

## 論文

# 言語学的カテゴリーに基づく翻訳QAスキームの分析

宮内拓也<sup>†</sup> 影浦峽<sup>†</sup>

<sup>†</sup> 東京大学大学院教育学研究科

「翻訳」をめぐる関連するアクター間の理解に齟齬をきたしているという問題が表出しており、その解決のためには翻訳プロセスを外在的なものとして高解像度でモデル化することが必須である。本稿では、その第一歩として、翻訳の出口側の最重要課題であると位置づけられる品質保証/品質評価について扱う3つのスキームを対象に、共通してみられる評価項目である「正確さ」、「流暢さ」、「用語」、「スタイル」という4項目の記述に関して、言語学的なカテゴリーとの対応を記述することで各スキームの記述解像度を分析した結果を報告する。その結果、「流暢さ」はきわめて多様な言語単位を対象としていたことから、さらに詳細に整理される必要があることが窺える。また、「正確さ」、「スタイル」についての記述は、本稿で想定するあらゆる言語単位を対象としており、「用語」についての記述は基本的に語句レベルを対象としていた。

キーワード：翻訳、品質保証、品質評価、翻訳プロセス

## 目次

- 1 はじめに
- 2 QA について
  - 2.1 QA の定義
  - 2.2 対象とする QA スキーム
- 3 評価視点について
- 4 言語学的カテゴリーとの対応
  - 4.1 言語学的カテゴリーについて
  - 4.2 「流暢さ」
    - 4.2.1 MQM
    - 4.2.2 ISO
    - 4.2.3 JTF
  - 4.3 「正確さ」
  - 4.4 「用語」
  - 4.5 「スタイル」
- 5 おわりに

## 1 はじめに

現在、翻訳をめぐるのは、翻訳論、翻訳者養成、実務翻訳、機械翻訳(MT)技術開発等の各領域において、知見や技術が充実してきている<sup>1</sup>。しかしながら、それらが適切に接続・統合できていないこと

から、翻訳実践や翻訳者養成において、これらの知見や技術が十分な理解のもとで適切に活用できているとは必ずしも言えない<sup>2</sup>。とりわけMT技術への理解が不十分なまま不用意な導入が進むと深刻な問題につながる恐れもある。例えば、クライアントが品質水準を理解しないまま翻訳サービスにMTの利用を求めた場合、重要な事項の誤りから事故を引き起こす恐れもある。また、ニューラルMT(NMT)が統計的MT(SMT)よりも「品質」がよいという単体の判断でMT+PE(Post Editing)にNMTを導入した結果、PEの作業負荷がかえって高くなるといった事態も考えられる。翻訳者養成における共通の質保証の不在から、一部で実務が要請するコンピテンスや品質とのミスマッチが生まれると、翻訳者の専門性に対する否定的な評価に繋がる恐れすらある。それは翻訳の「買い叩き」を加速させ、その結果として必要なコンピテンスと品質保証が業界全体として維持できなくなる事態さえ想定される。

こうした課題の背後には、単純だが根本的な問題が存在する。すなわち、そもそも、関与するアクターによる翻訳に対する理解が範囲も度合いも異なっており、「翻訳とは何か」についての共有理解が成立していないことである。その要因の一つに、翻訳のプロセスや規範・コンピテンスが共有

可能な外在的操作として具体的に記述されてこなかったことが挙げられる。

近年、翻訳への需要が増加の一途を辿る中で、Translation services – Requirements for translation services (ISO 17100:2015)<sup>3</sup>, Multidimensional Quality Metrics (MQM)<sup>4</sup>, Harmonized DQF-MQM Error Typology Template<sup>5</sup>, Standard Guide for Quality Assurance in Translation (F2575-14)<sup>6</sup>, JTF 翻訳品質評価ガイドライン<sup>7</sup>等、翻訳の品質管理基準がある程度具体的に整備され、活用されており、実務翻訳における品質の維持に貢献している。これらの品質管理基準は、主として実務翻訳従事者による利用を想定しており、翻訳に関与する概念を理解していることを前提とした記述・説明となっている。また、記述の解像度・具体性が、同一管理基準の中でも必ずしも同一でない。このことは、派生的に次のような問題を生む要因の一つとなっている。第一に、実務翻訳者養成において翻訳の品質管理基準を活用する場合、学習者の知識レベルに応じて品質管理基準の説明を展開する必要があるが、それを支える共通の枠組みや言葉が十分に共有されていないことである。第二に、MQMのように機械翻訳利用も考慮した枠組みがあるにもかかわらず、これらの基準がMTの評価にはあまり使われていないことである<sup>8</sup>。翻訳の品質管理が翻訳の出口側の最重要課題であると位置づけられていること<sup>9</sup>、翻訳者養成において実務翻訳やMTにおける品質管理の理解が重要課題の一つに挙げられていること<sup>10</sup>、MTの実利用においてBLEU<sup>11</sup>やNIST<sup>12</sup>といったMT向け自動評価ではなく実務翻訳の品質管理基準に基づく評価が重要であること<sup>13</sup>、日本語翻訳も視野に入れた標準的な翻訳品質評価手法が望まれていること<sup>14</sup>を踏まえれば、品質管理基準における記述の解像度・具体性を検討することは、確立した品質管理基準の適用範囲を広めるとともに一貫した適用を確保するための一歩として、重要である。

