資料

要約文章の多様性

——要約産出方略と要約文章の良さについての検討——

邑本俊亮*

THE VARIETY OF SUMMARIES

—— Strategies of summary production and the goodness of summaries ——

Toshiaki MURAMOTO

The present study investigated the variety of summaries and the determinants of its goodness. In Experiment 1, subjects were given two stories and asked to write a summary under either of the following two conditions, the referring-text condition in which subjects were allowed to see the original text while writing a summary, and the referring-memory condition in which subjects had to remember the story and summarize it without seeing the original text. Qualitative differences of summaries were found according to conditions, and the varieties of summaries were also observed among subjects. In Experiment 2, subjects were required to evaluate the summaries obtained from Experiment 1. Results showed that the summaries produced in referring-memory condition were evaluated poorer than the summaries in referring-text condition, because they tended to contain inadequate information and a lack of important information. Evaluations were divided concerning some summaries. Finally, four types of strategies of summary production considered to cause a variety of summaries, were proposed and the relationship between the strategies and the goodness of summaries were discussed.

Key words : summary production, strategies of summary production, goodness of summary, text comprehension, narrative text.

問題

人間は、文章を読んで理解するとき、そこに含まれている「要旨」なるものを把握し、求められればそれを言語化することができる。この「文章の要約」という作業は、人間の文章理解に関わる重要な認知活動の1つであり、その認知過程は我々が解明すべき大きな問題となっている。しかし、この認知過程には、「理解」と「産出」という、それら自体で大きな研究テーマとなりうる2つの認知的側面が含まれており、その解明は必ずしも容易なものではない。

文章理解研究における要約実験の多くは、文章理解結果としての心的意味表象モデルの検証のために用いられる（たとえば、Thorndyke, 1977; Omanson, 1982; van den Broek & Trabasso, 1986, など）。そこでは、要約は理解内容を端的に表わすものであり、要約文章中に現れる情報は文章の意味表象の中で支配的あるいは中心的な役割を果たしている、と仮定される。そして、実際に得られる要約実験データを提案されるモデルによってうまく説明できるならば、そのモデルの心理的妥当性が検証されたとするのである。しかしながら、この類の研究においては、要約産出時の心的処理過程そのものについては何ら言及されることはない。研究対象の中

* 北星学園大学 (Hokusei Gakuen University)

心はあくまで表象モデルであり、要約データそれ自体は補助的役割を果たしているにすぎない。

それに対して、要約の認知過程それ自体を明らかにしようとする研究もないわけではない。たとえば、Kintsch ら (Kintsch & van Dijk, 1978; van Dijk & Kintsch, 1983) の提案したマクロ規則 (macrorules) は、文章理解時の処理過程の 1 つとしての要約化の過程、すなわち要旨把握の過程における情報集約規則として提案されたものだが、そのような規則の使用が実際に産出される要約文章に現れていることが確認されており、規則の妥当性が支持されている (Brown & Day, 1983)。また、要約過程を理解過程の延長上に位置づけ、ある種の理解表象モデルに基づいて要約生成の規則やアルゴリズムを積極的に定式化しようとする研究 (たとえば、Rumelhart, 1977; Lehnert, 1982; 呂本・阿部, 1987; Alterman, 1986; Alterman & Bookman, 1990) もある。しかしながら、いずれの研究も要約現象の大まかな部分は捉えてはいるが、実際に人間が産出する要約文章そのものを完全に説明あるいは予測するにはかなり困難な点が多く、要約産出過程の決定的なモデルが提案されるには至っていない。

要約産出過程を解明していく上での困難さの大きな原因は、人間が産出する要約文章の多様性にある。人間はどのような長さの要約を求められたとしてもそれに対応することができるし、たとえもとのテキストを見ることができない状況でも要約を書くことができる。また、同じ文章を読んでもそれに対する要約文章にはかなり個人差が認められる。もちろん、同一の読み手でも、その時々によって要約の内容が異なることが考えられる。これらの要約産出時の外的制約や個人差による要約内容の変化については、これまでの研究においてはほとんど考慮されていない。これは、それらのほとんどが要約過程の一般性、普遍性を求めることを第 1 の目的としていたからであろう。

さらに、このような要約文章の多様性は、“要約文章の良さ”といった問題とも関連してくる。多様な要約文章のうちどのような要約文章が客観的に“良い”と評価されるのであろうか。

本研究では、要約の多様性は要約産出時に用いられる“方略”の違いを反映しているのではないかと仮定する。そして、それを明らかにするため、要約文章の個人差の検討に加えて、要約産出時の外的制約による要約内容の違いについても検討する。本研究でとりあげる要約産出時の外的制約は、もとのテキストを見ることができそれを参照しながら要約を書くことができ

る事態ともとのテキストを見ることができず記憶を頼りに要約を書かなければならない事態の 2 種類である。以下では、前者をテキスト参照事態、後者を記憶参照事態と呼ぶことにする。

本研究の目的は、2 種類の要約産出事態 (テキスト参照事態および記憶参照事態) でそれぞれ産出された要約文章の相違点を明らかにし、それに基づいてそれぞれの要約産出過程の違いについて考察すること、また、実際に産出される要約文章の多様性を確認し、要約産出過程において用いられる方略の個人差について考察すること (実験 I)、そして、要約文章に対する“良さ”の評価傾向を調べ、要約の良さを規定する要因について考察すること (実験 II)、である。なお、本稿では、要約産出の基礎となったもとのテキストのことを「物語文章」(あるいは単に「物語」)、要約産出の結果として得られた文章のことを「要約文章」(あるいは単に「要約」と呼び、両者を区別することにする。

