	66.21	19.03	30.68	3.63	31.01	4.02
リファレンス	51.15	11.13	-		31.25	5.25

表 4: MAD3 と MAD4 の訳文 1 発話あたりの平均単語数

	MAD3					
	JE		EJ			
			LJ		TJ	
	平均単語数	標準偏差	平均単語数	標準偏差	平均単語数	標準偏差
被験者	10.01	0.91	12.70	1.86	13.77	1.13
リファレンス	11.53	0.83	-		15.63	0.69

	MAD4					
	JE		EJ			
			LJ		TJ	
	平均単語数	標準偏差	平均単語数	標準偏差	平均単語数	標準偏差
被験者	9.54	0.80	11.20	0.97	11.56	0.83
リファレンス	10.69	0.74	-		13.09	0.46

表 5: MAD3 と MAD4 の DP 相関係数

	JE	EJ		
		LE	LJ	TJ
MAD3	0.625	0.848	0.795	0.770
MAD4	0.632	0.868	0.742	0.652

表 6: MAD3 と MAD4 の BLEU 相関係数

	JE	EJ		
		LE	LJ	TJ
MAD3	0.720	-	0.778	0.745
MAD4	0.664	-	0.682	0.659

表 7: 「システムの翻訳」とリファレンスの類似度

	DP		BLEU	
	JE	EJ	JE	EJ
MAD3	0.457	0.428	0.357	0.309
MAD4	0.543	0.532	0.450	0.442

表 8: MAD3 と MAD4 の被験者 TOEIC score からの機械翻訳 TOEIC score

	JE	EJ		
		LE	LJ	TJ
MAD3(DP)	326	-	258	-162
MAD3(BLEU)	633	-	282	-41
MAD4(DP)	592	-	274	39
MAD4(BLEU)	704	-	173	48

DP マッチングを用いて被験者の訳文のスコアを計算し、回帰直線を引いたものを図 9- 図 14 に示す。また BLEU を用いて被験者の訳文のスコアを計算し、回帰直線を引いたものを図 15- 図 20 に示す。これらの図における水平の直線は「システムの翻訳」のスコアを表している。

それぞれの場合において、回帰分析によってシステム TOEIC 換算点を求めた。その結果を表 8 に示す。ここで、LJ(「聞き取った英語の発話の和訳文」とリファレンスの類似度を用いてシステム TOEIC 換算点を求めるということは、音声翻訳システムの音声認識器の認識精度が 100 % という状態におけるシステムの TOEIC スコアを求めることに相当する。

6 考察

TOEIC テストというのは LISTEN と WRITE の 2 部からなり、TOEIC のスコアはこれらの合計点で表される。前章では TOEIC の TOTAL のスコアに対して相関を求めたが、今回のタスクは旅行会話の翻訳という、比較的会話能力に特化したタスクであるため、ここでは TOEIC の LISTEN のスコアに対しての相関を求める。

まず「聞き取った英語の発話」と「正解の英語の発話」の類似度を求め、それと TOEIC の LISTEN のスコアの関係性を調査した。この類似度の計算には DP マッチングを用いた。

次に「英訳文」、「聞き取った英語の発話の和訳文」、「正解の英語の和訳文」それぞれとリファレンスの類似度を求め、それと TOEIC の LISTEN のスコアの関係性を調査した。これらの類似度の計算には DP マッチングと BLEU を用いた。

また、これらの相関係数を表 9、表 10 に示す。なお、この表には比較のため TOEIC の TOTAL の相関係数も併記している。

これらの表 9、表 10 より、全ての値について TOTAL のスコアよりも LISTEN のスコアの方が相関が高いことが分かる。

また、TOEIC のスコアに対して、TOTAL と LISTEN、TOTAL と WRITE、LISTEN と WRITE それぞれの相関係数を表 11 に示す。

これより TOTAL と LISTEN の相関が一番高く、かつ非常に強い相関があることが分かる。よって、高い精度で LISTEN のスコアから TOTAL のスコアを導くことができる。

前章と同様に、DP マッチングを用いて被験者の訳文のスコアを計算し、回帰直線を引いたものを図 21- 図 26 に示す。また BLEU を用いて被験者の訳文のスコアを計算し、回帰直線を引いたものを図 27- 図 32 に示す。これらの図における水平の直線は「システムの翻訳」のスコアを表している。

それぞれの場合において、回帰分析によって LISTEN のスコアからシステム TOEIC 換算点を求めた。その

表 9: MAD3 と MAD4 の DP 相関係数

	JE	EJ		
		LE	LJ	TJ
MAD3(TOTAL)	0.625	0.848	0.795	0.770
MAD3(LISTEN)	0.634	0.891	0.850	0.801
MAD4(TOTAL)	0.632	0.868	0.742	0.652
MAD4(LISTEN)	0.676	0.889	0.789	0.712

表 10: MAD3 と MAD4 の BLEU 相関係数

	JE	EJ		
		LE	LJ	TJ
MAD3(TOTAL)	0.720	-	0.778	0.745
MAD3(LISTEN)	0.732	-	0.832	0.758
MAD4(TOTAL)	0.664	-	0.682	0.659
MAD4(LISTEN)	0.705	-	0.741	0.716

結果を表 12 に示す。相関係数が高いため、前章の結果よりこちらの方が信頼度の高い値と言える。

また、EJ_LE に関して MAD3 と MAD4 を比較したものを図 33 に示す。ただし、この図では MAD4 において著しくスコアの悪かった 2 つのデータを除いている。この 2 つの回帰直線を見比べると、LISTEN スコアの低い人の方が、高い人よりも MAD3 と MAD4 の差が大きい。これは、LISTEN スコアの高い人は、少し発話が難しくても聞き取れていることを表している。

7 まとめと今後の課題

松吉らとは異なるテストセットに対しても、日英方向の翻訳評価法は英日方向の翻訳評価に应用できることを確かめられた。また、今回の MAD3、MAD4 に対しては、TOEIC の TOTAL のスコアよりも LISTEN のスコアの方が相関が高いことが分かった。

MAD3 と MAD4 の機械翻訳の比較に関しては、話者に教示を与えることで機械翻訳の精度が上がるということが分かった。

英日方向の翻訳評価は人手による一対比較が行われていないため、今回の自動評価結果が妥当なものであるかどうかは分からない。実際に一対比較による評価を行ない、今回の結果の妥当性を検討することが今後の課題である。

謝辞

本研究を行なう機会を与えてくださいました 音声言語コミュニケーション研究所 菊井玄一郎 室長、京都大学 情報学研究科 佐藤理史 先生に感謝いたします。また、助言や討論をいただきました古澤一巳氏、林輝昭氏、安田圭志氏、持橋大地氏、その他研究所のみなさまに感謝いたします。

本研究は独立行政法人 情報通信研究機構の研究委託「大規模コーパスベース音声対話翻訳技術の研究開発」により実施したものである。

参考文献

- [1] 松吉俊、竹澤寿幸: 英日方向の音声翻訳能力評価実験 (TR-SLT-0051), ATR Technical Report(2003).

表 11: 被験者の TOEIC score 相関係数

TOTAL-LISTEN	TOTAL-WRITE	LISTEN-WRITE
0.954	0.795	0.836

表 12: MAD3 と MAD4 の機械翻訳 LISTEN score からの機械翻訳 TOEIC score

	JE	EJ		
		LE	LJ	TJ
MAD3(DP)	381	-	360	-88
MAD3(BLEU)	719	-	384	21
MAD4(DP)	708	-	379	167
MAD4(BLEU)	810	-	292	173

- [2] 菅谷史昭、竹澤寿幸、横尾昭男、山本誠一: 音声翻訳システムと人間との比較による音声翻訳能力評価手法の提案と比較実験, 電子情報通信学会論文誌, Vol.J84-D-II, No.11, pp.2362-2370(2001).
- [3] 竹澤寿幸、西野敦士、高寫浩司、松井孝典、菊井玄一郎: 機械翻訳を介した対話データ収集のための実験システム, FIT2003 一般講演論文集第2分冊, pp.161-162(2003).
- [4] 竹澤寿幸、菊井玄一郎、鈴木弥生、西野敦士: コーパスベース音声翻訳研究のための対話データ収集、第45回音声言語情報処理研究会, pp.71-76(2003).
- [5] 竹澤寿幸、菊井玄一郎: 翻訳システムを介した音声対話における教示と言語表現の関係, 情報処理学会研究報告音声言語情報処理, No.50-13, pp.81-86(2004).

