

視線追跡装置を利用した英語・日本語母語話者の読解過程の研究 -英語テキストと日本語テキストの読み方の特徴-

寺 朱美

北陸先端科学技術大学院大学 〒923-1292 石川県能美市旭台 1-1

E-mail: tera@jaist.ac.jp

概要 視線追跡装置を利用して日本語テキストの読解過程を観察記録し分析した。比較のために英語テキストの読解過程も調査した結果。その結果、日本語母語話者は英語と日本語の両テキストで読解時の視線移動時間が長い傾向が見られた。さらに視線移動の方向から日本語母語話者は戻り読みを多用する傾向があった。本稿では観察実験の分析結果を述べる。

The Difference of the Reading Processes Between English Text and Japanese Text —Analysis using Eye Tracking Recorder—

Akemi Tera

Japan Advanced Institute of Science and Technology: 1-1 Asahidai, Nomi-shi, Ishikawa 923-1292 Japan

E-mail: tera@jaist.ac.jp

Abstract: Using Eye-Tracking-Recorder, I had an experiment of reading processes in English and Japanese texts for 6 groups and analyzed the eye tracking data. During reading Japanese and English texts, Japanese native speaker has the tendency to have long distance of saccade and to use backward-jumping often. This report describes the results.

1. はじめに

語学学習において、基礎文法を修得すると文字情報から知識を獲得するレベルに達する。日本語教育においても中級レベルの学習者は徐々に長文読解の機会が増加するが、読解を苦手とする日本語学習者が多い。理由の1つとして漢字、カタカナ、ひらかな、ローマ字という4種の表記があげられる。近年、テキスト中の単語を辞書情報とリンクさせて提示する学習支援システムの開発が進み、多言語対応の学習支援が可能になるなど語彙レベルの学習支援が長足に進歩した。しかし、読解を苦手とする学習者を支援するために、読解過程そのものを詳しく調査する必要がある。

日本語教育分野においてもこれまで学習者の読解過程を調査する目的の研究が行われてきた。認知科学・認知心理学分野では、プロトコル分析、think-aloud法などにより読解過程を観察し、記録し、分析が行われ、多くの知見が明らかになった[1]。一方で、実験参加者自らが語るデータと

実験者による観察は人的環境に左右されやすく、客観性が低いという側面も持つ。

視線追跡装置を利用した研究による読解過程の研究も多く行われている。重松らは、中国人学習者が漢字に偏る視線停留を多用して読解を行うことを実証的に報告した[3]。鈴木は、初級レベルを終了した学習者の読解過程を観察し、未習事項を含む既習事項に注視が多いこと、既習項目と未習項目で注視時間に差があること、読解力の高い学習者に戻り読みがみられることなどを報告した[6]。Raynerは、当時注目されていた問題として、サッケード中に認知処理活動が中断されているかどうかをあげた[4]。また、Jinchoらは、子どもと成人の読解中の眼球運動を測定し、文字種による読解過程の違いを分析した[13]。

視線追跡装置を利用すると、テキストの読み方や読解課程を数値データとして取得でき、定量的な分析が可能となる。本研究では読解過程を調べるために視線追跡装置を利用した観察実験を行

い、視線データを取得し分析した。本研究では実験対象とする「読解」を『意味を持つ構成の文章で、眼球が文字を補足し、脳がテキストの意味内容を理解するまでのプロセス』と定義する。

古賀らは、読解中の眼球運動は視線停留と視線移動（サッケード）を交互に繰り返し、視線移動はスムーズではなく、ある視線停留位置から次の視線停留位置まで高速に移動し、サッケード速度は 500～600° /秒に達すると報告している [2]。

本研究では、母語や学習対象言語の読解における特徴を調査することを目的とし、視線追跡装置を利用して読解中の眼球運動を観察・記録し、分析を行った。また、実験参加者を 6 グループに分け、得られた視線データを、停留、移動（サッケード）、移動の方向について分析した。本稿ではその結果を報告する。

