

<論文>

## 誰がポストエディターになるのか？

Who will be post-editors?

山田 優

(大学非常勤講師)

### Abstract

Language technologies such as machine translation plus post-editing by human translators (MT+PE) are increasingly prevalent in the translation industry, and studies have investigated MT+PE about its greater efficiency, time savings, and higher quality. However, the one of the major issues that the translation industry (in Japan) faces is as to who will be the post-editors. Most professional translators who are currently actively working are under pressure of cost reduction; hence feel negative towards MT+PE, and even skeptical about the quality of MT raw output. This is in line an empirical research by Yamada (2011c) showing that the professional translators did not feel it advantageous to switch their production style to its MT post-editing mode.

This pilot study will investigate potentials of college language learners and novice translators to become post-editors. The investigation focuses on changes and improvement in participants' attitude/preference toward MT+PE. The study uses both questionnaires and experiments. The tentative results will be presented in this paper.

### 1. はじめに

近年の産業翻訳における翻訳関連テクノロジーの発展に伴い、従来のような人間の翻訳者が一から翻訳する手法 (human translation = HT、以下「人手翻訳」) に代わって、機械翻訳が起点テキストを下訳したあとに人間が修正を行う手法 (Machine Translation plus Post-editing = MT+PE、以下「ポストエディット」) が、産業翻訳界において使われるようになってきている。日本においては、効率性、品質などを含め、まだポストエディットの効用や有効性を問う声も多い。しかし、産業翻訳の現場では、ポストエディットの需要が増えており、すでに議論は、誰がポストエディットを行うべきかに焦点が合わせられている。というのも、ポストエディットの仕事は翻訳者への支払いが通常よりも低い、既存の翻訳者が消極的、翻訳会社としても翻訳単価を下げて発注しづらいことから、ポストエディターの不足が危惧されているためである。また、一般的に、ポストエディットは、人手翻訳よりも容易な作業であると考えられており、プロの翻訳者がやるべきではないという

意見もある。

そこで翻訳未経験者ないし翻訳の初心者が、ポストエディットを代行する案が考えられる。しかし、この場合も、最終的な翻訳品質が実務で耐えうるかという問題が残る。ただ最近では、ポストエディットや機械翻訳の使用が業界内に流布したことにより、翻訳ユーザの側でも「Fit for use」（目的に応じた翻訳の活用）という考え方が浸透し始めているのも事実だ<sup>1</sup>。結果的に、ポストエディットに対する品質期待値（翻訳の期待規範）は、人手翻訳のそれとは差別化されている。

このような背景全般を考慮した上で、本研究では、次の問を設定した。「学生はポストエディターになれるのか」。この仮説を検証するため、本稿<sup>2</sup>では、翻訳コースの学生のポストエディットに対する態度、品質、適正を調査し、その暫定結果を報告する。尚、現時点で、調査はパイロットレベルであり、今後の本調査に向けてデータを継続的に精査させていく予定であることを、はじめに断っておく。

## 2. 先行研究および議論の焦点

ポストエディットとは、機械翻訳の結果を人間の翻訳者や編集者が修正を行うこと（Veale and Way, 1997）、機械が生成した翻訳結果を、最小限の（人間の）手作業によって向上させるプロセス（TAUS report, 2010）、のように定義される。これ以外にも、翻訳品質要求ごとに、「フル・ポストエディット（full post-edit）」（文体・スタイルまで含めた高品質レベルまで修正する手法）、「ライト・ポストエディット（lite post-edit または rapid post-edit）」（意味の通じるレベルに修正する手法）、のように定義されることもある（Allen, 2003）。いずれも、機械翻訳の結果を人間が修正するという点においては同じだ。

ポストエディットが、従来の人手翻訳における修正（revising）や校正（editing）とどのように違うのかという点については、Mossoop（2001）の研究との比較から、ポストエディットでの修正箇所やエラーは人手翻訳とは種類が異なることが分かっている（McElhaney & Vasconcellos, 1986; Löffler-Laurian, 1985, p. 71）。また作業者の認知側面からも、Kring（2001, p. 360）らが示唆するように、翻訳者は心的に、複雑かつ非線状なプロセスで進行する人手翻訳と、ポストエディットは別物として受けとめられているようだ（山田, 2011b も参照）。

