

## J E I D A 機械翻訳システム評価基準 (利用者編)

--- ユーザのための評価項目の作成とアンケート調査による検証 ---

高山泰博	伊藤悦雄	八木沢津義	茂木健	野村浩郷
三菱電機 (株)	(株) 東芝	キヤノン (株)	三洋電機 (株)	九州工業大学

(社) 日本電子工業振興協会 (電子協; J E I D A) では、機械翻訳システムに関して (1) ユーザによる評価、(2) 開発者からの評価、(3) 翻訳品質の評価の3つの評価基準案を作成し、評価方法の検討を行なっている。

本報告では、このうちユーザによる評価基準について述べる。ユーザによる評価基準は (a) 「経済的な側面からの評価基準」 (b) 「技術的な側面からの評価基準」に大別される。この報告では、主として「経済的な側面からの評価基準」に対して実際にアンケート調査という方法で検証した結果の考察を述べる。この調査の結果、評価基準の改良のための基礎データとして有効な236件のユーザからの回答を得ることができた。また、「技術的な側面からの評価基準」に関しては、その改訂した部分について述べる。

## JEIDA's Proposed Method for Evaluating Machine Translation ( End user System Selection )

--- A System Selection Questionnaire for End Users ---

Yasuhiro Takayama	Etsuo Itoh	Tsuyoshi Yagisawa	Takeshi Mogi	Hirosato Nomura
Mitsubishi Electric	Toshiba Corp.	Canon	Sanyo Electric	Kyushu Institute of Technology

This paper describes a revised method for helping users select an appropriate machine translation system that was developed by the Japan Electronic Industry Development Association (JEIDA) Machine Translation Market and Technology Study Committee (a subcommittee of the Machine Translation System Research Committee).

This method approaches the problem of system selection from two perspectives; one technical, i.e., translation speed, etc., and the other economic. Our method consists of a questionnaire and scoring procedure that is designed to be easy to use and apply, represents results graphically, and is objective. To demonstrate the effectiveness of this method we present the results of 236 completed questionnaires.

## 1. はじめに

機械翻訳システムの評価は種々の観点から実施されなければならない。(社)日本電子工業振興協会(JEIDA)機械翻訳システム調査専門委員会(委員長:長尾真京都大学教授)の技術動向調査ワーキンググループ\*(主査:野村浩郷九州工業大学教授)では、機械翻訳システムの評価基準の検討を進めてきており、その第1版については既に報告した<sup>(1)</sup>。

以前に報告した評価基準の第1版では、評価基準を「ユーザ側からの評価基準」と「開発者側からの評価基準」<sup>(3)</sup>に大別していた。この度、「品質評価のための評価基準」<sup>(4)</sup>を付け加えるとともに、全体の評価基準を改良して第2版を作成した<sup>(2)</sup>ので報告する。

本報告では、このうち「ユーザ側からの評価基準」とその検証結果について考察する。ユーザによる評価基準は、更に、

- (a) 「経済的な側面からの評価基準」
- (b) 「技術的な側面からの評価基準」

に分かれている。この報告では、2章で「経済的な側面からの評価基準」(以下、経済評価基準)第2版の概要を述べ、3章で、主として経済的评价基準に対して実際にアンケート調査を行なった結果の考察を述べる。今回のアンケートの発送件数は3,827件(有効発送件数3,641件)、回答件数は242件(有効回答件数236件)であり、回収率約6.5%であった。この結果、評価基準の設問項目の改良に役立つ集計結果や、数値的なパラメータを修正するための基礎データを得ることができた。4章では、「技術的な側面からの評価基準」(以下、技術評価基準)第2版の枠組みを示す。

## 2. ユーザによる経済的な側面からの評価基準の概要

JEIDA機械翻訳評価基準については、第1版ですでに報告<sup>(1)</sup>してあるので、ここでは評価方法の概略と評価のための翻訳システムの分類について述べる。

### 2.1 経済的基準の評価方法

経済評価基準は、機械翻訳システムを導入しようとするユーザに、経済的な側面から導入が望ましいシステムタイプを提案するためのものである。ユーザのための経済評価基準(ユーザに対するアンケート)は大きく2つに分かれている。1つはユーザの現状の翻訳体制を判断する設問群(38項目)であり、もう1つはユーザの要求を判断する設問群(29項目、うち今回のアンケートに用いたのは11項目)である。

我々は機械翻訳システムを以下の14項目のパラメータから特徴付けている。

- A1: 翻訳量      A2: 文書の種類      A3: 翻訳の質
- A4: 翻訳言語    A5: 専門分野      A6: 期間
- A7: 電子化      A8: 体制          A9: コスト
- A10: 前編集     A11: 後編集        A12: 設置条件
- A13: 前修正    A14: 後修正

評価基準の各設問に対して、その回答がこれらパラ

メータに与える点数を決めてある。例えば、

用語や文体の統一にどう対処しているか?	A3	A8
① 特に考慮していない	小	小
② 1人で翻訳するので個人が注意する	中	中
③ 共通する資料を利用する	大	大
④ 最終的に1人が校正を行なう	中	中
配点	15	13

という具合である。この設問に対して③に回答すると、A3の「翻訳の質」やA8の「翻訳体制」の軸に大きな影響を与えることになる。表の下に示している配点に「大」1、「中」0.5、「小」0.2を掛けて、この設問への回答が各パラメータに与える影響を計算する。各パラメータへの配点は全体で100点になるように分配しており、回答から得た得点の和をそのパラメータの値とする。

すべての設問に対する回答を終了すると、2つのアンケートに対して、14項目のパラメータに点数が与えられたレーダーチャートが作成される。図1にレーダーチャートの例を示す。

一方で、機械翻訳システムをタイプ分類しておき、各タイプのシステムを利用するのに最も適していると考えられるユーザを想定し、その想定ユーザに対するアンケート結果のレーダーチャートを用意しておく。

