

機械翻訳システムにおける後編集過程のユーザ行動

荒井 良徳

嶋田 敦夫

成田 真澄

arai@ttc.rdc.ricoh.co.jp

shimada@ai.rdc.ricoh.co.jp

barbra@ttc.rdc.ricoh.co.jp

応用技術開発センター

知能工学研究センター

応用技術開発センター

(株) リコー 中央研究所

英日機械翻訳システムが出力した訳 (MT訳) を修正する過程 (後編集過程) におけるユーザの (認知的) 行動を知ることは、英日機械翻訳システムなどの設計において非常に重要となる。しかし、このような翻訳過程における人間の内的プロセスに関する心理学的な研究は、これまでほとんど行われていなかった。

ここでは、心理学実験及びプロトコル解析手法により、英日機械翻訳システムにとって有用となる User's Model を得ることを目的とした。この実験では、原文を与えた上に、MT訳をワープロ上に提示してこれを修正してもらった。その結果、MT訳の後編集過程においては、ユーザは基本的には提示された訳を何とか修正しようとし、1) 原文を理解する過程、2) 意味表象とMT訳を比較する過程、3) MT訳を修正する過程の3つの大きな過程により行動することなどが示唆された。

User Behavior During Post-Editing in Machine Translation

Yoshinori ARAI

Atsuo SHIMADA

Masumi NARITA

arai@ttc.rdc.ricoh.co.jp

shimada@ai.rdc.ricoh.co.jp

barbra@ttc.rdc.ricoh.co.jp

I.T. Department

K.E. Department

I.T. Department

R and D Center, **RICOH Company, LTD.**

16-1 Shin-el-cho, Kohoku-ku, Yokohama-shi, Kanagawa 223, JAPAN

Analysis of the user's thought processes and behavior during post-editing is important for the design of machine translation systems, yet this type of psychological research has yet to be conducted.

With the aim of creating a user's model useful for an English-Japanese machine translation system, we conducted psychological experiments and employed protocol analysis. We asked the subjects to correct the output from a machine translation system. The Japanese translation was displayed on a word processor screen, and the user could refer to a printed copy of the original English. The results indicate that three types of behavior are involved: 1) understanding the original, 2) comparison of the understood meaning with the machine translation output, and 3) editing the output.

1. はじめに

近年機械翻訳技術の発展はめざましく、翻訳の精度もある程度実用に耐えられるものになり、様々な製品が市場に現われるようになってきた。しかし、これらの機械翻訳システムは、ユーザが自然に作業を行える環境のユーザインタフェースになっているとは言い難い。それは、言語学としての機械翻訳技術は十分に研究されているが、人間が実際にどのような翻訳をし、またどのようなシステムを望んでいるかなどの、ユーザサイドに立った研究は進んでいないという事実に基づくものと考えられる。

認知心理学では、人間の記憶や理解などのメカニズムを解明するべく、様々な研究が進められている。また、ユーザインタフェースの重要性が認知されるようになり、User's Modelなどの認知心理学からのアプローチによる研究も行われるようになってきた。しかし、翻訳という行為の人間の内的プロセスに関する心理学からのアプローチはまだ少ない¹⁾。

一般的な対話型の機械翻訳システムを設計する上で翻訳の精度と同じく重要な要因として、ユーザインタフェースがあげられる。なぜならば、現状の翻訳技術は、完全翻訳のレベルにまで達しておらず、必ず人間が介入しシステムの不十分な点を人間が補充する必要が生じてしまうからである。

そこでここでは、まず、プロトコル解析という観察手法により、英日機械翻訳システムにおけるシステムの翻訳部が出力した訳 (MT訳) を人間が修正する過程 (後編集過程) において、ユーザの行動特性・認知的特性 (User's Model) を得ることを目的とする。

2. 実験

ユーザにとって使いやすいユーザインタフェースとは、ユーザが自然に作業を行なえる環境を機械翻訳システムが用意していることである。つまり、ユーザの行動や思考に適合したデザインがなされているということである。その