実験 I

方法

被験者 被験者は北海道大学文学部の学部学生および大学院生 16 名 (男性 10 名、女性 6 名) で平均年齢は 22.1 歳であった。

材料 2 種類の物語文章 (物語 I、物語 II) が用いられた。物語 I は、井上二美作『ハトとネコ』で長さは 2954 文字、物語 II は今江祥智作『まぼろしの海』で長さは 3356 文字であった。それぞれの文章は、縦長 A 4 版の大きさの用紙に、横書きの 40 文字×40 行の形式で印刷され、いずれの文章に対しても、改行などによる空白部分を含めると 3 ページからなる小冊子が作成された。

手続 被験者はまず、2 種類の小冊子のうちの 1 つを渡され、そこに印刷されている物語文章を読む。文章を読むペースは被験者の自由である。読み終わると、もう一方の小冊子が渡され、同じように文章を読む。なお、2 種類の物語を読む順序は被験者間で相殺されている。各物語の平均読解時間は、物語 I が 4 分 1 秒 (SD=1 分 12 秒)、物語 II が 5 分 41 秒 (SD=1 分 56 秒) であった。その後、被験者は、それぞれの物語に対して、テキスト参照事態または記憶参照事態のいずれかの要約産出条件で要約を書くように求められる。ただし、1 番目の物語と 2 番目の物語とは要約産出条件が替えられ、各被験者は必ず 2 通りの要約の仕方を体験することになる。なお、テキスト参照事態で要約を書く場合には被験者にもとのテキストが与えられる。

被験者は、要約の長さは自由であるが、いずれの要約を書く際にも「良い」要約を書くように強調される。要約産出に要した平均時間は、テキスト参照事態においては、物語Ⅰに対しては15分8秒(SD=3分28秒)、物語Ⅱに対しては18分52秒(SD=6分55秒)、記憶参照事態においては、物語Ⅰに対しては12分37秒(SD=4分25秒)、物語Ⅱに対しては17分37秒(SD=5分33秒)であった。

結果と考察

1. 量的分析

分析の対象は、アイディアユニット(以下ではIUと呼ぶ)とした。IUの認定に際してはTABLE 1に示すような基準を用いた。材料として用いられた物語文章のIU数は、物語Ⅰが163、物語Ⅱが173であった。次に、各要約文章のIU数を調べ、要約率(もとのテキスト中のIU数に対する要約文章中のIU数の割合)を産出した。各要約産出条件の平均要約率は、テキスト参照条件が0.174(SD=0.061)、記憶参照条件が0.170(SD=0.061)であった。要約率の逆正弦変換値を用いた分散分析の結果、2種類の要約産出条件の平均要約率の間に有意な差は認められなかった [$F_{(1,14)}=0.25$, n.s.]。

TABLE 1 アイディアユニット(IU)の認定の際に用いられた基準

- (a)基本的には単文を1つのIUとする。したがって重文は2つのIUからなる。また複文に関しては、従属節が時間的關係を表わす場合は2つのIU、それ以外は1つのIUとする。
- (b)連用修飾の句や節は原則として独立のIUとはしない。
- (c)連体修飾の句や節については、その修飾が非限定的(非弁別的)である場合にのみそれを独立のIUとする。すなわち、その修飾部分を取り除いても残りの部分の意味が変化あるいは抽象化しなければ、取り除いた部分は独立のIUである。
- (d)～ことを知る、～と思う、などのような埋め込み文は、埋め込まれている部分は単独のIUとはせず、それを含む全体を1つのIUとする。また、1か所に複数の文が埋め込まれている場合には、それらを分割して同一の述語によって別々の埋め込み文を作りうることに限り、それぞれ独立のIUとする。
- (e)会話文に関しては、一組のかぎ括弧でくられる内容はたとえ複数の文からなっていようと1つのIUとする。
- (f)文章中で複数回言及される対象に対する修飾句・節については、それらをすべて独立のIUとする。この理由は、文章中で複数回言及されるような対象は当然複数のIUに含まれ、その結果、それらが相互にその対象を限定しあうという性質を持つので、1つのIUの中でその対象を限定修飾する必要がないためである。
- (g)原因、目的、手段を表わす名詞はそれ自体を独立のIUとする。(この基準は、要約文章において要約化の1つの有効な方法として比較的頻繁に用いられる「IUの名詞化」に対処するための措置として設けられた。)

次に、要約率の個人差について検討を行った。各要約産出条件における要約率の最高値と最低値は、テキスト参照条件が0.277と0.081、記憶参照条件が0.276と0.086であり、いずれも3倍以上の開きがあった。また、いずれの被験者も2通りの要約産出条件を体験したが、各被験者が産出した要約の要約率は2つの条件間で極めて有意な正の相関を示した [$r=0.86$, $p<.001$]。したがって、要約率は、テキストを参照できるか否かという外的な条件に影響されることはなく、むしろ要約者個人の内的動機づけに強く依存することが示唆される。すなわち、比較的長い要約を産出する傾向のある人はテキストを参照できるときもできないときも比較的長い要約を産出し、逆に、比較的短くまとめようとする傾向のある人はテキストを参照できるときもできないときも比較的短い要約を産出する、と言える。

2. 質的分析

各被験者が産出した要約文章中では個々のIUはもとのテキストと比較するとさまざまな表現で表わされている。そこで、各要約文章中のすべてのIUに対して、それがもとのテキストと比べてどのように変化しているかについての分類を行った。分類に際しては以下の12種類の分類カテゴリーを用いた。

