

問題解決を中心とした翻訳タスク・評価法

井上 泉

(マッコーリー大学大学院 Lecturer)

Abstract

This paper primarily aims to propose and report the theoretical framework, construct, and specific designs of a comprehensive rubric as a means of translation quality assessment. Particular emphasis is placed on the development of learners' problem-solving expertise in translation. The present paper first discusses the ongoing issue of how students' translation quality should be assessed in relation to key developments of assessment approaches in Translation Studies. Building on research findings in my relevant empirical studies concerning novice-expert differences in problem solving in the context of translation (Inoue, 2008; 2013), this paper will then illustrate conceptual and practical framework underpinning the problem-based translation task and assessment approaches. The construct and design of both tasks and the assessment method will be illustrated in details. This paper concludes with the brief discussion of some pedagogical implications and limitations of the present model and the suggestions for future studies. This paper can be a useful reference particularly for translation educators, professional translators, and learners currently undertaking their studies in translation.

1. はじめに

翻訳教育において学生の翻訳評価を行う機会は、テストのみならず、翻訳タスクも含めて頻繁に生じ、なおかつ形式や目的も多岐にわたる傾向にある。しかし、翻訳教育に携わる者として、「何を」「どのような基準で」そして「どのような方法で」評価するべきかについては、常に頭を悩まされる課題であった。マッコーリー大学大学院の通訳翻訳課程は、国家通訳翻訳者認定機関である NAATI¹ の認可教育機関であり、NAATI での直接受験以外に、同大学での認定資格取得も可能となっている。同修士課程 2 学期目の期末試験結果が NAATI 認定取得の可否に大きな影響を与えることから、長くその採点は NAATI 指定の評価方法に準拠する形で行われてきた。しかし、同方法は translation product (すなわち Target Text) のみを対象とした減点方式であり、また評価対象項目および減点スコアのルールも客観性および現実の翻訳実務との関連性の面で課題があることは否めない。従って、学期を通して同方法を採用することは、翻訳実務における即戦力

INOUE Izumi, "The Development and Implementation of Problem-Based Translation Tasks and Assessment Approach," *Invitation to Interpreting and Translation Studies*, No.16, 2016. pages 66-83. © by the Japan Association for Interpreting and Translation Studies

養成の観点から考えると、不十分と言わざるを得ない。以上の問題および課題が、今回報告する実証的研究に基づく新たな翻訳評価方法の開発・導入の契機となった。

本稿では、一連の実証的研究 (Inoue, 2008; 2013) で明らかになった、プロ・アマ翻訳者²間の問題解決における諸相違に基づき、翻訳における問題解決で重要とみなされた全側面を包含した包括的タスクおよび評価法の内容、デザイン、設計プロセス有用性について報告し、課題についても触れることとする。次項では、これまで Translation Studies の分野で提唱されてきた主要な関連概念・研究について論ずる。

2. 翻訳評価方法の現状と課題

2.1 Product に特化した翻訳評価の特徴・課題

Colina (2008) が指摘するように、Translation Studies の分野では局所的な方法論 (単語もしくはテキストレベルの等価概念など) は存在するものの、翻訳評価における総合的な枠組みを模索している段階だと言える。特に、翻訳の「質」をどのように定義し捉えるべきか、ならびに「良い」または「一定水準」にある翻訳とは厳密に何であるかについては、議論が行われているものの、未だ明確な方向性は見えていない (Williams, 2001)。これらの課題により明確に対応する上で、その対象を translation process または product のどちらか一方にすべきであるのか、あるいは、両方を評価すべきなのか検討する必要があるだろう。筆者のスタンスは、後者である。というのも、学生のタスクならびにテストを評価する場合、product のみでは、どのような過程を経て、どのような根拠でたとえば product の特定の訳出に至ったのかについて理解することが困難であるからである。さらに、学期を通しての進捗状況を把握・評価する際に、product のみの評価では Source Text (ST)・Target Text (TT) 対照に基づく評価 (主に言語的な観点) のみに限定したものとなる恐れがあり、翻訳実務を取り巻くその他の諸相 (リサーチ力、対人関係・コミュニケーションスキルなど) を網羅することが困難であるからである。

2.2 既存翻訳評価法の種類・特徴

Turner et al. (2010) は、世界各地で実施されている 24 種類の通訳翻訳認定・資格制度において、どのような評価方法が採用されているのかについて調査を行った。その結果においては、大別すると以下の 3 種類が存在することが判明した。

- (1) エラー分析・減点法
- (2) 基準参照法 (Criterion-referencing) と (1) の併用
- (3) 基準参照法のみ

上記 (1) では、所定のスコアを満点とし、評価対象項目に関するエラーを特定したのち、一定のルール (すなわち特定エラーが何点マイナスなのか) に則り減点し、残存したスコアを最終結果とする方法である。一方、(3) の基準参照法では、各評価項目に通常 1-4 ないし 5 の scale (段階) を設定し、各段階がレベルに関する descriptor (説明) を参考にそれぞれ評価を行う方法である。上記 (2) は (1)、(3) の混合体であるため、(1) および (3) のみについて以下に論じることとする。

まず、上記 1)のエラー分析・減点方法であるが、Williams (2001), Hale et al. (2012)が指摘するように、これは従来から存在するもので、翻訳パフォーマンスを評価する上で、いまだに主流となっているものである。一例では、オーストラリアにおいては NAATI がこの方法を長年採用している。このテストでは、productのみを対象とし、所定のエラー種別および規定減点数のルールに基づいて減点を行う。Kim (2009)、Williams (2001)が指摘するように、減点対象となる基準の”too free, too literal, unjustifiable omission, and mistranslation” (Kim, 2009:126)などの定義が曖昧であり、詳細なガイドラインも欠如していることが疑問視されている。明確な定義・ガイドラインが欠如しているということは、結果的に採点者によってその解釈に顕著な幅が生じ、採点者間で評価にばらつきが生じる可能性が高いという課題が存在する。加えて、評価においてエラーのみを指摘するその性質上、学習者の強みについては指摘が行えず、また改善すべき点についても学習者に有用な情報・フィードバックが十分に行えるとは言いがたい³。