他方、ポストエディットを導入することによる翻訳作業の効率性の調査は、すでに数多くの先行研究が存在し、言語の組合せや作業者の経験により程度は異なるものの翻訳作業の効率が向上するのはほぼ事実として間違いない（山田, 2010 参照）。実際に翻訳実務や実践でポストエディットを活用するために、TAUS のガイドラインによれば（TAUS, 2010）、機械翻訳をカスタマイズするなどを奨励しているが、現状の汎用エンジン（Google 翻訳など）をそのまま使った場合でも、ポストエディットの効率性向上の効果が認められている。

大学の学生・大学院生を対象にするポストエディット作業の認知負荷と受容性に関しては、英語から日本語の翻訳において、人手翻訳と比べて約 20% 程度作業が楽になる（山田

2011a)。実務経験のあるプロの翻訳者と比較すると、学生翻訳者は機械翻訳結果の修正量が少なく、ポストエディットに対する受容性が高いように見受けられる（山田, 2011c)。Garcia (2010) の調査でも、Google 翻訳者ツールキットを使った実験で、9 割の学生が人手翻訳よりもポストエディットの方を好む (preferred) と回答しており、ポストエディットに前向きな態度が示されている。これらの調査結果から、学生 (翻訳未経験者) は、ポストエディターに向いていると考えることもできるのだ。

### 3. 目的と仮説

本研究では、「学生はポストエディターになれるのか」という問を検証する。そのために、学生のポストエディットへの態度、品質、適正を調査した。実際に測定したのは、以下の4つの指標である。

- A. 体感作業負荷 (Ease of PE)
- B. 機械翻訳の訳と最終翻訳物のテキスト類似度 (修正量)
- C. 翻訳品質
- D. クラス内の成績

A の体感作業負荷は、ALPAC 報告書 (1966)、山田 (2011b, 2011c) で用いた方法を踏襲した。調査対象者にアンケートを実施して、ポストエディットの体感作業負荷を数値化した。人手翻訳の体感負荷を 100 (基準) として、ポストエディットの体感作業負荷はどうかの回答してもらった。例えば、ポストエディットが人手翻訳よりも 2 割程度の負荷軽減を感じたならば 80 と記入 (100 マイナス 20)、逆に 2 割増えたと感じたら 120 (10 プラス 20) という具合である。機械翻訳を下訳 (ベース) として使って翻訳作業ができたのであれば、作業負荷は低減するだろうと予測されるので、この指標により、学生がどのくらいポストエディットを「使える」と感じているかを調べることができる。すなわち、学生のポストエディットに対する態度 (好み) の参考値になる。

B の機械翻訳の訳と最終翻訳物のテキスト類似度 (修正量) も、山田 (2011a; 2011b) で用いた手法だ。機械翻訳の訳と最終翻訳物のテキスト類似度を GTM で測定した。この指標により、学生が機械翻訳の訳をどのくらい修正する必要があったのかを知ることができる。修正量が少なければ、機械翻訳の訳をベースとして作業できた (機械翻訳が役に立った) ということの補強証拠にもなる。この指標 B 自体が学生のポストエディターへの適性を直接的に示すものではないが、修正量は上の A の指標の体感負荷の低減ともある程度関連し、また山田 (2011c) で見たように、プロで熟練の翻訳者は、どちらかと言えば機械翻訳の訳を使わないように訳出していく傾向が見られたことから、修正量を観察することで学生とプロとの違いを見ることができる。

C の翻訳品質は、対象テキスト (IT マニュアル、Firefox 手順書) の日本語版を参照文に

して、各翻訳者の目標テキストとの類似度を GTM 計測し、それから、筆者自身の実務翻訳の経験に基づくチェックにより、正確性、読みやすさ、文体について、減点法で評価した。読みやすさと文体は、評価基準が一定にならないことがあるため、優先順位としては正確性（誤訳など）よりは低めに設定した。例えば、「～してください」という機械翻訳の訳が「～します」というマニュアル文の規範的な文体に修正されていないとしても、前後の文章（一連の操作手順の中）で、一貫性が保たれている場合は減点していない。しかし読みづらさが著しく損なわれるような修正—例えば、機械翻訳の訳になかった「あなた」という言葉を、いちいち付加した翻訳者もいたのだが—そのような場合は、減点対象としている。原点幅は最大で 0.1 ポイントとした。