こうした状況を踏まえ、本研究では、翻訳プロセスにおける品質保証/品質評価(以下、QA)のスキームを対象に、その記述の分析を行う。記述を分析する際の基準としては、言語学において定式化されてきた概念とカテゴリーを用いる。翻訳は文書を対象とするものであり<sup>15</sup>、言語学的な概念で説明し尽くせるものではないが<sup>16</sup>、そうである

とはいえ、文書は(主に)言語表現から構成される。一方、言語学は、それが射程とする範囲で言語についての体系を構築しており、言語、およびその下位要素について語るための体系的かつ比較的明確なカテゴリーを提供している。したがって、QAスキームの検討と整理において言語学的なカテゴリーを用いることには大きく二つの意義が認められる。第一は、QAスキームの整理において一定の体系的な指針となることである。第二は、より翻訳という行為の内容に関わることであるが、言語学的なカテゴリーとの対応付けがどの程度可能であるかを見ることで、QAスキームの枠組みにおいて、翻訳がどの程度まで言語学が扱う意味での言語的なものでありどの程度がそうでないかを明確に把握する一助となることである。

以下、2ではQAについて、3では翻訳品質の評価の際の観点について、それぞれ概要を述べる。4では、各QAの記述がどのような言語学的カテゴリーを対象としているのかを記述する。5は本稿全体のまとめである。

## 2 QAについて

### 2.1 QAの定義

翻訳の品質管理に関する用語として、品質保証(Quality Assurance)と品質評価(Quality Assessment)がある<sup>17</sup>。MQM Definition<sup>18</sup>によれば、それぞれ以下のように定義される。

- 品質保証  
Quality assurance refers to ways of preventing mistakes or defects in manufactured products and avoiding problems when delivering solutions or services to customers. Quality assurance relies on continual assessment of quality.
- 品質評価  
Quality assessment (sometimes considered synonymous with “quality evaluation”) is the measurement of the extent to which a product complies with quality specifications [...].

これらの定義から、品質保証は翻訳において誤りや欠陥を防ぐ方法を指しており、実際に翻訳者がとる(べき)プロセスを扱うものである。一方、品

質評価は翻訳が仕様に準拠している程度の測定を指しており、あくまで翻訳成果物の評価について扱うものである。ただし、品質評価は、品質保証に包含され得る。翻訳成果物の仕様への準拠の程度を測定する、成果物中の誤りを同定、分類するといった品質評価の作業は、翻訳における誤りや欠陥を防ぐプロセスの内部、つまり品質保証のプロセスの内部に位置づけられ得るためである。

翻訳プロセスをモデル化するにあたっては、翻訳成果物の評価である品質評価の記述をプロセスとしての品質保証の記述に書き換えることが必要であり、翻訳プロセスモデルの構築という点から、これらの差異に意識を払うことは意義がある。しかし、本稿では、これらの差異は取り扱わず、同一の観点から分析を行う。本稿は QA 記述を言語学的カテゴリーと対応付けることを目的としており、その場合にはむしろ品質評価と品質保証を同一の観点から分析する必要がある。なぜなら、言語学的カテゴリーを軸に品質評価と品質保証を同様に扱うことで、それらのより正確な同異、およびそれらの対応関係の把握につながるからである。それは、品質評価の記述を品質保証の記述に書き換えるため、さらに翻訳プロセスをモデル化するために必要な過程であると捉えることができる。

## 2.2 対象とする QA スキーム

本稿では、代表的な QA スキームとして、具体的に以下に挙げた 3 つを分析対象とする。

- Translation services – Requirements for translation services (ISO 17100:2015)
- Multidimensional Quality Metrics
- JTF 翻訳品質評価ガイドライン

以下、それぞれ ISO, MQM, JTF と言及する。これら 3 つを取り上げるのは、それぞれ、ISO は翻訳サービスを包括的に扱うものとしてメジャーなものであるため<sup>19</sup>、MQM はより具体的に解像度の高いものを提供しているため、JTF は日本語の翻訳を評価する際に用いることを想定しているためである。

## 3 評価視点について

翻訳品質の評価については、大半の企業が独自に作成した評価モデルによる評価項目を設けていることからさまざまな評価項目が用いられているとされて

いる<sup>20</sup>。しかし、同時に、問題点の種類を表す用語として「正確さ (Accuracy)」、「流暢さ (Fluency)」、「用語 (Terminology)」、「スタイル (Style)」の 4 項目に関しては、ほぼすべての評価モデルで用いられていることが指摘されている<sup>21</sup>。本稿で分析対象とする ISO, MQM, JTF についても、すべてに共通する評価視点として上記の 4 項目が挙げられる。これらは、MQM Definition では、以下のように定義されている。

- Accuracy:  
The target text does not accurately reflect the source text, allowing for any differences authorized by specifications.
- Fluency:  
Issues related to the form or content of a text, irrespective as to whether it is a translation or not.
- Terminology:  
A term (domain-specific word) is translated with a term other than the one expected for the domain or otherwise specified.
- Style:  
The text has stylistic problems.