翻訳評価法の新たな試みとして、上記 3) の基準参照法に関する議論・提唱 (Lee, 2005; Angelelli, 2007; 2009; Jacobson, 2009 など) が近年なされ始めている。同方法は、一定基準のスキルや能力、知識(すなわち総称して *translator expertise*)が現時点でどのような水準にあるのかを特定することを目的としている。同評価法には、下記の利点があると考えられる (Angelelli, 2007; Hale et al., 2012)。

- 包括的な対象事項の評価 (product・process 両面のみならず、テスト・課題で評価が必要なあらゆる側面の包含が可能)
- より体系的な評価の実現
- 評価対象者の翻訳における強み・改善点の評価可
- 評価対象者および複数の採点者間に異議が生じた場合、評価結果の説明・理由提示がより容易

他方、この類の評価法には課題または条件も存在する。

- Test construct (テスト構成)が慎重に定義・分析され、なおかつ評価事項が実証的研究に基づくものでなければならない
- Descriptor が内容的に慎重かつ詳細に記載されていない限り、多様な解釈が生じ得てしまう
- 特に減点法による採点に慣れている採点者の場合、本方法はより広範な評価事項を包含し、採点方法がまったく異なるため、採点者が慣れるまで時間と労力を要する

2.3 評価観点と課題

上記大別して二種類の翻訳評価方法を論じたが、翻訳評価において考慮すべきいま一つの観点は、どのような切り口から翻訳評価を行うかである。翻訳評価に関する既存研究を参照すると、reader response アプローチ、textual and pragmatic-based アプローチ、Systemic-functional linguistics (SFL)アプローチが主要なものとして提唱されている (Colina, 2008)。まず、Reader response アプローチは、Source Language (SL)の読者によるSTへの反応と、Target Language (TL)

の読者による TT に対するそれとの間に、どの程度の **equivalence** (等価) または類似性が存在するかを評価するものである(例 Nida and Taber, 1969)。等価の概念と同じく、このアプローチにも疑問点が存在する。すなわち、TL 読者の反応をどのように把握すべきかという問題に加え、読者の反応(すなわち TL 読者にとっての読みやすさ)が、正確性ほど求められていないタスク・案件(例法律文書)も多々ある点である。この意味で、翻訳の目的を重視する **functionalist-based** アプローチも同様の課題を抱えると言える。さらに、同アプローチは、等価という翻訳の限定的な観点のみを対象としたものであるため、必然的に評価結果も限定的なものとならざるを得ない。2 点目の **textual and pragmatic-based** アプローチは、単語・センテンスレベルにおけるエラーの同定から、テキスト全体・翻訳の目的というより包括的な翻訳評価へと方向性の転換を提唱している面で、意義あるものと考えられる。反面、翻訳評価の現場では必ずしも広く浸透しているとは言えない側面もある。このアプローチには ST (Reiss 1971) または TT (Reiss & Vermeer, 1984; Nord, 1997 などが提唱する 'skopos' の概念など) への、ある意味重点が極端な傾向にある。House (1997, 2001) による **functional pragmatic model** は、ST-TT 間の言語的・文脈的特徴を比較し、どの程度類似点があるのかを評価するというものであり、両テキスト対比という意味では、バランスのとれたモデルとは言える。しかし、依然として **equivalence** の概念に基づいたものであり、**overt translation, covert translation** など House が用いる主要用語の意味するところが明確ではないという課題も存在する。また、相対性ならびに類似性を常に評価対象とすることにも、疑問が生じる。というのも、多岐にわたる翻訳案件を鑑みると、ST が書かれた目的と TT を必要とする目的ならびに TL 読者の特徴が大きく異なることも少なくないからである。このような場合には特に、同モデルの限界を呈してしまう恐れがあるわけである。3 点目の SFL アプローチは、筆者の言語的な選択から生じる意味をより広範な社会文化構造に体系的に関連付けるといった概念に基づくものである。例えば Kim (2009) は、SFL の概念を用いた **meaning-oriented assessment criteria** を提唱している。ここでは、文脈・状況的要素(レジスター、**translation brief** など)を考慮に入れつつ意味形態における **translation shift** の程度を判断することが重視されている。同モデルの実際の運用においては、なぜ減点法を採用するのかという理由については不明である。また、同モデルの導入後 NAATI 合格率が著しく増加したとの報告もなされているものの、調査対象となった学生の背景ならびにスキルレベルなどの諸要素に関する言及が限定的であり、また、NAATI の結果が翻訳実務にどの程度直結するのかについて、検討の余地があると言える。

上記 3 アプローチを対比すると、概ね評価対象となっているのは、**product** (TT) であり、**process** への考慮ならびに評価がどの程度なされているのかは不明と言わざるを得ない。仮に **product** のみを対象にしたという想定でも、その評価対象事項は依然として限定的であり、翻訳の包括的評価という観点では不十分だと言える。すなわち、翻訳における問題解決を主眼とする筆者の研究結果 (Inoue, 2008; 2013) では、情報収集、クライアントとのコミュニケーションにおいても、プロ・アマ英日翻訳者間で大きな相違がみられたが、上記いずれの方法においても、どの程度これらの側面の評価が可能なのか定かではない。さらに、一部の **functionalist** アプローチを除き、上記のモデルに共通するのが、テキストレベルを上限とした評価という限界である。テスト・タスクを対象とした評価のみに目的を据えるのではなく、学生の学習における進捗状況や、現時点で学習者が有する **expertise** とプロ翻訳者として求められる **professional expertise** との対比、および学習者の