D の成績は、調査対象者の翻訳クラスの最終評価（成績）である。この成績は、各学生の基本翻訳スキルおよび翻訳という営為に対する態度が反映されている。すなわち、成績優秀な学生は、クラス内において、比較的優秀な翻訳者であったといえる。後述するように当該クラスは、ポストエディットに特化したものではないので、このクラスの成績と本調査結果を比較することで、どのような学生がポストエディターとしての適性が高いのかを見ることができる。尚、成績の記載は、便宜的に「優」および「劣」の 2 段階のみで記した。

#### 4. 調査対象者のプロフィール

本調査は、某 R 大学の学部生 20 名を対象に行った。外国語学部の英語を専攻している 3・4 年生である。対象 20 名のほぼすべての学生が筆者が担当する翻訳クラスを 2 学期間（全 15 週 x 前期・後期）受講した。

授業は、英日・日英の翻訳の基礎演習が中心であった。毎週、様々な分野のテキスト（Web サイト、ニュース記事、ビジネスレター、IT マニュアル等）の翻訳練習を行った。実務翻訳の実践を想定した講義・演習となっている。学生達は、翻訳の基本を学び、実際に課題等をこなした。後期授業の後半で、3 回ほどポストエディットの演習を実施したが、カリキュラム全体からみると限定的で、学生たちがポストエディットの方法を習得したとは言いがたい。

#### 5. 実験デザイン

調査・実験は、2012 年度の後期授業の期末課題の一部として実施した。英日ポストエディットの問題を 1 問課題に含めた。起点テキストは、IT のマニュアル（取扱説明書）で、Firefox（オープンソースのウェブブラウザ）のオンラインヘルプから抜粋、ワード数は 115。Google 翻訳を使って、事前に起点テキストを翻訳しておいた。起点テキスト（原文）と機械翻訳を並べてある（下記の表 1 を参照）。

学生には、「あなたの訳」の欄に自分の修正後の翻訳を記入するように指示した。作業は、

筆記でなく、Word ファイルをダウンロードして、コンピューター上で行えるようにした。翻訳支援ツールを使用しなかったが、予め準備しておいた機械翻訳の訳をコピー&ペーストして、できるだけ通常のポストエディットに近い形で作業できるように配慮した。完成したファイルは、メールで提出してもらった。

表 1: 実際の課題の抜粋

原文	機械翻訳	あなたの訳
Deleting a single bookmark	単一のブックマークを削除する	
Go to the page in your bookmarks that you want to remove.	削除したいブックマークのページに移動します。	
Click on the star icon in the Location bar.	ロケーションバーのスターアイコンをクリックしてください。	
The "Edit This Bookmark" box will be shown.	[このブックマークを編集]ダイアログが表示されます。	
In the "Edit This Bookmark" box, click "Remove Bookmark".	[このブックマークを編集]ボックスで、[ブックマークを削除]をクリックしてください。	

今回検証に使用した起点テキストは、Firefox のマニュアルからの抜粋であるため、ネット上で検索すれば、日本語版を入手することができた。しかし、調査参加者でこれらを入手したものはいなかった。筆者は、事前にこれを入手し、Google 翻訳結果との GTM スコアを検証しておいた。結果は、GTM 0.621 と比較的高い数字だった。汎用の機械翻訳エンジン（Google 翻訳）をそのまま使った割には良好な結果だったと言える。上の例文からもわかるように、起点テキストがシンプルだったので機械翻訳でも良い訳ができたこと、また、Firefox というオープンソースの IT マニュアルなので翻訳コーパスが豊富であり、統計的機械翻訳エンジンを使う Google だからこそ、良好な結果だったと考えられる。

この結果をポストエディットするには、1) 正確性 (Accuracy)、根本的な誤訳がないかどうか、2) 読みやすさ (Clarity)、文章の流れを整える、3) 文体・スタイル (Style)、取扱説明書としての文体に揃える、という点に対して修正をすればよい。これら3つの点について、実験参加者は、授業内で翻訳の基礎として学んでいる。尚、3) の「スタイル」の修正までを行わなければならないポストエディットは、翻訳者（ポストエディター）の負担を増大させる（ポストエディットの効率性を低下させる）という研究もあるが（O'Brien, 2002）、本検証ではこの要素を品質確認項目に含めた。

## 6. 検証結果

### 6.1 体感作業負荷

では、結果をみてみよう。下の図 1 は、翻訳者 20 名の体感作業負荷、テキスト類似度、クラス成績、翻訳品質の一覧である。品質の一番高い者から、降順で並べてある。点線より上（翻訳者 F 以上）が、筆者が判定した品質合格レベルである。