そして、実際のユーザのアンケート結果からのレーダーチャートと我々の用意した各システムタイプ別のレーダーチャートと比較し、最も類似しているタイプを提案するか、またはそれぞれの典型的なシステムタイプとの類似の度合を棒グラフなどにより複数の候補の中からユーザが選択できる形で提案する。

各レーダーチャートは、現状の翻訳体制に関するものと、将来の要求条件に対するもの2種類があり、即時導入を検討しているユーザから、やや先の時点で導入を検討するユーザまでの意志決定を支援できる。

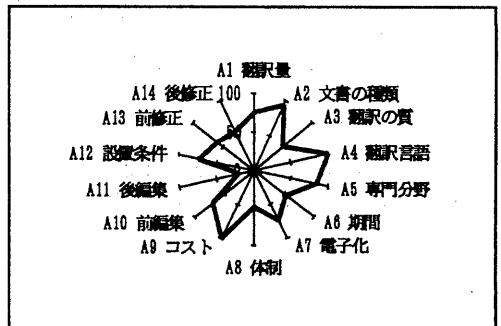


図1 経済評価基準のレーダーチャートの例

### 2.2 機械翻訳システムタイプの分類

経済評価基準第1版では、機械翻訳システムを7つのタイプに分類していたが、第2版では、より細かく9つのタイプ(S1~S9)に分類した。また、それぞれの各機械翻訳システムタイプに想定ユーザを設定してある。経済評価基準第2版の機械翻訳システムの各タイプと想定した典型的なユーザを表1に示す。

\* 平成5年度から自然言語処理応用システム専門委員会機械翻訳技術分科会。

機械翻訳システムのタイプ分類	典型的な想定ユーザ
<b>S1 パッチ型MT</b> 粗翻訳用、高速翻訳、大量翻訳、パッチ主体、WS [主な特徴] システムを社内に設置。チューニング可。サポート体制、OCRなどの周辺機器が充実している。専任管理者が必須 [翻訳言語] 英日、日英 [辞書など] 基本10万語以上(専門語辞書を分野毎10万語以上) [翻訳速度] 20000語/時以上 [価格など] 1000万円以上(システム価格、買い取り/リース)	巨大企業の翻訳部門。社内の翻訳の要求をほぼそこで受け持つ。製品の組立てマニュアルの翻訳などが中心の翻訳であるが、簡単な技術文書など種々雑多な翻訳要求も多い。翻訳量、翻訳文書の種類が多いのが特徴である。ブレンなテキストをもらい、下訳して要求部門に送る。最終訳まで作成する時は、用語の統一、一般にわかる文になる程度。図の挿入、訳のリライトは各部門で行う。外注も常時利用し、下訳をさせている。量をごなすためMTを導入する。
<b>S2 MT専用マシン</b> 高品質、分野限定、制限言語、専用WS [主な特徴] 分野を限定し、チューニングができる。OCRなどの周辺機器も充実している。専任の管理者をおくことが望ましい [翻訳言語] 英日、日英 [辞書など] 基本10万語(専門語辞書を各分野毎に10万語程度) [翻訳速度] 15000~20000語/時 [価格など] ~1000万円(システム価格、買い取り)	巨大企業ではないが、資本金50億程度の中規模の会社。現在、自社内に小規模な翻訳部門があり各自が人手で翻訳している。人数は20人くらい。輸出向けの図入りの使用説明書を翻訳する。量はほどほどにあるが、外注はなし。DTPを使う。基本的に1人が翻訳してネイティブがリライトする。徐々に製品の数が増え、翻訳量が増えてきた。できればリライトまでの部分を翻訳機にかけたい。多数の国に輸出するため、多言語へ翻訳する要求もある。
<b>S3 対話型の翻訳支援</b> 翻訳支援、辞書引き、対話主体、パソコンまたはWS [主な特徴] 辞書引き、翻訳支援機能が充実している理想システム。ユーザインタフェースが充実しており操作性が良い [翻訳言語] 英日、日英 [辞書など] 基本10万語(各分野の専門用語辞書を10万語程度) [翻訳速度] 10000~15000語/時 [価格など] 100~数百万円(システム一式/ソフトのみ、買取)	30人程度の翻訳会社。技術文書の翻訳依頼が多い。分野は電機関係という程度に限定している。文書の種類は特許、論文、マニュアルなどに限定。図や表なども多い。紙で原稿をもらい、訳の部分だけをつけて紙で返す。完全に図や表をレイアウトしたりしない。最終的な訳までこの会社で行ない、品質の高い訳が要求される。翻訳量は人数比では多くないが、全体的にはほどほどである。要求期間が短く、完訳までの時間を短縮したい。
<b>S4 パソコン上の翻訳ソフト</b> 粗翻訳、低速翻訳、少量翻訳、安価 [主な特徴] パソコン上で動く、安価なシステム。ファイル形式はテキストのみで図などは取り入れられない。一括翻訳は可。 [翻訳言語] 英日、日英 [辞書など] 基本5万語(専門辞書を各分野毎に5万語程度) [翻訳速度] 5000~10000語/時 [価格など] 10~数十万円前後(ソフトのみ)	フリーの翻訳者。技術文書を訳す場合には、下訳を機械翻訳でやらせたいと思っている。そこからのリライトは自分で行なうつもり。依頼は、パソコンのテキストファイルでもらうので機械翻訳にはすぐかけられるようになっている。時間短縮、翻訳量の拡大を希望している。
<b>S5 電子辞書</b> 専門用語、ターミノロジーバンク、パソコン、辞書引き主体 [主な特徴] 翻訳は行わず、用語の辞書引き結果を電子的に取込むワープロなどのソフトと並行して走らせ、画面上で辞書引きが可。 [辞書など] 英和、和英とも各5万語(~10万語)程度 [価格など] 10万円前後(ソフトのみ)	小さな翻訳会社で比較的機械化が進んでいる。複数人によって1冊の本などを訳す仕事が主の部門。ただし図が入ったテキストを電子媒体にし、そこから翻訳する形を想定する。非常に難しい専門分野、例えば医学書の翻訳などをする場合で、専門用語の統一や、すぐに訳を調べることを利用して、翻訳量はそれほど多くない。
<b>S6 英文ワープロ スペルチェッカ 文法チェッカ</b> [主な特徴] 通常の英文ワードプロセッサ。編集、印刷が容易。簡単な文法のチェッカやスペルチェッカがついている [辞書など] スペルチェックのための辞書 [価格など] 5~10万円(ソフトのみ)	個人レベルの翻訳会社(者)。機械翻訳などは導入するゆとりはない。また、翻訳文も技術的な文書だけでなく、小説など感情を訴えるものもある。原文が紙であり、WPにより翻訳文を作成していく。翻訳した文書の最終的なスペルチェックに利用する。翻訳量も多くない。
<b>S7 社内LANを利用した翻訳</b> 粗翻訳用、高速・大量翻訳 [主な特徴] 大量の文書を高速に粗翻訳するために社内LANを経由してシステムを利用する。パッチ主体でチューニングしにくい [翻訳言語] 英日、日英 [辞書など] 基本10万語(専門用語辞書を各分野毎に5万語程度) [翻訳速度] 10000語/時 [価格など] 千数百万円(但し既設の社内LANを利用すれば安価)	商社など比較的大規模な会社の一部門。海外動向の監視のためニュースなどを即時翻訳したい。出力に手を加えず理解可能な翻訳の質が望ましい。翻訳部門でないで機械翻訳の担当者を設けず、安価にすませたい。また、外注に出せないほど、即時性が要求される。できるだけ安価にすませたいので、自部門だけでは翻訳システムそのものの導入はしない。社内LAN利用可。
<b>S8 パソコン通信による翻訳</b> 粗翻訳用、パッチ主体、少量翻訳、(社外)ネットワーク、翻訳サービス、チューニング不可、安価 [主な特徴] パソコン通信による社外翻訳サービス。S1レベルの訳ユーザ側で翻訳システムの準備は不要。通常ユーザ辞書が使えない [翻訳言語] 英日、日英(最近、日韓や英・仏・独・西など多様) [辞書など] 基本10万語(専門辞書を各分野毎に5万語程度) [翻訳速度] (ターンアラウンドが)遅い [価格など] 数万/文字(A4版1枚1000円~2000円程度)	これはS7と比べると、より安価で手軽に使えるというところで、海外との取引のある数10人程度の会社や、趣味の個人ユーザが当てはまる。大学の研究室などもこの中に入る。翻訳の質は問わず、概略がつかめればよい。即時性は要求しない。翻訳には少ない。人手で翻訳できないこともないが、手軽に機械翻訳が使えるのであれば使いたいと思っている。パソコンは所有しており、パソコン通信、原文テキストの入力は問題ない。
<b>S9 WS上の翻訳ソフト</b> 粗翻訳、チューニング可、比較的安価 [主な特徴] WS上の翻訳ソフト。既存のWS上に載せて使用する。オプションツールによりチューニング可。特徴的にはS3とS4の間 [翻訳言語] 英日、日英 [辞書など] 基本5万語(専門辞書を各分野毎に5万語程度) [翻訳速度] 10000語/時程度 [価格など] 30~200万円(ソフトのみ)	中規模の会社の翻訳部門もしくは数十人程度の翻訳会社。技術文書の翻訳が主。そこそこの翻訳量があり効率化を図りたいと考えている。機械翻訳システムを導入することには積極的であるが、すでにWSがあり、その上で動作するシステムを希望している。予算は100万くらい。