強みと克服すべき側面を、翻訳実務という「現実」に照らし合わせて評価することを目的とするのであれば、包括的かつより翻訳実務に即した評価モデルを構築することが重要ではないかと考える。また、翻訳評価法および具体的な評価ツールを構築する上でのもうひとつの課題は、Waddington (2004), Anckaert & Eyckmans (2006), Anckaert et al. (2008) も指摘するように、実証研究に基づいた方法・ツールの開発である。しかし現状では、実証研究よりもむしろ、実務経験に基づいた評価法が採用されており、テスト・タスクおよび評価法の有効性・信頼性に関する実証的な検証も十分とは言えないのが現状である。その結果、主観性という課題への対応・改善が必要と言える (Anckaert et al., 2008)。

3. 実証研究に基づく包括的な翻訳タスク・評価方法の構築

上述した翻訳評価法の現状および課題を受けて、本セッションでは、包括的な翻訳タスク・評価構築の新たな試みについて報告する。まず、実証研究との関連性であるが、プロ・アマ英日翻訳者が直面する翻訳における問題および問題解決方法 (Inoue, 2008; 2013)をテーマとする質的・量的研究の結果に基づいた、翻訳タスクおよび評価法を構築することとした。上記研究の主要結果を端的に述べると、STの内容・論理的な構成の理解、対象読者に適したリーダビリティ(可読性)の実現、情報収集、クライアントなど翻訳実務に関与するステークホルダーとのコミュニケーションの4分野でプロ・アマ間の顕著な相違がみられた。この結果、下記の4点をタスクおよび評価法構築の基本原則とした。なお、本方法は、上記研究の後続研究(タスク観察、Problem-Based Learningを用いた翻訳教授法⁴など)での活用のみならず、筆者所属の通訳翻訳課程全体および筆者自身の翻訳教育の現場においても学期を通して採用することを前提とした。よって、テストにみられる単発的評価というよりは、研究参加者ならびに学習者のtranslator expertiseを継続的に把握し、その進捗の一助とすることを目的としたものである。

- 包括的な翻訳タスクおよび評価
筆者の実証的研究ではいずれも、翻訳を問題解決の観点からでき得る限り包括的に分析・解明することを基本原則としたので、タスク・評価においても、学習者の問題解決におけるtranslator expertiseを可能な限り包括的に把握できる点を重要視した。
- 問題・解決法・理由の正当化
実社会の翻訳においては、高度なexpertiseを求められるその他多くの領域と同様に、翻訳実務においても問題解決力が極めて重要な要素と考えられる。従って、この観点を中心とした翻訳タスクおよび評価法の構築・導入を行うこととした。具体的には、多様な問題の特定、各問題に対する適切な解決法、及びその解決法を選択した理由をいかに適切に正当化できるかに焦点を当てた。この三要素はいずれも、上述の実証研究および関連文献においてプロ・アマ間の相違が顕著にみられたものである。
- 翻訳processおよびproduct両面の評価
上述の本方法構築の目的および包括性という性格上、学習者の翻訳process・

product両面を評価対象にすることとした。特に翻訳processにおいては、問題・解決法・理由の正当化という上述の三要素が評価における重要な情報となる。

- 学習者のtranslator expertiseの現状および改善点の明示化
上述の如く、本方法は学習者のtranslator expertiseの継続的観察および向上をも目的としたものであるため、タスク・テスト毎に現状・改善点を学習者が明確に理解できる形で提示する方法を考案することとした。

3.1 翻訳タスク

翻訳評価と表裏一体をなすのが、どのような方針で翻訳タスクを構築するかという点である。NAATIに見られるような従来型のテストおよびタスクには、translation brief(cf. Holz-Mänttari, 1984)はあるものの、極めて限られた情報(例 STの出典、目的)のみが記載される傾向にある。特に同テストにおいて翻訳の目的およびTTで求められるテキストタイプは、STのそれと同様なものであるのが常である。本タスクにおいては、でき得る限り実務に近いタスク設定とするために、下記各点を網羅する、より詳細な情報をクライアントからの指示事項という形(すなわち、translation brief)で提示することとした。ただし、上記情報が一部欠落するケースも現実的にあり得るため、一定情報が欠落した形のタスクも意図的に適宜設定することとした。加えて、筆者の先行研究(ibid.)で浮上した、クライアントとのコミュニケーションをも評価対象とするため、学習者にはメールおよび翻訳者註などの方法で、適宜コミュニケーションを図ることを奨励することとした。⁵

- 本案件(すなわちタスク・テスト)の目的(想定される対象読者、TTが何のために、どのような形で使用されるのかなど)
- クライアントに関する情報(クライアントからの直接業務依頼または翻訳会社経由からの依頼、クライアントの所在地、企業の場合どのような業界のクライアントかなど)
- クライアントからの指示事項(フォーマット、体裁、表記、翻訳支援ツールを含む所定のソフト、スタイル面での指定など)

さらに、翻訳process評価を包含するため、TTに加えてannotation(注記)の提出を求めることとした。このannotationでは、各学習者が認識した問題およびその分析(すなわち、何が問題であり、なぜそれが問題であるのか)、具体的な解決策(どのような解決策を採用するのか)、ならびに解決策を選択した理由の正当化(理論・実践的な観点から、なぜその解決策が最善なものなのか、解決策に至るまでの情報収集を含む全プロセスなど)を問題ごとに説明する方法を採用することとした。

次に、業務全体の設定(すなわち translation brief) およびSTの選定についてであるが、ここで特に留意したのが、その難易度をどのように定めるかである。ここでは、上記同様、現実的な翻訳実務にでき得る限り近い難易度、すなわちプロ翻訳者が実務で取り組む、ある程度専門的な特徴を内包するtranslation briefおよびSTを選定するよう努めることとした。この根底には、authenticity(真正性)という概念が存在する。Angelelli(2009)が論じるように、authenticityの面で翻訳実務とタスク間の整合性が重要であるからである。また、認知心理学の分野でWard et al.(2006)が指摘するように、真正性および現実性を帯びたタスクを用いることで、学習者の現時点でのexpertiseをより