A グラフ

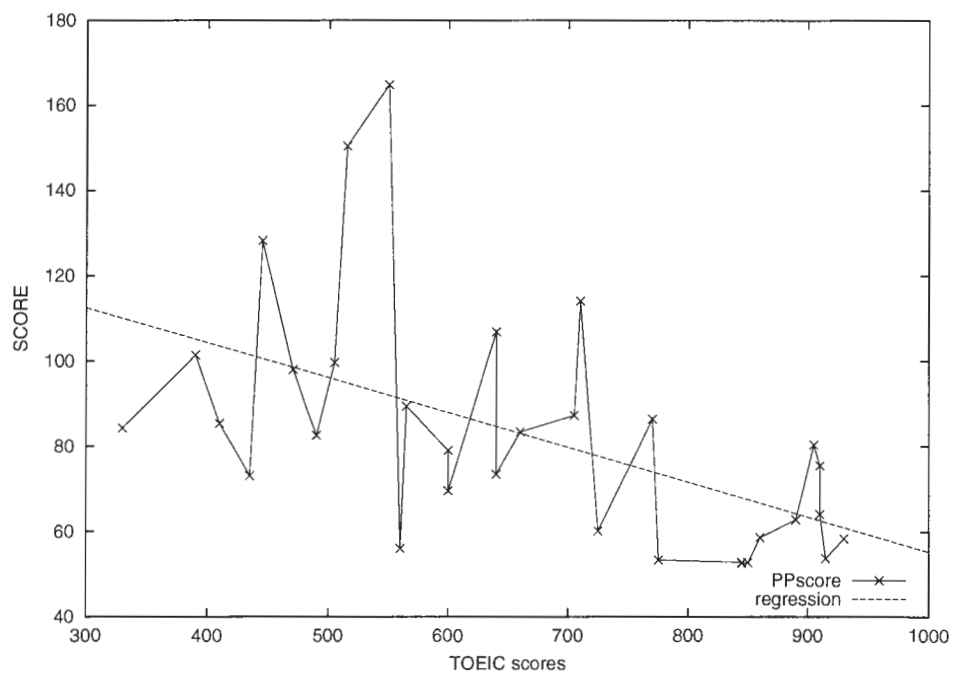
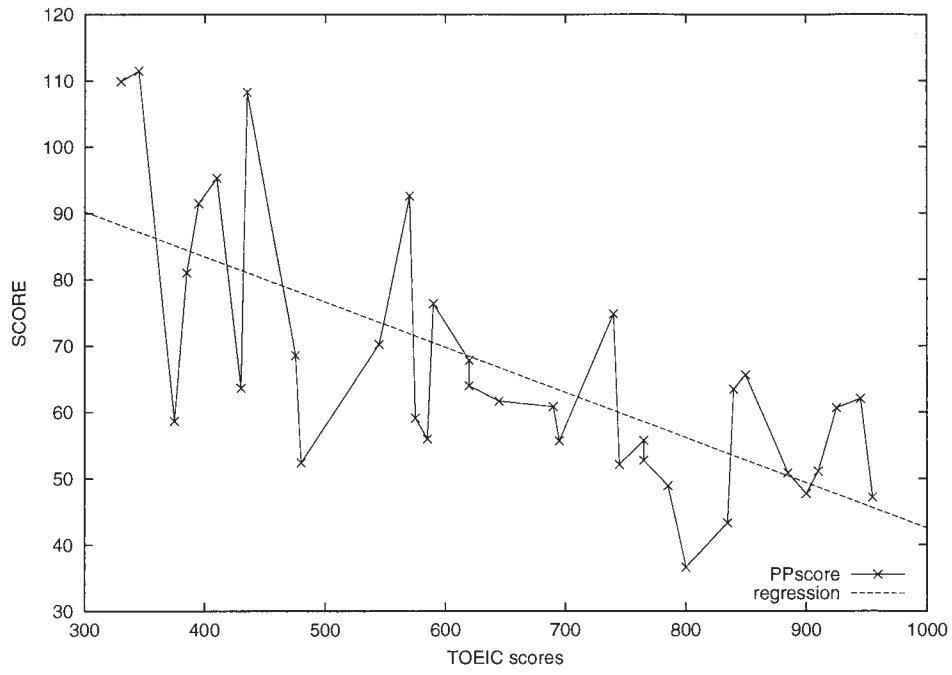
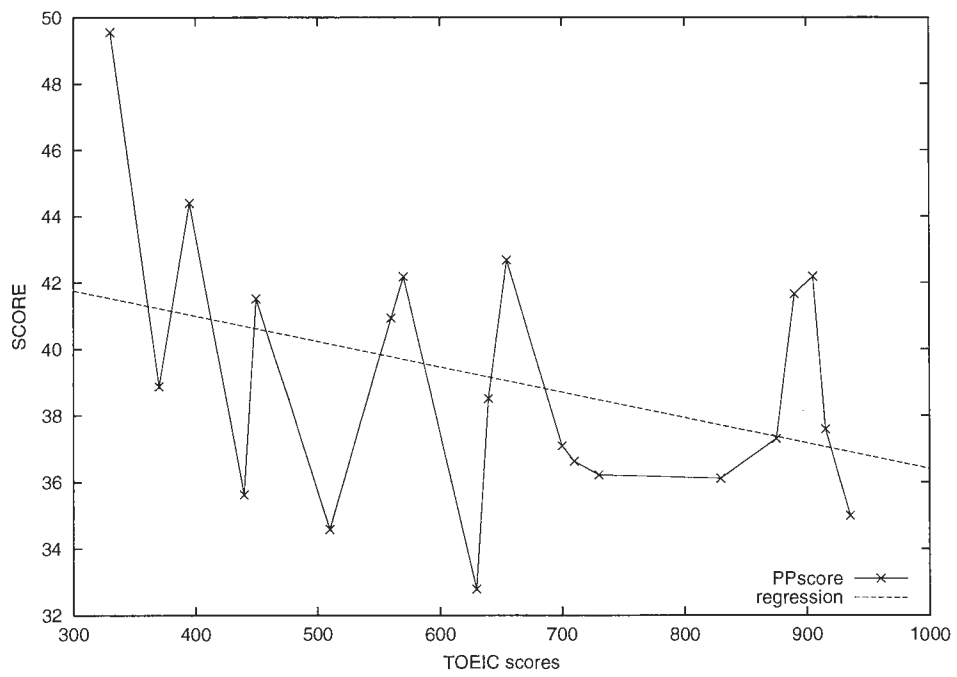


