

JEIDA 機械翻訳システム評価基準 (品質評価編)

—英日翻訳の品質評価項目の検討と評価用コーパスの作成—

井佐原 均	新納 浩幸	山端 潔	森口 稔	野村 浩郷
電子技術総合研究所	松下電器(株)	NEC	シャープ(株)	九州工業大学 現茨城大

(社) 日本電子工業振興協会機械翻訳システム調査専門委員会技術動向調査ワーキンググループでは、機械翻訳システムの翻訳出力の品質評価法の検討を進め、

- (1) システム開発者が自己のシステムの不備をチェックすること。
- (2) ユーザが、各システムの出力品質面での優劣を比較すること。

の二つの視点からの評価を目的としたコーパスを作成している。

本稿ではWGが昨年度に行った英日翻訳の品質評価項目の検討と評価用コーパスの作成について述べる。このコーパスでは、(1) 言語現象を網羅すること、(2) 機械翻訳システムにおいて取り扱うことが困難である言語現象を主に曖昧性の解消に注目して収集・分類すること、を試みた。

JEIDA's Proposed Method for Evaluating Machine Translation (Translation Quality)

--- A Proposed Standard Method and Corpus ---

Hitoshi Isahara	Hiroyuki Shinnou	Kiyoshi Yamabana	Minoru Moriguchi	Hirosato Nomura
Electrotechnical Lab.	Matsushita Elec.	NEC	Sharp	Kyushu Institute of Technology
Ibaraki Univ. at present				

This paper describes a method and associated corpus for evaluating the translation quality of English to Japanese machine translation systems that has been 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 achieves two goals:

- 1) It provides system developers and users with a sufficient variety of linguistic phenomena to identify areas of weakness in a system.
- 2) It enables potential buyers to make qualitative comparisons between different commercial machine translation systems.

1. はじめに

(社)日本電子工業振興協会機械翻訳システム調査専門委員会技術動向調査ワーキンググループ¹⁾では、機械翻訳システムの翻訳出力の品質評価法の検討を進め、評価用コーパスを作成している。本稿ではワーキンググループが昨年度に行った英日翻訳の品質評価項目の検討と評価用コーパスの作成について述べる。コーパス作成においては、

- (1) 言語現象を網羅すること、
- (2) 機械翻訳システムにおいて取り扱うことが困難である言語現象を主に曖昧性の解消に注目して収集・分類すること

を試みた。

ここで報告するコーパスを用いた評価法は、次の二つの視点からの評価を目的としている。

- (1) システム開発者が自己のシステムの不備をチェックすること。

このため、コーパスの翻訳結果から、直ちに不備の原因を把握できるようにしている。

- (2) ユーザが、各システムの出力品質面での優劣を比較すること。

複数のシステムによる、このコーパス全体の翻訳結果を参照し、相互に比較することにより、システム間の優劣を判断する。また、ユーザが、特定の言語現象を中心とする文書を翻訳しようとしているなら、このコーパスの翻訳結果のうちで、必要な部分についてのみ評価すれば良い。

本評価基準を開発者が使用する場合、開発者は次節以降で説明されるコーパスを自分のシステムで翻訳し、翻訳結果を参照しながら、各文に付加された設問に答えていく。各設問は、ある言語現象をシステムが処理できるかどうかを判断するためのものであり、客観的に（yes か no かで）答えられるようなものになっている。全ての質問に答えた後、その回答を見ることにより、そのシステムが対応できる

言語現象と対応できない言語現象とを、正確に把握することが出来る。

ある程度完成されたシステムにおいては、基本的な構文構造は当然全て解析できるであろう。その場合、開発者にとってシステム改良に有益であるものは、実際の文章に含まれる、基本構造が組み合わされることによって翻訳困難となっている文であろう。このような文を収集するために、ここでは、思考実験のみならず、既存の商用の機械翻訳システム3種を用いて、米国で実際に用いられている英語で書かれた数学と保健の教科書の文章を翻訳し、それらを逐一比較することも行った。