2. 研究方法

2.1. 実験環境

視線の測定に NAC 社製 eye-mark recorder EMR-8 を使用した。視野角レンズを 44°、サンプリング周波数を 120Hz に設定した。データ解析には NAC 社製 EMR-8 アイマークデータ解析システムと Windows NT、テキスト提示に 17 インチ液晶モニターを使用した。実験ではモニターと眼球の距離を一定に保つため、スタンド型顎台で額の位置を固定し、眼球からモニターまでの距離を 60cm に固定した（図 1）。また視線追跡装置で有効な数値を得るために、事前に参加者の視力を測定し、裸眼もしくはソフトコンタクトレンズ着用で 1.0 以上の者に限定した。実験は実験室で実施した。古賀らは、文字を認識する視線停留の長さは 150～500msec 程度で、ほとんどは 200～300msec であると述べている [2]。しかし、予備実験から、停留時間はさらに短いという印象を受け、本研究では 60msec 以上を視線停留と設定した。



図 1 実験機器と実験環境

2.2. 実験対象者

表 1 は、実験対象者の内訳である。英語母語話

者 2 名を除く全員が大学生、および大学院生で、年齢は 20～27 才である。英語母語話者 2 名は英語教師（32 才、39 才）で、英語テキストのみ読解実験を実施した。

母語または生活環境で漢字を利用する（または眼にする）という観点による寺ら [7] の分類に従って、対象者を、漢字圏、中間圏、非漢字圏グループに分け、さらに非漢字圏を、アジア、欧州、英語母語話者の 3 つの地域に分けた。すなわち、(1) 母語に漢字を用いるグループ (K)、(2) 母語に漢字を用いないが日常的に漢字に触れる機会があるグループ (M)、(3) 母語に漢字を用いないグループでアジア地域 (NA)、(4) 母語に漢字を用いないグループでヨーロッパなどアジア以外の地域 (NE)、(5) 英語が母語の地域 (NN)、である。上の外国人参加者の他、(6) 日本語母語話者グループ (J) を含め、全 6 グループである。

表 1 実験対象者：グループと人数

グループ ID	母語と漢字	人数	内訳
(K)	母語に漢字を用いる	4名	中国人(4)
(M)	母語に漢字を用いないが日常的に漢字に触れる機会がある	4名	韓国人(4)
(NA)	母語に漢字を用いないグループでアジア地域	4名	インドネシア、ネパール、ベトナム、タイ (各1)
(NE)	母語に漢字を用いないグループでヨーロッパ・その他の地域	4名	ハンガリー、ドイツ、スペイン、ベルギー (各1)
(NN)	英語母語話者	2名	オーストラリア、アメリカ (各1)
(J)	日本語母語話者	4名	日本人(4)

実験参加者の日本語能力測定として、日本語母語話者グループ (J) と英語母語話者グループ (NN) を除く外国人の実験参加者に対して日本留学試験の模擬テストを実施した [8]。模擬テストは、実験参加者として適切かどうかを選抜する目的で「読解」のみを利用して実施し、60 点未満の対象者を読解力が不足していると判断して排除した。

2.3. 実験に用いたテキスト

実験に用いたテキストは日本語と英語各 4 編で、新聞記事 [9] および Web ニュース [10] より選び、内容は自然現象、科学、社会、文化を選定した。

テキストの選定で、日本語テキストは、1) 漢字、カタカナ、ひらがなが含まれ、難解すぎない、2) 内容が普遍的、3) メディアで取り上げられた内容を基準に選び、英語テキストは、1) 難解すぎない、2) 内容が普遍的、3) メディアで取り上げられた内容を基準に選んだ。また、テキストは画面に収まる長さを限度とした。

表2は、テキストの概要である。表中の数字で、英語テキストは（英語文字数：296、単語数：55）、日本語テキストは（文字数：202）を表す。日本語の漢字含有率は23～28%である。

図2は、モニター上のテキスト例である。日本語の場合、1文字、1行に対応する視野角の平均値は、X軸方向1.6°、Y軸方向2.2°である。英語テキストは、1行の単語数から視野角の平均を割り出した。

テキストをランダムに提示し、提示順による誤差を排除した。また、テキストに各3問ずつ、合計24問の内容を問うテストを課題とした。これは実験参加者が真摯に読解に取り組み理解を高める目的で、実験前に告知し、実験後に実施した。最後にアンケートで難易度調査を行った。

表2 テキストの概要：日本語/英語

概要	日本語				英語			
	T-1	T-2	T-3	T-4	T-1	T-2	T-3	T-4
テキスト	T-1	T-2	T-3	T-4	T-1	T-2	T-3	T-4
字数	178	202	221	172	296/53	255/55	274/59	246/53
地域	USA	世界	日本	日本	USA	世界	日本	日本
内容	一般	科学	一般	一般	一般	科学	一般	一般
特徴	地名	専門	固有名	固有名	地名	専門	固有名	固有名