図 1: MAD3 JE PP



☒ 2: MAD4 JE PP



☒ 3: MAD3 EJ LJ PP

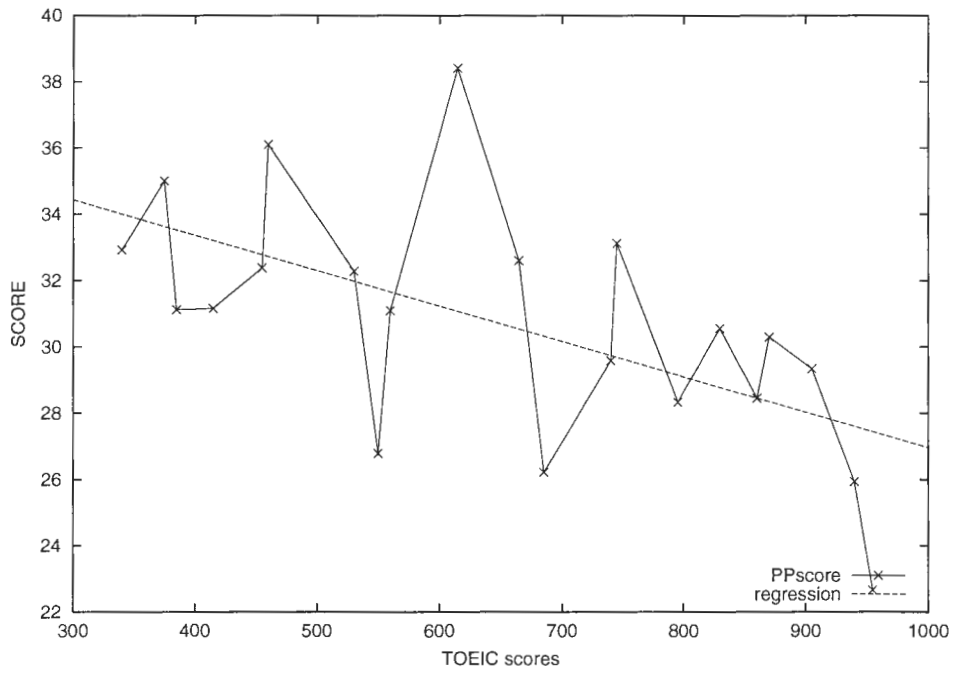


图 4: MAD4 EJ LJ PP

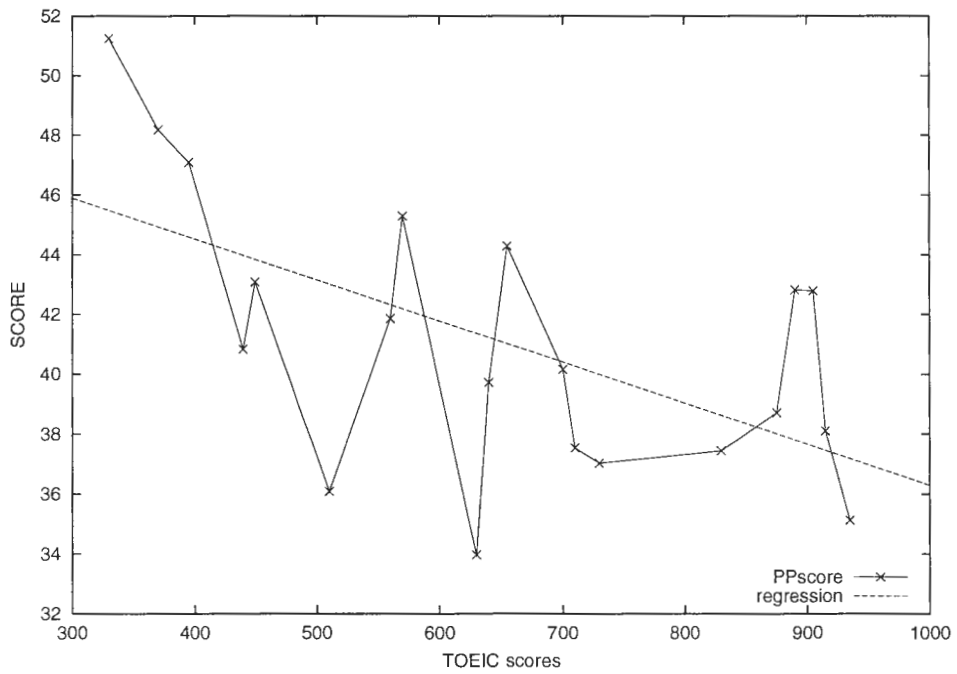


图 5: MAD3 EJ TJ PP

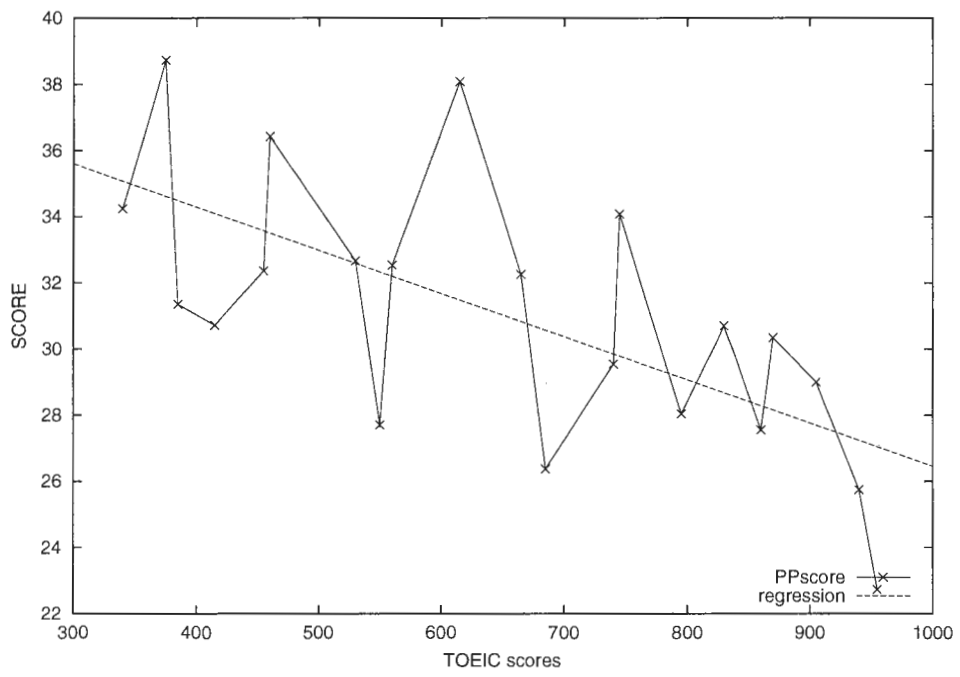
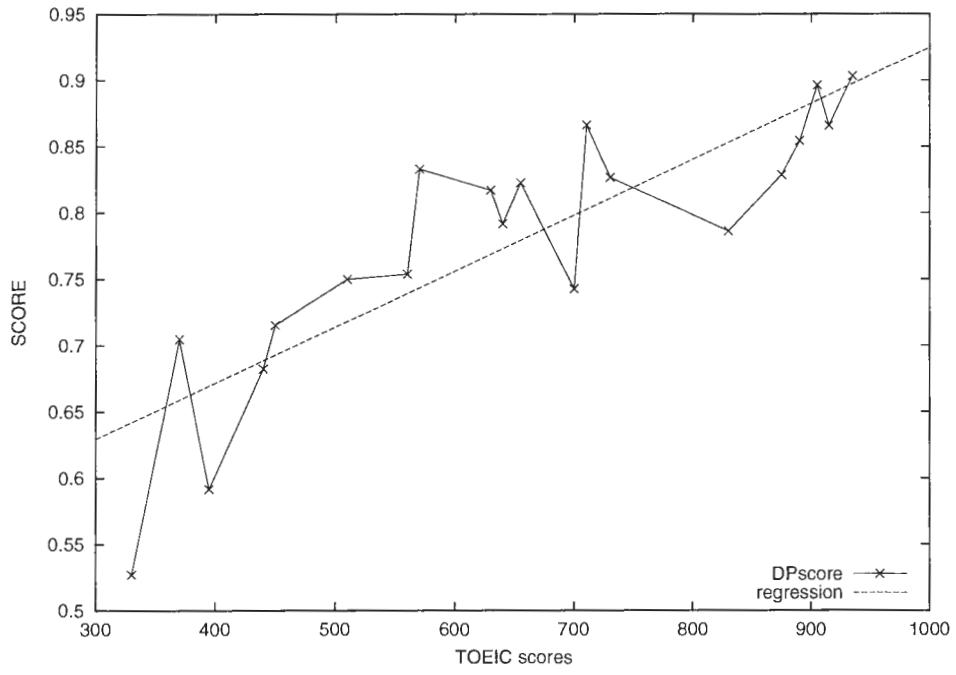
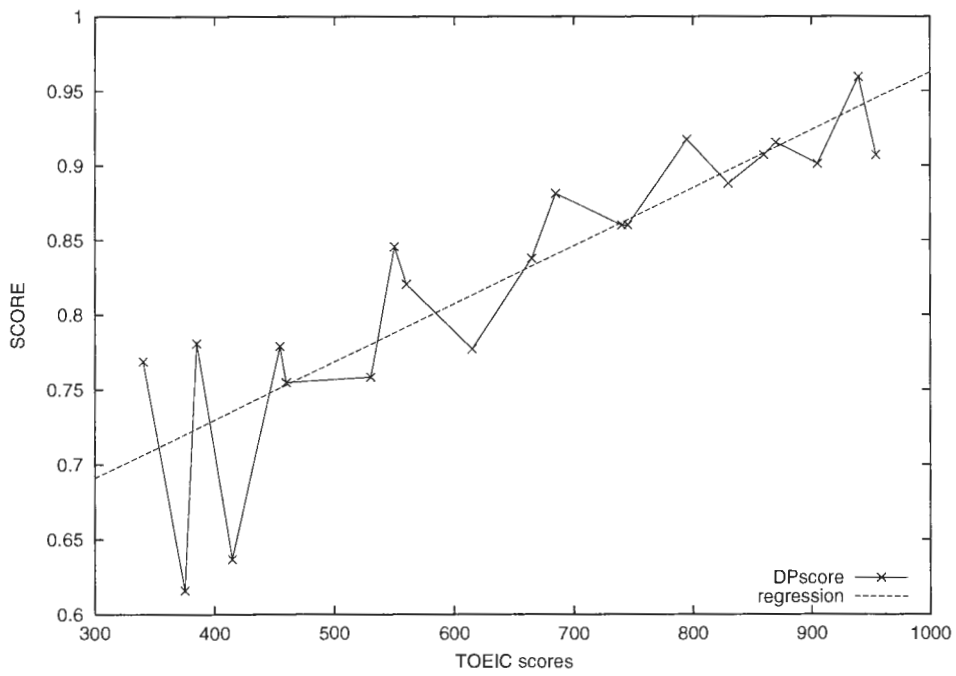


