コーパスコンコーダンサ『ChaKi.NET』の連続値データ型 (2)
―読み時間の表示―

浅原 正幸 (国立国語研究所コーパス開発センター) *
池本 優 (近畿大学)
森田 敏生 (総和技研)

Double Type Data on 'ChaKi.NET'
— Visualization of Reading Time —

Masayuki Asahara (Center for Corpus Development, NINJAL)
Yu Ikemoto (Kinki University)
Toshio Morita (Sowa Research Co., Ltd.)

1. はじめに

『ChaKi.NET』 (Matsumoto et al. (2005)) は、コーパスに付与されたメタデータ・形態論情報・係り受け情報を用いて文単位の検索を行ったり、形態論情報・係り受け情報に基づく頻度統計情報を取得したり、自動解析により付与された形態論情報・係り受け情報を修正したりすることができるコーパス管理システムである。開発当初は書記言語を対象にしていたためデータベース内部におけるデータ型のほとんどがテキストやアノテーションを格納する文字列型やアノテーションを抽象化した整数型が用いられてきた。前回のワークショップでは ChaKi.NET の連続値データ型 (浅原・森田 (2013)) の導入について発表し、利用法として話し言葉コーパスなどの音声コーパスの時刻情報を格納し、形態素単位の発話継続時間情報を可視化するデモを行った。

今回の発表では「現代日本語書き言葉均衡コーパス」の文の読み時間の可視化を行う。文の読み時間とは、被験者に文もしくは文章を呈示して文の一部分を読むのにかかる時間を計測したものである。読み時間の遅延を人の文処理機構を解明するための手がかりとして用いる。読み時間の測定方法として安価な機材で測定可能な「移動窓方式の自己ペース読文実験」や専用の実験装置を用いた「視線走査実験」などがある。

本稿では、移動窓方式の自己ペース読文実験や視線走査実験結果を元にどのような可視化が可能であるのか、また、ChaKi.NET 上でどのような分析が可能であるのかを議論するとともに、現在何ができないかを議論する。

2. 読み時間の測定方法

本節では読み時間の測定方法として「移動窓方式の自己ペース読文実験」や「視線走査実験」について概説する。

* masayu-a@ninjal.ac.jp

2.1　自己ペース読文実験

自己ペース読文実験は、スペースキーなどを押すごとにディスプレイ上に刺激文が単語もしくは文節単位に順次呈示し、被験者のペースでスペースバーをおしながら文を読み進め、その表示時間を記録することで読み時間を計測する実験である。Linger と呼ばれるソフトウェアが良く用いられる。自己ペース読文実験は、刺激文の呈示の方法が二種類ある。一つは中央表示方式で、もう一つは移動窓方式である。中央表示方式は、画面中央に単語または文節が表示され、スペースキーを押すごとに消えて別のものが表示する方式である。被験者から見ると同じ場所に新しい文字列が示される。移動窓方式は、最初に文字の数だけハイフンやアンダーバーなどを表示し、スペースキーを押すごとに部分的に単語または文節を表示する方式である。図 1 に移動窓方式の自己ペース読文実験の画面遷移を示す。被験者からみると見える文字列部分（窓）が移動していくように示される。調査対象の言語現象にもよるが、日本語の場合は通例文節単位に呈示し、何 msec かけて文節文字列を読むかを評価する。読文作業においては後戻りして読むことを許さないため、各文節は高々 1 回の注視時間が記録される。