表1 経済評価基準の機械翻訳システムの分類と想定ユーザ

### 3. ユーザによる経済的な側面からの評価基準の検証

ユーザによる機械翻訳システムの経済評価基準第1版の作成に際しては、仮想的なユーザの回答を作成して、思考実験によるシミュレーションで検証を試みた<sup>(2)</sup>。第2版では、実際のユーザに対し経済評価基準を用いてアンケートを実施した。ここでは、その集計結果と評価基準を改良すべき点についての考察を述べる。

#### 3.1 アンケート結果とその考察

この節では、個々の設問項目に対するアンケートの集計結果<sup>\*</sup>とその考察を述べる。アンケート回答者の業種別分類を以下に示す。

業種	件数(件)	構成比(%)
(1) 翻訳サービス業	66	28.0
(2) 製造業	97	41.1
(3) 卸売業	8	3.4
(4) 小売業	2	0.8
(5) 金融関係	18	0.8
(6) 情報サービス業	5	7.6
(7) 新聞社	5	2.1
(8) 出版社	4	1.7
(9) 特許事務所	1	0.4
(10) 官公庁	2	0.8
(11) 大使館	2	0.8
(12) 大学	2	0.8
(13) 研究機関	5	2.1
(14) その他	22	9.3

##### 3.1.1 現状の翻訳業務の実態

ユーザの現状を判断する設問群では、ユーザの現状の翻訳業務の実態を把握する。この設問群はQ1からQ38までの38項目からなる。

##### Q1 現在の翻訳量

	翻訳業	製造業	情報業 <sup>**</sup>	その他
① 月に数千ページ以上	27.3%	5.2%	11.1%	7.3%
② 月に数百ページ程度	57.8%	28.9%	50.0%	14.5%
③ 月に数十ページ程度	9.1%	48.5%	27.8%	52.7%
④ 月に数ページ以下	1.5%	15.5%	11.1%	25.5%

翻訳量に関しては、業種によりバラツキがある。当然のことながら、翻訳業が最も多い。

##### Q2 翻訳文書の形式

	翻訳業	製造業	情報業	その他
① 定型文書が多い	10.8%	17.5%	27.8%	25.5%
② 非定型文書が多い	50.0%	42.3%	44.4%	38.4%
③ 定型文書/ 非定型文書の両方	37.9%	40.2%	27.8%	38.2%