ようなユーザインタフェースを設計する指針を得るには先ずユーザの自然な振る舞いを記述することが必要である。ユーザの行動特性・認知的特性が何であるのかわかれば、後編集行動が認知的ストレスなしに自然に行なえるようなインタフェースのデザインが可能になる。そこで、ユーザの行動特性・認知的特性 (User's Model) を把握することを目的として実験を行った。

2.1. 実験計画

後編集過程におけるユーザの行動特性・思考特性を記述し得る実験を計画しなければならない。本実験では、プロトコル解析法という観察法の一種を用いた。

〔プロトコル解析法の概要〕

ワードプロセッサなどの機器操作を、操作者の認知過程から捉えると、問題解決過程と考えることができる。一般に、問題解決や思考などの人間の high-order 認知過程を詳細に計測する場合、心理学ではプロトコル解析 (protocol analysis) という定性的な手法がとられることがある。プロトコルとは被験者に課題を声に出して (thinking-aloud) 解かせることによって得られる発話データで、この発話データを分析することによって被験者の詳細な思考・推論のプロセスを記述することができる。プロトコルを収集するための手続きは、様々なバリエーションが考案されている。典型的なプロトコル収集の実験は、被験者に課題 (例えば、ワードプロセッサで会議の案内状を作成する) を与え、これを達成する様子を VTR 等で記録することでなされる。この時、被験者には課題を解いているとき何を考え、どのような疑問を持っているかなどをすべて口述するように教示される。また、実験者は、被験者の課題遂行に対して「考えていることを声に出してください」、「何をしようとしていますか」などの他はなるべく干渉しないようにしなければならない。

伝統的な実験心理学のパラダイムでは、主に課題達成時間やエラー発生率、などの定量的なデータが測定される。このようなデータのみではユーザの認知プロセスに直接触れることは難しい。しかしプロトコルのような定性的なデータは、後述するような問題点はあるが、ユーザがどのように課題を組立、何を手掛かりに、どのようなプラン・方略で、機器を操作するのかを表現してくれる。これは、機器の操作性をデザインする上で非常に有用である。

被験者のプロトコルはまず、実験対象の機器のどの部分に具体的な問題があるかを示してくれる。また、そのような問題がなぜ引き起こされたかを知ることができ、対応策や改善案などの具体的な設計変更の指針を導くことができる。この利点を利用し、商品化前の機器にプロトコル解析を施して問題点を洗い出しておけば、事前にある程度の製品の改善が行なえる。また、プロトコルはさらに、ユーザの認知的な側面や行動特性なども明らかにする。ユーザの考える機能の単位が推察できるとすれば、システムの機能の構造や、メニューの階層を決定できたりする。また同様に、ユーザが実験中使った言葉をそのままシステムの機能名称に利用することもできる。さらにユーザが課題や問題（トラブル）に直面したときにどのような行動をとるかを知ることができる。これによってトラブル時のユーザにどのような援助が効果的かが推測できる。プロトコルから得られるこのような情報は今後の機器設計に有効なデータである。

[プロトコル解析法の心理学的意味]

プロトコルのような認知プロセスの被験者自身による言語報告は、内観法と言われ伝統的な実験心理学では長く科学的なデータとしては扱われなかった。Ericsson, K.A. & Simon, H.A.^{2) 3)} は、プロトコルのような言語報告がデータとしての信頼性を持つための心理学的手法の確立を目指した。

被験者が言語報告したデータが被験者の認知プロセスを的確に捕えているか否かの第1の条件

は、言語報告が短期記憶内の情報に対して直接的に行なわれたかにかかっている。人間の記憶モデルによれば、現在注意している情報は短期記憶内で活性化（意識化）されていると考えられるので、なんらかの認知プロセスの活動に対して言語報告が必要な場合は、意識でまた直接検索が可能である短期記憶からの情報の言語報告が望ましい。従って、被験者が課題を解く過程で同時に、つまり短期記憶内で情報が活性化されている間に被験者が言語報告を行なえば、認知プロセスにとって正確で信頼性のあるプロトコルが収集できることになる。