品質については後述するとして、まず全体としての体感作業負荷と修正量（テキスト類

似度)をみる。平均で、体感作業負荷は 70.75、テキスト類似度は GTM 0.757 であった。これは、学生たちがポストエディットを通して人手翻訳よりも約 30%の作業軽減を体感したということである。またそれに伴うテキスト類似度 GTM 0.757 (修正量 0.24) が示すように、機械翻訳の訳に対する修正量もわずかであった。類似度が GTM0.464 以上であれば、翻訳者は 10~15%程度の作業負荷軽減を体感できるという先行研究通り(山田, 2011c)、今回の結果は、テキスト類似度 GTM 0.757 から想像がつくように、約 2 倍の平均で 30%の作業軽減が得られた。しかし、全ての翻訳者に当てはまるものではなく、翻訳者 S と R の 2 名は先行研究によるしきい値 GTM0.464 以上のテキスト類似度があったにもかかわらず(それぞれ GTM0.596 と 0.633)、作業負荷は 120 と 100 と回答している。一方、翻訳者 K の体感作業負荷が 125 と調査の中で最も高い数字になっているが、これは、この翻訳者の類似度 GTM 0.437 がしきい値 GTM0.464 以下であったことから、予測可能であったと言えるだろう。

このように、学生の翻訳品質や翻訳能力などとは関係なく、まずポストエディットという作業そのものが、全体として体感作業負荷軽減を促したことは、ほぼ間違いない。少なくとも 20 名中 17 名は、今回のポストエディットは人手翻訳よりも楽であったと回答していることから、学生はポストエディットを好意的に受け止めているといえる。

図 1: 結果一覧

翻訳者	体感作業 負荷	類似度 (修正量)	成績	品質	
A	80	0.654	優	0.726	
M	40	0.714	優	0.640	
I	30	0.955	優	0.613	品質レベル到達グループ
O	70	0.814	優	0.603	
E	50	0.824	優	0.584	
Q	85	0.748	優	0.570	
F	90	0.728	優	0.549	品質合格レベル
S	120	0.596	優	0.538	
R	100	0.633	優	0.528	
H	50	0.982	優	0.525	
T	50	0.787	劣	0.512	
P	90	0.978	劣	0.506	
J	70	0.882	劣	0.496	品質レベル未達グループ
D	85	0.881	優	0.496	
L	40	0.693	劣	0.493	
B	50	0.891	劣	0.487	
G	50	0.857	優	0.450	
N	50	0.616	劣	0.429	
C	90	0.461	劣	0.409	
K	125	0.437	劣	0.366	
平均	70.75	0.757		0.526	

## 6.2 品質

では、翻訳品質をみる。繰り返すが、品質の点数は筆者の経験則的な判断による要素も含まれることを断っておく。そのため、点数の順位が、必ずしも品質優劣の序列を意味するわけではなく、最終的には総合的に判断をして、合格ライン(点線)を設定して、合格

基準以上か以下かというように2つのグループに分けた。つまり、一番上の翻訳者 A がそのすぐ下の翻訳者 M よりも点数が高いから、品質的に優れているということは保証しないが、両者ともに合格基準以上にあることは確かである。

以上を踏まえて結果を見る。まず合格基準以上は7名のみであった。やや厳しい評価だが、逆にいえば7名が実務翻訳に近い品質であったとも言える。無論、起点テキストが非常に易しかったので、もう少し多い人が合格すると想定していたが、少なかったのには理由がある。今回の評価で使用した起点テキストのワード数が115ワードと限られていたことを予め考慮し、評価ポイントとなる文章を用意しておいた。合格基準を満たしたかどうかは、まず、これらのポイントをクリアしていたかどうかになる。具体的には、次に示すような訳文の「正確性」に関わるものだ。

評価ポイント例：

【起点テキスト】

In the right pane, click to select the item(s) you want to delete.

With the items to be deleted selected, click on the "Organize" button and then select "Delete".