本研究でもこれら 4 項目について取り上げ、分析をする。本稿では、紙幅の都合上、特に様々な言語学的カテゴリーとの対応が見られる「流暢さ」についての詳細を記す。「正確さ」、「用語」、「スタイル」については、比較的簡便に分析結果を紹介した上で、特筆すべきことを述べるに留める。

## 4 言語学的カテゴリーとの対応

### 4.1 言語学的カテゴリーについて

本稿で言語学的カテゴリーと呼ぶものは、統合関係 (rapport syntagmatique)<sup>22</sup> 的視点から見た諸単位を指すものとする。すなわち、各単位はその上位カテゴリーの単位に対し部分を構成し、下位カテゴリーの単位に対し全体を構成することになる。具体的には以下のカテゴリーの階層を想定する<sup>23</sup>。

- 音素 < 形態素 < 語 < 句 < 文 < 文章

ここで、本稿では翻訳の QA スキームを扱うということから、基本的に音声言語ではなく、書記

言語が想定される。そのため、本稿では音素は対象としない。また、実際に翻訳という行為について、その操作の対象となるのは、一般的に所与のものとして考えられる語というカテゴリー以上のものであると考えられることから、形態素も対象としない。よって、本稿において、QA スキームの対象として記述するカテゴリーの階層は以下の通りとなる：

- 語 < 句 < 文 < 文章

なお、必要に応じて上記の言語学的カテゴリー以外のもの(記号、文字<sup>24</sup>、文書構成要素)を含める場合もある。

以下、ここで示した言語学的なカテゴリーと QA スキームの記述との対応付けがどの程度可能であるかを見ることで、翻訳がどの程度まで言語的なものでありどの程度そうでないかを把握するという観点から、各 QA スキームの記述解像度を分析する。

#### 4.2 「流暢さ」

以下、「流暢さ」について各 QA スキームの記述、および言語学的カテゴリーとの対応をまとめ、それぞれに説明を加える。

##### 4.2.1 MQM

MQMにおいて、「流暢さ」に関してはいくつもの下位分類がある<sup>25</sup>が、その下位分類の記述の言語学的カテゴリーとの対応は表1の通りである<sup>26</sup>。表1には当該項目の判定の際の決定因、起点テキスト(source text; ST)と目標テキスト(target text; TT)のどちらを適用対象とするかについてもまとめている。なお、ここでは、MQM Definitionにおいて、Fluencyの子項目となっているものを挙げた<sup>27</sup>。以下、表1に沿って説明を加える。

ambiguity は以下の通り定義されており、意味にかかわるものであることがわかる。

- ambiguity:  
The text is ambiguous in its meaning.

意味を持つ最小単位は形態素であるため、形態素以上のすべての言語単位が対象となっている。また、曖昧性には、語彙的意味に起因する語彙的曖昧性(lexical ambiguity)<sup>28</sup>と統語構造に起因する構造的曖昧性(structural ambiguity)<sup>29</sup>がある。よっ

て、前者の決定因は語であり、後者の決定因は統語構造を持つ句、文である。

character-encoding は以下のように定義されている。

- character-encoding:  
Characters are garbled due to incorrect application of an encoding.

これは文字コードの問題であり、対象は文字のレベルである。決定因については、当該言語で用いられない文字が表記されている場合などは文字となるが、そうではない場合には語や句のレベルで判定されることになる。

coherence、および cohesion は以下のように定義されている。

- coherence:  
The text is not coherent.
- cohesion:  
Portions of the text needed to connect it into an understandable whole (e.g., reference, substitution, ellipsis, conjunction, and lexical cohesion) are missing or incorrect.

これらは文章のレベルで決定されるものであるため、文章を対象とした記述である。ただし、coherence に関して、その決定因は何らかの文や句、語といった単位となり得る。cohesion については、例えば、論理関係を示すための特定の語の不足などによって引き起こされることがあるため、その決定因となるものとしては、語や句なども挙げられる。

corpus-conformance は以下のように定義されている。

- corpus-conformance:  
The content is deemed to have a level of conformance to a reference corpus. The non-conformance type reflects the degree to which the text conforms to a reference corpus given an algorithm that combines several classes of error type to produce an aggregate rating.

これはあらゆる言語単位を対象とした記述であると考えられる。既存の参照コーパスに応じて判断

表 1: MQM における「流暢さ」の下位項目と言語学的カテゴリーとの対応

項目	記述対象	決定因	適用対象
ambiguity	語, 句, 文, 文章	語, 句, 文	ST, TT
character-encoding	文字	文字, 語, 句	ST, TT
coherence	文章	文章	ST, TT
cohesion	文章	文章	ST, TT
corpus-conformance	語, 句, 文, 文章	/	ST, TT
grammar	語, 句, 文	語, 句, 文	ST, TT
grammatical-register	語, 句, 文	句, 文, 文章	ST, TT
inconsistency	語, 句, 文, 文章	句, 文, 文章	ST, TT
index-toc	文書構成要素	/	ST, TT
link-crossreference	文書構成要素	/	ST, TT
non-allowed-characters	文字	文字	ST, TT
pattern-problem	句読点	/	ST, TT
sorting	文書構成要素	/	ST, TT
spelling	語	語, 句, 文, 文章	ST, TT
typography	記号, 文字 (, 語, 句)	語, 句, 文, 文章	ST, TT
unintelligible	/	/	ST, TT

されるため, その決定因は言語学的カテゴリーと対応付けることはできない。

grammar は以下のように定義されている。

- grammar:  
Issues related to the grammar or syntax of the text, other than spelling and orthography.