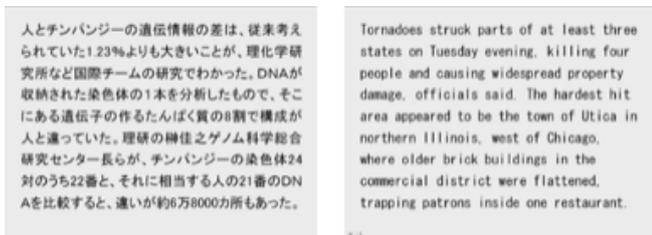


図2 表示テキスト（左：日本語、右英語）

2.4. 実験の手順

実験の手順は以下の通りである。

参加者には、事前に読解後に内容理解を確認するテストとアンケートの実施を知らせた。テキストを計算機のモニターで表示し、参加者はこれを読解した。参加者の時間制限による負荷を取り除き理解を高めるために読解時間の制限をしなかった。実験参加者はテキストを読み終えたところでマウスをクリックして次のテキストに進んだ。実験者は視線追跡装置で読解中の視線の移動・停留データを記録した。読解後、内容理解テストを実施した。実験後アンケートを実施した。英語母語話者グループ(NN)を除く5グループは日本語テキストと英語テキスト両方ともを読解し、英語母語話者グループは英語テキストのみ読解した。

3. 結果

3.1. 視線データ

図3は、日本語母語話者(J-1)と非漢字圏欧州グループ(NE-3)の日本語と英語テキスト読解時の視線データを表す。円の大きさは視線停留時間の長さ、直線はサッケード（視線移動）の軌跡を表わす。J-1は日本語テキストで視線停留時間が短く、英語テキストの視線停留時間が長い。一方、NE-3は日本語テキストで視線停留時間が長く、英語テキスト読解時の視線停留時間は短い。また、J-1は上下に移動する視線データが多く、NE-3は左右に移動する視線データが多いことがわかる。

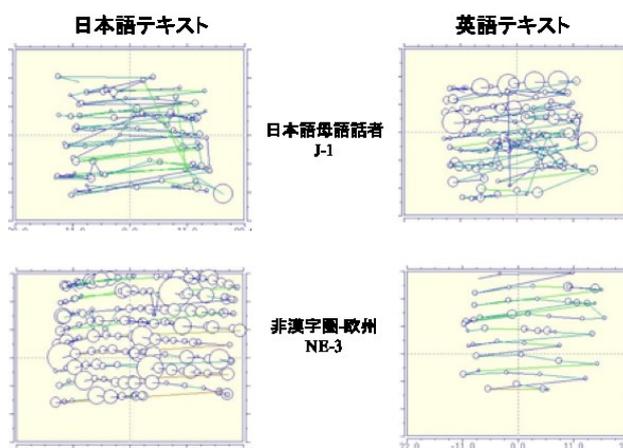


図3 視線データ (J-1, NE-3)

図4は、日本語テキスト読解時の視線停留の様子を比較している。非漢字圏欧州グループ(NE)、アジアグループ(NA)共に視線停留時間が長いことがわかる。一方、漢字圏グループ(K)は、重松らが指摘したように[3]、特に漢字に注目して読み進む傾向がわかる。日本人の視線停留は短い。

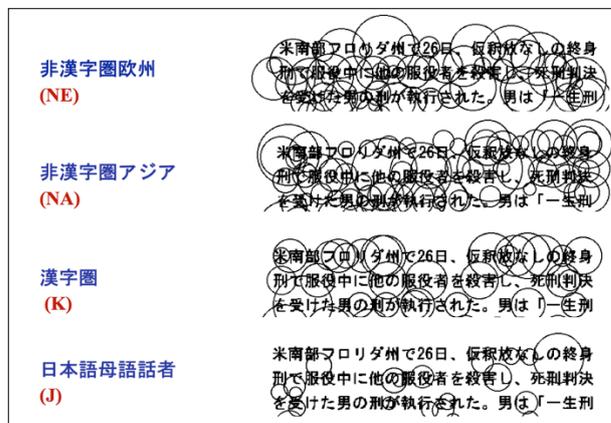


図4 視線停留の比較

3.2. 難易評価

表 3 は、日本語テキストの難易評価である。もっとも難しいと評価したのは T-4 (12 名選択)、難しいと評価したのは T-2 (7 名選択) であった。一方、もっともやさしいと評価したのは T-1 (12 名選択) であった。T-2 と T-3 は、6 名ずつがやさしいと評価した。