的確にとらえることができるという利点もその背景にある。ただし、上記真正性・現実性のみでは具体性に欠けるため、より具体的にどのようなtranslation briefおよびSTの難易度および言語・語用的特徴に則したものとすることのかを考慮した結果、日本でフリーランスのプロ翻訳者として実務に携わるうえで適用されることの多い、「トライアル」を参照することとした。トライアルは、主に翻訳会社にプロ翻訳者として登録する際に課される、言わばテストである(近藤 2005)。筆者が研究・教育対象とする学習者は、主に日本語を母語とし、学習期間の終了後は日本にてフリーランスまたは正社員として翻訳業務に携わることを志望する者が大多数である。従って、学習者の現実的なキャリアプランの面からも、トライアルに準拠した難易度のST⁶を選定することとした。また、テキストタイプについても、できる限り広範なものを網羅するため、ビジネス、法律、医療、IT、メディア、特許などの分野とした。6つのST候補選定およびタスク設定のうち、英日翻訳実務の経験(日本のクライアントから依頼の案件を含む)を5年以上有するプロ翻訳者5名を対象に、本構想の説明後フィードバックを求めた。ここでのフィードバックは、タスク設定(上記translation briefを含む)およびST自体がどの程度上述の条件を満たすものを中心としたものであった。結果的に、いずれのタスク候補においても、ST自体およびタスク設定両面で条件を満たすという肯定的なフィードバックを得ることができた。さらに、先行研究(Inoue, 2008; 2013)におけるSTの正確な理解の側面で複数のプロ翻訳者から指摘された、STが必ずしも言語的・理論的に適切ではない場合が少なくないという点をも考慮に入れることとした。その結果、日英翻訳実務を専門とし、英語を母語とするプロ翻訳者1名に依頼し、選定STに適宜、言語的・理論的な瑕疵を意図的に包含した(一例として、巻末資料1を参照のこと)形で修正を加えた。

3.2 翻訳評価法

上記の基本方針に基づき、本方法ではルーブリック、特に分析的ルーブリック(analytic rubric)をツールとして用いた評価を行うこととした。この種のルーブリックは、言語運用能力の評価でこれまで広く採用されてきた方法である(Mertler, 2001)。上記セクション2と重複する点もあるが、本ツールの特徴として下記の諸点が挙げられる。まず、特定の側面について学習者の長所・改善点両面を特定できる(Huges, 2003)のが大きな利点と言える。特に、改善点は学習者のexpertise向上において内省を促進する意味で重要な要素であると言えよう。さらに、他のルーブリック同様、本ツールは多角的な側面の評価事項を包含できるという利点もある(Angerelli, 2009; Hale et al., 2012)。その結果、既存の翻訳評価法とは異なり、process・product両面のみならず、テキストレベルを超えた諸側面をも評価の対象とすることが可能となるわけである。

ルーブリックを具体的に設計・構築する際の留意点は、概ね下記の通りである。各点ともタスク・評価の目的および対象範囲によるため、ケースバイケースとなるが、本件では、上述のタスクおよび評価法構築の基本原則を最も適切に反映できる形式を採用することとした。

- 評価項目数をどの程度とするか

翻訳process・produce両面の評価ならびに包括的な翻訳の評価という性質上、採点者にとっては煩雑と感じられる恐れはあるものの、実証研究で特定された全側面を網羅するため、13の評価事項数とすることを決めた。例えば、STの正確な理解を

評価する際には、ST・TTを対比することにより行える一方、processの適切性についてはannotationを、また、クライアントとのコミュニケーションについては、クライアントとのメールのやり取りおよび翻訳者駐を参照にすることで、評価が行えるものと判断した。

- 各評価項目に対するscale数をどうするか

ルーブリックでは通常、各評価項目に対して、複数のscaleが用いられる。ただし、そのscale数の決定は、ルーブリック設計者の教育的信条および目的に即したものとなる傾向にある。本件では、1～4の4つのscaleを採用することとした。この中で、1・2に該当するscaleは、特定の評価項目でexpertiseが不十分と認識される場合であり、1が2に比べ、より不適切とみなされるということになる。他方、3は平均を超えるケースで、4が最も要件を満たすレベルということになる。ルーブリックの中には、5つのscaleを設け、3を可もなく不可もなしとする例も存在するものの、同scaleが大多数となった場合、学習者にとっても教員にとっても、professional expertiseにおける学習者の現在の立ち位置が不明瞭になる恐れがあるため、この方法は採用しないこととした。またこの4つのscaleでは、4がプロ翻訳者と同等レベルと設定することとした。これにより、学習者にとってプロ翻訳者に求められるレベルと自らの現状レベルの対比が明確になると考えたことに起因する。
- Descriptorをどのように記載するか

ルーブリック利用に際しては、scaleに加えdescriptorを併用することが多い (Hale et al., 2012)。採点者がどのscaleを選択するかを判断する上で重要な参考情報となり、また学習者にとっても、特定の評価項目における現在の立ち位置ならびに強み・改善点がより明確になるという利点がある。従って、本評価法でもdescriptorを各scaleに併記するとともに、できる限り明確かつ詳細な情報を盛り込むこととした。
- 選定したscaleとdescriptorの提示で、教育的な目的上十分か否か

4 scaleのうち1つを選択して、descriptorとともに提示することのみで、なぜ特定の評価項目についてその判断となったかが十分に伝達できるかについても検討を行った。結果として、必要に応じてMicrosoft Wordのコメント機能を用いて追加コメントを記載するという形式を採用した。
- 最終評価をどういう形で算出するか

学期中のタスクならびにテストへの本評価法導入を想定すると、各タスク・テストとも最終評価を量的に算出する必要性が生じてくる。本評価法では、各評価項目を同じ点数配分(すなわち4が最高点)とし、その総計を算出したのち、総計を100%に換算することとした。なお、全項目の加重を同じにしたのは、前出の実証研究結果において、いずれの分野も同様にプロ・アマ間における相違が同程度みられたことに起因する。