【機械翻訳の訳】

右ペインで、削除する項目（複数可）をクリックします。

項目が選択削除すると、[整理]ボタンをクリックし、[削除]を選択します。

上例の下線部に着目したい。【起点テキスト】の「With the items to be deleted selected」に対応する【機械翻訳の訳】が「項目が選択削除すると」のように、誤訳に近い訳出になっている。このすぐ前の文で「削除する項目を選択する」とあるので、文脈から考えても、この下線部の訳は「削除する項目を選択した状態で・・・」のような訳に修正する必要がある。このような評価ポイントをクリアできたのが、7名であったという結果だ。

上のように誤訳が修正されずに見落とされる現象は、統計的機械翻訳でのポストエディットや翻訳メモリを使った時の翻訳作業で典型的に見られる「エラーの伝播（error propagation）」である。機械翻訳の訳や翻訳メモリ内の既存訳にエラーや誤訳があると、修正されずにそのまま残ってしまうケースがある。エラーが翻訳者によって検出されない（ことが多い）理由としては、すでに O'Brien (2003) や山田 (2011b) が指摘するように、翻訳者（や読者）の意識なかで、「スタイル（文体）」の要素に対する注意のレベルが高くなり、「正確性」への注意が低くなることがある。すなわち、目標言語がこなれた訳（自然な訳）に見受けられる場合、細かな誤訳や数字などのエラーが検出しづらくなるというのが原因と考えられている。

今回の場合の「項目が選択削除すると」という訳が、「自然な訳」だったから翻訳者が検出できなかったのかどうかは、判断が分かれるところだろう。この訳は日本語としても不自然であるからだ。ただ、全体として、機械翻訳の訳が良質であったことから、翻訳者が油断をして、当該エラーを見過ごしてしまった可能性は否めない。

一方で、今回の検証対象者が学部学生だったことに鑑みると、実はこのエラーは、そもそも学生の語学力に起因しているかもしれない。「With .... selected」の構文解析ができて

いなかったか、このフレーズの訳し方が分からなかったことによる誤訳かもしれないのだ。

### 6.3 品質と成績(「劣」成績者)

これらを踏まえて、ポストエディットの品質と学生の成績との関係を見てみると、合格基準以上の翻訳者の中に、「劣」成績者がいないのがわかる。合格基準以上の7名すべてが「優」成績者であった。「劣」成績者の学生は、上図の下方に固まっている。つまり、自明ではあるのだが、やはりポストエディットをする場合でも、翻訳の基礎能力を備えている必要があるということである。

### 6.4 品質と成績(「優」成績者)

この結果よりも、成績との関係で着目すべきは、成績優秀な学生(「優」成績者)が、合格基準レベル以下にも散見されることだろう。5名の「優」成績者が合格基準以下にいる。これら「優」成績者が、もし同じ起点テキストを人手翻訳していたならば、同様のエラーや誤訳を犯さなかったかもしれない。ならば、やはり「エラーの伝播」が、ポストエディット時の弊害となってしまっているのかもしれない。

### 6.5 学生のコメント

これに関連して、学生たちにポストエディットの感想を記述してもらっているので、以下でみる。まず合格ラインに達していない翻訳者のコメントだ。

#### 翻訳者のコメント (合格ライン未達者)

翻訳者 K (「劣」成績者) :

もともとポストエディットのように訳された文を見てしまうと、その解釈を踏まえた訳しか浮かばなくなってしまう。その訳が全く間違っても一度解釈してしまうと、ゼロからの思考が出来なくなるため、どこか直訳のような文になってしまうし、日本語の語順も危うくなってしまう。

翻訳者 C (「劣」成績者) :

翻訳する上で容易になる部分もあるが、機械翻訳につられてしまうことも多かった。

翻訳者 R (「優」成績者) :

ポストエディットの場合、先に機械翻訳が出されると、その訳に引っ張られてしまうため、直すべきか直さないべきか迷ったり、直したいけどどのような文法・単語を当てはめればいいのかアイデアが浮かびにくくなることが時々あった。

翻訳者 S (「優」成績者) :

ふつうの翻訳よりも翻訳しにくく感じます。どのような言い回しが適切なのか判断する



のが難しく感じました。

学生からの回答は、そのまま記載した。ここで興味深いのは、「機械翻訳に引きずられてしまう」という回答が多いことである。成績 A レベルの学生たちのコメントにも共通してみられる。

機械翻訳の訳や他人の下手な翻訳を見てしまうと、それに（自分の訳が）引きずられてしまうというのは、よく言われることでもある。翻訳をある種の創作的行為と捉えるならば、悪訳に触れることは、翻訳者にとって悪影響となる危険性をはらんでいる。翻訳業界やプロ翻訳者の間でも言われるように、長期的には、「ポストエディットをしても、一流の翻訳者にはなれない」とか「(自分の)筆が荒れる」<sup>3</sup>といった類の話とも関係する。翻訳者の意識のなかに、ポストエディットに対する、ある種、悲観的な態度が見え隠れする。