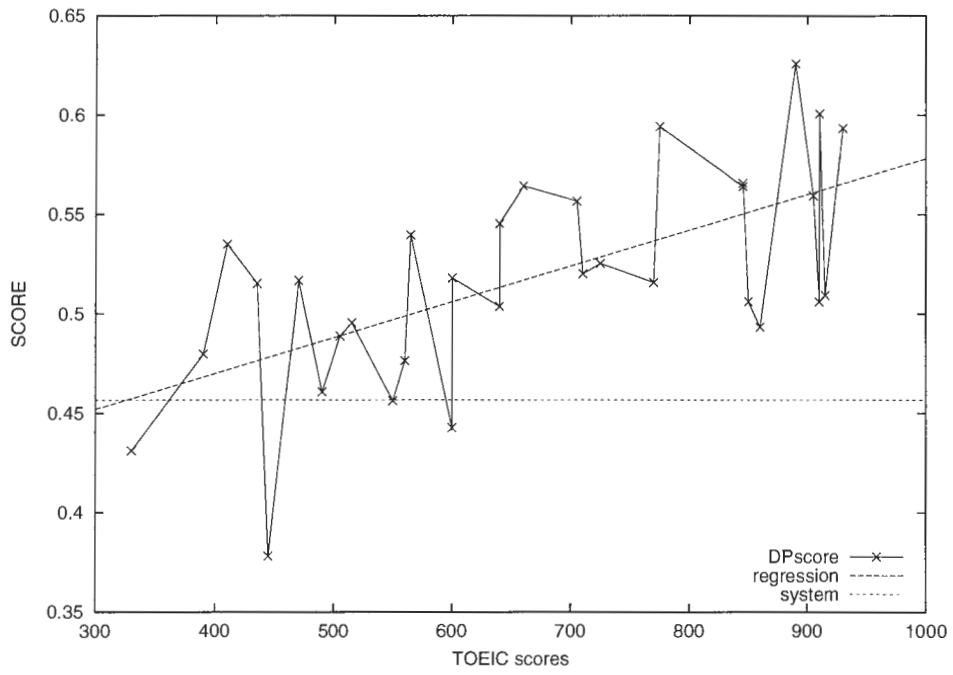
图 6: MAD4 EJ TJ PP



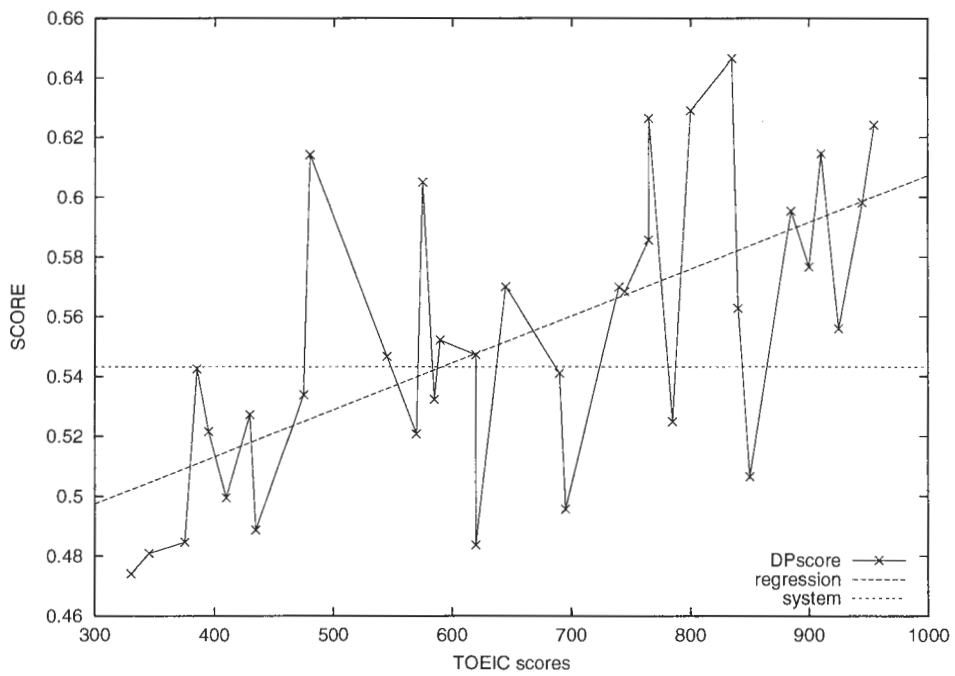
☒ 7: MAD3 EJ LE DP



☒ 8: MAD4 EJ LE DP



☒ 9: MAD3 JE DP



☒ 10: MAD4 JE DP

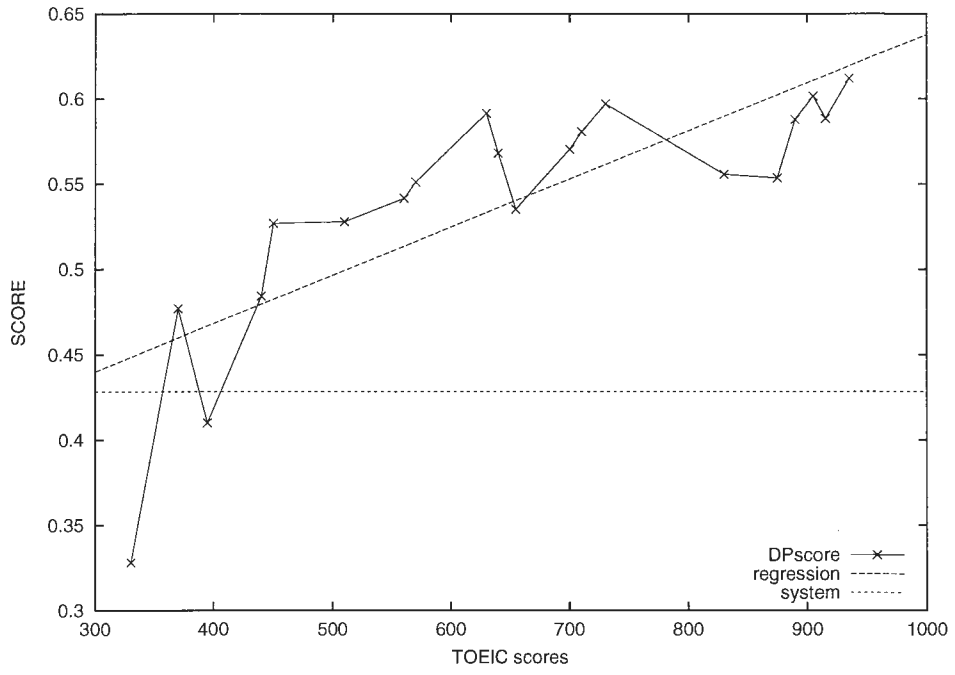


图 11: MAD3 EJ LJ DP

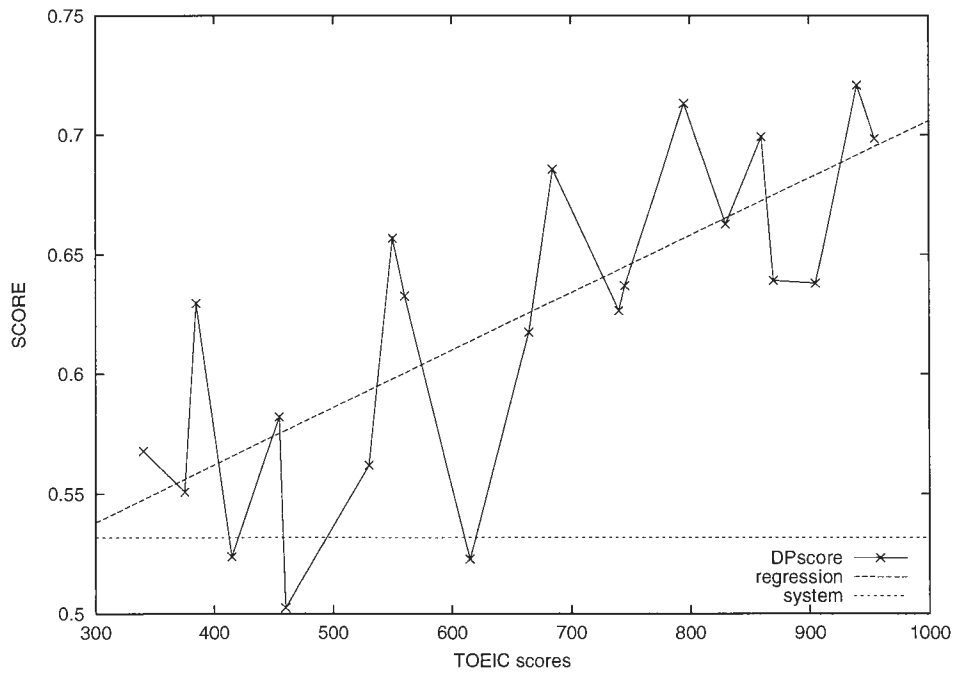


图 12: MAD4 EJ LJ DP