表 3 難易度評価

難易評価-「難しい」：グループ別集計					
ID	T-1	T-2	T-3	T-4	合計
J	2	2	0	0	4
K	0	3	0	5	8
M	0	0	1	3	4
NA	0	0	1	4	5
NE	1	2	1	6	4
合計	3	7	3	12	25

難易評価-「やさしい」：グループ別集計					
ID	T-1	T-2	T-3	T-4	合計
J	1	1	1	1	4
K	8	0	0	0	8
M	1	3	0	0	4
NA	1	1	3	0	5
NE	1	1	2	0	4
合計	12	6	6	1	25

3.3. 視線停留時間

図 5、図 6 は、日本語テキストと英語テキストの視線停留時間の平均値と標準偏差値を表す。

難易評価と対応させてみると、読解時の視線停留データは、テキストの難易による影響はほとんど見られず、各グループの特徴が現れている。

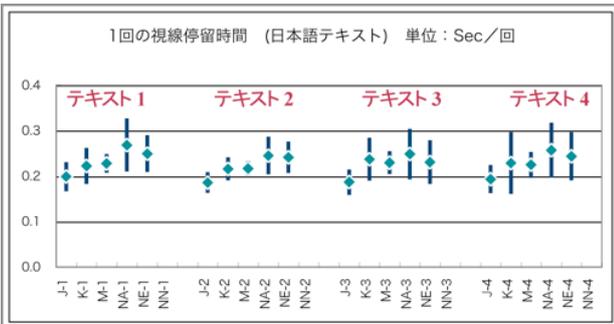


図 5 視線停留時間 (日本語：テキスト別)

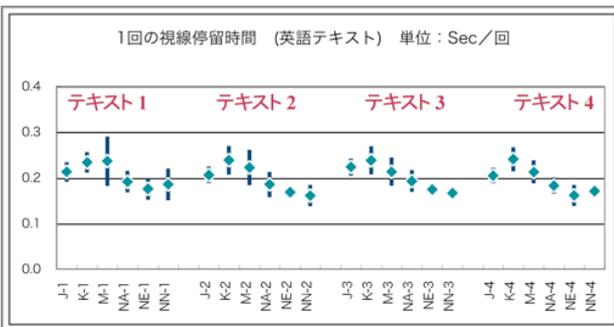


図 6 視線停留時間 (英語：テキスト別)

図 7、図 8 は、日本語と英語のグループ別の視線停留時間の平均値と標準偏差値を表す。

日本語テキストでは、日本語母語話者グループの視線停留が最も短く、次に漢字圏グループ(K)と中間圏グループ(M)が近似の値を示し、最も長い視線停留は非漢字圏の 2 つのグループ (NA) (NE) であった。一方、英語テキストの読解で、3 つの非漢字圏グループ (NA) (NE) (NN) が短い視線停留を行った。中間圏グループ (M) はもっとも長い視線停留を行い、日本語母語話者グループ (J) と漢字圏グループ (K) はその中間であった。

日本語テキストと英語テキスト共に、母語や親密な言語の読解時の視線停留時間は短い傾向がみられた。

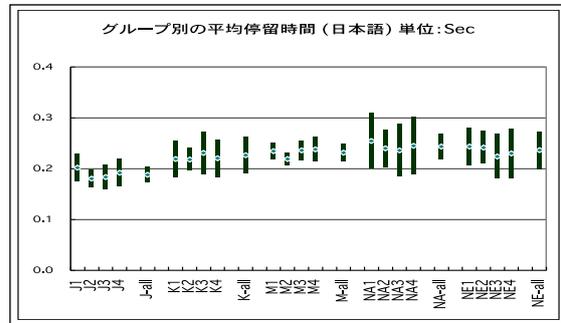


図 7 視線停留時間 (日本語テキスト)