- ①同一表現IU もとのテキスト中の表現とまったく同一の表現が用いられているIUである。ただし、「ヒロシはいじらしくなった」を「ヒロシはいじらしくなって……」と記述する場合のように、文を接続するための述語の語尾変化は許容する。
- ②類似表現IU もとのテキスト中の表現とは、用いられている述語自体は同じであるが、態の変化や様相的な違いが見られるIUである。「ヒロシはハトを一羽、のらネコにさらわれた」を「のらネコはヒロシのハトを一羽さらった」と記述するような場合がこれにあたる。
- ③言い換えIU もとのテキスト中のIUと意味的にはほぼ同じと考えられるが、用いられている述語が異なるIUである。「金吾は胸をわくわくさせた」を「金吾は胸をおどらせた」と記述するような場合がこれにあたる。
- ④脱特定化IU もとのテキスト中ではかなり特定化されているIUが、その特定情報を欠いて表現されているものである。項(argument)の脱特定化と述部(predicate)の脱特定化の両方がある。前者は連体修飾語句の欠落によって、後者は連用修飾語句の欠落または修飾の程度の弱化によって、それぞれ生じる。「リング

箱をバタツとかぶせ」を「箱をバタツとかぶせ」と記述するような場合、あるいは「金吾は長い間かめめの群れを眺め続けた」を「金吾はかめめの群れを眺め続けた」と記述するような場合がこれにあたる。

⑤抽象化 IU もとのテキスト中のある IU に対し、その述語が抽象化され、もとの IU のもつ具体的意味をもはや伝えていないものである。「おそろしい重さの箱が、金吾の腰の上に落ちた」を「事故が起きた」と記述するような場合がこれにあたる。

⑥具体化 IU もとのテキスト中にある IU よりも具体性が増している IU である。修飾語句の付加や具体的述語の使用によって生ずる。「のらネコは鳴き続けた」を「のらネコはものすごい声で鳴き続けた」と記述するような場合がこれにあたる。

⑦合成 IU もとのテキスト中に存在する複数の IU を、それぞれの具体的意味を保持させたまま 1 つの IU にまとめて表現したものである。「船に乗り込んだのは夜だった。……船を下りたのも夜だった」を「船の乗下船は夜に行われた」と記述するような場合がこれにあたる。

⑧統一 IU もとのテキスト中に存在する複数の IU を、それらすべての IU の具体的意味を保持せず多かれ少なかれ抽象化して 1 つの IU にまとめて表現したものである。「ヒロシは走り寄って戸を開け石を取り、リンゴ箱をガバツと持ち上げた。黒ネコはさっと走り出し……」を「ヒロシは黒ネコを解放してやった」と記述するような場合がこれにあたる。

⑨推論 IU もとのテキスト中に明示されていないが、テキストより容易に推論可能な IU である。たとえば、もとのテキストが「金吾の一生の望みは、海を心ゆくばかり眺めてみたいということだった。金吾の生まれた国には海はなかった。小さい頃から村の桶屋に丁稚奉公に出された。……」と続くとき、そこに明示されていないが推論可能な「金吾は生まれてから一度も海を見たことがなかった」という情報を記述した場合がこれにあたる。

⑩個性的 IU もとのテキスト中に明示されておらず、テキストの内容と矛盾はしていないが容易には推論できないと思われる IU である。要約者の個性的な（客観性を欠く）推論情報である。

⑪物語外 IU 物語の筋そのものではなく、要約者の物語や登場人物に対する評価やコメントを表わす IU である。もちろん、もとのテキストに明示されていない IU である。「これはある少年の物語である」や「なんと無情なことであろうか」などという表現がとられる。

⑫誤り IU 物語の内容と明らかに矛盾する IU である。たとえば、「リンゴ箱をかぶせた」を「ミカン箱をかぶせた」と記述するような場合や「少しも退屈しなかった」を「退屈だった」と記述するような場合である。

なお、登場人物など複数の IU に共通して含まれる項の誤りや脱特定化については、そのうちの最初に現れる明示 IU に対してのみそれを誤り IU あるいは脱特定化 IU とした。

次に、各要約文章毎に各 IU カテゴリに属する IU の要約文章全体に占める割合を算出した。これを各 IU カテゴリの占有率と呼ぶことにする。TABLE 2 に、要約条件毎の各 IU カテゴリの平均占有率を示す。各要約文章の各 IU カテゴリ占有率の逆正弦変換値を用いた分散分析によって、各 IU カテゴリ毎に 2 つの要約条件間での当該カテゴリの平均占有率の違いを検定したところ、「同一表現」「合成」の 2 カテゴリはテキスト参照条件の方が、「抽象化」「具体化」「個性的」「誤り」の 4 カテゴリは記憶参照条件の方が平均占有率が高かった[それぞれ、 $F_{(1,14)}=14.20$, $p<.01$; $F_{(1,14)}=5.88$, $p<.05$; $F_{(1,14)}=5.54$, $p<.05$; $F_{(1,14)}=6.36$, $p<.05$; $F_{(1,14)}=5.06$, $p<.05$; $F_{(1,14)}=6.86$, $p<.05$]が、それ以外のカテゴリは 2 つの要約条件間で平均占有率に有意な差は認められなかった。

情報の抽象度のレベルの変化（抽象化、具体化）や不適切な情報（個性的情報、誤り情報）の混入は、文章の記憶現象の 1 つであり、記憶参照事態の要約においては生じやすい現象であろう。一方、テキスト参照事態の要

TABLE 2 各要約産出条件における各 IU カテゴリの平均占有率

IU カテゴリ	平均占有率 (SD)		
	テキスト参照条件	記憶参照条件	条件差 (F 値)
同一表現	.239 (.171)	.066 (.064)	14.20**
類似表現	.030 (.047)	.032 (.042)	0.01
言い換え	.153 (.054)	.174 (.067)	1.02
脱特定化	.134 (.071)	.117 (.050)	0.85
抽象化	.076 (.056)	.121 (.051)	5.54*
具体化	.014 (.022)	.042 (.048)	6.36*
合成	.039 (.047)	.007 (.013)	5.88*
統一	.120 (.065)	.124 (.076)	0.08
推論	.160 (.095)	.220 (.114)	3.49
個性的	.019 (.029)	.057 (.077)	5.06*
物語外	.008 (.019)	.003 (.011)	1.51
誤り	.009 (.016)	.038 (.039)	6.86*