では、「合格ライン到達者」たちのコメントはどうだったのだろうか。以下を見ると、少し事情は異なっている。

翻訳者のコメント（合格ライン到達者）

翻訳者 I:

一つ一つの単語の意味を自分で調べる必要がない。 出来あがった訳をより日本語的に直すだけなので楽でした。

翻訳者 H:

普通の翻訳では、わからない単語をひとつひとつ辞書で調べるけど、ポストエディットはその手間が省けて楽だと思いました。

翻訳者 A:

普通の翻訳では自分で知らない単語を調べながら少しずつ翻訳しなければなりません、それよりは、ポストエディットで元となる文があってそこから文のたまかな意味をとらえてから訂正していく方がやりやすかったように感じます。

Translator O:

同じ単語を繰り返し（何回も）使っているから混乱することもないし、気にする必要もない。長い文章だとどこに修飾されているのかとか、どこに繋がっているのかを考えるのに時間を取られがちだが、読んだまま訳せば良さそうだ。この場合は機械翻訳も悪くないと思った。

合格ライン到達者グループは、先とは対称的に「一つ一つの単語の意味を調べる必要がない」といったコメントが多い。「機械翻訳につられてしまう」という意見がなかったわけではないが、それでも好意的にポストエディットを受け入れている印象がある。

しかし、一見したところ、「単語の意味を調べる必要がない」、というコメントは、翻訳者としての怠慢な態度にも受け取れるし、どちらかと言えば、合格ライン未到達者にもまして、翻訳の初心者的なコメントに見受けられる。

## 6.6 翻訳単位と品質

訳出時の翻訳単位から考えてみると、初心者翻訳者は、熟練翻訳者に比べて、単語レベルに近い小さな翻訳単位で、翻訳をしていくことが多いといわれる (Dragsted, 2005)。これに対して、熟練翻訳者は、起点テキストの難易度 (複雑性等) が下がれば下がるほど、文レベル以上の大きな単位で、訳出する傾向がある。

本検証対象の合格レベル到達者は、当然、Dragsted (ibid) のいう初心者翻訳者に分類されるだろうから、単語レベルに近い小さな翻訳単位で、翻訳作業を進めていたと想像できる。つまり、「単語の意味を調べる必要がない」というコメントが意味するのは、起点-目標テキスト間の単語レベルの対応づけを容易に行えた。逆の言い方をすれば、(初心者翻訳者なので) 単語レベルしか見ていなかった、ということにもなりうる。つまり、小さな翻訳単位での対応付けが確実にできたことが、エラー抽出につながった可能性があるのだ。