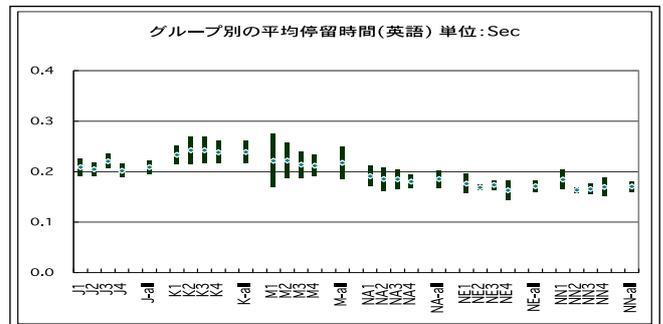


図 8 視線停留時間 (英語テキスト)

3.4. 視線移動 (サッケード) の平均時間

図 9、図 10 は、両テキスト読解時の視線移動 (サッケード) 時間のグループ別の分析結果である。

日本語テキストで、視線の移動時間の平均がもっとも長いのは日本語母語話者 (J) と中間圏グループ (M) である。中間圏グループは偏差が大きい。もっとも短いのは非漢字圏アジアグループ (NA) と漢字圏グループ (K) で、非漢字圏欧州グループ (NE) はそれよりやや長い (図 9)。

英語テキストで、視線移動時間の平均がもっとも長いのは日本語母語話者 (J) と中間圏グループ

(M)で、中間圏グループは偏差が大きい。もっとも短いのは英語母語話者グループ(NN)で、漢字圏(K)、非漢字圏アジア(NA)、非漢字圏欧州(NE)のはそれより長い(図 10)。

視線移動時間は、両方のテキストともに、各グループが特徴的であるといえる[11]。

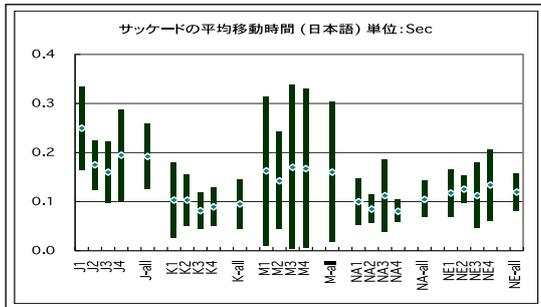


図 9 1回のサッケードの平均時間（日本語）

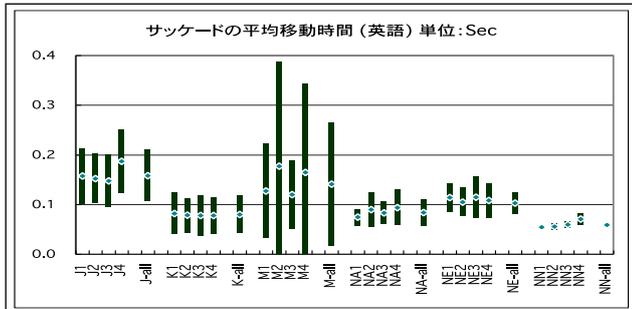


図 10 1回のサッケードの平均時間（英語）

3.5. サッケードの移動距離の平均

図 11 はサッケード（視線移動）の移動距離を表す。日本語テキストで、日本語母語話者グループ(J)と漢字圏グループ(K)が長い距離を移動し、中間圏グループ(M)と非漢字圏グループ(NA)(NE)が短い移動を行った。英語テキストでは、非漢字圏グループが永井距離を移動し、両テキストで対象的な結果であった。

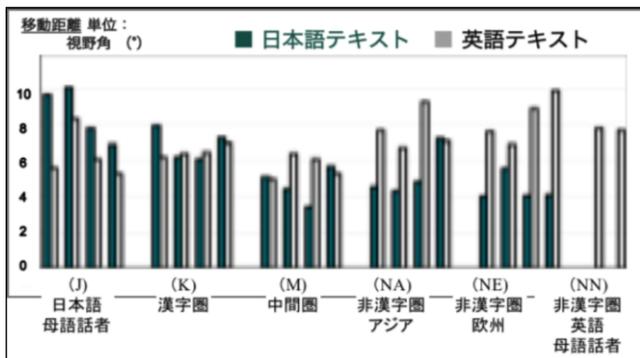


図 11 サッケードの移動距離（両テキスト）

図 12 はサッケード（視線移動）1秒間の移動速度を表す（単位：視野角°/sec）。

日本語テキストでは、漢字圏(K)がもっとも高速で移動した。英語テキストでは、非漢字圏グループ(NA)(NE)と漢字圏グループ(K)が高速に移動した。日本語母語話者グループ(J)はもっとも速度が遅い結果であった。