これは文法, および統語法に関するものであり, 語, 句, 文を対象とした記述である。

grammatical-register は以下のように定義されている<sup>30</sup>。

- grammatical-register:  
The content uses the wrong grammatical register, such as using informal pronouns or verb forms when their formal counterparts are required.

これについては, MQM は具体例として語に関するものを提示しているが, 文法上のレジスターにかかわる問題は句や文の単位でも起こり得る。

inconsistency は以下のように定義されている。

- inconsistency:  
The text shows internal inconsistency.

矛盾をきたすことは, 理論的にはあらゆる単位で起こり得ることから, 語, 句, 文, 文章が記述の対象となっている。矛盾を判定する際には, 語単体で判定されることはなく, 複数の語, およびそれ以上の単位を参照する必要がある。そのため, 決定因は語より大きな句, 文, 文章となる。

index-toc, link-crossreference, sorting は以下のように定義されている。

- index-toc:  
Issues related to an index or Table of Contents (TOC).
- link-crossreference:  
A link or cross reference points to an incorrect or nonexistent location.<sup>31</sup>
- sorting:  
A list is not in the appropriately collated sequence.

これは目次, リンクおよびリストに関わるものであり, 言語学的カテゴリーではない文書構成要素<sup>32</sup>を対象とした記述である。

non-allowed-characters は以下のように定義されている。

- non-allowed-characters:

The text includes characters that are not allowed.

これは文字、記号を対象とした記述であるが、実際にこれについて判断するためには、社会文化規範や仕様書などの言語外要素を参照する必要がある。

pattern-problem は以下のように定義されている。

- pattern-problem:

The text contains a pattern (e.g., text that matches a regular expression) that is not allowed.

これは句読点などの記号を対象とした記述である<sup>33</sup>。

spelling は以下のように定義されている。

- spelling:

Issues related to spelling of words.

語の表記上の問題を扱っており、記述の対象は語である。ただし、ある語の表記の誤りを指摘する場合、その語が実在しないものであれば語単体で判定が可能となるが、その語が実在するならば句や文、文章を決定因として判定せざるを得ない<sup>34</sup>。

typography は以下のように定義されている。

- typography:

Issues related to the mechanical presentation of text. This category should be used for any typographical errors other than spelling.

これは spelling に含まれない表記上の問題を扱っていると言え、記述の対象は記号や文字であるが、その判定には語や句、文、文章を参照しなければならない。

unintelligible は以下のように定義されている。

- unintelligible:

The exact nature of the error cannot be determined. Indicates a major break down in fluency.

これについては、判読不能部分の特定ができない限り、対象も決定因も特定することはできない<sup>35</sup>。

#### 4.2.2 ISO

ISO から「流暢さ」に関する記述を抽出すると、具体的には syntax, spelling, lexical cohesion の 3 項目についての評価となっている。これらの記述の言語学的カテゴリーとの対応は表 2 の通りとなる。

表 2: ISO における「流暢さ」についての記述と言語学的カテゴリーとの対応

項目	記述対象	決定因
syntax	句、文	句、文
spelling	語	語、句、文、文章
lexical cohesion	語	句、文、文章

syntax は統語法について記述したものであることから、記述の対象は語、句、文である。

spelling は語の表記の問題であるため、対象は語である。

lexical cohesion については、基本的に語を対象としているが、語の結束性は語単体では定義され得ず、必ず句や文、文章を参照する必要がある。

#### 4.2.3 JTF

JTF のエラーカテゴリーには、概要レベルと詳細レベルの 2 レベルがあり、「流暢さ」に関する詳細レベルの記述と言語学的カテゴリーとの対応は表 3 の通りである<sup>36</sup>。以下、表 3 に沿って説明を加える。

「誤入力」は以下の通り定義されている。

- 「誤入力」:

入力時に不注意で発生したと思われる誤りがある。

これは入力時の誤りについての記述であり、その対象となるのは語や記号であると判断できる<sup>37</sup>。誤入力であることの決定は語や記号単体で行われることもあり得るが、それ以上の単位、すなわち句や文を参照しなければならない場合もある。

「誤字」は以下の通り定義されている。

- 「誤字」:

ひらがな、カタカナ、漢字、句読点の表記や、欧文のスペルに誤りがある。

これはスペルミスを含む誤字について記述しており、対象となる単位は基本的に語である。ただし、

表 3: JTF における「流暢さ」の詳細レベル項目と言語学的カテゴリーとの対応

項目	記述対象	決定因
誤入力	語, 記号	語, 句, 文, 記号
誤字	語, 記号	語, 句, 文, 記号
同音異義語誤り	語 (, 文字)	句, 文, 文章
文法誤り	語, 句, 文	語, 句, 文
誤用	語, 句, 文	句, 文, 文章
コロケーション誤り	語 (, 句, 文)	句 (, 文, 文章)
待遇表現誤り	語, 句	文, 文章
不整合	語, 句, 文, 文章	句, 文, 文章
あいまい	句, 文	句, 文
読解不能	/	/
その他	/	/

ここには句読点などの記号の誤りも含まれる。誤字であることの決定も語や記号単体で行われることもあり得るが、句や文を参照しなければならない場合もあり得る。

「同音異義語誤り」は以下の通り定義されている。

- 「同音異義語誤り」:  
同音異義語の誤りがある。

これについては、基本的に語のレベルでの問題であると考えられるが、単なる文字表記上の問題と本質的には区別できない。「誤字」や「誤入力」とは異なり、この場合は誤った語も同音異義語として存在するため、この誤りを認定するためには、当該の誤りを含む文、文章を参照する必要がある。