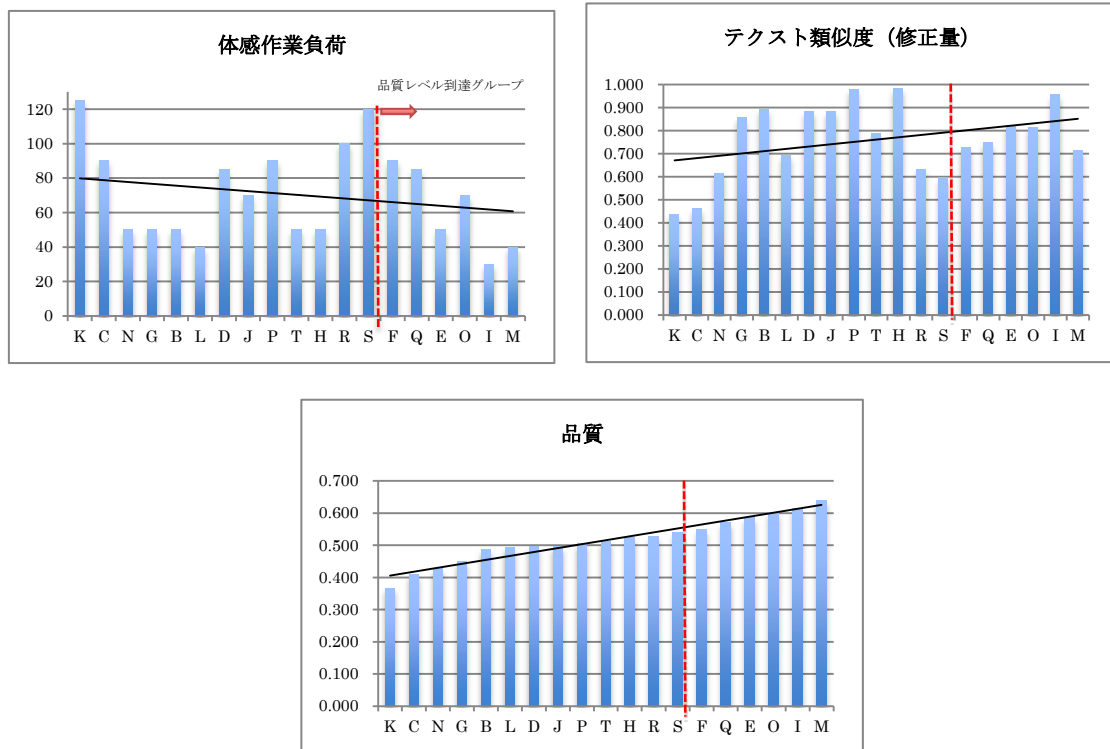
他方、合格ライン未達者、中でも「優」成績者は、単語レベルを越えてもっと大きな単位 (文など) のレベルで、機械翻訳に引きずられないように、別の訳語を考えようとしたがゆえに、小さな単語レベルでの対応付けを怠り、エラーを見落とした可能性がある。すなわち、憶測の範囲を脱しないが、これらが示唆するのは、単語レベルに近い小さな単位で訳出しようとする初心者翻訳者の特徴や態度は、実はポストエディットという作業における誤訳などのエラー検出に適しているといえるかもしれないということ。

しかし、優秀がゆえに機械翻訳に引きずられないようしようとする学生は、単語レベルを超えた創造的な訳出を試みようとするために、逆に単語レベルに近い小さな翻訳単位での確認ができなくなってしまったと考えられる。(筆者の) 翻訳クラスの中では、市販の辞書の訳語をそのまま書き写すだけが翻訳ではない、すなわちテキストの繋がりや文脈 (コンテキスト) などの単語を超えたレベルから考慮して積極的かつ創造的に訳語を考えることが場合によっては必要である、というような指導を行なっている側面もあり、こういった指導方針とも関係しているのであろう。

## 7. グラフから見えること

では最後に、上の表の体感工数負荷、テキスト類似度、品質の数値を品質点数の順に並べてグラフ化し、全体の傾向をみておく。各要素内の翻訳者の値のバラツキが大きいので、要素間に相関性があるとまでは言えないが、グラフ内には近似曲線を入れて、およその傾向を見るようにした。

図 2: 検証結果のグラフ



この近似曲線の傾きをみると、品質が高かった翻訳者ほど、体感工数負荷が低く、テキスト類似度が高くなる（すなわち修正量が少なくなる）傾向が見られる。これは、品質基準到達グループの翻訳者たちにとっても、ポストエディットが作業負荷低減に貢献していたことの証とはなる。無論、本結論を一般化するにはサンプル数を増やすか、他の要素も含めて検証する必要があるようだ。

その他で特筆すべきは、類似度のグラフで、品質合格者と不合格者の両方に、修正量の多い翻訳者とそうでない翻訳者が、混在していることだ。換言すれば、沢山の修正をして品質を高めた者もいるし、劣化させてしまった者もいる、また少ない修正しかしていないのに、高品質な翻訳に仕上げられた者とそうでない者がいるということだ。ポストエディットという作業が目指すところは、少ない修正量で高品質を担保するのが理想なので、この辺りについては、今後の研究とポストエディットの指導面を含め改善の余地が残る。

## 8. まとめ

以上、本稿では、学生がポストエディターになれるか、ないポストエディターとして機能できるのかを考察してきた。暫定結論を以下にまとめる。

- ・学生はポストエディターになれるのか。

その可能性はある。20人中7人（35%）は実務翻訳の品質レベルに達した。しかし、検証に使用した起点テキストの難易度が低かったので、分野やテキストタイプによるので一

概には言えない。またポストエディットを行う場合も、基礎的な翻訳スキルを身につけていることが必須となる。

- ・ 学生のポストエディットに対する態度はどうか。

体感作業負荷は、全体的に低減した。20人中17人(85%)はポストエディットのほうが人手翻訳よりも作業が楽になると回答した。学生はポストエディットを好意的に受け止めているようだ。