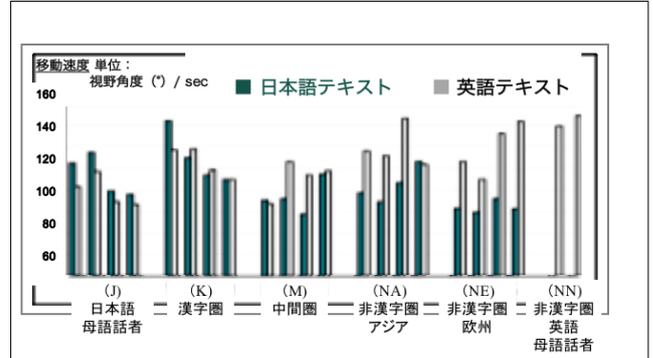


図 12 サッケードの移動速度（両テキスト）

3.6. 戻り読み

読解における視線移動は、前方だけではなく既読の場所へ数文字、あるいは数行前に戻る場合がある。逐次的な読み、戻り読み、飛ばし読みの詳細を知る目的で、画面上に提示されるテキストと視線移動の視野角から、視線移動の方向と種類を分析した[12]。

表 4 は、テキスト読解時の視線移動を X 軸、Y 軸により分類した視線移動の区分を表す。日本語の表記と英語の表記は異なる形態であることから、読解における視線移動の単位として、日本語では文字数で表示し（表 4 左）、英語では単語数で表示した（表 4 右）。それぞれ、赤枠は後方への長い移動（戻り読み）、青点線の枠は、前方への長い移動（飛ばし読み）を表す。

表 4 視線移動の種類

日本語テキスト上の視線移動の分類：		英語テキストの視線移動の分類：	
A	Y 軸上の戻り読み	A	Y 軸上の戻り読み
B	X 軸上の戻り読み：3 字以上	B	X 軸上の戻り読み：3 語以上
C	2 字戻り読み	C	2 語戻り読み
D	1 字戻り読み	D	1 語戻り読み
E	1 字前方読み	E	1 語前方読み
F	2 字前方読み	F	2 語前方読み
G	3-4 字前方読み	G	3-4 語前方読み
H	X 軸上の飛ばし読み：5 字以上	H	X 軸上の飛ばし読み：5 語以上
I	Y 軸上の飛ばし読み	I	Y 軸上の飛ばし読み
J	その他：改行など	J	その他：改行など

図 13 は、日本語テキストにおける視線移動方向を表す。また、図 14 は、英語テキストにおける視線移動方向を表す。非漢字圏グループ(NE)

(NA) と中間圏グループ(M) で、1字ごと、または2字ごとに前方に読み進む割合が高いことがわかる(グラフに赤、ピンクの帯の割合が高い)。一方、漢字圏グループ(K)は、3~4字ごとに前方に読み進む割合が高く、日本語圏グループ(J)は戻り読みと飛ばし読みの割合が高い傾向がみられた。日本語圏グループは英語テキストでも戻り読みと飛ばし読みの割合が高い傾向があり、特にY軸方向の戻り読みが多い傾向がみられた。

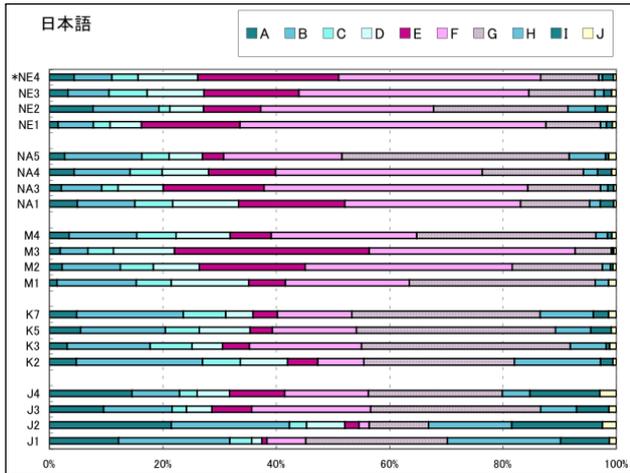


図 13 視線移動の種類 (日本語テキスト)