一方、実験後に認知プロセスについて言語報告させようとする、情報は既に長期記憶内に貯蔵されているので、情報の再構成や言語報告するための推論という、課題を解いていた認知プロセスとは違うものを報告してしまう。

ところで、課題を解く過程と同時に言語報告させるのに問題がないわけではない。例えば、本来のプロセスに解釈のための中間的処理が介在する場合も生じ、本来のプロセスが歪曲され信頼性のあるプロトコルを得られない可能性が高い。このようなことを防ぐために、被験者への教示で、短期記憶から直接アクセスできるものだけを言語報告させるようにし、さらに、実験者の被験者への介入は必要最小限に留めるなど、注意が必要となる。

また、被験者に自らの思考を口述させるという状況は不自然であり、被験者によっては発話そのものが困難な場合がある⁴⁾。従って、一般に実験を行なう際は被験者に発話訓練が必要であるとされる。

最後に、プロトコルは定量的なデータと比べて、データとして多量の情報を含んでいるが、解析方法には確立された最適の方法があるというわけではない。従って、情報量の多い多量のデータに埋もれてしまわないためにも、プロトコルを収集する際には、十分な事前の問題分析、適切な分析の枠組み、それに対応した適切な課題等の実験計画が必須である。これらを十分にコントロールすればプロトコル解析はデザイン

に有用なデータを提供してくれる強力な測定技術である。

2.2. 実験の概要

【目的】

MT訳を修正し訳文を作成する被験者の自然な後編集過程を、被験者の行動特性及び認知的特性から記述する。さらに、全ての被験者に共通の特性をもとに後編集過程のモデル化を試みる。

以上によって得られた結果が、機械翻訳システムのユーザインターフェース設計の指針の基礎データとなり得ること。

【方法】

刺激文： 文章の種類としては、製品のユーザーズマニュアルとした。また、被験者による当該製品の事前知識の差を均質化する狙いから、自動車を選定した。以上から訳出対象文書はフランス製の自動車ルノー5のドライビングマニュアルとした。

上記マニュアルをリコーの英日機械翻訳プログラムを用いてMT訳を作成した。これを実験の被験者が普段使用するワードプロセッサフォーマットに予め変換したものを実験刺激とした。

被験者： 被験者は企業内で翻訳業務に携わっている者6名に対し実験を行った。

材料： 8mmビデオカメラRICOH R-630S、三脚、ワードプロセッサ

教示： 被験者に対しまず次のような質問を行ない、続いてその答えに応じた教示を行なった。

「まず、あなたは普段、英語を日本語に訳される時、どのようになさっていますか。一度、一気に直訳を作ってから、日本語としてこねた訳にしますか。それともいきなり最終的な訳を作ろうとしますか。」