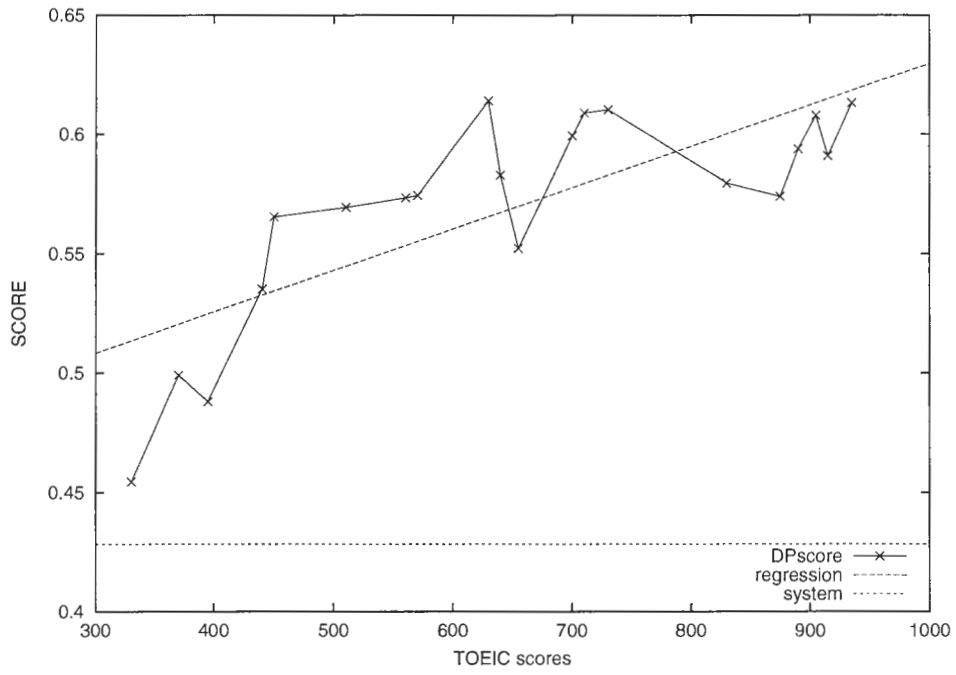


图 13: MAD3 EJ TJ DP

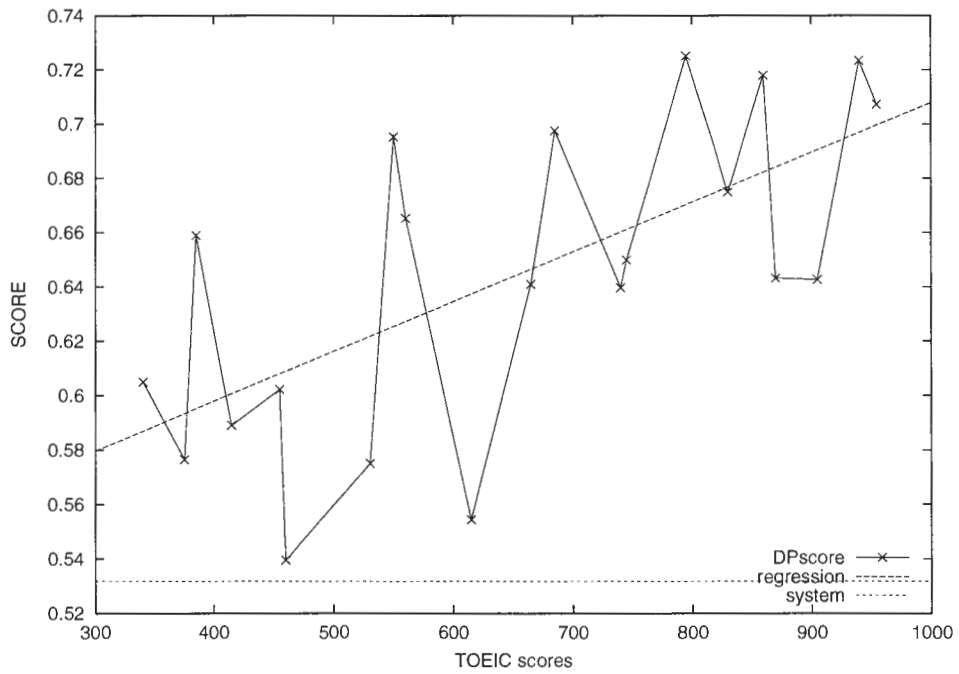


图 14: MAD4 EJ TJ DP

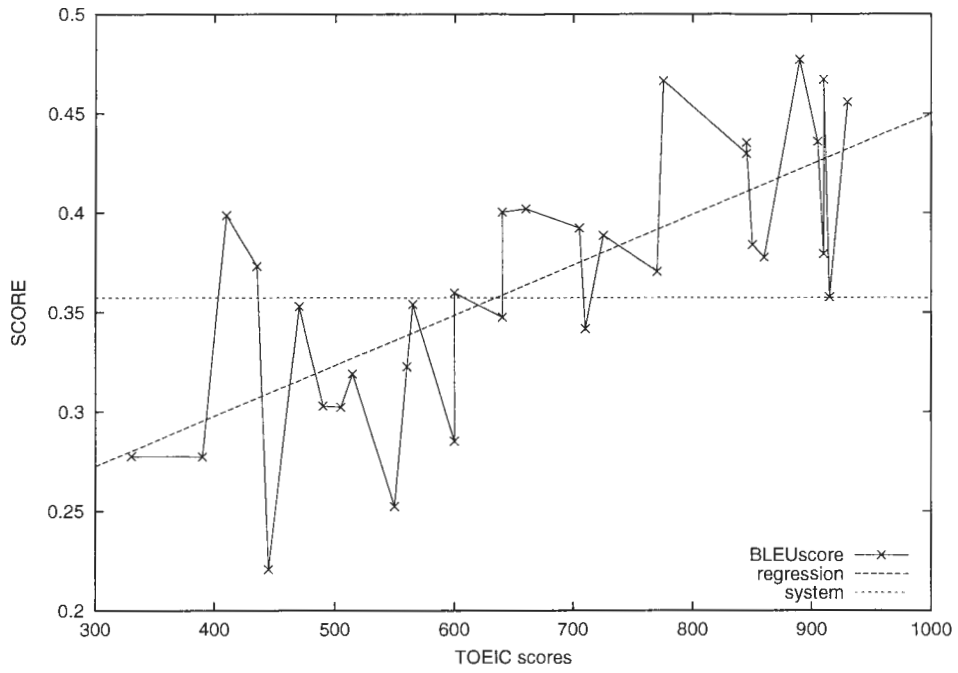


图 15: MAD3 JE BLEU

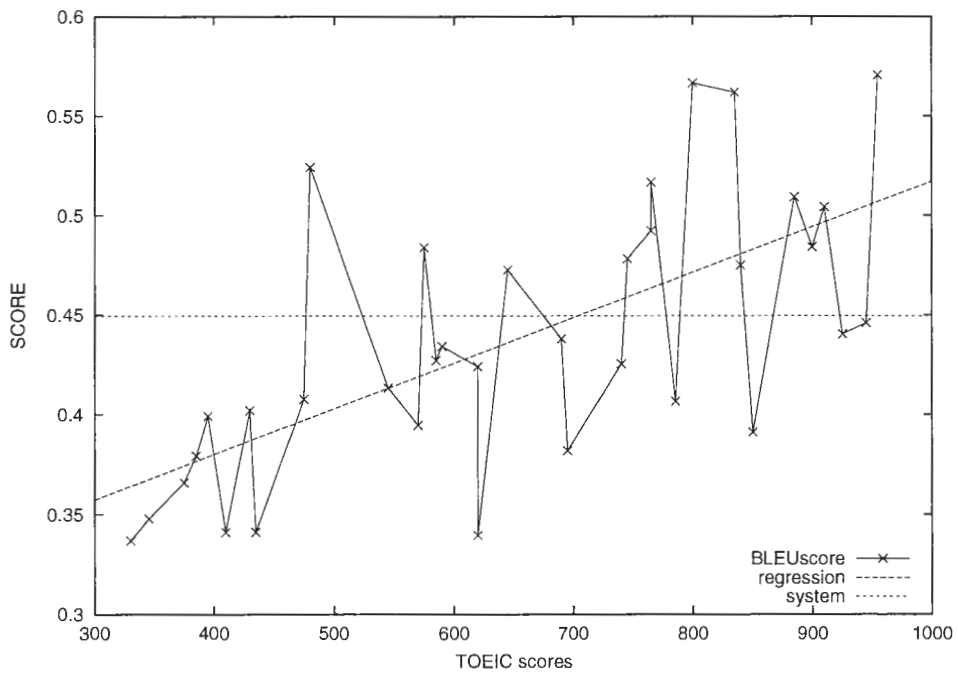
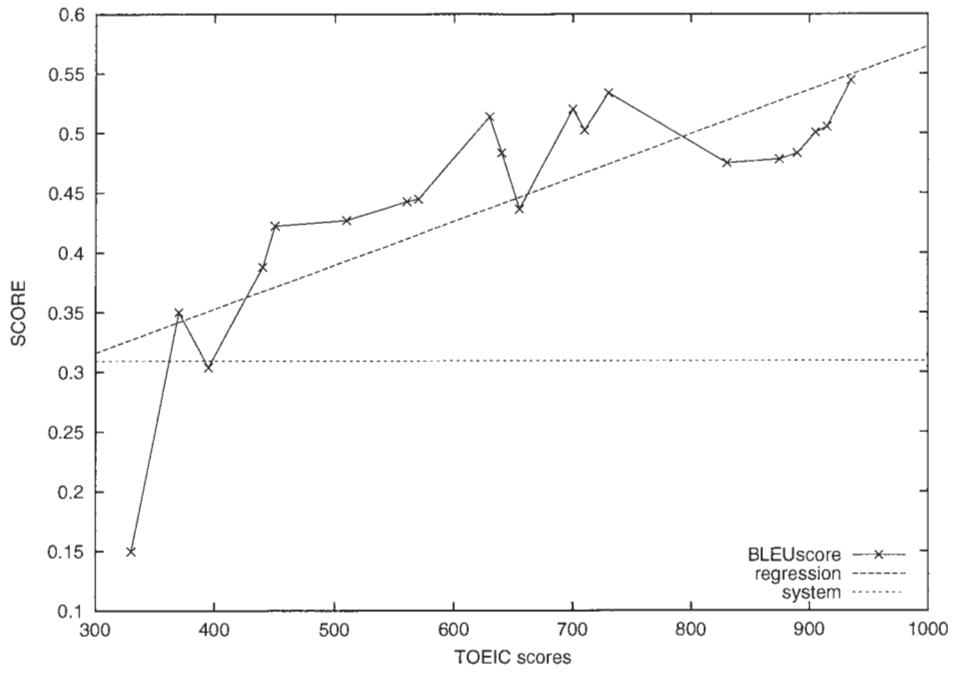
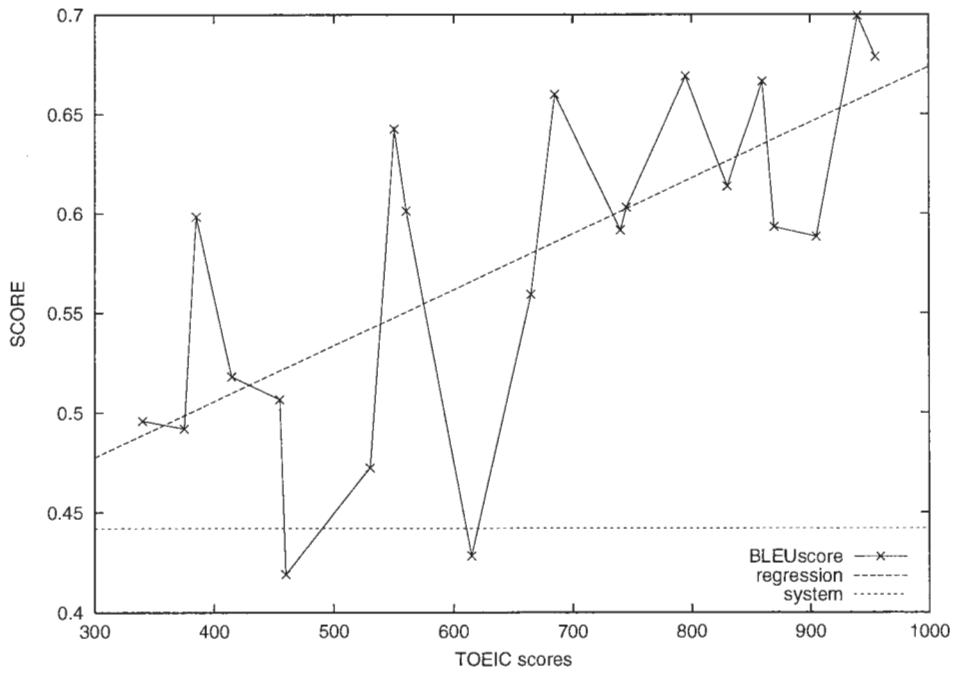