注) * $p<.05$ ** $p<.01$

約においてもとのテキストとまったく同じ表現のIUが多く現れる傾向が強かったのは、テキストを参照できる場合には要約者はそれに束縛されてしまう傾向があることを示している。また、情報を集約しようとする際にすべての情報を漏らさないように綿密に吟味した上で置換すべき別の表現を決定する「合成」傾向も、テキストを十分に参照できるときに出現しやすい現象であると理解できる。

次に、要約内容の個人差を調べるため、各被験者毎に要約文章の内容について調べてみた。FIG. 1に、各被験者が2種類の要約条件下で産出した要約文章における各IUカテゴリーの占有率をそれぞれ示す。この図において最も顕著なのは、テキスト参照事態の要約における同一表現IUの割合が被験者毎に著しく異なっている、ということである。たとえば、被験者S12の要約においては同一表現IUが7割以上を占めているの

□ 同一表現 ■ 類似表現 ▨ 言い換え ▩ 脱特定化 ▧ 抽象化 ▦ 具体化 ▤ 合成 ▥ 統一 ▧ 推論 ▨ 個性的 ▩ 物語外 ■ 誤り

FIG. 1 各被験者によって産出された要約文章中の各IUカテゴリー占有率

に対して、被験者 S2 の要約にはほとんど含まれていない。このことは、テキスト参照事態における要約文章の質的多様性を如実に表わしており、要約産出方略の多様性を示唆しているように思われる。また、被験者毎に 2 つの事態での要約を比べてみると、各 IU カテゴリーの占有率が著しく異なっている被験者もいるが（たとえば、前述の S12 など）、その一方で、それらの占有率がほとんど変わらない被験者もあり（たとえば、S9）、前者は 2 つの要約産出事態で異なる産出方略を用いた可能性があり、後者はいずれにおいても同じ要約産出方略を用いたことが推測される。

2 つの要約産出事態における最大の違いは、もとのテキストを参照できるかできないかである。テキスト参照事態ではもとのテキストを参照できるため、もとのテキスト中の表現を要約文章においても利用することができる。そして、その利用傾向の強さは、記憶参照事態の要約に比べてテキスト参照事態の要約で同一表現 IU の占有率が有意に増加しているか否かで、ある程度判断することができる。そこで、各被験者毎に同一表現 IU の占有率のテキスト参照事態での増加傾向を Fisher の直接法によって検定したところ、S5, S7, S12, S13, S14, S16 で有意な増加がみられ [S5, S7, S12 は $p < .01$, その他は $p < .05$]、S8 および S11 で増加の傾向が認められた [$.05 < p < .10$]。このうち、S5, S7, S12, S13 はテキスト参照事態の要約の同一表現 IU の占有率も相対的に高く、かなりもとのテキストの表現に依存した要約産出方略を用いていたと考えてよいのではないだろうか。

また、たとえまったく同一の表現を用いなかったとしても、もとのテキストに明示された情報を取捨選択し、場合によっては自分のことばに置き換えたり修飾部分を落としたりすることで要約文章を作り上げるといった方略もテキストを参照できる事態に特有な方略の 1 つであろう。この明示情報依存度は同一表現 IU、類似表現 IU、言い換え IU、脱特定化 IU の合計の割合と考えることができる。この割合がテキスト参照事態の要約で増加する傾向があるかどうかを Fisher の直接法によって検定したところ、S7, S8, S10, S12, S16 で有意な増加がみられ [S7 と S12 は $p < .01$, その他は $p < .05$]、S3, S5, S15 で増加の傾向が認められた [$.05 < p < .10$]。

3. 記憶参照事態とテキスト参照事態における要約産出過程

記憶参照事態における要約の場合、まず、文章の内容を想起することから始まる。想起されうる内容の多

くは、文章中で重要な役割を果たしている情報であろう。その中には、テキスト中には明示されていないが重要な役割を果たす情報もあるかもしれない。そのような情報は、文章読解時に読み手によって推論され、その結果として記憶に残っているのである。次に、想起された情報に基づいて文章を再構成する過程がある。想起された情報をつなぎ合わせ、全体として一貫した文章を作り上げるわけである。このとき、想起された情報が少なければそれらを関係づけるために多くの推論を必要とする。それらの中には、もちろん、もとのテキスト中に明示されていた（が、想起できなかった）情報もかなりあるであろう。しかし、一方で、もとのテキストには明示されていない新たな（具体的な）情報が生成される可能性もあるわけである。

以上のような要約産出過程を考えるならば、この種の要約における個人差は、主として、読解時に理解された内容の個人差、要約時に想起されうる内容の個人差、および文章の再構成の過程において推論される内容の個人差によって生ずると考えることができる。

一方、テキスト参照事態における要約産出過程は記憶参照事態のそれに比べて極めて多様性に富んでいる。どの程度もとのテキストを参照するかは、要約者個人に任されているからである。もとのテキストに強く依存し同一の表現を多く用いる者、もとのテキストに明示されている情報に基づきそれをするだけ自分のことばに置き換えて要約を産出する者、もとのテキストにはほとんど依存せずに記憶参照事態と同じようなやり方で要約を産出する者、などさまざまな方略が考えられる。

以上のように、同じ文章に対する要約であっても、その長さや内容は被験者によってさまざまに異なり、特にテキストを参照できる事態においては、個人個人の要約が質的に非常に異なることが分かった。では、これらのさまざまな要約文章は、どのような評価を受けるのであろうか。実験 II では、このことについて検討する。