最初から本訳を作成すると答えた被験者に対

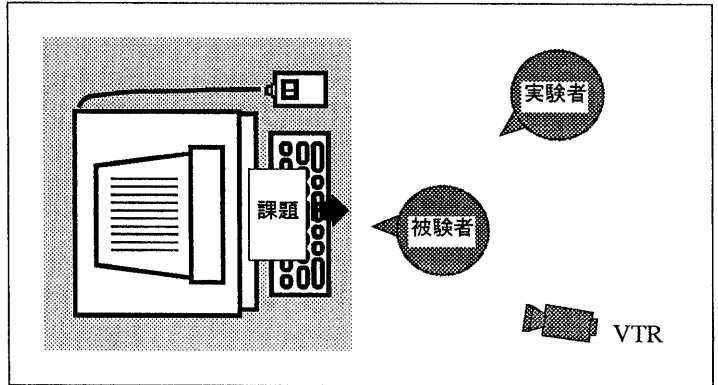


図1.実験状況

するインストラクション： 「画面には、フランスのルノーサンクという車のドライビングマニュアルを訳したものが表示されています。普段他の人に見せるためにあなたが作っていらっしゃる程度になるようにこのワープロを使って修正もしくは編集してください。適当なところで終了ですといえますから、その合図があるまで作業を続けてください。それから作業をしながら、ちょっと難しいかもしれませんが、その時、頭の中で考えていることを全て口に出して言うよう心掛けてください。私達はあなたの考えていることに興味がありますのでよろしくお願いいたします。」

先に下訳を作成すると答えた被験者に対するインストラクション： 「画面には、フランスのルノーサンクという車のドライビングマニュアルを訳したものが表示されています。普段あなたがするようにこれを元にこのワープロを使って修正もしくは編集し下訳を作ってください。適当なところで終了ですといえますから、その合図があるまで作業を続けてください。それから作業をしながら、ちょっと難しいかもしれませんが、その時、頭の中で考えていることを全て口に出して言うよう心掛けてください。私達はあなたの考えていることに興味がありますのでよろしくお願いいたします。」

課題： 上記マニュアルについてのMT訳を普段和訳するレベルの文章に編集する。

表1. 被験者属性及び実験実施状況

被験者	業務内容	実験日時	実験場所	ワードプロセッサ
KMN	社内文書の翻訳	17.Oct.90	所内実験室	Fuji Xerox JStar8083
TIS	技術文書の翻訳	18.Oct.90	H社内	PC9801 一太郎
YTK	翻訳と英文作成	19.Oct.90	R社内	Ricoh RS8000
HTK	機械翻訳調査	23.Oct.90	J社内	PC9801 一太郎
QKT	機械翻訳調査	23.Oct.90	J社内	PC9801 一太郎
NMY	MTSを用いた翻訳サービス	29.Oct.90	所内音響室	PC9801 一太郎

手続き：被験者に上記のように教示し、図1に示すような実験を行なった。MT訳は被験者が普段使用しているワードプロセッサの画面に表示され、原文のコピー及び辞書を併用して課題を遂行させた。また、実験を始める前に、各被験者にいつもの翻訳スタイルを聞き、通常下訳を経過する人であるか否かを確かめたうえで、いつものように翻訳をしてもらった。尚、実験者は被験者の作業に介入しないように被験者の側に付き口述を促した。被験者の課題遂行過程は全てVTRに記録した。課題は、被験者の疲労を考慮して30分で終了させた。但し、予め被験者に30分程度で実験が終了することを知らせると、通常より、丁寧に作業を行なう可能性がある（実験を行なうことによる効果）、被験者には知らせなかった。

また、実験終了後、データ解析時の補助として、機械翻訳についての若干のインタビューを実施した。

2.3. 心理実験実施状況

表1に、実験に参加した被験者の属性及び、被験者が使用したワードプロセッサの機種を示す。

また、6名全ての被験者は下訳を作らずに最初から本訳を作成すると答えた。

2.4. 行動発話抽出

VTRに記録された6名に対し、プロトコルを作成した。

このプロトコルの作成は、3段階に分けて作業を行った。まず、第1段階としては、複数人の評定者が、VTRから被験者が行った行動及び発話をあらかじめ用意した比較の詳細な行動パターンを使用して、記録した。

次に第2段階としては、被験者の典型的な行動や思考等のボタン（認知単位）を基に、個々の被験者がMT訳を修正する過程の認知的メカニズムが推定できるように、整理した表2の認知

表2. 抽出した認知単位 (cognitive units)