ルーブリックのみならずいかなる評価法においても、導入前にその有効性を検証することは極めて重要である。ここでカギとなるのが、validityとreliabilityという両概念である。Validity(妥当性)

は、タスクのスコアより得られる特定の推定事象における適切性・意義・有効性を指す(Angelelli, 2009)ものである。タスクおよび評価の妥当性を保証するためには、想定され得るバイアスを検証する必要がある。すなわち、評価におけるスコアがタスクで定義・特定されるさまざまなexpertiseを反映しているか否かを判断することが重要となる。本タスク・評価法のvalidityは、下記の3点より検証を行うことができると考える。まず、タスク・評価法ともプロフェッショナルとしての翻訳実務という「現実」をでき得る限り反映するよう設計されている点。2点目として、多角的な実証研究結果を反映した結果、実証的に裏付けられた妥当性(empirically-underpinned validity)が実現されている(Fulcher and Davidson, 2007)こと。3点目は、実証研究結果に基づき、さらに第三者(すなわち複数のプロ翻訳者パネル)の見解・合意を得たこと。これらの理由により、一定のvalidityが確保されたと考える。

一方、reliability (信頼性)とは端的に述べれば、ある評価法においてどの程度評価に整合性があるかを指すものである (Bachman, 1990; Bachman & Palmer, 1996)。換言すれば、例えばある特定の受験者が同一テストを2回の異なる時期に受験した際に、その結果にどの程度の類似性があるか、また、複数の評価者がある受験者の同一テストを採点した場合、その結果がどの程度類似しているかなどが問われるものである。特に、後者のいわゆる ‘inter-rater reliability’ (e.g. Bachman, 1990; Bachman & Palmer, 1996; Angelelli, 2009)は、検討の必要がある重要な要素だと言える。というのも、翻訳の質的評価においては、でき得る限り評価者の主観性を抑制する必要があるからである。従って、ある翻訳評価において、複数の評価者が評価を行うことも重要であり、さらに、評価者が一堂に会して、あるタスクの評価を各自行ったのち、全員で討議・検証する機会を設けることに大きな意義がある (Alderson et al., 1995; Weir, 2005)。本タスク・評価法においても、その導入前に、前出のプロ翻訳者パネル全員を招き、同様のワークショップを開催した。ここでは、2件のタスクに事前に4名の学習者の協力のもと取り組んでもらい、それぞれの翻訳processおよびproductをワークショップの各参加評価者に、本ルーブリック(巻末資料2参照)およびルーブリックで用いられる主要用語の説明(巻末資料3参照)を用いて、まずは個別に評価してもらうこととした。その後、全体でのディスカッションという形式で各評価項目におけるscaleの比較、その理由、および本評価法で改善すべき点などについて話し合った。これにより、本評価法の一定のreliabilityを確立できたと考える。

4. 今後の課題・展望

本稿で報告した、問題解決を主眼とする翻訳タスクおよび評価法の開発は、実証研究の結果を反映したものであり、また、翻訳というプロフェッショナル分野のみならず、多様な分野で中心的な役割が指摘されている問題解決という側面から、タスクおよび評価法を構築する新たな試みである点に意義があると考えられる。しかし、今後取り組むべき課題も存在する。

まず、同タスク・評価法の開発の根底にあるのは、問題解決を中心とした翻訳教授法(Task-Based Problem-Focused Learning)であり、翻訳における問題解決力養成を実現することであった。従って、同方法を用いた場合の継続的かつ長期的な(例えば1ないし2年間)学習・教育効果を検証することの必要性を認識している。筆者は、教授法の試験的な導入を通じた研究(Inoue & Candlin, 2015)ならびにマッコーリー大学において同方法を翻訳の理論・実習教科に導入したが、その経験則のみならず、今後はあらたに実証的研究を考案し、実施する必要があると

考えている(例 一定の学習者グループの質的・量的な進捗観察、同方法を用いた学習者グループ・用いない学習グループにおける translator expertise の比較、同方法を用いる学習促進者としての筆者自身によるアクションリサーチ、他教員によるピア・オブザベーション及び提言)。

次に、筆者の翻訳教育における対象学習者層は、概ね日本人をはじめとするアジア圏の学生であることから、同方法をはじめとする learner-centred approaches(学習者主体の学習法)への経験が十分ではないくらいがある。また、上記1点目の学習・教育効果の観点をも含めて、本方法の中心的な役割を担うべき学習者が、本方法にどのような反応・対応を示し変化を呈し、また学習者サイドからどのような提言が存在し得るのかについても検証する必要がある。継続的な研究の一例としては、対象学習者からの定期・継続的なジャーナル・ダイアリーの収集・分析、対象学習者全員または複数グループによるフォーカスグループ形式でのインタビューが考えられる。

第三点目の課題は、いかに急速に変容を遂げる実社会における翻訳実務の要件を把握し、本方法に反映していくかである。すでに、IT を中心とした技術革新の影響で、ローカライゼーションなど、業界やコミュニティが求める翻訳者の役割に大きな変容が見られている。学習者を実社会における問題解決・翻訳のプロフェッショナルとして継続的に活躍できる人材に育成するうえで、このような変化を翻訳学習・評価法に時宜に応じて盛り込んでいくことも、極めて重要だと考える。

.....

【著者紹介】

井上 泉 (INOUE Izumi) オーストラリア・マッコーリー大学大学院言語学部 Lecturer。専門は翻訳理論、翻訳教育、エキスパート論。

.....