「文法誤り」は以下の通り定義されている。

- 「文法誤り」:  
文法上の誤りがある。

これは文法上の誤りを扱うものであるため、記述の対象は語、句、文である。

「誤用」は以下の通り定義されている。

- 「誤用」:  
本来の意味とは違う意味で使われている。

これは意味に関するあらゆる誤用を含むものであり、意味を持つ最小単位である形態素以上の語、句、文を対象としている<sup>38</sup>。

「コロケーション誤り」は以下の通り定義されている。

- 「コロケーション誤り」:  
コロケーション (語どうしの慣用的なつながり) に誤りがある。

コロケーションについての記述であるため、その対象は語であるが、本質的にコロケーションに関して判定をする際は句や文を参照しなければならない。コロケーションに慣用表現を含むと考えるなら、当該表現が慣用的に用いられているのか、そうでないか (構成的意味計算可能な意味で用いられているのか) の判定には、必ず文脈を参照しなければならないことになる。

「待遇表現誤り」は以下の通り定義されている。

- 「待遇表現誤り」:  
敬語などの待遇表現が文脈上適切ではない。

これは待遇表現についての記述であることから、語、および句が記述の対象である。ただし、適切な待遇表現が用いられているか否かの判定には、文や文章を参照する必要がある。

「不整合」については以下の通り定義されている。

- 「不整合」:  
同一のテキスト内に、記述の不整合がある。

不整合性は、語、句、文、文章といったあらゆる単位で確認され得る。

「あいまい」は以下の通り定義されている。

- 「あいまい」:  
係り受けなどが原因で、さまざまに解釈可能されうる部分がある。

これは曖昧性についての記述である。ここでは「係り受けなどが原因」となるような統語的曖昧性について記述されているため、その対象となるカテゴリは句、文となる。

「読解不能」については以下の通り定義されており、判読不能である部分の対象の特定はできない<sup>39</sup>。

- 「読解不能」：  
言葉として読めない部分がある。

また、「その他」以下の通り定義されているが、これについてもその他のあらゆる問題が含まれるため、対象の特定はできない。

- 「その他」：  
流暢さに分類するのが妥当なエラーで、上記のいずれにも該当しないもの。

以上、MQM、ISO、JTFにおける「流暢さ」の下位項目の記述と言語学的カテゴリの対応を見たが、各項目に応じてきわめて多様な語単位を対象としていたことが分かる<sup>40</sup>。

以下、「正確さ」、「用語」、「スタイル」について比較的具体的に記述されているMQMとJTFの記述をまとめ、特に言及が必要なものについて若干の説明を加える。

#### 4.3 「正確さ」

MQM、JTFの「正確さ」についての下位項目と言語学的カテゴリとの対応は表4、および表5の通りである。

これらは基本的に文章を対象とした記述であり、対応する言語学的カテゴリは基本的には文章となると判断できる。しかし、実際にはすべての項目に関して、語、句、文を含むこともあり得る。また、「正確さ」に関してはすべてSTを参照することで判断されるため、決定因となるのは、基本的にST(内の語、句、文、文章)となる。

#### 4.4 「用語」

「用語」については、すべての記述が専門用語、固有名詞に関するものであるため、言語学的カテゴリとしては、語、ないしは句に対応する。MQM、JTFの「用語」についての下位項目と言語学的カテゴリとの対応は表6、および表7の通りである<sup>41</sup>。

これらは、語、句に対応する記述である。term-inconsistency、および「用語不統一」は当該の文、文章内での不統一の問題であることから、その決定因は文、文章となる。termbase、domain-terminology、および「指定用語違反」、「特定分野用語違反」については、仕様書や当該分野といった言語外要素を参照することで判断されるため、その決定因は言語学的カテゴリと対応づけられない。

#### 4.5 「スタイル」

「スタイル」は基本的に文章全体から決まるものであるため、その記述は文章を対象としている。MQM、JTFの「スタイル」についての下位項目と言語学的カテゴリとの対応は表8、および表9の通りである。

これらの記述において、基本的に対応する言語学的カテゴリは文章であるが、unidiomaticについてはより小さい単位である語、句、文も対象となっている。求められるスタイルは基本的に仕様書や当該分野などの言語外要素によって決まるため、決定因については言語学的カテゴリと対応しない。ただし、awkwardについては、文章内の各言語学的単位が決定因となる。inconsistent-styleについても、文や文章から統一性が判断される。

### 5 おわりに

以上、本稿では、翻訳プロセスを高解像度で語るための第一歩として、ISO、MQM、JTFの3つのQAスキームを対象に、「正確さ」、「流暢さ」、「用語」、「スタイル」という4つの項目に関して各記述と言語学的カテゴリとの対応を記述することでQAの記述解像度を分析した。本稿で特に重点的に記述した「流暢さ」については、きわめて多様な言語単位を対象としていたことから、さらに詳細に整理される必要があることが窺える。「正確さ」、「スタイル」についての記述は、あらゆる言語単位を対象としており、「用語」についての記述は基本的に語、句レベルを対象としていた。翻訳プロセスについて解像度が高いモデルを構築していくために、今後は扱うQAスキーム、および評価視点を追加し、非言語学的カテゴリも含め、詳細に分析していく必要がある。また、本稿をある種の叩き台として、実務翻訳のQA記述について議論が惹起され、今後より活発化されることが