認知単位	内容	認知単位	内容
???	不明	辞書	辞書を引く
読む-原文	原文（コピー）を音読する/見る	保留	わからない箇所を後回しにする
読む-MT	画面のMT訳を音読する/見る	解決	思考して問題が解決した
読む-訳文	自分で修正した訳文を音読する/見る	未解決	思考しても問題が解決しなかった
思考-表現	文体や語、語句等の表現を考える	不明	問題が解決したか否かわからない
思考-内容	原文の表す内容について考える	入	訳文を入力する
思考-用語	語や語句について意味を考える	削	MT訳、訳文を削除する
対応探し	MT訳（部分）に対応する原文を探す	削入	訳文をMT訳にオーバーライトで書く
MT-誤り	MT訳の誤訳を指摘する	置	語、語句を置換する
発話-訳	訳文を発話する	写	語、語句、文をコピーして入力する
発話-内容	原文の内容について言及する	移	語、語句、文を移動して入力する

単位表を基に、行動発話データを整理していった。

そして第3段階として、上記で得られた各評定者のデータを総合的にまとめた。この時、被験者がどのような推移で最終の訳文までたどり着いたかが分かるように、訳文の途中経過も記していった。

表2に示される認知単位を用いて加工したプロトコルの例を図2に示す。

3. 結果及び考察

3.1. 典型的なユーザ行動パターン

加工されたプロトコルによると、被験者は概ね<読む-原文>から<読む-MT>を経て、<読む-原文・読む-MT・入・削>サイクルを繰り返す。換言すれば、被験者は先ず原文を読み次にMT訳を読んでから、ある単位毎にMT訳修正を進めている。例えば、図2の被験者TISの場合、原文を読

んだ後、「ステアリングをロックするには」という発言があり、続いてMT訳を読んでいる。その後、もう一度原文、MT訳を読み、最初の単位「進むことに錠を掛けて」を「ハンドルをロックするには」に修正する。以下同様に、「鍵を移し」、「そして回転するために」、「進むまでの車輪は錠が掛かるのを聞けます」の順に<読む-原文・読む-MT・入・削>で修正が進む。

以上のように、被験者のMT訳修正過程は、原文を読む過程、MT訳を読む過程、ある単位毎に原文・MT訳を読み修正する過程の3つに分解することができる。また図2の訳文の遷移と編集行動における下線部がMT訳の左から右へ規則正しく移動していることから、MT訳の修正の単位は、被験者が規定する意味単位ではなく、多くがMT訳の「、」から「、」の間の文である。

次に各過程をそれぞれ詳細に見てみる。第1の

文番号	3	被験者	TIS	評定者	荒井・嶋田	シート番号	1
時間	プロトコル	認知的行動	解決/未解決	訳文の遷移と編集行動			
	「ステアリングをロックするためには」	読む-原文 読む-MT 発話-訳		To lock the steering, remove the key and turn the wheel until the steering can be heard to lock.			
	「掛けてではなく掛けるためにですね」	読む-原文 読む-MT MT-誤り		進むことに錠を掛けて、鍵を移し、そして回転するために、進むまでの車輪は錠が掛かるのを聞けます。			
				☒ 文節入替 (誤訳訂正) ☒ ハンドルをロックするには、 <u>鍵</u> を移し、そして回転するために、進むまでの車輪は錠が掛かるのを聞けます。			
	「これは、はずしてですね」	読む-原文 読む-訳文 発話-内容		☒ 文節入替 (誤訳訂正) ☒ ハンドルをロックするには、 <u>キーを外し</u> 、そして回転するために、進むまでの車輪は錠が掛かるのを聞けます。			
	「これは何だろう？不定詞の使い方がメチャクチャだから訳が全部おかしくなっています。全部消します」	読む-原文 MT-誤り		☒ 残り			
	「ロック音が聞こえるまで」	読む-原文 発話-内容		☒ ハンドルをロックするには、キーを外し、 <u>ロック音が音で確認できるまで</u> 。			
	「ホイールを回す。ホイールは車輪、車輪を回すということはハンドルを回す」	読む-原文 発話-内容 思考-表現	解決	☒ ハンドルをロックするには、キーを外し、 <u>ロック音が音で確認できるまで、ハンドルを回してください。</u>			