【注】

1. National Accreditation Authority for Translators and Interpreters. 詳細は、www.natti.com.au を参照のこと。
2. 上記研究におけるプロ翻訳者の定義は、NAATIプロ翻訳者認定資格を保有し、実務翻訳経験を5年以上有する者とした。一方、アマ翻訳者は、同資格を有さず、なおかつ実務経験が皆無で、本研究実施時点において、オーストラリアの大学にて通訳翻訳課程に在籍中の学生を指す。本稿で用いられるプロ・アマ翻訳者の定義もいずれも上記に準じる。
3. ただし、NAATI の諮問パネルによる、タスクおよび評価法改善に向けた取り組みは進行中である。詳しくは Hale, et al. (2012)を参照のこと。
4. 詳しくは Inoue & Candlin (2015)を参照のこと。
5. メールでのコミュニケーションでは、筆者がクライアントとしての役割を担い、translation brief におけるクライアント(英語または日本語を母語とする)として対応することとした。
6. 近藤 (2005)で示されたトライアルの課題例には、translation brief なし。

【引用文献】

- 井上 泉 (2008) 「英日翻訳者が直面する問題点—翻訳学習者とプロ翻訳者にみる差異を中心に」『翻訳研究への招待』第3号: 159-175.日本通訳翻訳学会
- 近藤哲史 (2005) 『トライアル現場主義!』丸善

- Alderson, J. C., Clapham, C., & Wall, D. (1995). *Language test construction and evaluation*. Cambridge: Cambridge University Press.
- Anckaert, P., & Eyckmans, J. (2006). Ijken en ijken is twee. Naar een normgerelateerde ijkpuntenmethode om vertaalvaardigheid te evalueren. In C. v. d. Poel & W. Segers (Eds.), *Vertalingen objectief evalueren. Matrices en ijkpunten* (pp. 53-67). Louvain: Voorburg: Acco.
- Anckaert, P., Eyckmans, J., & Segers, W. (2008). Pour nen evaluation normative de la competence de traduction. *ITL International Journal of Applied Linguistics*, 155, 53-76.
- Angelelli, C. V. (2007). Assessing Medical Interpreters: The Language and Interpreting Testing Project. *The Translator*, 13(1), 63-82.
- Angelelli, C. V. (2009a). Using a rubric to assess translation ability. *Testing and Assessment in Translation and Interpreting Studies: A Call For Dialogue Between Research and Practice*, 14(13).
- Bachman, L. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.
- Bachman, L., & Palmer, A. (1996). *Language testing in practice*. Oxford, New York: Oxford University Press.
- Colina, S. (2008). Translation quality evaluation. *The Translator*, 14(1), 97-134.
- Fulcher, G., & Davidson, F. (2007). *Language testing and assessment*. London and New York: Routledge.
- Hale, S., Garcia, I., Hlavac, J., Kim, M., Lai, M., Turner, B., & Slatyer, H. (2012). *Improvement to NAATI testing: Development of a conceptual overview for a new model for NAATI standards, testing and assessment*. Retrieved from Sydney, Australia:
- Holz-Mänttari, J. (1984). *Translatorisches handeln: Theorie und methode*, Helsinki: Suomalainen Tiedeakatemia.
- House, J. (1997). *Translation Quality Assessment: A Model Revisited*. Tübingen: Gunter Narr.
- House, J. (2001). Translation Quality Assessment: Linguistic Description versus Social Evaluation. *Meta*, 46(2), 243-257.
- Hughes, A. (2003). *Testing for language teachers*. Cambridge: Cambridge University Press.
- Inoue, I. (2013). *Novice-expert differences in identifying and addressing translating challenges: the development of an effective pedagogical approach to translator education* (Doctoral dissertation). Macquarie University, Sydney.
- Inoue, I. & C.N. Candlin (2015). Applying Task-Based Learning to translator education: assisting the development of novice translators' expertise in identifying and addressing translating challenges [Special issue: T&I Pedagogy in Dialogue with Other Disciplines]. *Translation and Interpreting Studies*, 10 (1): 58-86.
- Jacobson, H. (2009). Moving beyond words in assessing mediated interaction: measuring interactional competence in healthcare settings. In C. V. Angelelli & H. Jacobson (Eds.), *Testing and Assessment in Translation and Interpreting*. Amsterdam: John Benjamins.
- Kim, M. (2009). Meaning-oriented assessment of translations: SFL and its application for formative assessment. In C. V. Angelelli & H. Jacobson (Eds.), *Testing and assessment in translation and interpreting studies*. Amsterdam/Philadelphia: John Benjamins.

- Lee, J. (2005). *Rating interpreter performance*. (Unpublished Master's dissertation), Macquarie University.
- Mertler, C. A. (2001). Designing scoring rubrics for your classroom. *Practical Assessment, Research & Evaluation*, 7(25).
- Nida, E. & Taber, C. (1974). *The theory and practice of translation*. Leiden: Brill.
- Nord, C. (1997). *Translating as a purposeful activity: Functional approaches explained*, Manchester: St Jerome.
- Reiss, K. (1971). *Möglichkeiten und grenzen der Übersetzungskritik*, Munich: M. Hueber.
- Reiss, K. & H.J. Vermeer (1984). *Grundlegung einer allgemeinen translationstheorie*, Tübingen: Niemeyer.
- Turner, B., Lai, M., & Huang, N. (2010). Error deduction and descriptors—A comparison of two methods of translation test assessment. *Translation & Interpreting*, 2(1), 11-23.
- Waddington, C. (2004). Should student translations be assessed holistically or through error analysis? *Lebende Sprachen*, 49(1), 28-35.
- Ward, P., Willaims, A., & Hancock, P. (2006). Simulation for performance and training. In K. Ericsson, N. Charness, P. Feltovich & R. Hoffman (Eds.), *The cambridge handbook of expertise and expert performance* (pp.243-264). Cambridge: Cambridge University Press.
- Weir, C. (2005). *Language testing and validation*. Hampshire, New York: Palgrave Macmillan.
- Williams, M. (2001). The Application of Argumentation Theory to Translation Quality Assessment. *Meta*, 46(2), 326-344.

巻末資料 1 タスク例

Translation Brief: The following text was taken from a novel entitled “Cricket Kings” by William McInnes. The client is an Australian publishing company.

A game of cricket, in its strictest sense of classical clichéd description, should ebb and flow like a tide. There should be wings and roundabouts. You should take the good with the bad.