現在の多くの機械翻訳システムでは、「定型文書」を対象とすることが多いが、現実には非定型文書を翻訳しなければならない場合も多いことを示している。

##### Q3 翻訳文書の種類

	数千頁	数百頁	数十頁	数頁 <sup>Q1</sup>
① 技術文書が多い	89.0%	49.4%	42.5%	28.1%
② ビジネス文書が多い	0.0%	9.8%	20.7%	43.8%
③ 技術文書/ ビジネス文書の両方	31.0%	41.0%	34.5%	28.1%

\* 各設問に対する回答の集計結果の合計が100%に満たないものは、その設問に無回答のものが含まれていた場合である。

\*\* 翻訳業86件、製造業97件、情報サービス業18件、その他55件

やはり技術文書の場合にページ数が多いことが確認できる。全体では、技術文書46.6%、ビジネス文書16.9%、技術文書とビジネス文書の両方35.6%である。

#### Q4 翻訳文書の内容

① 事実伝達型文書が多い	69.5%	(164件)
② 感情伝達型文書が多い	5.1%	(12件)
③ 事実伝達型/ 感情伝達型文書が多い	25.0%	(59件)

今回のアンケートは、企業等が対象であるため、翻訳文書の内容が事実伝達型の文書である割合が高い。

#### Q5 翻訳文書の文体

	翻訳業	製造業	情報業	その他
① 書き言葉が多い	87.9%	78.4%	83.3%	80.0%
② 話し言葉が多い	1.5%	1.0%	0	3.8%
③ 書き言葉/ 話し言葉の両方	10.8%	20.8%	11.1%	16.4%

翻訳文書は書き言葉の場合が8割以上である。製造業で、「書き言葉と話し言葉の両方」という回答が若干多くなっているのは、口語調の説明書などが含まれることなどによるものと推定される。

#### Q6 翻訳する文書における専門語の量

① 多い	90.3%	(213件)
② 少ない	9.7%	(23件)

専門語の量が多いと感じている場合が9割を超える。特に、情報サービス業では100%が多いと回答した。

#### Q7 翻訳文書の専門分野

	翻訳業	製造業	情報業	その他
① 限定されている	39.4%	68.0%	72.2%	50.8%
② 限定されていない	60.6%	34.0%	27.8%	49.1%

全体での比率は限定されている場合が55.5%、そうでない場合が44.5%である。現行の機械翻訳システムは、何等かの形で分野限定を要求することが多いが、実際には翻訳専門分野が限定されていない場合が多い。

#### Q8 翻訳文書の言語対

	外注無	外注有 <sup>Q14</sup>	全体	
① 日本語→英語	17.4%	27.3%	26.3%	(62件)
② 英語→日本語	41.3%	18.2%	22.5%	(53件)
③ 日本語<=>英語	39.1%	37.9%	43.8%	(103件)
④ 日英以外の2ヶ国語間	2.2%	9.1%	3.8%	(9件)
⑤ 多国語間	0	8.1%	3.4%	(8件)

全体では、日本語と英語との間の翻訳が92.4%を占める。日本語から英語への翻訳の場合に外注比率がやや高くなる。

#### Q9 過去に翻訳した文書の再利用

① 頻繁にある	9.7%	(23件)
② ときどきある	45.8%	(108件)
③ ほとんどない	44.5%	(105件)

翻訳文書を再利用しているのは半数を超える程度で、頻繁に利用しているのは1割程度である。

#### Q10 翻訳する文書の利用目的

① 法的テキスト(原文と同じ重み有り)	25.8%	(61件)
② 法的テキスト(原文と同じ重み無し)	18.1%	(38件)
③ 規約、既定を書いたマニュアルなど	57.8%	(138件)
④ 事実を伝える報道記事など	28.8%	(68件)
⑤ 内容熟知者向けの寿命の長いテキスト	29.7%	(70件)
⑥ 内容熟知者向けの寿命の短いテキスト	25.8%	(61件)
⑦ 一般の人向けの一般情報	19.5%	(48件)
⑧ 専門家がざっと目を通すためのもの	30.9%	(73件)
⑨ 読者に何かを説得するためのもの	14.4%	(34件)

**Q11 同一種類の文書の翻訳量**

① 多い	54.2%	(128件)
② 少ない	44.8%	(106件)

約半数は同一種類の文書の翻訳量が少ないと答えており、これがQ9の翻訳文書の再利用に影響を与えていると考えられる。

**Q12 専門語の対訳表**

① 蓄積されている	44.1%	(104件)
② 蓄積されていない	55.9%	(132件)

機械翻訳利用への移行を容易にする要素として、専門語の対訳表が既に準備されているかどうかが重要であるが、現状では蓄積されていない場合が半数を越える。

**Q13 翻訳専門の人員**

① 0~30名	91.5%	(216件)
② 31~100名	3.0%	(7件)
③ 101名以上	4.7%	(11件)

9割が30名以下である。項目①は今後の評価基準の改訂で細分する必要がある。

**Q14 外注の利用**

	翻訳業	製造業	情報業	その他
① 全く利用しない	10.8%	14.4%	38.9%	32.7%
② 通常利用する	50.0%	17.5%	27.8%	20.0%
③ 翻訳量により外注する	18.2%	28.9%	11.1%	21.8%
④ 対象文書により外注する	19.7%	39.2%	22.2%	25.5%

翻訳作業を請け負う側である翻訳業においても通常5割は外注化されている。個人翻訳者やフリーランスの翻訳者に外注されているものと推定される。情報サービス業では、全く外注化しない割合がやや高い。

**Q15 年間翻訳量に対する外注の割合**

	翻訳業	製造業	情報業	その他
① 0~30%	42.4%	60.8%	66.7%	61.8%
② 31~50%	13.6%	12.4%	11.1%	9.1%
③ 50%以上	42.4%	23.7%	22.2%	25.5%