なお、昨年度は、対象を英日翻訳とし、例文の収集は単文レベルのものに限定した。より長い文、すなわち文が接続するようなもの（重文）や並列構造等は今年度に収集する予定である。また、日英翻訳についても、同様のコーパスを作成する予定である。

2. 品質評価に要求されるもの

機械翻訳システムの品質評価においては、何が評価されるのであろうか。当然のことながら、翻訳が出来るかどうか第一の評価基準であろう。さらに、翻訳が出来るものについて、その質が良いかどうかは次の評価基準となる。品質評価は、本質的にシステムの技術的達成度を評価するものであり、これらの評価基準は、システムに正しい知識が与えられているか、システムはそれらの知識を適切に選択することが出来ているかという、より詳細な技術的視点から分類できるであろう。

一方、機械翻訳システムの品質を評価する立場には、ユーザがシステム導入に際し、そのシステムの性能を（相対的に）判断する場合と、開発者が自分のシステムの性能を（絶対的に）判断する場合とがある。前者においては、総合的な品質評価と同時に、自分が処理したいと思う言語現象をシステムが処理できるかどうかを判断することが要求されよう。後者においては、そのシステムがどの言語現象に対応できていないかを正確に把握することが要求される。

¹⁾ 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

また、評価を客観的にくだせることも、汎用の評価法においては、重要な要求のひとつである。ここには二つの要求があり、その一つは、評価過程が客観的であることであり、もう一つは、評価結果が客観的に判断できることである。この点において、特定のシステムを対象に、評価者が、その言語的直観に頼って評価点をつける方法は、評価法としては不十分なものといえよう。

これまで、品質評価を行ないたいという要求に対して、忠実度・理解容易度等を各文に対して人間が評価し数値化するという、ALPACレポート等に代表される方法や、原文の構造と訳文の構造とをグラフ的に照合し、二つの構造の合致度を評価点とする方法等が提案されてきた。前者は、評価結果を数値で表現することにより、評価結果の判断においては客観性を満足している。しかし、評価の軸として「忠実度」「理解容易度」といった、その解釈が評価者の言語的直観に基づく主観的判断に依存する基準を採用しており、評価結果が評価者によって大きく異なってしまうという問題がある。この評価のばらつきは、特に、不完全な翻訳結果を評価する際に顕著に表れるが、現実のMTシステムを評価し、開発側にフィードバックする際には、翻訳に成功した場合よりも、失敗した場合についての検討が重要であることを考えると、この評価法は、実質的には原始的な主観評価の域にとどまっており、不完全なものであると言わざるを得ない。すなわち、この手法では、「なぜ評価者がそのような評価に至ったか」が明らかにならないため、評価結果をシステムの改良などにつなげ、発展的に利用することは困難である。特に、開発者においては、どの言語現象が、どの原因によって解析できなかったのかを知ることが重要であり、言語現象が複雑に絡みあった文の翻訳結果を単純に得点化するだけでは有効とはいえない。

一方、グラフ理論的アプローチの問題点は、訳文の品質が、本当に原文の構造と訳文の構造の照合により計算できるのかどうか、必ずしも明らかでないことにある。もちろん、原文と訳文の構造が対応することは、評価における重要なポイントの一つで

はある。しかし、機械翻訳システムの品質評価において解決すべき本質的な問題は、多様な現象を持つ言語表現の間の対応をどのようにして構造間のマッチング問題に還元するかにあるのであって、構造間マッチング問題に還元した後にあるのではないであろう。そして、実際の多様な言語現象を対象として、一般法則を導き出すことは困難であり、この点を無視して研究をすすめても、「機械翻訳における品質評価」に十分貢献することができないのではないかと恐れる。