図2. 加工されたプロトコルの例 (被験者TIS: 文番号: 3)

過程は原文を読む過程であるが、この過程で被験者が言及する話題は、「St.というのが何のことだかわからない。専門用語なのでしょうか(YTK)」のような専門用語に関するものと、「・・・略・・・鍵を差し込むという意味でしょうかね(TIS)」や「・・・略・・・これはキーを抜くことなんですかね・・・略・・・(KMN)」のような原文の意味に関するものに集中している。また、この過程に費やされる時間が比較的短いのも特徴である。このことから、被験者は、先ず専門用語に注意してあまり時間をかけずに、原文の大まかな意味表象を捉えようとする傾向があると考えられる。

第2の過程はMT訳を読む過程である。被験者の言及は、「全面的に変えたいですね(YTK)」や「日本語で何処か使えるところがないか読んでいます(NMY)」、「休み・・・略・・・はちょっと意味がおかしい(QKT)」のように、MT訳に関する方針や誤りの指摘に関するものが多い。MT訳を読む過程とは、MT訳が使えるか否かを、前の過程で得た大まかな意味表象とMT訳を比較して判断する過程と推察できる。前の過程同様、この過程も費やされる時間は非常に短

い。

上記2つの過程は、主に1文に対して行なわれる処理で、実験では複数の文にまたがって処理をした被験者は1名も観察されなかった。被験者は1文単位で作業を行なっているものと考えられる。

第3の過程は、実際にMT訳を修正する過程である。まず被験者はMT訳の一番最初の部分(先頭から最初の「、」まで)に対応する原文を読み意味を把握し、再びMT訳を読んで修正すべき場所や正しい表現を考える。従って「鍵を移すとは言いませんね(YTK)」や「自由なままでしようというのは自由に動くです(HTK)」、「エンジンをかけることって何て言います?(YTK)」等のMT訳の誤りの指摘や表現に関する発話が主となる。また、ここでは「ある程度の英語は日本語になっているので(YTK)」や「鍵をキーに直します(QKT)」のように、表現を決定するためのルールが形成されていることを示唆する発話が数多く観察される。これは、MT訳の修正方法がパターン化しルールとなるものと考えられる。実験で確認されたルールは、わからない場合は英語をカタカナにする「ignition-