翻訳会社における外注比率が高いのが目立つ。

**Q16 翻訳作業の分業**

	翻訳業	製造業	情報業	その他
① 1人で最終訳まで作成	50.0%	68.1%	50.0%	63.0%
② 原文を修正する人がいる	4.5%	10.6%	16.7%	13.0%
③ 訳文を修正する人がいる	66.7%	35.1%	44.4%	40.7%
④ 訳文を消す人がいる	22.7%	2.1%	5.6%	9.3%

この設問は複数回答可である。翻訳向けに原文に手を入れている比率は低い。一方、訳文は修正される場合が多い。全体では、一人で翻訳作業を行なう場合が約60%を占める。

**Q17 用語や文体の統一**

① 特に考慮していない	25.0%	(59件)
② 1人で翻訳するので個人が注意する	34.7%	(82件)
③ 共通する資料を利用する	18.2%	(43件)
④ 最終的に1人が校正を行なう	20.8%	(49件)

7割以上が翻訳文書用語や文体の統一に注意を払っている。Q23の訳語の整合性では8割以上である。

**Q18 部門の年間予算に対する翻訳予算の割合\***

	翻訳業	製造業	情報業	その他
① 大きい	56.1%	2.1%	18.7%	7.3%
② ある程度を占める	16.7%	6.2%	27.8%	12.7%
③ 小さい	19.7%	88.7%	55.6%	80.0%

\* アンケートに示した参考値

大きい:50%以上,ある程度:25~50%,小さい:25%未満  
高い:5000円以上,普通:3000~5000円,安い:3000円未満

**Q19 翻訳コストのうち、外注コストが占める割合\***

	翻訳業	製造業	情報業	その他
① 大きい	39.4%	28.9%	27.8%	32.7%
② ある程度を占める	21.2%	9.3%	5.6%	5.5%
③ 小さい	34.8%	58.8%	61.1%	54.5%

**Q20 平均的なケースでの外注翻訳コスト\***

	翻訳業	製造業	情報業	その他
① 高い	7.6%	15.5%	5.6%	10.9%
② 普通	39.4%	50.5%	27.8%	40.0%
③ 安い	43.9%	24.7%	44.4%	30.9%
無回答	9.1%	9.3%	22.2%	18.2%

**Q21 平均的なケースでの自部門の翻訳コスト\***

	翻訳業	製造業	情報業	その他
① 高い	12.1%	18.6%	5.6%	21.8%
② 普通	50.0%	40.2%	50.0%	27.3%
③ 安い	30.3%	30.9%	44.4%	38.2%
無回答	7.6%	10.3%	0	12.7%

Q20, Q21をみると、若干外注翻訳コストを安いと捉えている場合が多い。また無回答の割合が高い。

**Q22 翻訳の品質**

	翻訳業	製造業	情報業	その他
① 完璧な正確さと適切な文体	62.1%	20.6%	22.2%	50.9%
② 完璧な正確さを要求するが 文体の適切さは重要でない	13.6%	18.6%	22.2%	7.3%
③ ユーザを誤らせない正確さ と適切な文体	21.2%	44.3%	38.9%	25.5%
④ ユーザを誤らせない正確さ 文体の適切さは重要でない	1.5%	12.4%	5.6%	7.3%
⑤ ラフな翻訳でかまわない	0	3.1%	11.1%	9.1%

**Q23 訳語の整合性**

① 重視する	84.3%	(199件)
② 重視しない	14.0%	(33件)

**Q24 即時性の要求**

	数千頁	数百頁	数十頁	数頁
① 要求される	89.7%	81.9%	72.4%	59.4%
② 要求されない	10.3%	14.5%	26.4%	37.5%

**Q25 A4版で10ページ程度の完訳までの期間**

	翻訳業	製造業	情報業	その他
① 即時	1.5%	3.1%	5.6%	1.8%
② 1日以内	48.5%	10.3%	11.1%	20.0%
③ 1週間以内	47.0%	72.2%	66.7%	63.6%
④ 1か月以内	3.0%	12.4%	16.7%	10.9%
⑤ 1か月以上	0	1.0%	0	0

**Q26 A4版で10ページ程度の粗訳までの期間**

	翻訳業	製造業	情報業	その他
① 即時	10.6%	5.2%	5.6%	7.3%
② 1日以内	50.0%	34.0%	22.2%	41.8%
③ 1週間以内	33.3%	52.6%	66.7%	43.6%
④ 1か月以内	0	3.1%	5.6%	3.6%
⑤ 1か月以上	0	1.0%	0	0

翻訳業が最も即時性の要求が高く、製造業ではやや期間をかけても良いという回答になっている。

**Q27 前編集に時間がかかっているか**

① かなり時間がかかっている	8.1%	(19件)
② あまり時間がかかっていない	25.0%	(59件)
③ 消書だけをワープロを利用	8.9%	(21件)
④ 訳文を作るところからワープロを利用	21.6%	(51件)
⑤ 翻訳作業はコンピュータ上で行なう	7.6%	(18件)
⑥ 前編集は行なっていない	25.8%	(61件)

**Q28 後編集に時間がかかっているか**

① かなり時間がかかっている	30.5% (72件)
② あまり時間がかかっていない	38.1% (90件)
③ 後編集は行っていない	28.0% (68件)

**Q29 原文の手直し(前修正)に時間がかかっているか**

① かなり時間がかかっている	20.8% (49件)
② あまり時間がかかっていない	75.4% (178件)

**Q30 訳文の手直し(後修正)に時間がかかっているか**

① かなり時間がかかっている	36.4% (88件)
② あまり時間がかかっていない	61.4% (145件)

Q27, Q29, Q36からみると原文の前編集、前修正を行わない場合が多くを占める。Q28, Q30と比べると後編集、後編集を行なう場合の方が若干数値が高い。

**Q31 原文の供給形式**

① 紙	49.6% (117件)
② FAX	29.7% (70件)
③ 磁気テープ	0.4% (1件)
④ フロッピーディスク	13.6% (32件)
⑤ 光ディスク	0.4% (1件)
⑥ コンピュータネットワーク	5.9% (14件)