3. 品質評価法の提案

さて、提案する品質評価法は、基本的には、test-suiteを用意するコーパスベースのアプローチである。そのポイントは、以下の2点にある。

- (1) 現在の機械翻訳システムの能力を勘案し、効率的な評価用コーパスを作成する。そのために、現在の機械翻訳システムの出力を分析し、共通して問題となっている項目を分析し、これに対して特に密になるようコーパスを用意する。
- (2) 項目を十分細分化することにより、各文に対する設問を、具体的かつ単純なものとし、評価の客観性を保持する。

(1) においては、コーパスベースの品質評価における評価用コーパスは、問題となる言語現象をすべて網羅していること（網羅性）と、評価対象として意味のない現象を含まず、効率的な評価ができる（効率性）という、ともすれば背反する要求を満たしていなければならない。

システムが実際の言語現象を網羅的にカバーしているかどうかを判断するには、全ての言語現象を、組合せ的に求めていくトップダウン的なアプローチが必要であろう。ただ、この場合、最終的には単語毎にその単語の持つ役割の数だけの例文を用意することになったり、全ての言語現象の組合せについての例文を機械的に作成することになり、コーパスを収集し使用することに、多大の労力を必要とし、そ

れを用いた評価を行なうことが非常に困難となることも考えられる。網羅性を考える場合は、基本的な現象を要領良く集積することが必要となる。

また、機械翻訳システムが、ある言語現象を処理できるかどうかを評価するためには、その言語現象を含む文を単独で翻訳するだけでは不十分である。曖昧性がない文を翻訳できるということは、当然の要求であるが、機械翻訳システムの性能を評価するためには、現実の機械翻訳システムにおいて正しく処理することが困難な状況、すなわち、曖昧性の解消に注目する必要がある。

そこで、このコーパスの作成においては、「構造の把握」と「構造の選択」という二つの視点から考えることにした。「構造の把握」では、ある(構文)構造をシステムが把握することが出来るかどうか(すなわち、必要な規則を持っているかどうか)を問う。これは網羅性を保証するトップダウンアプローチである。基本的な文の構造をそのシステムが処理することが出来るかどうかを、他の要素や文型とは独立して調べるものであり、翻訳システムとしての必要条件を判断している。「構造の選択」では、曖昧性のある文から、正しいものを選択できるかどうかを問う。これは、機械翻訳における問題点を実際の翻訳結果から抽出し、その問題性によって例文の粗密を決定するボトムアップなアプローチである。このボトムアップアプローチは、次のような考えで構成されている。機械翻訳における自然言語処理技術は多岐に亙るが、その中で、解析・生成といった表層の言語現象に直接関与する部分は、言語現象を軸として分類し木構造に表現することが可能であろう。コーパスの各文は、このような木構造の末端節点に付随している。別の言い方をすれば、この木構造は言語現象を機械翻訳における困難度という観点から階層化したものである。木構造中の各節点をより詳細な下位節点に分割し、最も下位の節点に、その節点に対応する現象を処理できるかどうかを判断する文を付随させる。

簡単な例を述べると、SVO および SVOO 構文について、noun verb noun という構造の文はもち

ろん SVO 構文を処理する規則がシステム中にあるかどうかを調べるという「構造の把握」からの視点で扱われるが、noun verb noun noun という構造の文は、SVOO 構文を処理する規則があるかどうかを調べる「構造の把握」の視点と、この文が、実は、動詞の後の二つの名詞が一つの名詞句を作る SVO 構文であった場合に、それを適切に判断し、SVOO 規則ではなく、SVO 規則と名詞連続規則とを選択し適用しているかどうかを調べる「構造の選択」の視点で、このコーパスに含まれることになる。

こうして作られたコーパスの各文には、対応する言語現象が扱えているかどうか、すなわち、その構造を正しく認識した翻訳結果が出力されるかどうかを直接的に問う質問が付いている。設問は十分に具体的かつ、原則として yes/no で答えられるようなものとする。なお、このような、設問の具体化・明確化は、上述のボトムアップアプローチによる言語現象の重要性の差別化と詳細化によってはじめて行なえるものであることを注意しておく。