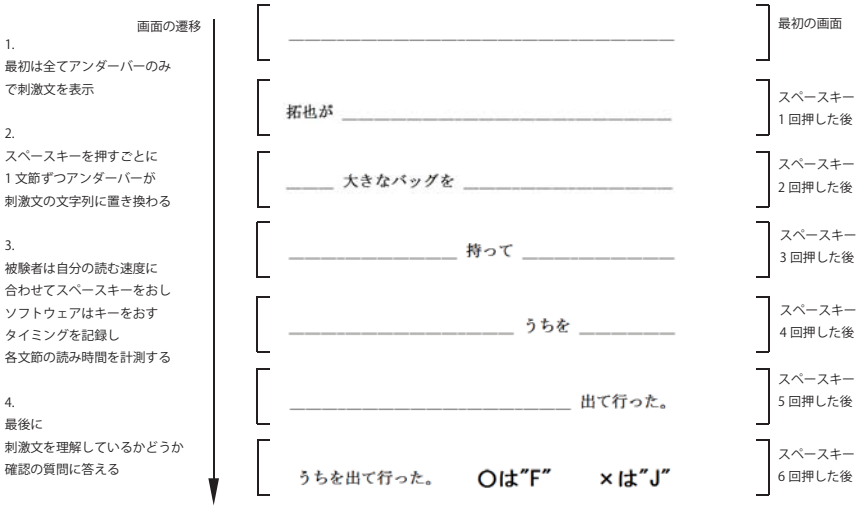


図 1　移動窓方式による画面の遷移例

現在、「現代日本語書き言葉均衡コーパス」(BCCWJ) のサンプルを用いて移動窓方式による自己ペース読文実験による読み時間評価を実施している。表 1 に BCCWJ に対する読み時間評価結果を示す。1 画面に最大 5 文呈示する。横の最大文字数は約 40 文字で、40 文字を超えた文節の次で改行して呈示する。表 1 の例は、サンプル名 PN1c_00001 の 2 画面目の一部の結果を示しており、文節 ID 4, 5, 14 の直後で改行している。文単位は基本的に BCCWJ の基準に従うが、あまりに不自然なものについては適宜修正している。

通例、呈示文節の「呈示順」「文字数」「モーラ数」などの情報で回帰を行ったうえで、（標準）残差を算出しどの部分で読み時間の遅延が発生するかを識別する。

表 1　BCCWJ に対する読み時間付与結果（移動窓方式による自己ペース読文実験）

画面 ID	文節 ID	呈示文節	停留時間 (msec)
2	0	子育ての	1428
2	1	かたわら	248
2	2	テレビ、	208
2	3	ラジオなどで	191
2	4	活躍中。	206
2	5	３３歳。	203
2	6	幼稚園から	203
2	7	大学まで	230
2	8	通った	199
2	9	青山学院では、	170
2	10	とにかく	193
2	11	活発で、	109
2	12	目立つ	192
2	13	生徒だったと	204
2	14	いう。	205
⋮	⋮	⋮	⋮

図 2　視線走査装置

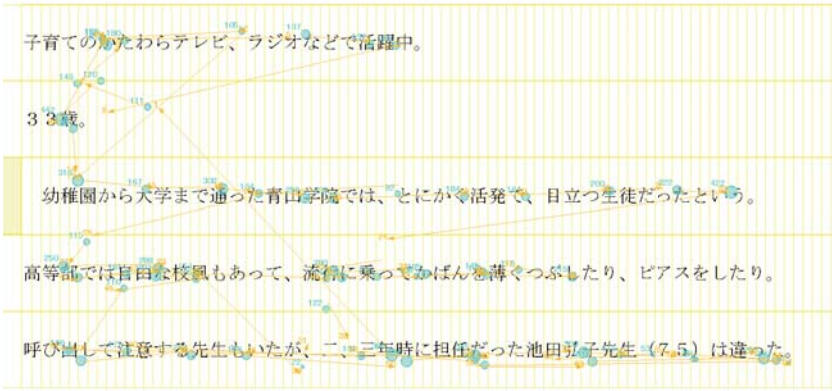


図 3　視線走査実験で得られるデータ例

2.2　視線走査実験

視線走査実験とは、視線走査装置を用いて画面上のどの文字を注視しているかを注視時間とともに計測することにより、文字列の文の読み時間を測定する方法である。心理言語実験に適した視線走査装置として SR Research 社の Eyelink CL を利用する。図 2 に機材の写真を示す。誤差を最小限にするためにハーフミラー越しに正面から瞳孔に弱い赤外線をあて、その反射をカメラでとらえるタワーマウントを用いる。あご台で顔を固定することで精度の高いデータを取得することができる。Eyelink CL は最大 1000Hz のサンプリングレートで瞳孔を走査し誤差が視野角 0.2 度未満としている。これは 50cm 離れてみる場合のモニター上の誤差が約 2mm となる。我々の実験では 21 インチの 1920x1080 ピクセル画面上の文字を 22 ポイントで呈示する。等幅フォントの文字幅は約 8mm である。