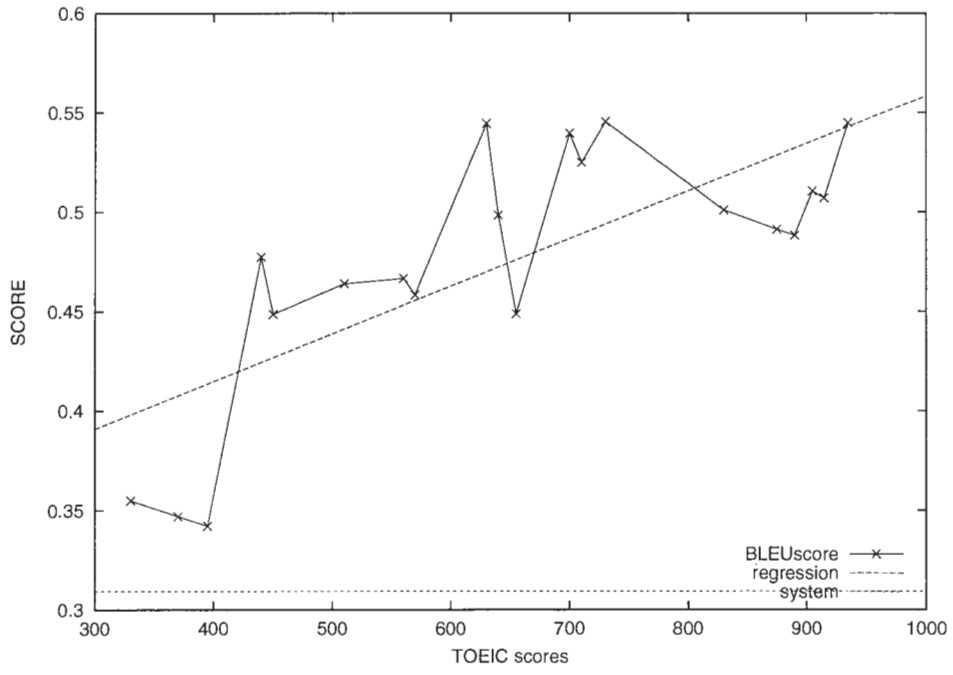
图 16: MAD4 JE BLEU



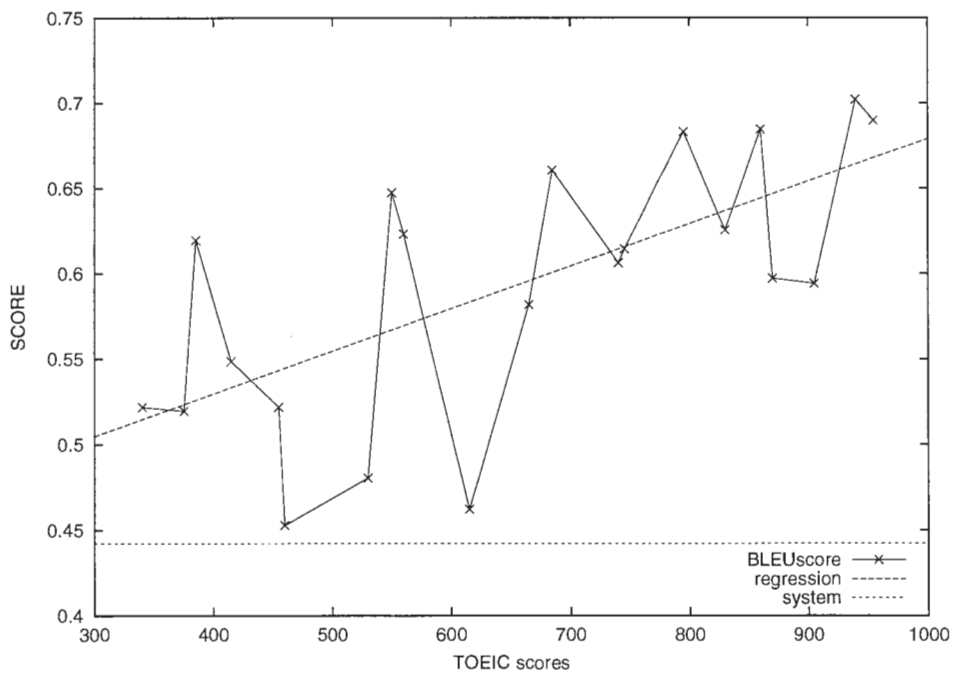
☒ 17: MAD3 EJ LJ BLEU



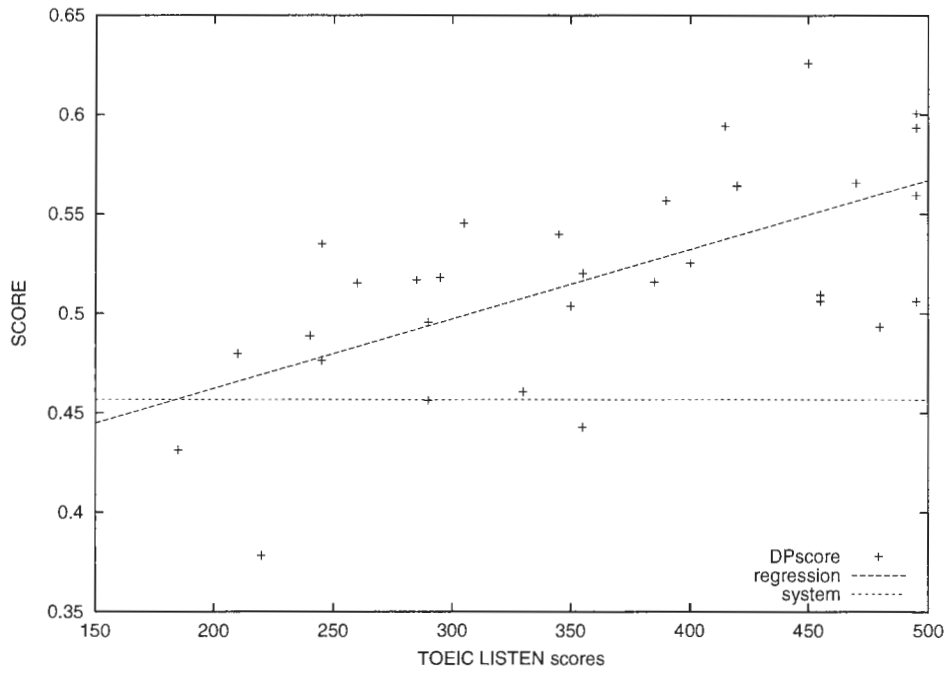
☒ 18: MAD4 EJ LJ BLEU



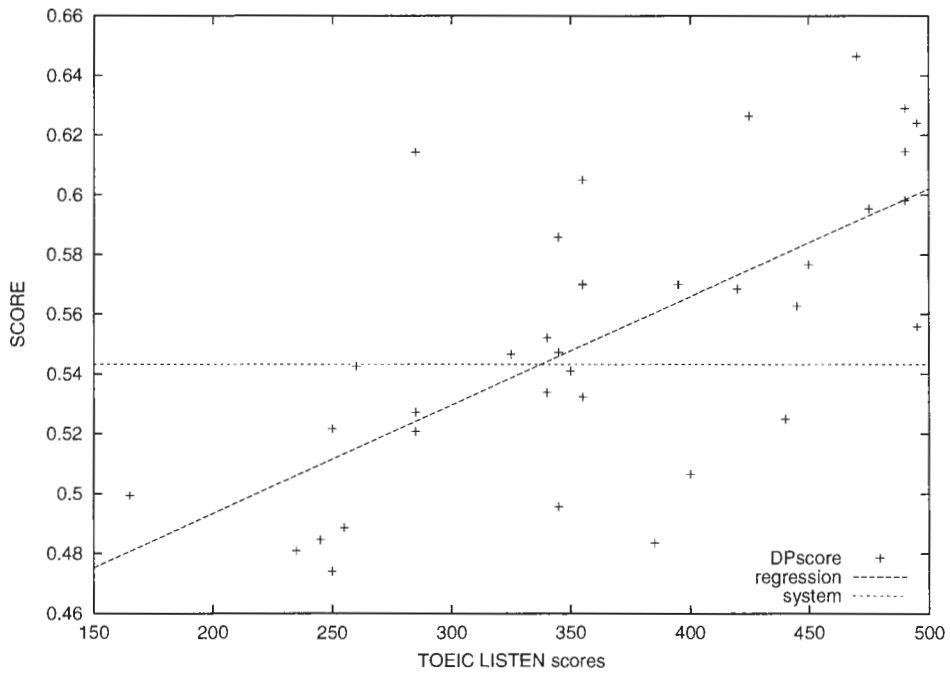
☒ 19: MAD3 EJ TJ BLEU



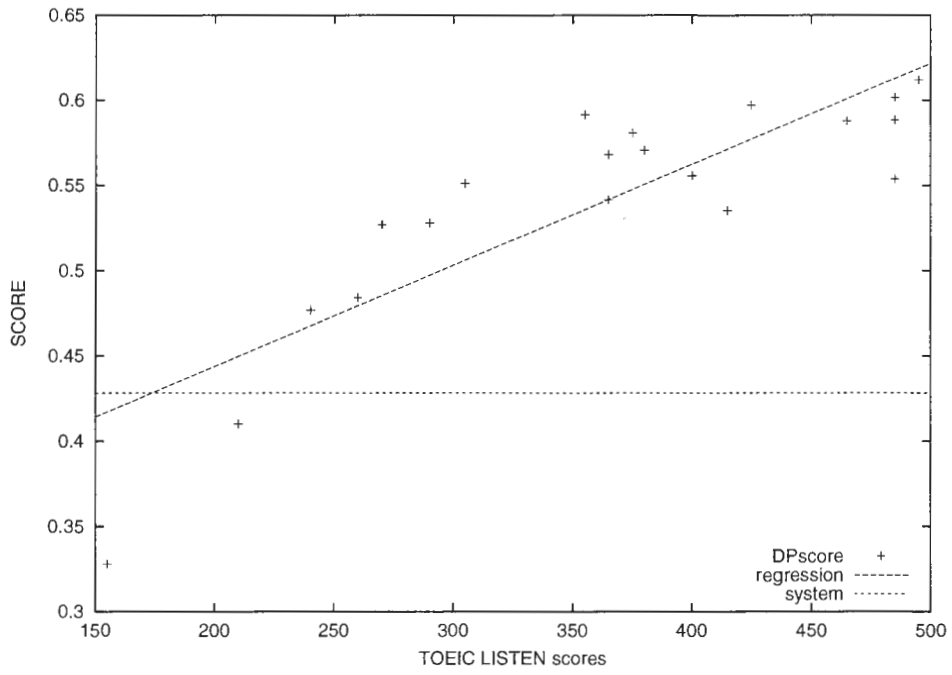
☒ 20: MAD4 EJ TJ BLEU



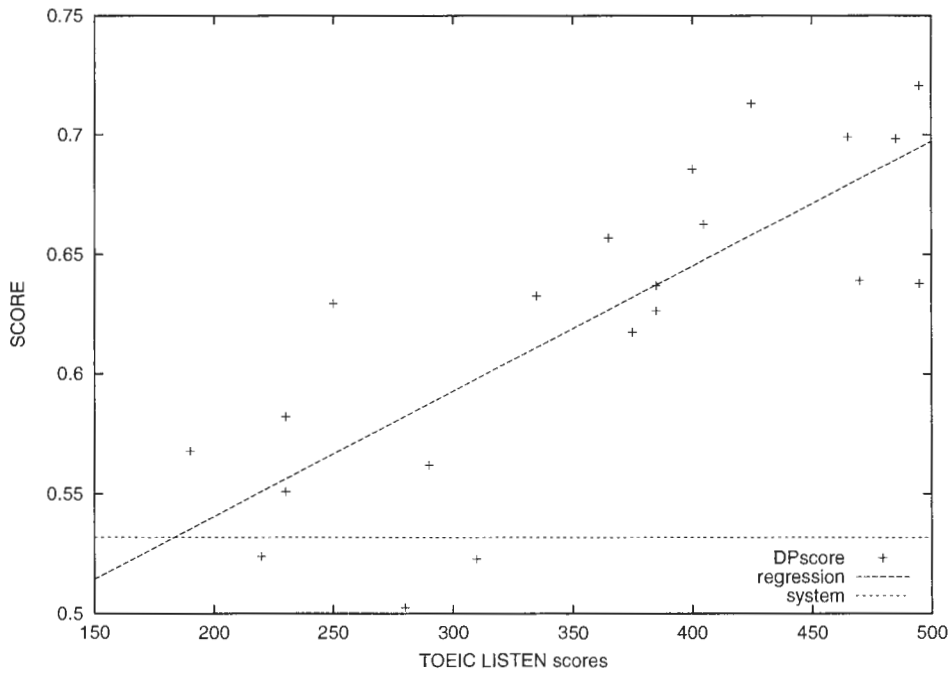