このような、言語現象を分類した木構造を作成する場合には、機械翻訳システムの品質評価において重要である言語現象とさほど重要でない言語現象とを同列に取り扱うことをせずに、重要な言語現象については、木構造を詳細に分枝してゆき、例文を付加していく。これによって、言語現象の集積、さらにはこのコーパスの利用が実用的に効率良く行なえるようになる。このような作業は、機械翻訳システムを使用する過程で見つけられた現象を着実に収集していくというボトムアップ的アプローチとなる。

そして、既に述べたように、以上のボトムアップ的収集とは別に、基本的な言語現象を網羅的に分類した木構造をトップダウンに作り、これに沿ったコーパスも作成しておく。ボトムアップ的収集で得られた言語現象が、この網羅的トップダウン木構造の、どの位置にあるかを常に把握し、これを指針として文収集作業を進めることにより、常に網羅性を意識したコーパス作成が可能となる。

このように、この品質評価法では、「機械翻訳における困難さ」という観点から得られたボトムアッ

ブ情報と、「文法規則で記述される言語現象」という観点から得られるトップダウン情報とを利用することにより、網羅性と効率性を両立できる可能性を持つ。また、コーパスを翻訳した出力は、本評価基準における客観性保持の方針によって定量的に評価されるが、個々の文について評価するだけでなく、コーパス全体に対する得点化を行えば、その機械翻訳システムの品質の(他のシステムと比較可能な)総合評価とすることができよう。このような総合評価であれば、ユーザにとってもシステムの一般的な能力を知る助けとなり、導入時の指針となろう。

4. コーパスの構成

既に述べたようにコーパスは大きく(1)構造の把握と(2)構造の選択の2つに分類される。現在、コーパスには約400文が含まれている。図1に木構造(一部省略してある)を、図2に実際のコーパスの例を示す。なお、実際のコーパスの構成においては、このように2分化することは困難なものがあリ、そのような場合には、「構造の把握」部に含ませている。本稿では、コーパスの詳細を説明する余裕がないので、いくつかのコメントを述べるに止める。詳細は、参考文献[1]を参照されたい。

(4.1) 構造の把握

ここでは英文の構造を網羅するために、a)文法現象をホーンビーの英文構造の分類法に沿った文型に関する視点、b)時制に関する視点、c)助動詞とその代用形に関する視点、d)文種からの視点の4つの視点から眺め、各々の視点から細分類を行なった。a)は、ホーンビーのパターンに基づいているが、機械翻訳システムの品質評価において不必要な区別については、分類を省略してある。

(4.2) 構造の選択

ここでは翻訳で実際に問題となる言語現象を、a)構文構造の曖昧性に関するものと、b)訳し分けに関するものに分類した。
a) 構文構造の曖昧性に関するもの

英語解析においては、多品詞語の曖昧性解消(品詞認定)が重要な課題となる。解消の鍵には、たとえば以下のようなものがある(ここに挙げるのはごく一部である)。

- ・主語名詞句と述語動詞句における数と人称の一致
- ・選択ヒューリスティックス(文献参照)
- ・動詞格フレームの優先順位

また、多品詞語と品詞認定のキーが離れている場合、局所的な手がかりだけでは必要な情報が得にくいので、決定的な構文解析を行なうシステムでは問題が起こる可能性がある。そうでないシステムでも、途中段階で保持する曖昧性の数が増加するので、困難が増す。

動詞型認定の曖昧性は、ある場合には構文構造の曖昧性として現れ、ある場合には意味認定や訳し分けの問題として現れるが、その問題の重要性に鑑み、ここに独立させる。具体的な細項目は今後設定していく。

b) 訳し分けに関するもの

ここでは訳し分けの問題を生じさせる文法現象を、その中心となる文法事項からコロケーション、不定詞、動名詞、分詞、名詞、代名詞、冠詞、ハイフン、その他の9つに分類した。