図 3 に視線走査実験で得られるデータ例を示す。表 1 と同じ PN1c_00001 の 2 画面目の結果

表 2　Dundee Eye-Tracking Corpus

WORD	TEXT	LINE	OLEN	WLEN	XPOS	WNUM	FDUR	OBLP	WDLP	LAUN	TXFR
Are	1	1	3	3	1	1	216	1	1	0	351
tourists	1	1	8	8	6	2	156	2	2	-5	3
enticed	1	1	7	7	17	3	227	4	4	-11	1
these	1	1	5	5	25	5	187	1	1	-8	73
attractions	1	1	11	11	33	6	182	3	3	-8	2
threatening	1	1	11	11	44	7	96	2	2	-11	3
Blink	1	-99	11	11	-99	-99	82	-99	-99	-99	-99
threatening	1	1	11	11	52	7	232	10	10	-99	3
very	1	1	4	4	62	9	335	2	2	-10	56
their	1	1	5	5	57	8	168	3	3	5	225
existence?	1	1	10	9	65	10	173	0	0	-8	4
existence?	1	1	10	9	71	10	188	6	6	-6	4
existence?	1	1	10	9	72	10	88	7	7	-1	4
enticed	1	1	7	7	19	3	174	6	6	53	1
these	1	1	5	5	29	5	168	5	5	-10	73
these	1	1	5	5	28	5	170	4	4	1	73
attractions	1	1	11	11	36	6	271	6	6	-8	2
attractions	1	1	11	11	34	6	88	4	4	2	2
threatening	1	1	11	11	46	7	232	4	4	-12	3
very	1	1	4	4	61	9	202	1	1	-15	56
existence?	1	1	10	9	72	10	222	7	7	-11	4
existence?	1	1	10	9	74	10	157	9	9	-2	4

例文 “Are tourists enticed by these attractions threatening their very existence?”

であり、●が注視点を表現し、その大きさが注視時間を表現している。注視点間を結ぶ線が、注視点の移動を表現している。画面から何度か後戻りして読むことが行われており、注視点も文節境界とは独立に分布しているように見える。

以前は Tobii や SR Research 社などの一部のベンダーが高価な視線走査装置を販売していたが、最近 Tobii 社が一般向けの視線走査装置を発売している。例えば、Tobii REX が 249 ユーロ、Tobii EyeX が 170 ユーロで発売される。精度では高価な機材に劣るが、一般に手に入る価格になり、言語研究において視線走査装置の利用が加速することが考えられる。

3.　視線走査実験データの表現手法

本節では視線走査実験データの表現手法について議論する。まず数少ない読み時間情報つきコーパスの先行研究である “Dundee Eye-tracking Corpus” Kennedy and Pynte (2005) について述べ、現在収集している日本語視線走査実験データの表現手法についても示す。

3.1　Dundee Eye-Tracking Corpus

読み時間を付与したコーパスの先行研究として Kennedy らの Dundee Eye-Tracking Corpus (以下 “Dundee Corpus”) がある。対象言語を英語とフランス語、それぞれ 10 人の母語話者を被験者とし、視線走査装置を利用して 20 ファイルの新聞社説に対する視線走査情報を記録している。各ファイルは 5 行ごとからなる 40 画面により構成されており、研究用途に一次情報が公開されている。Dundee Corpus は特定の言語現象分析を目的とせずに作成されたコーパス

に基づいており、心理言語学におけるさまざまな仮説の客観的な検証に用いられている。

表 2 に Dundee Corpus の例を示す。WORD が注視点が含まれた単語で時系列に表示される。TEXT が呈示画面番号、LINE が呈示行番号、OLEN が句読点を含めた文字列（オブジェクトと呼ぶ）長、WLEN が句読点を含めない単語文字列長、XPOS が注視点になる文字位置、WNUM が単語のインデックス、FDUR が停留時間 (単位は msec。以下時間・時刻については msec とする。)、OBLP がオブジェクト中の注視点位置、WDLP が単語中の注視点位置、LAUN が注視前のサッケード（視点を移動する眼球運動）長を相対位置で示したもの、TXFR がテキスト中の単語頻度である。尚、WORD の Blink はまたたきを表現する。