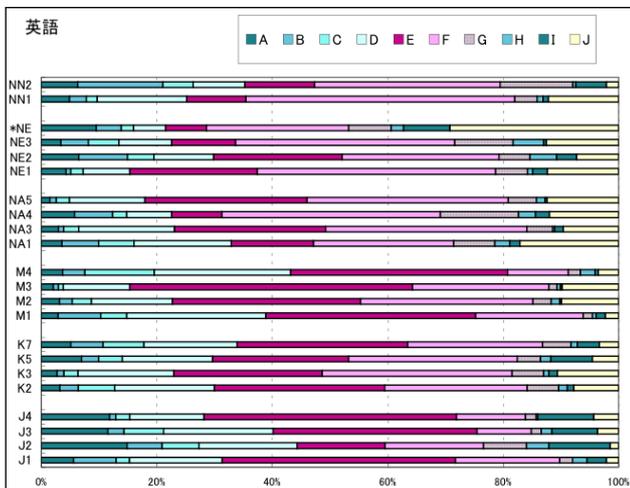


図 14 視線移動の種類 (英語テキスト)

4. 考察

視線追跡装置を利用して日本語母語話者(J)と日本語非母語話者(K) (M) (NE) (NA) (NN)、合計6グループを対象にして、日本語と英語テキストの読解過程を観察する実験を行い、その視線データを分析した。

漢字圏グループ(K)は、重松らが指摘したように漢字の上で視線停留を行う傾向が顕著であっ

た[3]。日本語母語話者グループ(J)は、日本語テキストにおいて視線停留が少なく停留時間も短い。一方で、英語テキストにおいて、視線停留が長くまた、視線停留時間は、英語と日本語の両方のテキストで母語または母語に近い言語で短い傾向がみられた。

視線移動時間で、両テキストで日本語母語話者(J)と中間圏グループ(M)に長い傾向がみられた。

視線移動の方向で、日本語母語話者に戻り読みと飛ばし読みが多い傾向がみられた。

戻り読みについて、Kintschは、「読み手はある単語に対し複数の意味・意義を保持しながら文章を読み進めるが、その意味・意義と逸脱した使い方を見出すとき戻り読みが起り、上級学習者に多い傾向がある」と述べている[5]。しかし、日本語母語話者が日本語テキスト読解で戻り読みと飛ばし読みを多用したことは、習熟した学習者に戻り読みが多いという理論と合致しない。なぜなら、日本語母語話者は日本語学習者ではない。

本実験では、英語テキストの読解で英語母語話者グループ(NN)でも戻り読み、飛ばし読みがみられたが、日本語母語話者ほど多くなかった。また、日本語母語話者(J)は、日本語テキストでも英語テキストでも戻り読みと飛ばし読みを多用する傾向がみられた。このことから、日本語母語話者(J)に特徴的な読解の特徴である可能性がある。

5. まとめ

本研究で、視線追跡装置を利用して日本語テキストと英語テキストの読解過程の観察実験を行った。比較のために、日本語母語話者(J)と非母語話者(K)(M) (NA) (NE) (NN)、合計6グループの読解時の特徴を分析した。その結果、視線停留時間で、母語または母語に近いグループの平均停留時間が短かった。すなわち、習熟している言語の読解では、視線停留時間は短い傾向であった。

テキストの難易評価と視線停留時間では、どのテキストでも各グループの視線データはグループの特徴が現れた。このことは言語習得において母語が影響を及ぼすことを示唆している。

視線移動時間では、日本語と英語の両テキストで日本語母語話者グループ(J)と中間圏グループ(K)の視線移動時間が長く、他のグループは短い傾向がみられた。すなわち、両テキストともに母語によって同様のサッケードを行う結果がみられた。

視線移動の方向を調べたところ、日本語母語話者の視線移動で「戻り読み」が顕著に多かった。

日本語は、「膠着語系」「Head-Final 文法」「省略が多い」などの特徴を持ち、文脈の逸脱を起ししやすい言語である。いたずらに長いテキストは短期記憶や長期記憶などワーキングメモリの負荷についても考えなければならない。これが日本人に戻り読みを多くさせ、かつ、その方法を知らない外国人が日本語の読解を困難と考える原因の1つである可能性が考えられる。