☒ 21: MAD3 JE DP LISTEN



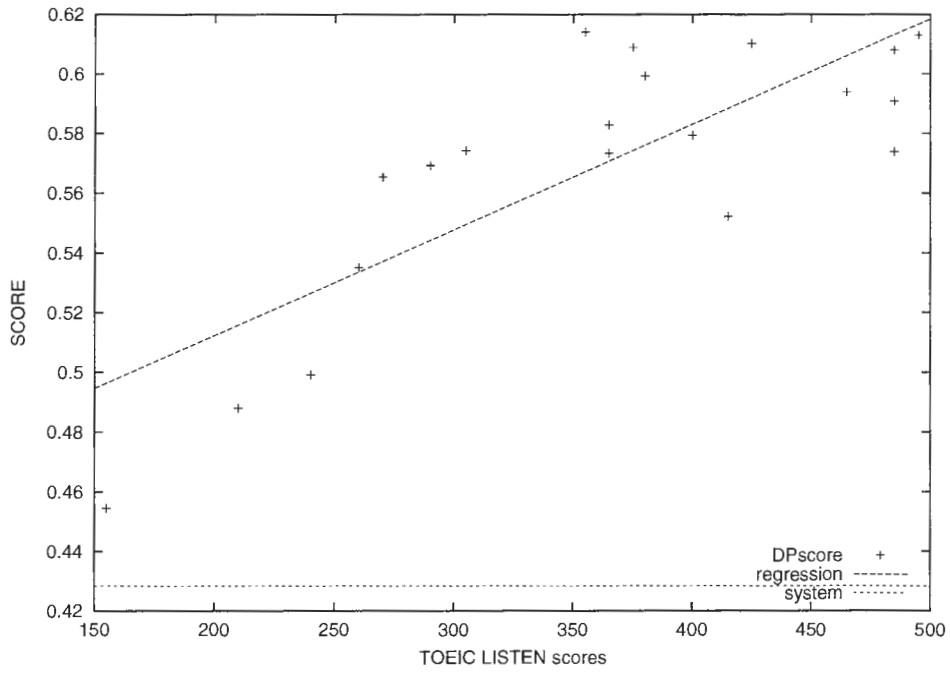
☒ 22: MAD4 JE DP LISTEN



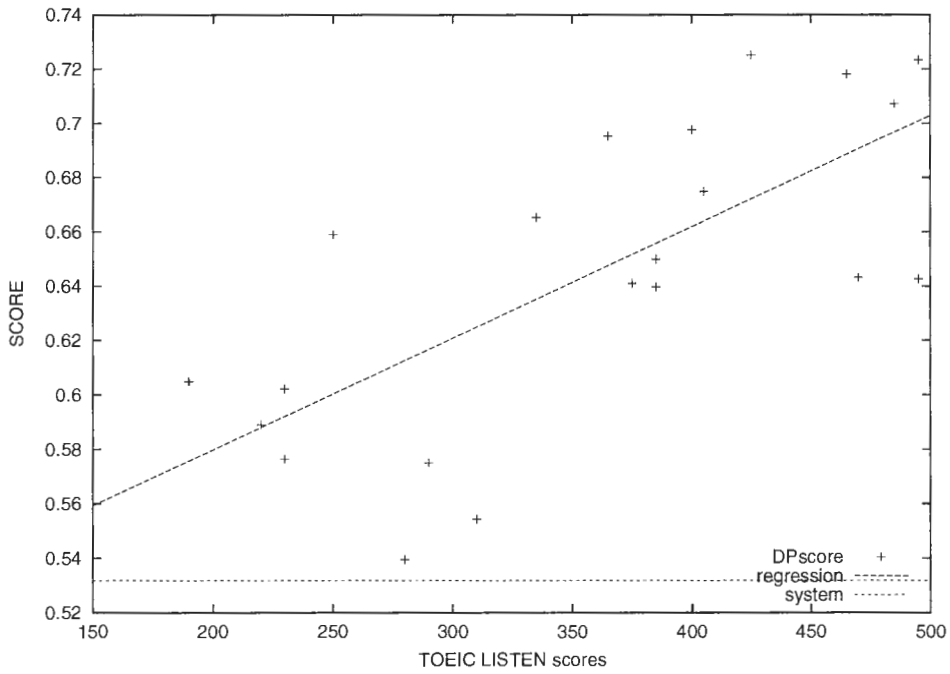
☒ 23: MAD3 EJ LJ DP LISTEN



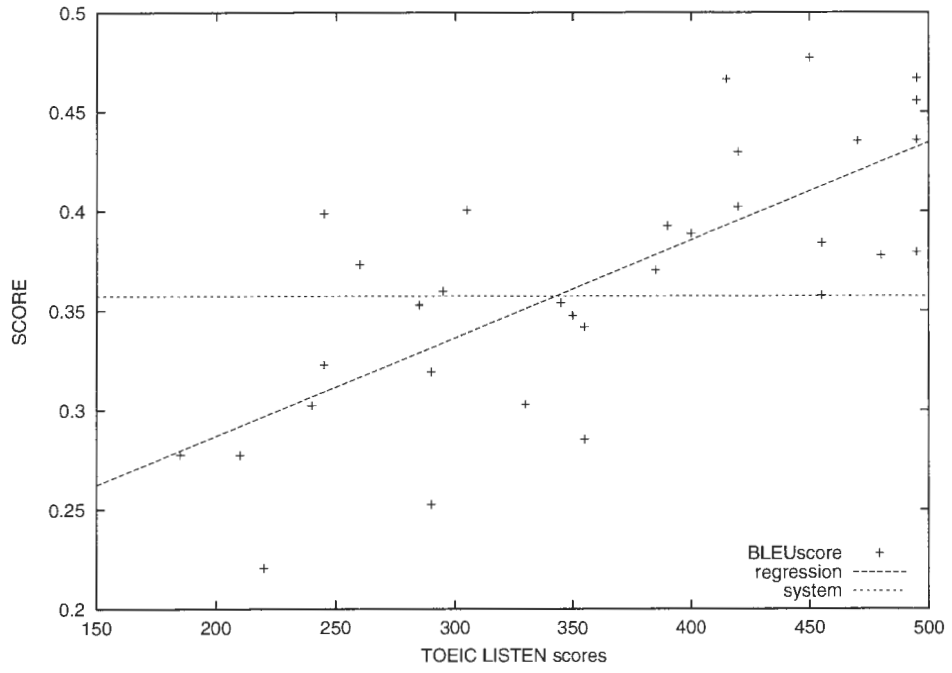
☒ 24: MAD4 EJ LJ DP LISTEN



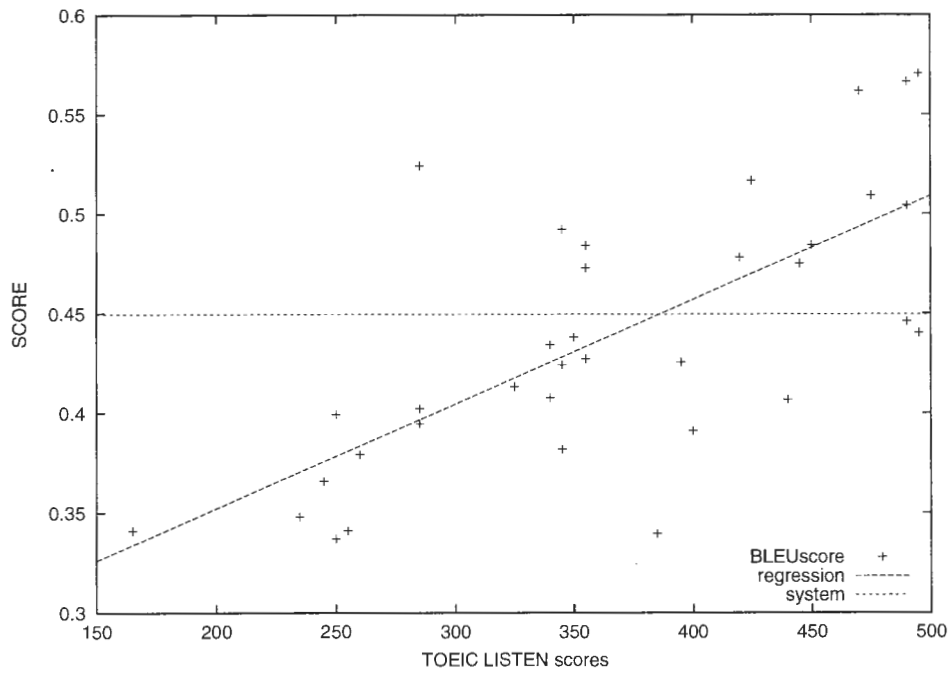
☒ 25: MAD3 EJ TJ DP LISTEN



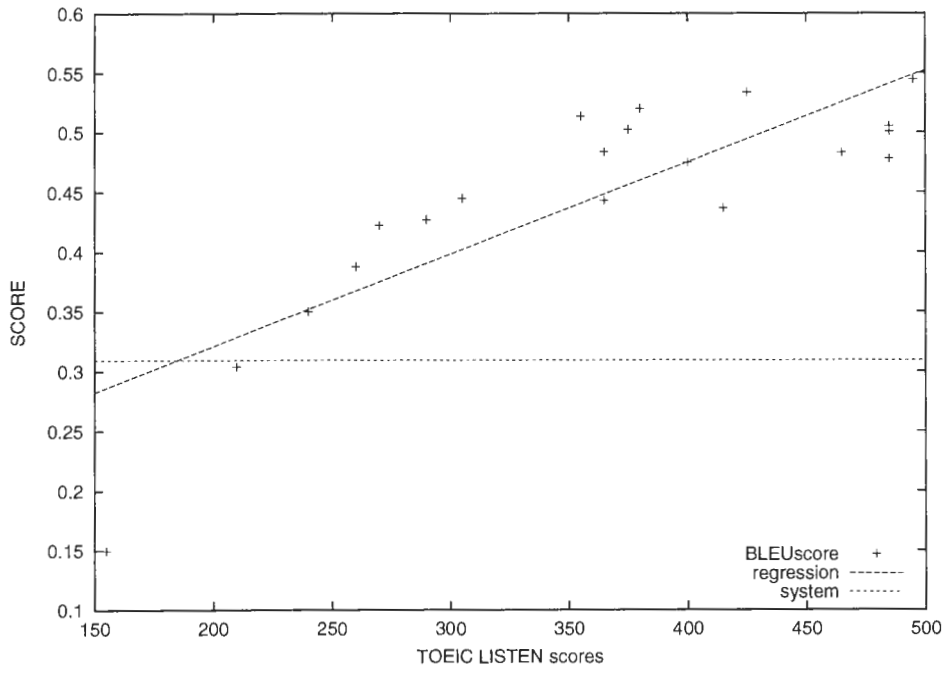
☒ 26: MAD4 EJ TJ DP LISTEN