- ・ どのような学生がポストエディターに適しているのか。

基礎的な翻訳スキルを身につけている学生が、ポストエディターに適している。しかし、実践翻訳の授業で成績優秀な学生すべてに適正があるわけではない。この原因は断定できないが、少なくとも人手翻訳のスキルが高いことが、ポストエディットのスキル向上に直結しない。むしろ、翻訳初心者にありがちな、単語レベルから小さな翻訳単位にフォーカスした訳出手法は、ポストエディットの作業に向いている可能性がある。機械翻訳教育面からは、従来の人手翻訳の指導とは異なるカリキュラムを導入する余地があるかもしれない。

.....

【著者紹介】山田優 (YAMADA, Masaru) 立教大学大学院異文化コミュニケーション研究科博士後期課程修了。異文化コミュニケーション学博士 (翻訳学)。麗澤大学、青山学院大学、神戸女学院大学などで講師を務める

.....

#### 【註】

1. 立教 SFR 翻訳研究プロジェクト(2013).『翻訳「革命」期における翻訳者養成：公開シンポジウムの報告と今後の取組み』(2013年7月1日  
<http://www.apple-eye.com/rikkyo/report.pdf> から取得)を参照
2. 本稿は、2013年4月の *TAUS Tokyo Executive Forum* で発表した *A pilot investigation on possibilities for novice translators to be post-editors in MT+PE settings* の内容に基づき執筆した。
3. 井口耕二氏 (翻訳者、立教大学講師) との個人談話による。同様の談話は、立教 SFR 翻訳研究プロジェクト (2013) にも見られる。「...MT のような技術を使うと...かえって混乱してしまう傾向にあります。MT で翻訳をして自分で post-edit すると、変な日本語に引きずられてしまう。それなら最初からやったほうがいいと考える人が多いと思います」(pp. 27-28)。

#### 【参考文献】

ALPAC. (1966). *Languages and machines: Computers in translation and linguistics. A report by the automatic language processing advisory committee, division of behavioral sciences, national*

*academy of sciences, national research council*. Washington, DC: National Academy of Sciences, National Research Council.

Retrieved October 1, 2010 from [http://www.nap.edu/openbook.php?record\\_id=9547&page=R1](http://www.nap.edu/openbook.php?record_id=9547&page=R1)  
Dragsted, B. (2005). Segmentation in translation: Differences across levels of expertise and difficulty.

*Target*, 17(1), 49-70.

García, I. (2010). Is machine translation ready yet? *Target*, 22(1), 7-21.

Krings, H. (2001). Repairing texts: Empirical investigations of machine translation post-editing processes (G. S. Koby, G. M. Shreve, K. Mischerikow, & S. Litzer, trans, G. S. Koby, Ed). Ohio: Kent State University Press.

Löffler-Laurian, A. (1985), Traduction automatique et style, *Babel* 31(2), 70-76.

McElhaney, T., & Vasconcellos, M. (1988), The Translator and the Postediting Experience, in *Vasconcellos (1988)*, 140-148.

Mossop, B. (2001). *Revising and editing for translators*. Manchester: St. Jerome.

O'Brien, S. (2003). Teaching post-editing: A proposal for course content. EAMT proceedings, 2002.

Retrieved July 1, 2013 from <http://mt-archive.info/EAMT-2002-OBrien.pdf>

Taus. (2013). MT post editing guidelines. TAUS

Retrieved July 1, 2013 from

<http://www.translationautomation.com/postediting/machine-translation-post-editing-guidelines>

立教 SFR 翻訳研究プロジェクト(2013).『翻訳「革命」期における翻訳者養成：公開シンポジウムの報告と今後の取組み』立教大学（2013年7月1日

<http://www.apple-eye.com/rikkyo/report.pdf> から取得)

山田優 (2010) 「機械翻訳+ポストエディットの実証研究：先行研究レビュー」『AAMT ジャーナル』第48号: 38-43.

山田優 (2011a) 「機械翻訳+ポストエディットの実証研究（その2）：英日翻訳での実験結果」『AAMT ジャーナル』第49号: 30-35.

山田優 (2011b) 「作動記憶と訳出プロセス」『翻訳研究への招待』第6号: 1-14.

山田優 (2011c) 「Revising text: An empirical investigation of revision and the effects of integrating a TM and MT system into the translation process」(未刊行) 博士論文:立教大学