コロケーションについては、動詞、名詞、形容詞、副詞、前置詞を組合せ、各組においてどちらかの品詞が2通りの訳語を持つようなフレーズを考えた。たとえば、動詞+名詞の組み合わせにおける、play tennisとplay the pianoのようなものである。ここで、 $5 \times 5 = 25$ の組合せが考えられるが、動詞+動詞等、通常の英語では出現しないと思われる組合せは割愛した。また、前後の品詞の一方しか固定できない場合は2通りの文を、前後の品詞の両方が固定できる場合は4通りの文を出した。上記の例で言えば、play tennisのtennisを固定し、like tennisとした時、likeとplayという動詞の違いによって、tennisを訳し分けてくるかどうかとも評価する。この他、be動詞または前置詞を間に挟んだ2つの名詞

構造の把握		名詞／限定詞／関係代名詞／接続詞
文型		名詞／副詞
動詞型		動詞／助動詞
VP1型		現在分詞／動名詞
VP2A型		形容詞／前置詞
(略)		前置詞／接続詞
VP24C型		限定詞／接続詞
VP25型		代名詞／接続詞
名詞型		限定詞／関係詞
NP1A型		接続詞／関係詞
NP1B型		係り先認定
(略)		前置詞／前置詞相当語句
NP2B型		現在分詞／過去分詞
NP3型		自由不定詞
形容詞型		副詞
AP1A型		名詞内部構造
AP1B型		動詞型認定
(略)		訳し分け
AP2型		コロケーション
AP3型		Verb+noun
時制		Verb+adj
現在		Verb+adv
現在完了		Verb+prep
(略)		Noun+verb
未来完了進行		Noun+noun
未来完了進行		Noun+prep
助動詞と、その代用形		Adj+noun
shall/should		Adj+prep
will/would		Adv+adj
(略)		Adv+adv
dare		Prep+noun
have_to		Noun+noun
文種		Noun+prep_p
疑問文		Verb+prep_p
yes-no 疑問		不定詞
一般動詞型		形容詞用法
be 動詞型		副詞用法
助動詞型		動名詞 Gerund
Wh 疑問		意味上の主語
Wh 必須格要素型		日本語の名詞への転換
Wh 自由格要素型		分詞
Wh 形容詞型		分詞構文
Wh 前置詞目的語型		形容詞用法
Wh 前置詞句の修飾型		Presp-Keiyō
付加疑問		Pastp-Keiyō
選択疑問		名詞
命令文		無生物主語構文
感嘆文		文頭固有名詞
What 型		名詞連続の「の」生成
How 型		同族目的語
特殊構文		代名詞
その他		人称代名詞の不訳
個別		一般人称の we, you, they, one
呼応		名詞の反復を避けるための one
記号		名詞の反復を避けるための that
括弧		天候・時間・状況等の it
引用		再帰代名詞の不訳
セミコロン		冠詞
構造の選択		定冠詞
構文構造の曖昧性		不定冠詞
品詞認定		ハイフン
名詞／動詞		その他
名詞／助動詞		命令文 + and/or
		関係詞非制限用法

図1 木構造

◎Sentaku-Kobun-Kakarisaki-Presp/pastp

【例文】 The magazine mentioned that the Japanese are more hypochondriac compared to the Germans.

【質問】 「日本人はドイツ人と比較してより心気症的だ」という意味に訳されていますか？

【ポイント】 "compared to" の係り先が "are more hypochondriac" であることが認定できているか。

【解説】 文の主動詞、"mentioned" にかけてしまつては誤り。

【例文】 That is a situation observed on workstations running X11R5.