紙によるものが約半数を占める。また、FAXの場合もそのまま直接には電子媒体に変換できないため、オフライン形式で受け取る場合が全体の約80%になる。

**Q32 原文の電子化状況**

① 原文は電子化しない	74.2% (175件)
② 電子化されたものを提供してもらう	8.1% (19件)
③ キーボードから打ち込んで電子化	9.3% (22件)
④ OCR(文字読み取り装置)を利用	6.4% (15件)

3/4近くの回答が「原文は電子化しない」と答えている。Q16では「原文に手を入れる人がいる」が10%以下であり、Q27やQ29で前編集や原文の手直しに時間がかかっていないことの原因と推定される。一方で、Q39で「文書を電子化していく」要求が60%近くある。

**Q33 翻訳した文書の供給形式**

① 紙	49.6% (117件)
② FAX	12.3% (29件)
③ 磁気テープ	0.4% (1件)
④ フロッピーディスク	28.8% (68件)
⑤ 光ディスク	0.0% (1件)
⑥ コンピュータネットワーク	8.9% (21件)

訳文をフロッピーディスクで供給する場合が原文の場合の2倍を占める。最終的な印刷を受け取り側で行なう場合もあるためと考えられる。

**Q34 翻訳作業における電子的な文書入力装置の使用**

① 電子装置は使わない	13.8% (32件)
② 消音だけワープロを利用	21.8% (51件)
③ 訳文を作るところからワープロを利用	45.8% (108件)
④ 翻訳作業はコンピュータ上で行なう	18.6% (44件)

**Q35 翻訳作業時における電子的な資料の利用**

① 電子化資料は使わない	64.8% (153件)
② 使う時もあるが減多に使わない	15.7% (37件)
③ 時々使う	10.8% (25件)
④ 頻繁に使う	8.5% (20件)

現状では電子化資料を使わない場合が6割を超える。

**Q36 翻訳向けに原文の前編集を行なう習慣**

	翻訳業	製造業	情報業	その他
① 専門要員が行なう	13.8%	9.3%	33.3%	5.5%
② 文書作成者が行なう	18.7%	21.8%	18.7%	29.1%
③ 行なわない	69.7%	67.0%	50.0%	65.5%

**Q37 部門内の翻訳作業** | 翻訳業 | 製造業 | 情報業 | その他

① 専門要員(ネイティブ)	25.8%	4.1%	18.7%	10.9%
② 専門要員(日本人)	59.1%	38.1%	44.4%	25.5%
③ 文書作成者が行なう	6.1%	45.4%	22.2%	41.8%
④ 部門内では行なわない	9.1%	9.3%	18.7%	18.4%

**Q38 訳文の後編集** | 翻訳業 | 製造業 | 情報業 | その他

① 専門要員(ネイティブ)	37.9%	8.2%	11.1%	21.8%
② 専門要員(日本人)	43.9%	27.8%	50.0%	21.8%
③ 文書作成者が行なう	3.0%	40.2%	22.2%	32.7%
④ 部門内では行なわない	15.2%	19.8%	16.7%	23.8%

翻訳業ではネイティブの専門要員が翻訳作業にかかわっている場合が多い。

### 3.1.2 翻訳に対する要求条件

ユーザの要求条件を判断する設問群には18項目の設問があるが、今回のアンケート調査には、評価基準第1版で作成したQ39～Q49の11項目を用いた。

**Q39 文書を電子化していく要求**

① 電子化していく	59.3% (140件)
② 電子化する予定はない	39.4% (93件)

現状でQ32「文書を電子化しない」74.2%に対して、文書を電子化していく要求が60%近くもある。

**Q40 翻訳工程の確立の要求**

① 確立していく	51.3% (121件)
② 確立する予定はない	45.8% (108件)

翻訳工程を確立する要求は約半数と、意外と少ない。

**Q41 翻訳を時間を短縮する要求**

① 短縮していく	66.5% (157件)
② 現状のまま	32.2% (76件)

当然のごとく翻訳時間を短縮したいという要求は多い。

**Q42 翻訳の質の向上 [複数回答]**

① 現状の質で満足している	9.8% (23件)
② 用語の統一	49.6% (116件)
③ 専門用語の正しい訳	70.1% (164件)
④ 的確な文書スタイル	41.9% (98件)
⑤ 意味を考慮した意識	46.2% (108件)
⑥ 小説のような表現も考慮した 高度な訳	9.4% (22件)

用語の統一(約50%)や専門用語の正しい訳(約70%)といった語彙面での改善要求が高いことが注目される。

**Q43 今後の翻訳量**

① 増加する	69.1% (163件)
② 変わらない	25.4% (60件)
③ 減少する	4.7% (11件)

今後の翻訳量は増加するとの回答が70%を占める。

**Q44 専門分野拡大の要求**

① 要求がある	42.8% (101件)
② 今程度の専門分野を今後も扱う	52.5% (124件)
③ 扱う専門分野を減らして行きたい	2.1% (5件)

専門分野の拡大の増加要求については、拡大要求がある企業と現状維持の企業とで半々である。

**Q45 文書の種類の増加の要求**

① 要求がある	45.8% (108件)
② 今程度の文書の種類を今後も扱う	47.0% (111件)
③ 扱う文書の種類を減らして行きたい	4.7% (11件)

文書の種類の増加要求では、専門分野の拡大と同様に拡大要求がある場合と現状維持とで半々である。

#### Q46 言語対の種類を増やす要求

① 要求がある	34.7% (82件)
② 今程度の専門分野を今後も扱う	58.8% (134件)
③ 扱う専門分野を減らして行きたい	3.0% (7件)

言語対の種類に関しては、専門分野や文書の種類に比べて拡大要求が若干弱い。

#### Q47 翻訳システム導入のために確保できる予算

① ～100万円まで	83.1% (149件)
② ～1000万円まで	22.0% (52件)
③ それ以上	2.1% (5件)