表 4: MQM における「正確さ」の下位項目と言語学的カテゴリーとの対応

項目	記述対象	決定因	適用対象
addition	語, 句, 文, 文章	ST	TT
improper-exact-tm-match	語, 句, 文, 文章	/	TT
mistranslation	語, 句, 文, 文章	ST	TT
omission	語, 句, 文, 文章	ST	TT
over-translation	語, 句, 文, 文章	ST	TT
under-translation	語, 句, 文, 文章	ST	TT
untranslated	語, 句, 文, 文章	ST	TT

表 5: JTF における「正確さ」の詳細レベル項目と言語学的カテゴリーとの対応

項目	記述対象	決定因
抜けと余分	語, 句, 文, 文章	ST
誤訳	語, 句, 文, 文章	ST
抜けと余分	語, 句, 文, 文章	ST
未翻訳	語, 句, 文, 文章	ST
その他	/	ST

期待される。

## 謝辞

本研究は JSPS 科研費 (課題番号: 19H05660, 19K23073) の助成を受けている。

また、本稿の内容は「言語処理学会第 26 回年次大会」(2020 年 3 月 17 日)における口頭発表「翻訳における QA 記述の分析: 言語学的カテゴリーを手掛かりに」に基づき、その内容を修正、増補したものである。発表の際には参加者の方々から貴重なご意見を頂いた。ここに感謝申し上げる。

## 注

- 1) 例えば、翻訳論においては、翻訳規範とコンピテンスに関する理論的知見が蓄積されてきているし、翻訳者養成に関しては、世界的に大学院レベルでの翻訳者養成コースは増えていることが挙げられる。翻訳実務においては、翻訳産業の拡大に伴い、実務プロセスが MT 等の利用を含め多様化している。MT 技術については、ニューラル MT (NMT) の精度と流暢さの向上が挙げられる。
- 2) MT の不適切な利用、翻訳教育が依然として

主に実習と講師による朱入れで構成されていることなど。

- 3) International Organization for Standardization. “Translation services – requirements for translation services (ISO 17100:2015),” 2015.
- 4) Deutsches Forschungszentrum für Künstliche Intelligenz GmbH and QT-LaunchPadn. “Multidimensional quality metrics,” 2015.
- 5) 以下を参照のこと。  
TAUS. “Harmonized DQF-MQM Error Typology,” available from <https://www.taus.net/qt21-project/#harmonized-error-typology>. (accessed date: 2020-12-25)
- 6) ASTM International. “Standard guide for quality assurance in translation (ASTM F2575-14),” 2014.
- 7) 日本翻訳連盟 “JTF 翻訳品質評価ガイドライン” 2018.
- 8) 例えば、以下を参照のこと。  
Lommel, Arle and Alan Melby. “Tutorial: MQM-DQF: A good marriage (translation quality for the 21st century),” in *Proceedings of the 13th Conference of the Association for Machine Translation in the Americas (Volume 2: User Track)*, Boston, MA, Association for Machine Translation in the Americas, 2018.  
むろん、MT の評価に実務翻訳の品質管理基準が使われていないのは、第一義的には品質管理基準の記述の問題ではなく MT 研究側の

表 6: MQM における「用語」の下位項目と言語学的カテゴリーとの対応

項目	記述対象	決定因	適用対象
term-inconsistency	語, 句	/	ST, TT
termbase	語, 句	/	ST, TT
domain-terminology	語, 句	/	ST, TT

表 7: JTF における「用語」の詳細レベル項目と言語学的カテゴリーとの対応

項目	記述対象	決定因
用語不統一	語, 句	文, 文章
指定用語違反	語, 句	/
特定分野用語違反	語, 句	/
その他	/	/

問題である。しかしながら、翻訳の品質管理基準が MT 研究者にとって使いにくい理由として手間だけでなく一貫した適用の困難がある。これについては以下を参照のこと。

Way, Andy. “Quality expectations of machine translation,” in Moorkens J., Castilho S., Gaspari F., Doherty S. eds. *Translation Quality Assessment: From Principles to Practice*. Cham, Springer, 2018, p. 159–178.

- 9) 日本翻訳連盟 “翻訳品質評価方法に関する業界アンケート: 結果報告” 2016.
- 10) European Master’s in Translation. “Competence framework 2017,” 2017.
- 11) Papineni, Kishore, Salim Roukos, Todd Ward and Wei-Jing Zhu. “BLEU: a method for automatic evaluation of machine translation,” in *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics*, Philadelphia, PA, Association for Computational Linguistics, 2002, p. 311–318.
- 12) Doddington, George. “Automatic evaluation of machine translation quality using N-gram co-occurrence statistics,” in *Proceeding of the Second International Conference on Human Language Technology Research*, San Francisco, CA, Morgan Kaufmann Publish-

ers, 2002, p. 138–145.