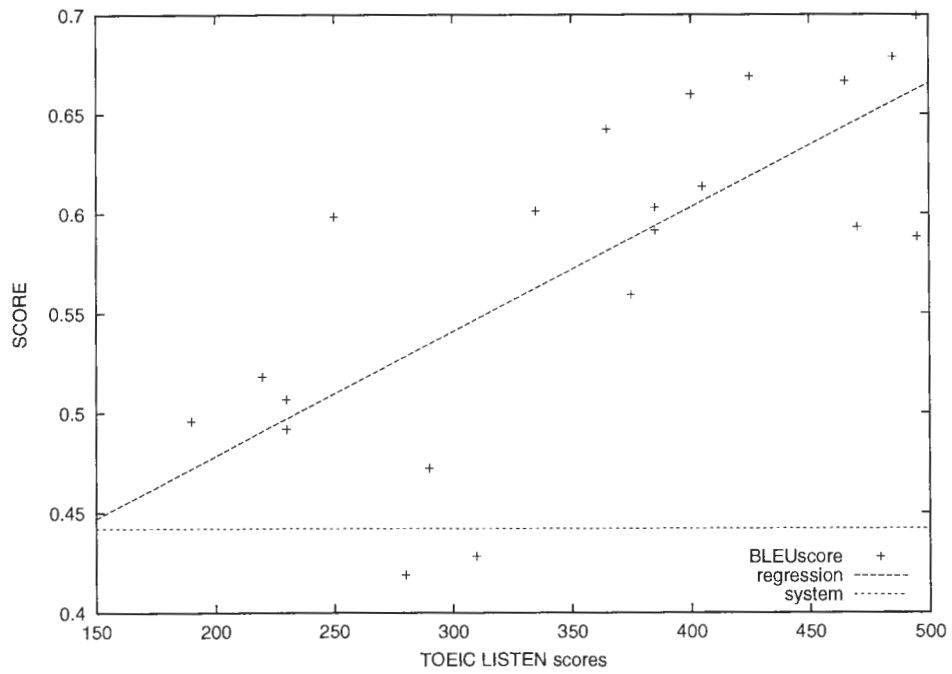
☒ 27: MAD3 JE BLEU LISTEN



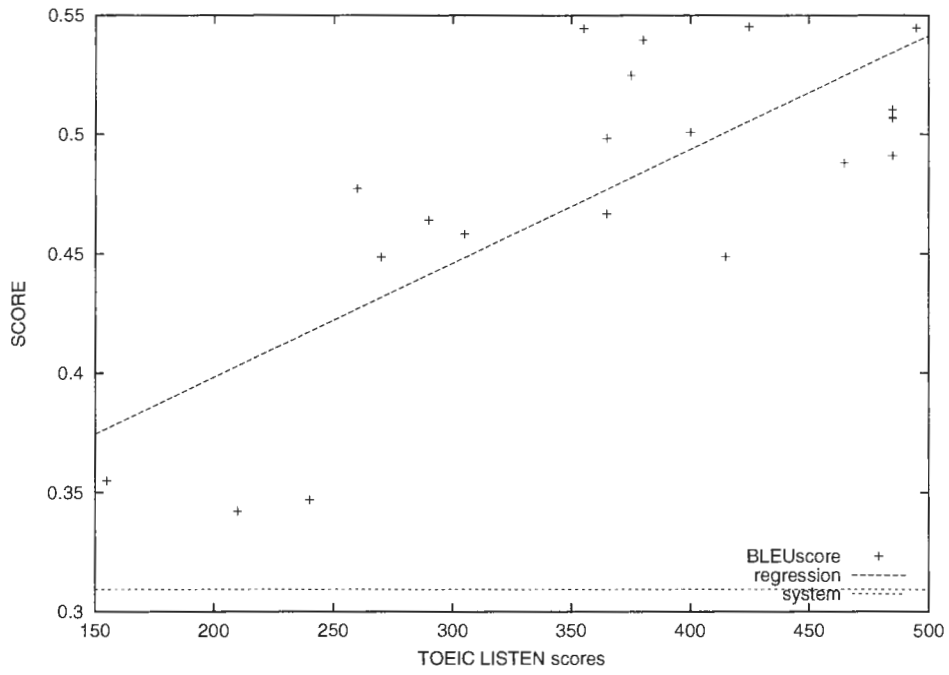
☒ 28: MAD4 JE BLEU LISTEN



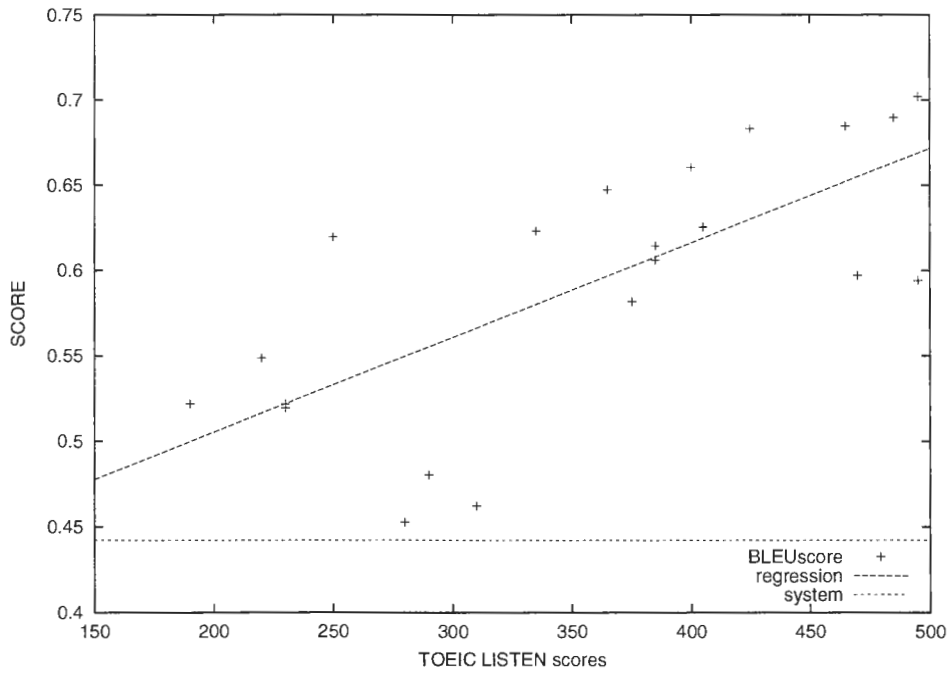
☒ 29: MAD3 EJ LJ BLEU LISTEN



☒ 30: MAD4 EJ LJ BLEU LISTEN



☒ 31: MAD3 EJ TJ BLEU LISTEN



☒ 32: MAD4 EJ TJ BLEU LISTEN

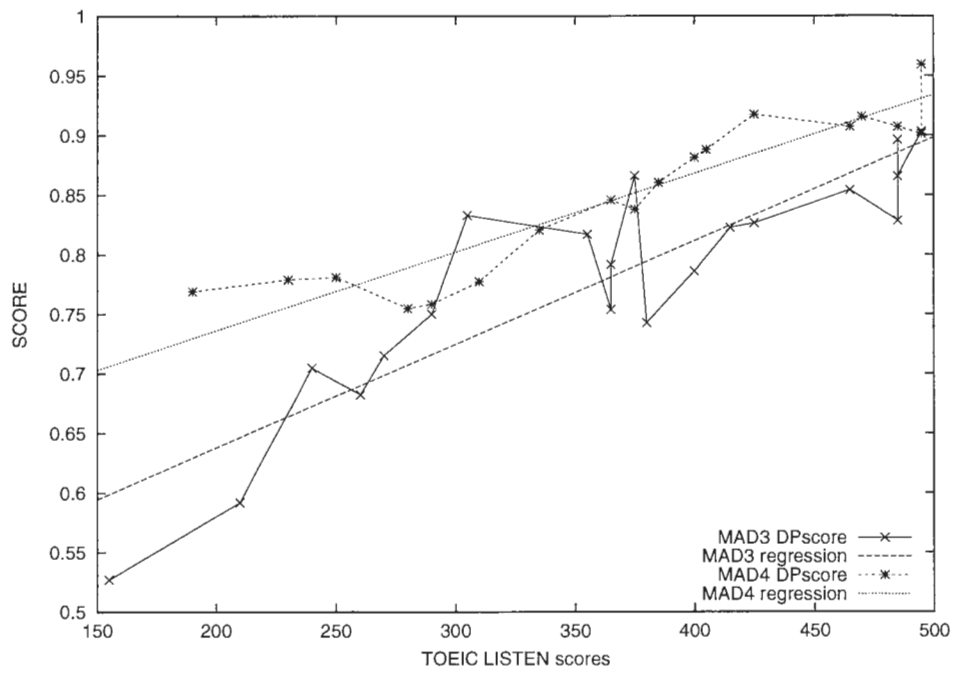


図 33: MAD34 EJ LE DP LISTEN