Ron Sparrow threw the ball to Christopher Andersen. Well, thought the old umpire, there hadn't been much ebbing and flowing. The swings had no roundabouts and from the Yarraville West Fourths' viewpoint there had been a grain of food to take with the sacks full of bad.

It was now very hot and that quiet, which is a characteristic of park cricket, began to fall.

Standing there in the middle of the oval on that hot summer afternoon, Ron Sparrow thought of how quiet it was out there. There was the talk and the noise of the players, but that was only close in. The slips and perhaps the cover fieldsmen. Yarra West always seemed to chat amongst themselves. It was to there in the outfield that the quiet would settle.

He look as Chris Andersen threw the ball to the chunky little right-armer called Matthew Halley.

The new batsman was taking guard. He asked for middle and Ron Sparrow gave it to him. The new batsman was the scorer for the Trinity team and so a new scorer had to be found amongst the Trinity contingent.

If the new scorer had any sense he wouldn't ask the name of the bowler. But most people who have never scored before or who are relatively new to the whole idea of filling in the cryptic sheets will cling to the set routine of the scorer's guidelines. Namely, to ask for the bowler's name when the new over commences.

This should be a pretty straightforward system. Cricketers by and large are a relatively conservative lot and closely adhere to what the appropriate behaviour should be. But there are some teams that, in a time-honoured tradition, will always give in to the temptation to, as the great expression readily describes, take the piss. Though these same teams will remain deadly serious in their endeavours on the field.

Ron Sparrow knew that Yarraville West Fourths were just this side. He called out to the batsman, 'Right an over, batsman.'

The batsman nodded. Matt Halley came and gave him his hat.

'What are you bowling, Matthew?' asked the old umpire. Matt Halley Smiled. He looked down at Livey. 'What am I bowling, Livey?'

Lively snorted. 'Give us a bit of the Kenny Loggins-cum-Kevin Bacon, mate.'

Word count: 315 words

Source: "*Cricket Kings*" by *William McInnes*

[注]

- Translation brief では、クライアントからの指示情報の大半を意図的に削除してあります
- 上記 ST のフォーマットは、原著に従った形になっております
- プロ日英翻訳者パネルとの協議の結果、本 ST には言語・論理的瑕疵を意図的に盛り込んだ部分がございます(下線部参照)

巻末資料 2 ルーブリック

Items	4	3	2	1
1) In referring to the student's translation, how precisely did the student reflect instructions by the client ?	The student <u>reflected all the requirements</u> given in the translation brief	The student <u>mostly reflected the requirements</u> given in the translation brief	The student <u>did not reflect part of the requirements</u> given in the translation brief	The student <u>could not reflect the requirements at all</u>
2) How well did the student identify problems posed by the task as shown in the relevant materials including their annotations.?	The student <u>identified all</u> the key problems posed by the task	The student <u>identified many of</u> the key problems posed by the task	The student <u>did not identify the majority of</u> the key problems posed by the task	The student <u>identified very few or no</u> key problems associated with the task
3) How acceptably did the student provide solutions for identified key problems derived from the task?	The student <u>dealt with all</u> the key problems in an acceptable way	The student <u>dealt with many of</u> the key problems in an acceptable way	The student <u>dealt with a few</u> key problems derived from the task in an acceptable way	The student <u>did not deal with any</u> key problems derived from the task in an acceptable way
4) How accurately did the student comprehend the content of the ST?	The student accurately comprehended <u>all the part of the ST</u>	The student accurately comprehend <u>most part of the ST</u>	The student <u>did not accurately comprehend many parts of the ST</u>	The student <u>did not accurately comprehend most or all of the ST</u>
5) How accurate was the student's translation in terms of the understanding of the text as a whole?	The student <u>fully comprehended the logical argument of the ST as a whole</u> in an accurate manner	The student <u>mostly comprehended the logical argument of the ST as a whole</u> in an accurate manner.	The student <u>did not comprehend many parts of the logical argument of the ST as a whole</u> in an accurate manner.	The student <u>did not comprehend the majority or all the parts of the logical argument of the ST as a whole</u> in an accurate manner
6) How well did the students maintain the stylistic features of	The student demonstrated <u>masterful domain-specific</u>	The student demonstrated <u>proficient domain-specific</u>	The student demonstrated <u>weak domain-specific knowledge</u>	The student demonstrated <u>poor domain-specific knowledge</u>

the ST in TT where necessary?	<u>knowledge</u> of the TT which resulted in <u>achieving an exceptionally readable translation</u>	<u>knowledge</u> of the TT which resulted in <u>achieving an acceptably readable translation</u> in the TT	of the TT with the result of <u>many parts</u> of the TT <u>being inappropriate</u> in relation to <u>readability</u>	of the TT with the result of <u>most parts</u> of the TT <u>being inappropriate</u> in relation to <u>readability</u>
7) How grammatically satisfactory was the student's command of the TL ?	The translation displayed <u>exceptional command</u> of <u>LOTE/English</u>	The translation displayed <u>acceptable command</u> of <u>LOTE/English</u>	The translation displayed <u>frequently inappropriate command</u> of <u>LOTE/English</u>	The translation displayed <u>poor command</u> of <u>LOTE/English</u>
8) Overall, how acceptably did the student achieve readability in the TT?	The student <u>fully achieved</u> the <u>readability</u> of the translation in the TL	The student <u>mostly achieved</u> the <u>readability</u> of the translation in the TL	The student <u>did not achieve</u> the <u>readability</u> of the translation in the TL and <u>many parts</u> of the translation <u>posed difficulty</u> in achieving <u>appropriate readability</u>	The student <u>did not achieve</u> the <u>readability</u> of the translation in the TL and <u>dominant portion</u> of the translation <u>posed difficulty</u> in <u>achieving appropriate readability</u>
9) How adequately did the student achieve readability for target readers in terms of the translation as a coherent text ?	Most or all of the parts of the translation <u>did not pose difficulty</u> in achieving a coherent TT.	<u>Many parts</u> of the translation <u>did not pose difficulty</u> in achieving a coherent TT	<u>Many parts</u> of the translation <u>posed difficulty</u> in achieving a coherent TT	<u>Most or all parts</u> of the translation <u>posed difficulty</u> in achieving a coherent TT
10) Overall, how appropriately did the student try to communicate with the client ?	The student <u>fully achieved</u> appropriate communication with the client	The student <u>very frequently achieved</u> appropriate communication with the client	The student <u>rarely achieved</u> appropriate communication with the client	The student <u>did not achieve</u> appropriate communication with the client <u>at all</u>
11) How effectively did the student attempt to justify translation decisions in a form of, for instance, translator's	The student demonstrated a <u>masterful ability</u> in attempting to justify translation decisions to the client	The student demonstrated a <u>proficient ability</u> in attempting to justify translation decisions to the client	The student demonstrated a <u>weak ability</u> in attempting to justify translation decisions to the client.	The student <u>did not justify any of</u> <u>translation decisions</u> that were necessary to do so for the client