100万円までの回答が3割を占める。項目①に関しては、評価基準の改訂に際して細分する必要がある。

#### Q48 翻訳システム導入のために確保できる人数

① 1人だけ	53.4% (126件)
② ～5人まで	32.2% (76件)
③ それ以上	1.3% (3件)

Q48から、翻訳作業は比較的個人的な作業であることがわかる。項目②は細分する必要がある。また、複数人の場合、作業分担も設問として設定することで、詳しく評価ができる。「機械翻訳システムのために新規に人を確保できない」という項目も必要であろう。

#### Q49 翻訳システムを導入した後の試行期間

① 不可能。導入後、すぐ使用して効果を出したい	25.0% (59件)
② ある程度は可能。	81.9% (146件)

現状の機械翻訳についての知識がある程度ひろまっていることによるためか、試行期間を置くことに対する抵抗は強くないようである。

Q47～Q49では1割以上が無回答であった。

### 3.2 経済的側面からの評価基準全体に対する考察

この節では、ユーザからみた経済評価基準の評価方法全体について考察する。今回のアンケートは、評価基準の設問項目の改良や数値的なパラメータを調整するための基礎データを得るためのものであり、個別の結果は今後の改良の後に検証するべきであるため、ここでは今回の集計の全体的な傾向について述べる。

今回のアンケートの各回答に対して、最適な機械翻訳システムの種類 (表1参照) として、経済評価基準から計算された比率を以下に示す。

タイプ	特徴	現状 (%)	要求 (%)
S1	バッチ型MT	1.7	0.0
S2	MT専用WS	28.4	22.9
S3	対話型翻訳支援	8.1	0.4
S4	パソコンの翻訳ソフト	3.0	24.2
S5	電子辞書	1.7	13.6
S6	英文ワープロ	19.5	9.7
S7	社内LAN利用MT	15.3	3.4
S8	パソコン通信のMT	1.3	2.1
S9	WS上の翻訳ソフト	21.2	23.7

[現状の翻訳業務の実態から想定されるシステム]

アンケートの結果では、ユーザの現状の翻訳作業の実態に対して、評価基準から計算された、推奨すべき機械翻訳システムのタイプは、多いものから順に、表1のS2、S9、S6の順となった。

システムタイプS2は、システムをチューニングし、対象文書の分野限定を行ない、入力文を前修正したうえで翻訳にかけるという現在の標準的なシステムであ

る。システムタイプS9は、現状では一般的となってきたワークステーション上の翻訳ソフトウェアの典型システムとして設定したものである。したがって、推奨するシステムの割合としては比較的妥当な結果であると考えられる。また、現状の体制では、システムタイプS6の英文ワードプロセッサのみを翻訳作業に利用するのが妥当な場合もまだ多いと推定される。

[ユーザの要求条件から想定されるシステム]

アンケートの結果では、ユーザの要求条件から想定される推奨システムは、多いものからS4、S9、S2の順となった。

S4はパソコン上の機械翻訳システムである。これは、体制的にも、価格的にも、導入しやすいものから徐々に機械翻訳を使っていきたい、というユーザの要求の反映であると考えられる。

S3はある意味では理想的な翻訳支援システムとして設定したものであるが、要求条件に対する設問項目への回答からS3を提案した場合が少ない。これがどの要因によるものかは今後検討しなければならない。

要求条件からの推奨システムでは、関連する設問の数が少ないため、個別の回答が大きく影響してしまう。そこで、今回のアンケート調査には加えていないが第2版には技術評価基準の初期質問のために4章に示す18項目の設問 (Q50～Q67) を追加した。ユーザの導入・運用条件に関する要求をより詳細に判断し、最適なシステムタイプを提案するため、これらの設問の一部を経済評価基準に反映させることを検討したい。

評価基準全体として当初設定した目標は、(1) 数値的な集計の容易さ、(2) 評価結果の視覚的な提示、(3) 客観的評価の3つであった。今回のアンケート調査において、経済評価基準では容易に数値評価が可能であることが確認できた。また、アンケート調査結果をユーザに提示する際に、最適なシステムのレーダーチャートを示すと共に、棒グラフにより各典型的なシステム間の類似度の差を示すことで、複数の推奨システムを視覚的に提示した。したがって、(1)、(2) についてはある程度達成されたと考えている。しかし、(3) の客観的評価、言い換えれば評価の適切さに関しては更に以下の課題も残されている。

レーダーチャートの類似度を計算する際に14個の特徴パラメータに対して、特徴ベクトル<sup>(2)</sup>と呼ばれる係数を掛けて計算しているが、この重み付けが計算結果に与える影響が大きいため検討の余地がある。また、3.1節のアンケート調査結果をもとにして、設問項目 (質問文)、特徴パラメータの分類や設問との対応、数値的なパラメータを見直して行きたい。機械翻訳システムタイプの分類や仮想ユーザの設定は、技術的側面との関連を更に密接にし、ユーザのシステム選択に役に立つものにしていかなければならない。

### 4. ユーザによる技術的側面からの評価基準

技術評価基準の目的は、機械翻訳システムを導入しようとしているユーザにとってシステム選定の指針を与えることである。特長は、機械翻訳システムを単体として評価するのではなく、ユーザの要求と組み合わせ「機械翻訳システムに対するユーザの満足度」を技術的に評価する点にある。技術評価基準第2版は、

経済評価基準との設問の共通化と技術評価基準自身の簡略化により、評価の所要時間を第1版の2/3以下とし、実用性を高めたものである。

機械翻訳の導入においては、翻訳品質ばかりでなく、例えば以下のような項目にも検討・評価が必要である。

- ・導入コスト、維持コスト
- ・前編集、後編集の機能
- ・基本辞書の充実度
- ・ユーザ辞書作成の手段
- ・専門用語辞書の充実度
- ・ユーザインターフェース
- ・入出力形態
- ・システム形態
- ・メーカーのサポート体制
- ・ユーザに対する教育体制等