実験 II

方法

被験者 北海道大学文学部の学部学生 18 名（男性 10 名、女性 8 名）で、平均年齢は 21.6 歳であった。いずれの被験者も実験 I には参加していない。

材料 実験 I に参加した 16 名の被験者によって産出された 32 種類の要約文章が用いられた。各要約文章は、明らかな誤字や脱字を訂正され、語句の表記形態

(漢字・仮名)や送り仮名などが統一された。また、句読点の用法に関しても明らかに誤解を招くと思われるものは訂正された。各要約文章はそれぞれ縦長A5版の用紙に1行40文字の横書きで印刷された。また、各用紙の右上には、その要約文章の採点結果および順位を記入する欄が設けられた。

手続 実験は個別に行われる。被験者は原文章(物語文章)とそれに対する16種類の要約文章を受け取り、それぞれの要約文章に対して“要約としての良さ”を100点満点で採点すること、および16種類の要約文章の順位づけを行うことを求められる。要約文章が印刷された16枚の用紙はランダムな順に重ねられている。制限時間はなく、採点・順位づけの方法や基準は被験者の自由であるが、評価にあたってはすべての要約文章に対して厳密かつ公正を期するよう強調される。実験は2種類の物語文章(物語Ⅰ, 物語Ⅱ)とそれぞれに対する要約文章(各16種類)について順に行われる。実験の順序(いずれの物語に対する要約文章群の評価を先に行うか)に関しては、被験者間で相殺された。被験者がすべての要約文章の評価に要した時間は、おおよそ40分～60分であった**。

結果と考察

いずれの物語に関しても、各要約文章の順位について被験者間に有意な一致度が認められた[物語Ⅰに対しては、ケンドールの一一致係数 $W = .228$, $\chi^2 = 61.58$, $df = 15$, $p < .01$, 物語Ⅱに対しては, $W = .298$, $\chi^2 = 80.46$, $df = 15$, $p < .01$]。このことは、被験者間に共通の、要約文章の“良さ”を規定する何らかの要因が存在することを示唆している。

まず、各要約文章の要約率および各IUカテゴリー占有率と平均評価点の相関を調べたところ、「抽象化」「個性的」「誤り」の3つのカテゴリーの占有率と平均評価点の間に有意な負の相関が認められた[それぞれ, $r = -0.49$, $p < .01$; $r = -0.44$, $p < .05$; $r = -0.43$, $p < .05$]が、要約率やその他のカテゴリーの占有率との間には有意な相関は認められなかった。また、明示情報依存傾向(「同一表現」「類似表現」「言い換え」「脱特定化」の4つのカテゴリーの合計占有率)と平均評価得点との間にも有意な相関は認められなかった。

IUの「抽象化」は、その程度にもよるが、「合成」や「統一」のような情報の凝集性がない上に情報を曖昧にしまうため、その多用は要約としての評価を下げてしまうようである。また、客観性のない自分勝手

な解釈(個性化)やもとのテキストと矛盾する内容(誤り)は、当然のことながら要約の評価を下げる原因となることが確認されたと言えよう。

次に、分散分析の結果、テキスト参照条件で産出された要約文章の方が記憶参照条件で産出された要約文章よりも有意に良い評価点を得ていることが分かった [$F_{(1,541)} = 36.25$, $p < .001$]。これは、後者においては記憶の歪みによって生じると考えられる抽象化IU, 個性的IU, 誤りIUの占有率が高い(TABLE 2参照)ことや忘却によって必要とされるIUが欠落している***ことが主な原因であると思われる。

では、記憶現象に由来しない要因で、要約文章の良さ評価に影響を及ぼす要因は存在しないのであろうか。そこで記憶要因の影響が比較的少ないと思われるテキスト参照事態の要約のみを対象として分析を行うことにした。順位づけデータに基づいて、テキスト参照事態の要約のみの順位を求め、それによって被験者間の順位の一致度を求めたところ、物語Ⅰに対してのみ一致度が有意****で [$W = .128$, $\chi^2 = 16.10$, $df = 7$, $p < .05$], 物語Ⅱに対しては一致度は有意ではなかった [$W = .098$, $\chi^2 = 12.35$, $df = 7$, n.s.]。そこで、2種類の物語に対するテキスト参照条件の要約文章の評価(要約条件別につけ直された順位)に基づいて、被験者(評定者)間の評価の類似性についてのクラスター分析を行い、18名の被験者を2群(評定者群A 11名, 評定者群B 7名)に分けた。そして、評定者群別に各要約文章の平均評価点および標準偏差を算出した。その結果をTABLE 3に示す。なお、TABLE 3には各要約文章の有するいくつかの特徴も記されている。各要約文章毎に2つの評定者群間の平均評価点の差の検定を行ったところ、要約文章T3, T9, T11の3つで有意差が認められ、要

*** 必要とされるIUが何であるかを決定することは極めてむずかしい問題であるが、いずれの物語の記憶参照条件の要約文章においても物語の後半部分の比較的重要であると思われる部分が欠落している傾向が認められた。このことは、要約として必要とされる情報が必ずしも記憶にも残りやすいものであるとは限らないということを示唆しているように思われる。

**** 物語Ⅰに対する要約文章の中には、テキスト参照条件の要約でありながら誤りIUを含んでいるものが3種類もあり、それらが他の要約文章よりも悪い評価を得たために、物語Ⅰの要約文章に関する順位の一致度が有意となった可能性がある。この3種類を除く5種類の要約文章の順位の被験者間の一致度を求めたところ、有意な一致は認められなかった ($W = .111$, $\chi^2 = 8.00$, $df = 4$, n.s.)。実際のところ、その3種類のうちの1種類はそれほど悪い評価を得ておらず、それ以外の2種類の影響が特に大きいようである(その2種類のみを除くと一致度は $W = .087$, $\chi^2 = 8.00$, $df = 5$, n.s.)。