今回の実験で、日本人は日本語と英語の両テキストで戻り読みを多用する傾向がみられた。このことから、戻り読みの多用は日本語文法の特徴によるのではないかという仮説が生まれる。日本語は「Head-Final 文法」であり、主要部はテキストの最後に現れる。「膠着語系文法」の言語は1文が長い傾向がある。これは、文法的に似通っている中間圏グループ(M)でも特徴的に出現した。

英語は関係代名詞や不定詞などの役割が明瞭であり、これをよく理解している英語母語話者や英語に近い母語の人々が、戻り読みを多用せずに読解を行っている可能性がある。英語圏の人々の場合、戻り読みは「読み手はある単語に対し複数の意味・意義を保持しながら文章を読み進めるが、その意味・意義と逸脱した使い方を見出すとき戻り読みが起りやすい」という Kintch の理論が当てはまる[5]。すなわち、Kintch の母語である英語テキスト読解上の理論である可能性がある。

日本語母語話者(J)は、英語テキストでも戻り読みを多用した。日本語母語話者(J)は日本語の読解方法を英語で利用した可能性が考えられる。

5.1. 今後の展望

本研究で、学習対象言語の読解において母語の影響がどのように現れるかを調査する目的で、視線追跡装置を利用して視線データを取得し、視線停留とサッケードから分析した。実験参加者は6グループ、22名で、対象人数として充分とは言えず、今後人数を増やして実験を行い、さらに詳細な分析を行う必要がある。

今回、両方のテキストにおいて母語によるグループが同様のサッケードを行う結果を得た。今後、視線移動の観点から各グループの読解スキルを調査する必要がある。

また、視線移動の方向を分析した結果、日本語母語話者の視線移動で「戻り読み」が顕著に多かった。戻り読みは、現時点で日本語母語話者の特徴的な現象の1つである可能性が高いという結果を得たが、今後、複眼的な観点からさらに分析を進める必要がある。日本語と英語のテキスト読解

では、それぞれの母語話者の他に、両方を理解するバイリンガルの読解過程を調査することも有益であると考えられる。

日本語のように膠着語系言語や Head-Final 文法の特徴を持つ言語では1文が長い傾向がある。このような言語の理解において、ワーキングメモリにどのような負荷があるかを調べることも、今後の課題の1つである。

文 献

- [1] 谷口すみ子, 日本語学習者の読解過程の調査, 日本語教育学会大会予稿集, 1992
- [2] 古賀一男, 中澤幸夫, 苧阪良二 編, 眼球運動の実験心理学, 名古屋大学出版会, 1993.
- [3] 重松, 鴻巣, 福田, 「アイカメラ」による Non-Native の「読み」の実証的研究-第2報-, 日本語教育学会大会予稿集, 1994.
- [4] Keith Rayner, "Eye Movements in Reading and Information Processing: 20 Years of Research, Vol. 124, No. 3, pp. 372-422, psychological Bulletin (1998).
- [5] Walter Kintsch Comprehension, Cambridge university press, 1998.
- [6] 鈴木美加, 初級後半の学習者は文章をどう読むのか -アイカメラによる文章読解中の眼球運動の記録-, 東京外国語大学留学生センター紀要, 1998
- [7] 寺朱美, 北村達也, 奥村学, 日本語読解支援システム「DL」の検証 -- 日本語学習者の読解プロセスの研究 --, 日本語教育方法研究会会誌, Vol. 6, No. 1, 1999.
- [8] 日本留学試験予想テスト, アルク, 2003.
- [9] 中日新聞, 2004.
- [10] 石田健, 毎日1分! 英字新聞, 2004.
- [11] 寺朱美, 杉山公造, 視線追跡装置による日本語学習者の文章読解過程の研究 (II) -視線停留とサッケード-, 電子情報通信学会技術研究報告, TL2005-18-26, pp.31-36, 2005.
- [12] 寺朱美, 杉山公造, 視線追跡装置による日本語学習者の文章読解過程の研究 (IV) -視線追跡データから分析した戻り読みと飛ばし読み-, 電子情報通信学会技術研究報告, TL2005-35-48, pp.43-48, 2006.
- [13] Jincho, N., Feng, G., & Mazuka, R., Development of text reading in Japanese: An eye movement study, Reading and Writing, Volume 27, Number 8, 27, pp.1437-1465, (2014).