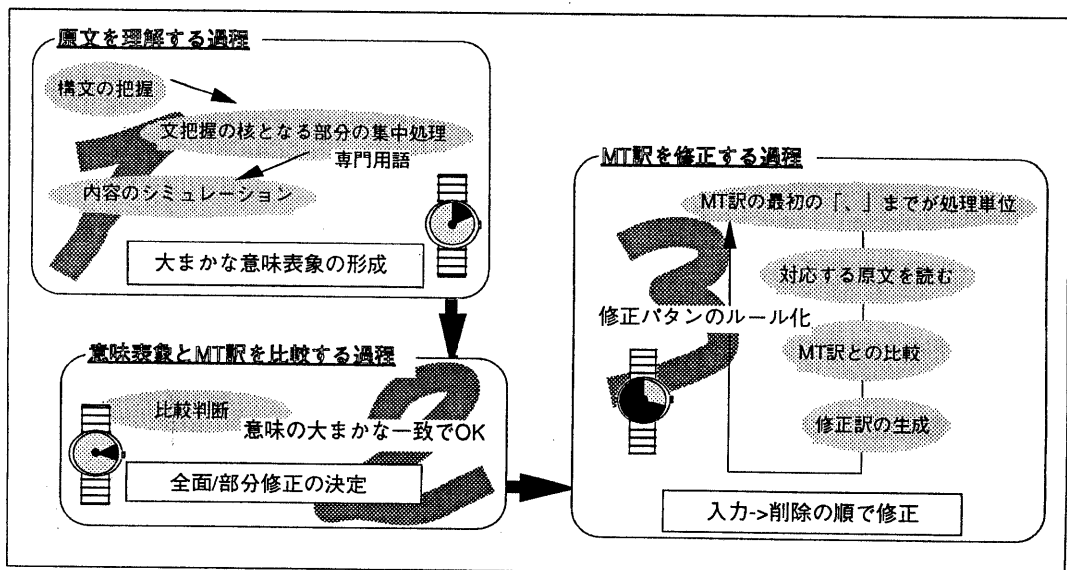


図3.訳文修正過程のラフなUser's Model

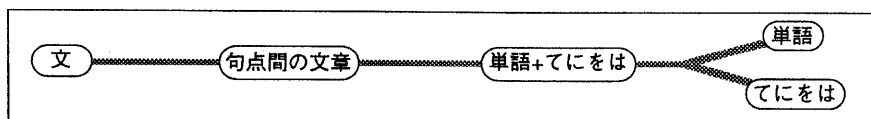


図4. 処理単位の階層構造

>イグニッション (YTK等)」や英単語と日本語を1対1対応させる「start->始動 (YTK)」等である。

さらに、第3の過程では<MT-誤り>が顕著に多い。「鍵は移すではなく回す (YTK)」、「放ってではなく放してが正解です (QKT)」等と基本的に処理単位毎の誤訳の指摘が多く文全体のバランス等の複数の処理単位をまとめて対象とした、誤りの指摘があまり見られないのも特徴と言える。

3.2. まとめ

図3に、以上から考えられるMT訳修正過程のラフな認知モデル(User's Model)を示す。

結果をもとに、先ず被験者の訳文生成の処理単位の点について考察を行なう。被験者の処理単位は結果が示すように、ほぼ全ての文で、被験者は句点と句点の間の文章を1回の訳文生成単位としている。さらに、訳文生成単位での編集単位は、単語、単語+助詞、てにをはで、これらについて入力・削除が行なわれている。以上から被験者の処理単位は図4のようになることがわかる。

次に、訳文生成過程で獲得される修正ルールについて考察を行なう。獲得されたルールは、「わからない言葉は英語をそのままカタカナにする」や「同じ単語の訳語は文書全体に共通にする」、「受動態の文章を能動態にする」等比較的単純で定型的なものが多い。また、これらルールの適用は、第1の過程でも、頻繁に行なわれていることから、被験者にとってはかなり優先度の高い編集行為であると考えられる。

4. おわりに

MT訳修正過程のユーザの認知的行動は、

- 1：原文の理解のステージ、
- 2：下訳と原文内容の比較のステージ、
- 3：下訳の修正のステージ

の3ステップに分けられることが示唆された。

ユーザの顕著な行動としては、提示された下訳の一文のみに注視し、ある単位(”、”読点間)で提示されている順番に従って、その場で考えられるより自然な言い回しに修正していく。これらの行動は、比較的機械的にルールとして覚えられて行くようである。また下訳を修正する時点では、あくまでもその下訳が主体となり、原文の内容を見直しながら修正をしていた。

今後は、User's Modelのより深い検討、考察を行い、対話的MTシステムのコンセプトの検証を行う。また、これらに基づきユーザにとって使いやすいMTシステムとの設計指針を抽出していく。

謝辞

本研究の実験にご協力をいただいた、被験者の方々に感謝致します。

参考文献

- [1] 杉本：”翻訳における内容的知識と言語的知識の相互作用－理解のレベルと訳語選択－”、日本認知科学会テクニカルレポート、No. 16
- [2] Ericsson, K.A. & Simon, H.A.: "Verbal Reports as Data", Psychological Review, Vol.87, No3, pp.215-251, 1980
- [3] Ericsson, K.A. & Simon, H.A.: "Protocol Analysis", The MIT Press, Cambridge, Mass, 1984
- [4] 加藤：”質問記録分析法－計算機ユーザの行動理解をはかる－”、Human Interface Symposium, pp.211-216, 1986