** 例外的に約2時間もの評価時間を要した被験者が1名いた。

TABLE 3 テキスト参照事態の要約文章の特徴とその評定者群別の評価

要約文章 ¹	要約率	同一表現 IU		明示性命題の IU ³		個性的 IU + 誤り IU の占有率	要約産出方略	評定者群 A の 平均評価点 (SD)	評定者群 B の 平均評価点 (SD)	評定者群間の 評価点の差 (t 値 ⁴)
		占有率	増加傾向 ² の有意性	占有率	増加傾向 の有意性					
T1	.117	.211	n.s.	.421	n.s.	.053	凝縮型	64.2 (17.3)	68.3 (13.5)	-0.53
T2	.196	.031	n.s.	.344	n.s.	.031	具現型	69.7 (19.0)	72.9 (17.7)	-0.35
T3	.160	.115	n.s.	.577	$p<.10$	.115	換言型	43.3 (18.0)	64.4 (16.2)	-2.53*
T4	.166	.185	n.s.	.593	n.s.	.000	換言型?	62.6 (12.9)	59.9 (13.6)	0.44
T5	.184	.333	$p<.01$	.667	$p<.10$	.000	複写型	71.2 (22.1)	72.9 (13.2)	-0.18
T6	.129	.048	n.s.	.429	n.s.	.000	凝縮型	59.9 (19.5)	70.1 (13.2)	-1.22
T7	.252	.390	$p<.01$	.780	$p<.01$	.000	複写型	72.8 (16.8)	63.3 (11.5)	1.31
T8	.270	.318	$p<.10$	.750	$p<.05$	.045	複写型	52.8 (21.8)	58.6 (16.6)	-0.60
T9	.081	.143	n.s.	.357	n.s.	.000	凝縮型	44.6 (15.4)	84.4 (10.5)	-5.97***
T10	.145	.120	n.s.	.440	$p<.05$	.000	換言型	55.3 (15.4)	69.3 (21.1)	-1.63
T11	.277	.167	$p<.10$	.542	n.s.	.063	具現型	87.6 (10.0)	71.6 (12.7)	2.99**
T12	.231	.725	$p<.01$	.825	$p<.01$	.000	複写型	72.3 (21.0)	76.3 (11.0)	-0.46
T13	.098	.412	$p<.05$	.588	n.s.	.000	複写型?	68.8 (17.8)	76.0 (12.4)	-0.93
T14	.150	.154	$p<.05$	.462	n.s.	.000	具現型	80.2 (12.8)	77.3 (14.1)	0.45
T15	.116	.250	n.s.	.650	$p<.10$	.050	換言型	60.0 (28.3)	74.1 (13.9)	-1.22
T16	.208	.222	$p<.05$	.472	$p<.05$	.083	具現型	68.0 (17.9)	60.1 (19.1)	0.88

注1) T1～T8は物語Ⅰに対する要約文章，T9～T16は物語Ⅱに対する要約文章である。

注2) 記憶参照事態で同一被験者によって産出された要約文章中の占有率と比較したときの占有率の増加傾向を表わす（有意性の検定は Fisher の直接法による）。

注3) 明示性命題の IU とは、「同一表現」「類似表現」「言い換え」「脱特定化」のいずれかのカテゴリーに属する IU を指す。

注4) * $p<.05$ ** $p<.01$ *** $p<.001$

約文章 T11は評定者群 A の方で $[t=2.99, p<.01]$ 、要約文章 T3および T9については評定者群 B の方で有意に高い評価点を得ていることが分かった〔それぞれ、 $t=-2.53, p<.05$ ； $t=-5.97, p<.001$ 〕。しかし、その他の要約文章においては有意な差は認められなかった。

要約文章 T3, T9, および T11は、2つの評定者群の評価傾向の違いのある一面をいわば典型的に反映した材料と言える。要約文章 T3は、テキスト参照事態の要約であるにもかかわらず個性的 IU および誤り IU が含まれており、2つのカテゴリーの合計占有率は16種類のテキスト参照事態の要約の中で最も高い。このことが評定者群 A での評価を低下させた原因ではなからうか。また、要約文章 T9は、要約率が.081と最も低く、内容的には統一 IU の占有率が16種類のテキスト参照事態の要約の中で最も高い。したがって、非常に簡潔にまとめられているがもとのテキストに比べるとかなり具体性を欠く内容となっており、このような特徴が評定者群 A の評価を低下させ評定者群 B の評価を最高にした原因であると考えられる。それと対照的

に要約文章 T11は、要約率が.277で最も高く、相対的に具体性の高い要約文章である。この要約文章に対しては評定者群 A で評価が高く評定者群 B で評価がやや低い。これらのことから、要約文章の簡潔性と内容的具体性は相反する関係にあり、2つの評定者群で最も評価の分かれる点になっていると言える。

全体的考察

本研究では、人間の要約産出過程の解明を目的として、2つの異なる要約産出事態（テキスト参照事態と記憶参照事態）を設定し、そこで産出される要約文章についてそれぞれ具体的に分析を行った。記憶参照事態においては、要約者は読んだ文章の内容を想起し、それに基づいて文章を再構成して、要約文章を産出しなければならない。この認知過程自体にはほとんど個人差はないと思われるが、実際に産出される要約文章は、内容を正しく想起できなかったり再構成の際に誤った推論をしたりすることによって、ある重要な部分がそっくり抜け落ちたり間違った情報が入り混んだりして、多様なものとなることが観察された。また、このよう