各評価項目の重要度は、すべてが均等ではなく、ユーザがどのような体制で機械翻訳システムを使用したいかに依存する。例えば、外国語文書の概要把握のために粗翻訳結果のみが欲しい場合には、生の翻訳品質の高さが重要であり、編集能力は重要でない。逆に、翻訳業務に使用する場合には、翻訳出力の品質の高さに加えて前・後編集の支援機能やユーザ辞書整備機能の充実、更には、顧客に対する翻訳文書の受渡に便利な入出力機能を持っていることなども重要であろう。

従って、技術的内容を評価する場合でも、各種の機能の内、どれを重視し、どれを考慮しないか、といった重み付けはユーザの要求により変更する必要がある。また、ユーザ自身、その要求を明確にできない場合があるため、どのような体制においては、どのような要求をするべきであるかということも自然に導かれるような評価基準である必要がある。

このため、技術評価基準は以下の構成になっている。

- (1) ユーザの要求・状況分析のための設問項目
- (2) システムの仕様を調べるための設問項目
- (3) ユーザの要求とシステム仕様から、満足度を評価するためのテーブル

(1) 「ユーザの要求・状況分析のための設問項目」は機械翻訳の導入を検討しているユーザが回答することにより、ユーザ自身が自覚していない機械翻訳システムへの要求を明確化するものである。つまり、機械翻訳の導入を検討しているユーザは、経済評価基準と技術評価基準の両者の評価を行ない、より適したシステムを選択することになる。そこで、第1版の両者の設問事項に重複する部分を経済的評価基準と共通化した。ただし、経済評価基準と技術評価基準では目的が異なるため、技術評価独自の設問もある。

- Q50 導入のポイントは何か？
- Q51 どんな人が機械翻訳システムを実際に使うか？
- Q52 どういう形態で機械翻訳システムを使いたいか？
- Q53 一次出力にどの程度の翻訳品質を期待するか？
- Q54 機械翻訳システムで翻訳予定の文書の特徴は？
- Q55 同時に複数の人間が機械翻訳システムを使用したいか？
- Q56 複数の端末で辞書・文書を共有する必要があるか？
- Q57 端末はどのような構成を希望するか？
- Q58 端末同士は何で接続したいか？
- Q59 端末はどの程度分散して置けるか？
- Q60 どの程度の台数の端末を接続しようと考えているか？
- Q61 他端末(ユーザ)から文書や辞書をアクセス不可にするか？
- Q62 どのような機器で使用したいと考えるか？
- Q63 現在所有しているハードウェアは？
- 名称、OS、メモリ、ディスク容量、周辺機器、拡張可能な機器
- Q64 購入するマシンの機種についての希望は？
- Q65 1台目として購入するマシンの設置場所についての希望
- Q66 2台目以降に購入するマシンの設置場所についての希望
- Q67 設置する場所の電源は？

(2) の「システムの仕様を調べるための設問項目」は、ユーザの要求とは独立して、システム提供者が回答する。この項目の解答によりシステムの仕様を明確化することができる。各項目は、それぞれ機械翻訳システムに要求される項目に類別し、評価に用いる。

(3) の「(1)と(2)から満足度を評価するためのテーブル」では「ユーザの要求・状況分析」の回答と「システム仕様」の回答とを組み合わせて、ユーザ別の翻訳システムの評価を行なう。その中間データとして以下の項目の点数が算出される。

- ・言語対応致度
- ・翻訳速度
- ・前編集、前修正能力
- ・未知語処理能力
- ・文書適合性
- ・辞書記述内容
- ・専門辞書分野
- ・翻訳失敗時の処理・一次品質
- ・辞書・文法拡張能力
- ・学習能力
- ・チューン後品質
- ・訳文編集能力
- ・レイアウト反映能力
- ・文書互換性
- ・操作性
- ・システム構成合致度
- ・ランニングコスト
- ・イニシャルコスト

これらの値はシステムの仕様を単純に点数化しただけの値となっている。これをユーザの要求とクロスさせることにより最終的な評価を図2に示す10軸のレーダーチャートを用いて行なう。すなわち、ユーザの要求に関する設問からどの項目を重視するべきかを決定し、先に求めた点数と表形式で集計する。

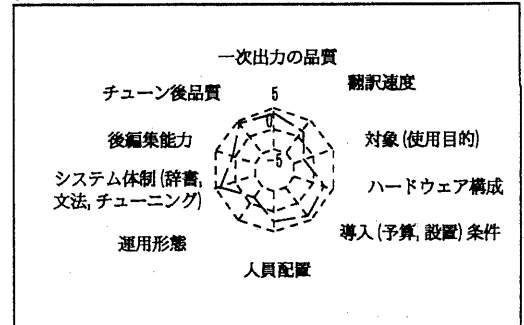


図2 技術評価基準の希望満足度レーダーチャート例

## 5. おわりに

ユーザによる評価基準の枠組みは、機械翻訳導入時に、経済的・技術的なさまざまな要因に基づく定量的評価を可能とすることを目的としている。今回の経済的評価基準を用いたアンケート調査は、JEIDA機械翻訳システム評価基準の最初の適用実験であり、236件の有効な回答を得た。評価基準第2版では、評価の枠組の提案から一歩進めて、ユーザによる評価のために必要な項目の改良案の提示と基礎データの収集ができた。今後、実際にユーザに役に立つ評価基準にするために、技術評価基準案に対しても適用実験を行ない、更に改良していきたいと考えている。

### 参考文献

- (1) 野村, 井佐原: 機械翻訳の評価基準について, 情知研報NL89-9 (1992. 5).
- (2) 電子協: 機械翻訳システムの実用化に関する調査研究, (社) 日本電子工業振興協会, (1993. 3).
- (3) 中岩ほか: JEIDA機械翻訳評価基準 (開発者編), 情知研報NL96-10 (1993. 7).
- (4) 井佐原ほか: JEIDA機械翻訳評価基準 (品質評価編), 情知研報NL96-11 (1993. 7).