- 13) Panić, Milica. “DQF-MQM: Beyond automatic MT quality metrics,” TAUS Blog, 2019, available from <https://blog.taus.net/dqf-mqm-beyond-automatic-mt-quality-metrics>. (accessed date: 2020-12-25)
- 14) 日本翻訳連盟, *op. cit.*, (2018)
- 15) これについては以下を参照のこと。  
影浦峽 “人間の翻訳と機械の翻訳 (1): 翻訳者は何を翻訳しているか?” 『AAMT ジャーナル』 no. 71, 2019, p. 14–19.
- 16) Oustinoff, Michaël. 『翻訳: その歴史・理論・展望』 [*La traduction*. Paris, Presses Universitaires de France, 2003] 服部雄一郎訳, 白水社, 2008.  
Kyo Kageura. “The status of documents and related concepts in translation and in library science,” in *Proceedings of the 9th Asia-Pacific Conference on Library and Information Education and Practice*, 2019.  
なお, European Master’s in Translation, *op. cit.*, (2017) においては, 翻訳は言語的プロセスを前提とすることが示されている。具体的には, 翻訳する言語ペアのそれぞれに関して CEFR (Common European Framework of Reference for Languages) の C1 レベルが要求されている。CEFR の C1 レベルとは, 例えば, 国際コミュニケーション英語能力テスト (TOEIC) において, Listening で 490 点, Reading で 455 点に相当する。また, 国内の試験に関しては, 実用英語技能検定 (英検), 実用フランス語技能検定試験 (仏検), ロシア語能力検定試験 (露検) で 1 級に相当するとされる。CEFR については以下を参照のこと。  
Council of Europe. *Common European*

表 8: MQM における「スタイル」の下位項目と言語学的カテゴリーとの対応

項目	記述の対象	決定因	適用の対象
awkward	文章	語, 句, 文, 文章	TT
company-style	文章	/	ST, TT
inconsistent-style	文章	文, 文章	ST, TT
register	文章	/	ST, TT
third-party-style	文章	/	ST, TT
unidiomatic	語, 句, 文, 文章	/	ST, TT

表 9: JTF における「スタイル」の詳細レベル項目と言語学的カテゴリーとの対応

項目	記述の対象	決定因
指定スタイル違反	文章	/
特定分野スタイル違反	文章	/
スタイル不統一	文章	/
その他	/	/

*Framework of Reference for Languages: Learning, teaching, assessment.* Cambridge, Cambridge University Press, 2001.

また, CEFR と各言語の検定試験との対応は, 以下を参照のこと。

Tannenbaum, J. Richard and E. Caroline Wylie. *Linking English-Language Test Scores Onto the Common European Framework of Reference: An Application of Standard-Setting Methodology.* Princeton, NJ, Educational Testing Service, 2008.

Dunlea, Jamie. “英検と CEFR の関連性について Part 1”『STEP 英語情報 11・12月号』, 2009, p. 6-10.

Dunlea, Jamie. “英検と CEFR の関連性について Part 2”『STEP 英語情報 1・2月号』, 2010, p. 4-6.

富森伸夫 “フランス語能力検定試験 (DELTA / DALF, TCF, DAPF) と日本におけるフランス語教育” 〈富森伸夫編『EU および日本の高等教育における外国語教育政策と言語能力評価システムの総合的研究 (科学研究費補助金基盤研究 B 研究プロジェクト報告書)』東京外国語大学, 2012.〉 p. 171-186.

中澤英彦 “ロシア語検定試験について” 〈富森伸夫編『EU および日本の高等教育における外国語教育政策と言語能力評価システムの総合的研究 (科学研究費補助金基盤研究 B 研究プロジェクト報告書)』東京外国語大学, 2012.〉 p. 149-169.

17) 本稿では, 「QA」という用語はこれらのカバータームとして用いる。

18) 以下を参照のこと。

Deutsches Forschungszentrum für Künstliche Intelligenz GmbH. “Multidimensional Quality Metrics (MQM) Definition,” 2015, available from <http://www.qt21.eu/mqm-definition/definition-2015-06-16.html>. (accessed date: 2020-12-25)

19) 例えば, 以下を参照のこと。

市村美樹子 “ISO17100 の衝撃: 知らなければ世界から取り残される『翻訳の国際常識』”『日本翻訳ジャーナル』 vol. 267, 2013, p. 10-14.  
井上孝, 長田孝治, 石崎俊 “翻訳サービスの品質管理 ISO17100 にみる翻訳プロジェクトの国際標準化”『情報管理』 vol. 58, no. 9, 2015, p. 666-672.

20) 田中千鶴香 “実務翻訳における翻訳品質評価プロセス”『言語処理学会第 23 回年次大会発表論文集』2017, p. 489-492.

21) *loc. cit.*

22) 要素がひとまとまりの連鎖として配列される際における要素間の関係を指す。この関係は言語の線条性 (linearity) からの要請である。以下を参照のこと。

Ferdinand de Saussure. *Cours de linguistique général.* Paris, Payot, 1916 [1972].