な重要な情報の欠落や誤り情報の混入は、その要約文章の評価に致命的な影響を及ぼすことが分かった。一方、テキスト参照事態においては、要約産出過程そのものかなりの個人差があることがうかがわれた。すなわち、もとのテキストにどの程度依存するかで要約産出過程そのものが大きく異なってくると考えられるのである。そして、その結果産出される要約文章も非常に多様性に富むことが観察された。また、それらの良さ評価は、すべての評定者に共通な一定の傾向は認められなかったが、評定者を2群に分けることで各評定者群に内在する共通の評価傾向を多少明らかにすることができた。

要約文章の多様性は要約産出時の認知的方略の違いを反映しているのではないか、というのが本研究の仮定である。このことについて考察しておきたい。テキスト参照事態では、もとのテキストにどの程度依存して要約文章を産出するかは要約者の自由である。実験結果を参考にすると、テキスト参照事態において要約産出者がとりうる認知的方略は大まかには以下のように分類できるのではないかと考える。

(1)表象依存型要約方略 テキスト依存度の低い要約産出方略であり、記憶参照事態の方略に比較的近い方略である。すなわち、まず記憶から情報の検索を行い、基本的にはそこで検索された情報に基づいて要約文章を組み立てるやり方である。もとのテキストは、主として誤りや自分勝手な解釈がないかをチェックするために参照される。記憶参照事態の方略に近い場合、要約文章中の各IUカテゴリーの占有率は記憶参照事態と大きく変わることはなく、推論IUや抽象化IU、合成IU、統一IUなどテキスト中には明示されていないIUの占有率が高くなる。文章内容をどの程度具体的に再現するかによって、以下の2つの型に下位分類できよう。

①凝縮型 比較的集約化された表現を用いて文章内容をコンパクトにまとめる方法であり、要約文章の長さは短く(IU数は少なく)なり、その中でも統一IUや推論IUの割合が多い。T9がこの型の要約文章の典型例である。また、テキスト中に明示されたIUの占有率がやや高めだがT1やT6もこの型に含まれると思われる。

②具現型 文章内容をある程度具体的な表現を用いて記述する方法である。凝縮型に比べて要約文章の長さは長く(要約率が高く)なる。また、テキストの参照頻度も凝縮型よりも多いと考えられるので、記憶参照事態の要約に比べて、推論IUの出現傾向が若干減少し

たり、テキストに明示されたIUの出現傾向が若干増加することもある。しかし、テキストに明示されたIUの占有率が後に述べるテキスト依存型の方略による要約文章ほど高くなることはない。T2, T11, T14, T16がこの型の要約文章の例であろう。

(2)テキスト依存型要約方略 テキスト依存度の高い要約産出方略であり、与えられたテキストに依存して、その中に述べられている情報を取捨選択的に抜き出していき、というやり方である。同一表現IUや類似表現IU、言い換えIU、脱特定化IUの占有率が高くなり、推論IU、抽象化IU、合成IU、統一IUなどの占有率は低くなる。さらに、テキスト中の表現を要約文章中にそのまま用いるか否かで、以下の2つの型に下位分類できよう。

③複写型 テキスト中の各IUに対してそれを要約文章中に含めるか否かを決定し、含める場合にはテキスト中の表現をそのまま要約文章中に複写する方法で、テキストに明示された情報を表わすIUとりわけ同一表現IUの占有率が著しく高くなる。また、どちらかといえば要約文章の長さも長くなる。なお、この方略はすでにBrownら(Brown & Day, 1983; Brown, Day, & Jones, 1983)によって報告されているcopy-delete方略に相当する。この型の要約文章の典型例はT12であり、T7やT8もこの型に入るとされる。また、推論IUが若干多いがT5もこの型に含まれよう。T13は、各IUの占有率はこの型に近いが、要約率がやや低く、判断がむずかしい。

④換言型 テキスト中のIUをそのまま複写するのではなく、自分なりに解釈して自分なりの表現で記述する、という方法である。したがって、同一表現IUよりも類似表現IUあるいは言い換えIU、脱特定化IUなどの割合が多くなる。T3やT15がこの型に相当すると思われる。これらの他に、T10は、同一表現IU、類似表現IU、言い換えIU、脱特定化IUの合計占有率がやや低いが、同一被験者の記憶参照事態での要約文章に比べてそれらの合計占有率は有意に増加しており、この型に含まれるのではないかとと思われる。また、T4は、テキスト参照事態の要約だけをみるとこの型に近いように思われるが、記憶参照事態の要約との比較ではあまり各IUカテゴリー占有率の変化は見られず、判断がむずかしい。

このように、テキスト参照事態で産出された16種類の要約文章をその要約産出方略から以上の4つのタイプに分類することで、要約の際に用いられた方略とそれによって産出された要約文章の良さ評価との関係を

対応させてみる事が可能となる。前述の TABLE 3 の中には、各要約文章が産出された際に用いられたと推測される要約産出方略を示してある。そこで、各要約産出方略を TABLE 3 に示した要約文章の特徴と関連させながら、要約の良さの評価との関係をもてみると、以下のようなことが明らかとなる。

凝縮型の要約文章は、評定者群 A よりも評定者群 B において高い評価を得るようである。特に、要約率が低く典型的な凝縮型といえる T9 では、両群の評価差が極めて有意である。具現型の要約文章はいずれの評定者群でもそれほど評価は低くないが、同一表現 IU が非常に少ない、いわば表象依存度の高い具現型 (T2) よりも、多少同一表現 IU が含まれ若干のテキスト依存傾向のある具現型の場合に、評定者群 A での評価が上がるようである (T11, 有意ではないが T14 や T16)。複写型の要約文章は、個性的 IU や誤り IU が含まれない限り、いずれの評定者群においても悪い評価を得ることはない。むしろ、評価は高い方である (T5, T7, T12)。換言型の要約文章に関しては、全体的に低めの評価であるが、評定者群 A に比べると評定者群 B の方がやや評価が高い傾向があるようである。