【質問】 「X11R5を走らせているワークステーション上で」という意味に訳されていますか？

【ポイント】 現在分詞 "running" の主語が "workstations" であることが認定できているか。

【解説】 "a situation" と "observed" の間の関係が正しく認定できているかどうかポイントである。

◎Sentaku-Kobun-Kakarisaki-Inf

【例文】 We talked about a new method to build an electronic dictionary.

【質問】 「電子辞書を作る新しい技術」という意味に訳されていますか？

【ポイント】 TO不定詞句が "method" に支配されていることが認定できているか。

【解説】 不定詞句が目的を表す意味で "talked" にかかる解釈も可能だが、"method" 自身、TO不定詞を下位範疇化する名詞であり、こちらにかかる解釈のほうが自然だと思われる。本質的には、Hornby型NPICに対する設問と同じである。

【例文】 All items are distributed with permission to copy.

【質問】 「コピーする許可とともに」という意味に訳されていますか？

【ポイント】 "to copy" が "permission" にかかっているか。

【解説】 "permit to" の共起をとらえているか。本質的には、Hornby型NP1Aに対する設問と同じである。

【例文】 They won't have more to spend after they finished supporting the Government.

【質問】 「これ以上使うお金を持っていない」という意味に訳されていますか？

【ポイント】 "to spend" が "more" にかかっていることが認定できているか。

【解説】 "more" と "to do" の共起と関係の認定ができているかを問う。誤りとして、例えば不定詞句を "have" にかけて目的と解釈する可能性がある。

◎Sentaku-Kobun-Kakarisaki-Adv

【例文】 We can write English simply by writing Japanese.

【質問】 「日本語を書くだけで」「書きさえすれば」という意味に訳されていますか？

【ポイント】 "simply" が前置詞句 "by writing Japanese" を修飾しているか。

【解説】 限定の副詞が前置詞句の直前に現れた場合、前置詞句を修飾することが多い。

【例文】 These bolts differ from the standard ones only in their head form.

【質問】 「ヘッドの形においてだけ」あるいは「ヘッドの形だけにおいて」のように訳されていますか？

【ポイント】 "only" が前置詞句 "in their head form" を修飾しているか。

【解説】 (同上)

◎Sentaku-Kobun-Kakarisaki-Noun_p

【例文】 We like 10 mile runs.

【質問】 「10マイル競走」という意味に訳されていますか？

【ポイント】 まず "10 mile" がまとまることの認定ができているか。

【解説】 10の「マイル競争」という解釈もあり得る。この場合、局所的な数量表現を優先してまとめあげる必要がある。

◎Sentaku-Kobun-Bunkei

【例文】 It provides everything you want to make a UUCP connection.

【質問】 「UUCP接続をつくるために必要なすべてのものを」のように訳されていますか？

【ポイント】 "want" と "to" の間のギャップが認定できているか。

【解説】 動詞 "want" の二つの格パターン (VP6とVP7) の選択。

【例文】 These parts need painting.

【質問】 「これらの部品を塗る必要がある」という意味に訳されていますか？

【ポイント】 "require" (VP6E) の動詞に従う動名詞は受身の意味をもつ。

【解説】 主語 "parts" の助詞を「は」で生成して格解釈に曖昧性を持たせておけば、"need" を VP6A (通常の他動詞) として (誤って) 解釈しても、「これらの部品は塗ることを必要とする。」のように正しく解釈可能な日本語が生成される。

【例文】 The parts that require painting shall be painted within 24 hours.