notes?				
12) How appropriate did the student address culturally-challenging problems in the translation task?	<u>All of the student's solutions in addressing culturally-challenging problems were appropriate</u>	<u>Most of the student's solutions in addressing culturally-challenging problems were appropriate</u>	The student <u>frequently failed to address culturally-challenging problems</u> in an appropriate manner	The student <u>did not address culturally-challenging problems at all</u>
13) Overall, how appropriately did the student use available resources to collect necessary information to assist with the translation?	The student demonstrated a <u>masterful ability</u> in using available resources appropriately	The student demonstrated a <u>proficient ability</u> in using available resources appropriately	The student <u>did not use</u> available resources appropriately, and he/ she <u>frequently depended on his/her own decisions</u> <u>without reference to resources</u>	The student <u>did not use</u> available resources appropriately, and he/ she <u>very frequently depended on his/her own decisions</u> <u>without reference resources</u>

/52= /100

巻末資料 3 ルーブリック内使用の主要用語説明

An accurate comprehension of a ST: The extent to which a student comprehended a ST accurately. A failure to do so is likely to result in a mistranslated translation which does not reflect the content of a ST in a precise manner.

Client: A person or an organisation who wish to offer translation jobs to a translator. They can be an individual, company or a translation agency.

Command of LOTE/English: a translator's knowledge of LOTE/English and his/her ability to use such knowledge.

Communication with a client: ways in which the student attempted to communicate with a client. This would include translator-client negotiation (e.g. concerning a deadline, and the stylistic features of a TT), confirmation (e.g. the meaning of a term in abstract nature, client's stylistic preferences, and information about proper nouns) and the methods of communication (e.g. when and how the student communicated with the client). If such communication is failed, this is likely to have harmful effects on translator-client relationships.

Comprehension of the logical argument: The extent to which a student precisely comprehended the logical argument of an ST as a whole. The logical argument can be defined as logical relationships between sentences and paragraphs, forming a logically coherent text as a whole. Failure to do so would result in the misrepresentation of the content of a ST in a TT.

Culturally-challenging problems: problems that are derived from cultural differences between an ST and a TT, and where efforts are required to achieve the TT audience's comprehension of the translation. Examples would include culturally-specific terms (e.g. 'mate' in English) and culturally-specific concepts (e.g. *wabi* and *sabi* in Japanese). Failure to solve such problems is likely to result in a situation where a translator cannot accurately comprehend culturally-specific information in the ST and target readers find such information difficult to understand in the TT.

Domain-specific knowledge: Specialised knowledge with which a translator is required to be familiar in order to achieve the accurate understanding of the content of a ST and/or to achieve a translation which is legitimate for a specific kind of target reader. An example of domain-specific knowledge would be medical knowledge necessary in translating a ST in the medical domain. Lack of such knowledge is likely to result in the miscomprehension of the content of the ST.

Instructions by the client: Usually, a client provides a translator with various instructions regarding the translation job at the time of offering a job. Instructions would include a deadline, purposes of the translation job, kind of target readers, and formatting issues. In the case of assignments, students should be given a 'translation brief' for a task. If a student did not precisely reflect the instructions regarding kind of target readers, for example, this may result in producing a TT that makes it difficult for target readers to comprehend it easily.

Justification of translation decisions: providing a client with legitimate reasons and evidence of translator's decisions in translating. Such decisions would be typically realised by the use of translator's notes and/or emails and phone calls to the

client. Failure to do so is likely to result in giving unreliable impression of a translator to the client due to failure to make reasonable efforts in reaching such decisions.

Problem: any matter in a translation job which can be explicitly identified as a challenge for a translator. Unless a solution can be found, such a matter will remain, or deteriorate. In the context of translation, examples of such problems can include: difficulty in achieving an accurate translation to the ST, a translator being unfamiliar with the domain of the ST and/or TT, the content of a ST being contrary to translator's beliefs, or socially-accepted norms, or where ST contains culturally-specific references which do not exist in the TT culture.

Readability: The extent to which a TT can be read and understood by target audience naturally and easily. If target readers are medical specialists, for example, terms used in the TT need to be suitable terms used by the specialists. Failure to achieve appropriate readability is likely to result in a situation where target readers find the TT difficult to understand or the TT being unsuitable for expectations of readability held by target readers.

Stylistic feature: apart from the content of a text, its style is likely to provide readers with writer's intentions. Examples of stylistic features would include the presentation of a text, choices of words, and the developments of the argument of the text. Failure to maintain stylistic such features of a ST in a TT is likely to result in the misrepresentation of writer's intentions and therefore an inaccurate translation of a ST.

Use of available resources: any tool which provides a translator with useful information in gaining domain-specific knowledge required by an ST and/or a TT, and in finding appropriate equivalences between the ST and the TT. Examples of this would include a dictionary, Google search engine and various other websites (e.g. a company's website). Failure to use available resources, where necessary, is likely to result in the inappropriate choices of equivalences and mistranslations. Also, this could result in giving a negative impression of his/herself to the client.