本研究は、文章の要約産出過程の解明に向けての 1 つの探索的研究であり、今後究明してゆくべき課題は山積している。まず、本研究では 4 種類の要約産出方略を提案したが、その心理的妥当性の検証が最も重要な課題となろう。また、発達の観点からすれば、要約産出方略の発達・熟達の問題も興味深い。Brown et al. (1983) は、copy-delete 方略 (本研究における複写型方略に相当) は年少の子供たちが好んで用いると述べているが、本研究の被験者である大学生においてもその使用が認められ、この方略使用の発達の傾向についてはさらに詳しく調べていく必要があろう。また、その他の 3 種類の方略についての発達の知見はまだ得られておらず、それも今後の課題と言える。熟達という意味では、要約の熟達者のデータを得る必要も出てくる。Kay & Black (1986) は、要約の熟達者とそうでない人々 (素人) の産出した要約文章を質的に分析・比較し、素人はできるだけ多くの情報を要約文章内に押し込もうとする傾向があるが、熟達者は中心的な情報をうまく取り出し、それに関する推論も含めて簡潔にまとめる能力があると述べている。しかし、熟達者がもとのテキストにどの程度依存して要約を産出しているのか、その依存傾向についてはあまり明確にはされていない。さらにこの問題は、“要約の良さ”と密接に関わってくる。通常、要約の熟達者の産出する要約は“良い要約”

であるとされているが、果たして本当に彼らの要約文章が良い評価を得るかどうかは実験的に確かめてみる必要がある。熟達者の要約産出方略が明らかになり、彼らの産出する要約が本当に良い評価を得るということが確かめられて初めて、良い要約を書くための方略が解明されたと言えるであろう。そしてそれは言うまでもなく教育的に非常に有効で貴重なものとなるにちがいない。

本研究では、テキスト参照条件の要約文章の評価がすべての評定者で必ずしも一致を見ず、評定者を 2 群に分けて分析を行ったが、この評価の個人差ということに関しても今後の課題と言えよう。これに関しては、要約における個人差と同様にいくつかの類型が考えられるかもしれない。そして、たとえば、凝縮型の要約をする人は凝縮型の要約を“良く”評価し、複写型の要約を産出する人は複写型の要約を“良く”評価する、といったことがありうるかもしれない。いずれにせよ、要約文章に限らず、普通の文章やあるいは絵画や音楽などの芸術作品から無意味なパターンに至るまで、人間の“対象評価”の認知過程については、極めて難しいが探究に値する認知心理学的問題であろう。

文章の「要約産出」は、人間の言語活動においてはどちらかといえば特殊な認知過程と言えるかもしれない。しかし、そこには「理解」と「産出」の 2 つの重要な過程が含まれており、その解明へ向けての探究は一般的な「理解」あるいは「産出」の過程の解明につながることは言うまでもないであろう。

引用文献

- Alterman, R. 1986 Summarization in the small. In N. Sharkey (Ed.), *Advances in cognitive science*. Chichester, England: Ellis Harwood.
- Alterman, R., & Bookman, L.A. 1990 Some computational experiments in summarization. *Discourse Processes*, 13, 143—174.
- Brown, A.L., & Day, J.D. 1983 Macrorules for summarizing texts: The development of expertise. *Journal of Verbal Learning and Verbal Behavior*, 22, 1—14.
- Brown, A.L., Day, J.D., & Jones, R.S. 1983 The development of plans for summarizing texts. *Child Development*, 54, 968—979.
- Kay, D.S., & Black, J.B. 1986 Explanation-driven processing in summarization: The interaction of content and process. In J.A.

- Galambos, R.P. Abelson, & J.B. Black (Eds.), *Knowledge structures*. Hillsdale, N.J. : Lawrence Erlbaum Associates.
- Kintsch, W., & van Dijk, T.A. 1978 Toward a model of text comprehension and production. *Psychological Review*, 85, 363—394.
- Lehnert, W.G. 1981 Plot units and narrative summarization. *Cognitive Science*, 5, 293—331.
- 邑本俊亮・阿部純一 1987 物語要約における情報集約の過程 阿部純一(編), 昭和61年度科学研究費補助金(特定研究2)研究成果報告書「人間の言語理解過程についての認知論的研究」, 北海道大学文学部.
- Omanon, R.C. 1982 The relation between centrality and story category variation. *Journal of Verbal Learning and Verbal Behavior*, 21, 326—337.
- Rumelhart, D.E. 1977 Understanding and summarizing brief stories. In D. Laberge, & S. Samuels (Eds.), *Basic processing in reading, perception, and comprehension*. Hillsdale, N.J. : Lawrence Erlbaum Associates.
- Thorndyke, P.W. 1977 Cognitive structures in comprehension and memory of narrative discourse. *Cognitive Psychology*, 9, 77—110.
- van den Broek, P., & Trabasso, T. 1986 Causal networks versus goal hierarchies in summarizing text. *Discourse Processes*, 9, 1—15.
- van Dijk, T.A., & Kintsch, W. 1983 *Strategies of discourse comprehension*. New York : Academic Press.

付 記

本論文の作成にあたり御指導いただいた北海道大学文学部 阿部純一助教授に深く感謝申し上げます。

本研究の一部は、日本教育心理学会第33回総会において発表された。なお、本研究は、文部省科学研究費補助金〔平成2年度奨励研究(特別研究員), 課題番号01790025, 研究代表者: 邑本俊亮〕の補助を受けた。

(1991年12月24日受稿)