【質問】 「塗装を必要とする部品は24時間以内に塗られなければならない」という意味に訳されていますか？

【ポイント】 "require" (VP6E) の動詞に従う動名詞は受身の意味をもつが、生成結果はそうのように解釈できるか。

【解説】 日本語の連体修飾節には、動詞と係り先体言の間の格表示が (通常) ないため、"require" の格パターンを VP6A と (誤って) 解釈して翻訳しても、ほぼ正しく解釈できる日本語になるとと思われる。

図2 コーパス (一部)

が互いの関係によって訳し分けが必要な場合もある。
(例：名詞 + be 動詞 + 名詞、名詞 + 前置詞 + 名詞、
動詞 + 前置詞句) 補語が形容詞の場合も考えられる
が、上記のパターンと重複すると思われるので割愛
する。

不定詞には名詞用法、形容詞用法、副詞用法があ
るが、名詞用法については特に訳しわけが必要な所
は多くないと思われるので割愛している。動名詞に
ついては、これを日本語のサ変名詞に変換する場合
と「〇〇すること」というように訳す場合がある。

代名詞に関して、人称代名詞を辞書の通り訳さず
に、それが指している名詞を使って代用することが、
人間の翻訳においてはしばしば見られる(人称代名
詞の前方照応)。しかし、これは1文を越えての照
応であり、現在の商用システムの限界を上回ってい
るため、本コーパスでは扱わない。定冠詞の前方照
応も同様の理由で割愛している。

ハイフンについてもコロケーション同様 25 の
組合せを作成するが、全体的に見た時、ハイフンは
必ずしも主となる問題点ではない。また、意味解析
及び生成というよりは、辞書でハイフン語の処理を
する方が実際的であると考えられる。このため、ハ
イフンについては、さほど突っ込んだ考察をしてい
ない。

5. 検証

このコーパスの実用性の検証と、設問の修正を行
なうために、ハードウェアタイプの異なる3種のシ
ステムにおいて、シミュレーションを行なった。3
85文全体に対する成功率は、44%から65%ま
での開きがあった。このコーパスは、作成途中で
あり個々のシステムの成功率自体には意味がないが、
各システムの成功率を相互に比較すると、把握部
においても選択部においても順序関係が存在し、シ
ステムの翻訳できる言語現象の範囲に差があることを
表現できている。したがって、ここで提案したコー
パスベースの評価法により、機械翻訳システムの品
質評価を行なうことが可能であると考えられる。

一方、このシミュレーションによって、このコー

パスにおけるいくつかの問題点が明らかになった。
たとえば、各設問は訳を示して、そのように訳され
ているかを聞くものが多いが、設問のポイントと直
接関係のない単語の訳に問題があったり、一部の
みが成功しているものがあった。また、システムは必
ずしも詳細に言語現象を理解してはいないように思
われるが、翻訳結果の日本語の持つ曖昧性のために、
人間にとっては、設問の条件を満たしているよう
にも読めるという場合もあった。

6. おわりに

(社)日本電子工業振興協会において開発中の機
械翻訳システムの品質評価法およびそこで用いるコー
パスについて説明した。コーパス自体は、まだ、
作成途中であり、今後は、上記のシミュレーション
によって得られた問題点を解決するとともに、現実
の翻訳例を収集検討することにより、機械翻訳シ
ステムの翻訳品質の評価を的確に行なえるようなコー
パスを作成していく予定である。

【参考文献】

- [1]「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)
- [2] Robert Milne, "Lexical Ambiguity Resolution in a
Deterministic Parser", pp. 45 -- 71, in *Lexical Ambiguity Resolution*, Morgan Kaufmann Publishers, Inc.
(1988).
- [3] Graeme Hirst, "Semantic Interpretation and the Reso-
lution of Ambiguity", Cambridge University Press,
(1987).
- [4] Jerry R. Hobbs et al, "Two Principles of Parse Prefe-
rence", in *Proc. COLING-90*, pp. 162 -- 167, (1990).
- [5] A. S. Hornby, 「英語の型と語法(第2版)」, オ
ックスフォード大学出版局(東京), (1977).
- [6] 小川芳男他, 「よくわかる英文法[再訂新版]」,
旺文社, (1991).
- [7] 江川泰一郎, 「英文法解説(改訂新版)」, 金子
書房, (1964).