- 23) ここで「<」の記号は階層における上下関係を示すものとする。以降も同様である。
- 24) 文字論を言語学の下位分野に位置づけるか否かについては議論があるが、ここでは取り扱わない。
- 25) 各下位項目の定義は MQM Definition を参照のこと。
- 26) 表中の「/」は当該要素が決定できないこと、および言語学的カテゴリーではないことを示す。以降の表についても同様である。
- 27) MQM では、項目間の親子関係に不整合が見られる部分がある。MQM Definition において、例えば Fluency の部分ではその子項目として duplication は記されていないが、duplication の部分ではその親項目として Fluency が挙げられている。本稿では、Fluency の部分で子項目として記述されている項目について記す。ただし、Fluency の部分で子項目として Style も挙げられているが、Style は Fluency と同じレベルの項目として扱うのが適切であると考えられるため、Fluency の子項目としては扱わない。
- 28) 例えば、以下の文には語彙的曖昧性がある。
- そこに穴が開いている。
- この文は「底に穴が開いている」と「その場所に穴が開いている」の両方の意味に解釈可能であるが、それは「そこ」という語が2つの意味に解釈可能であるために起こっている。
- 29) 例えば、以下の文には統語的曖昧性がある。
- 自転車で走る君を追いかけた。
- この文は以下の2つの統語構造が想定される。
- [自転車で [[走る君を] 追いかけた]]  
[[[自転車で走る] 君を] 追いかけた]]
- この統語構造に応じて自転車に乗っているのが「君」であるのか、話者であるのかという2つの解釈が生じる。
- 30) grammatical-register の項目では親が Accuracy になっているが、Fluency の項目では子として grammatical-register が含まれている。
- 31) これは厳密には broken-link の定義である。ただし、Fluency の子項目として link-crossreference が挙げられているものの、link-crossreference の項目は MQM Definition にはない。broken-link の項目は親項目を Fluency としているが、Fluency の子項目として挙げられてはいない。内容から考えれば、broken-link と link-crossreference は同一視してよいと思われる。
- 32) 文書構成要素がその下位にいかなる要素を包含するかについての具体的な詳細は、まだ未整備であるが、起点言語文書(source document) について文献レビューを進めており、その中で取り扱っている。なお、これについては以下を参照のこと。
- 宮田玲, 宮内拓也, 影浦峯 “翻訳のための起点文書分析: 文献レビューの枠組み” 『言語処理学会第 26 回年次大会発表論文集』2020, p. 533-536.
- 33) ただ、定義から考えれば、音素配列論的な (phonotactic) パターンまで含まれることになり、spelling と本質的に区別できなくなる。実際には、あくまで記号(等)についての記述であると考えられる。
- 34) 実際には、語は閉じたクラスをなさないため、(特定の辞書内にといた条件がない限り) 原理的に「実在しない語」と「実在する語」の区別はできず、spelling の問題の指摘は句や文、文章を参照せざるを得ない。
- 35) 判読不能部分自体はあらゆる言語的単位である可能性がある。ただし、この unintelligible についての記述(定義)には対象とする言語単位についての記述は明示的にも暗示的にも含まれておらず、また、ある判読不可能部分があった際、当該部分がどの言語単位であるかの判定は(判読不可能であるがゆえに)できないと考えられる。よって、ここでは「特定はできない」とした。その他の項目については、暗示的であったとしても、その記述において対象となる言語単位が示されている。この点が unintelligible と他の項目の違いであり、こういった意味で unintelligible についての記述は他の項目の記述と比して解像度が低いと言

える。例えば, inconsistency については, 「理論的にはあらゆる単位で起こり得る」としているが, これは記述(定義)に言語対象として“text”(文章)が明示的に示されており, 実際には文章の矛盾は, 文章を構成するあらゆる要素が引き起こし得ることを考えた上でのことである。

- 36) なお, JTF は, MQM と一定の対応がある。
- 37) 理論的にはあらゆる単位で誤入力が起こり得るが, 現実的には語より大きな単位においてすべての文字列が不注意により誤入力されることは考えづらいため, 語や記号を対象としているとした。
- 38) JTF には「煮詰まる」を行き詰まる意味, 「おもむろ」を突然の意味で使っている, という語の誤用のみが挙げられているが, 句, 文においてもここで言う誤用はあり得る。例えば, 「犬も歩けば棒に当たる」という文を本来の意味とは違う意味で用いるといった誤りが考えられる。
- 39) 注 35 と同様の理由である。
- 40) このことからさらに詳細な整理が必要となるが, それは今後の課題としたい。
- 41) 参考までに, ISO から「用語」についての記述を抽出すると, 以下の通りとなる。

- up-to-date
- appropriate
- compliance with specific domain terminology
- compliance with client terminology
- ensuring terminological consistency
- accuracy (including terminology)

# An Analysis of QA Schemes in Translation Based on Linguistic Category

Takuya MIYAUCHI <sup>†</sup> Kyo KAGEURA <sup>†</sup>

<sup>†</sup> Graduate School of Education, the University of Tokyo

It can be observed that the actors involved in translation do not necessarily share the same understanding of the concepts referred to by the word “translation.” This issue can be remedied by developing a formal and sharable “translation process model” with high granularity. As the first step in developing such a model, this paper reports the results of analyzing the descriptive resolution of three schemes for quality assurance/assessment, which is regarded as an essential task. The analysis was done by describing correspondence to linguistic categories, focusing on descriptions in the four evaluation categories common to the three schemes: “Accuracy,” “Fluency,” “Terminology,” and “Style.” A wide variety of linguistic categories were associated with each description of “Fluency,” and thus it needs to be organized in more detail. “Accuracy” and “Style” were related to every linguistic category assumed in this paper and descriptions in “Terminology” were basically targeted at the word and phrase level.

Keywords: Translation, Quality Assurance, Quality Assessment, Translation Process