この例では 3 単語目 (1-origin) の “by” が読み飛ばされたり、 8 単語目の “their” が 9 単語目の “very” より後に読まれたり (単語を超えても LAUN が正の値) する。10 単語目の単語 “existence” は文字列長が長いために単語内で複数個所注視点が含まれる。

3.2　日本語視線走査実験データの表現手法

本節では Dundee Corpus を元に日本語視線走査実験データをいかにして表現するかについて議論する。

再度図 3 を参照されたい。分析の際には細長い長方形のグリッドを等幅フォントの半分の幅かつ文字列を中心に行間の高さで稠密に配置し、半文字単位で注視点の位置を割り当てる。視線走査装置はこのグリッド単位に注視されたかどうかを視線の停留時間とともに記録する。

分析の際にはこのグリッド上の注視点の停留時間を文字単位、形態素単位（国語研短単位 or 国語研長単位)、文節単位など用途に応じた単位に割り当てる必要がある。さらに読み戻しなどをどう扱うかを決める必要がある。

現在、データの表現方法として、最初の 1 回目の視線走査のみを表示する方法（First Pass と呼ぶ)、全ての視線走査を表示する方法の 2 種類を検討している。以下 2 種類の方法を示すが、確定したものではなく意見の募集を行っている。

3.2.1　1 回目の視線走査のみを表示する方法（First Pass）

一つ目の方法は、1 回目の視線走査のみを表示し、他の情報を捨てる方法である。以下 “First Pass” と呼ぶ。この場合、原文の文字列順序に沿って視線走査記録を示すことができる一方、読み戻しなどの一部の情報は捨てられてしまう。また読み戻しを文字・形態素・文節のどの単位で認定するかにより、捨てられる情報が変わってくる。

表 3 に文字単位の日本語の視線走査情報（First Pass）の例を示す。LINE が何行目か、CIDX が原文文字列上の文字位置、FDUR が 1 回目の注視時間、DSUM が複数回の視線停留時間の合計、FSTT が 1 回目の注視開始時刻、FEND が 1 回目の注視終了時刻、CHAR が文字、SUW が形態素、BPH が文節である。SUW は基本的に国語研短単位に準拠しているが、数値のみだけ元の書字形で記述している。実データにおいては、図 3 のように注視点は文字幅の半分単位で計測しているが、これを文字単位に集計したものである。

FSTT を若い順にたどることによって被験者がどの順で文を読んでいるかがわかる。この例では、前画面から復帰するための途中で CIDX 26 の注視点を経て CIDX 6 の注視点で読文を開始し、CIDX 5 に推移している。また CIDX 6 と CIDX 5 の関係をみると、周辺視野でとな

表 3　日本語の視線走査情報（First Pass）

LINE	CIDX	FDUR	DSUM	FSTT	FEND	CHAR	SUW	BPH
1	1	.	0	.	.	子	子育て	子育ての
1	2	.	0	.	.	育	子育て	子育ての
1	3	.	0	.	.	て	子育て	子育ての
1	4	.	0	.	.	の	の	子育ての
1	5	168	554	1055	1222	か	かたわら	かたわら
1	6	180	180	857	1036	た	かたわら	かたわら
1	7	.	0	.	.	わ	かたわら	かたわら
1	8	.	0	.	.	ら	かたわら	かたわら
1	9	.	0	.	.	テ	テレビ	テレビ、
1	10	.	0	.	.	レ	テレビ	テレビ、
1	11	.	0	.	.	ビ	テレビ	テレビ、
1	12	105	105	3120	3224	、	、	テレビ、
1	13	.	0	.	.	ラ	ラジオ	ラジオなどで
1	14	.	0	.	.	ジ	ラジオ	ラジオなどで
1	15	.	0	.	.	オ	ラジオ	ラジオなどで
1	16	187	187	1280	1466	な	など	ラジオなどで
1	17	.	0	.	.	ど	など	ラジオなどで
1	18	.	0	.	.	で	で	ラジオなどで
1	19	167	167	1490	1656	活	活躍	活躍中。
1	20	.	0	.	.	躍	活躍	活躍中。
1	21	186	186	1674	1859	中	中	活躍中。
1	22	.	0	.	.	。	。	活躍中。
2	23	.	0	.	.	３	３３	３３歳。
2	24	.	0	.	.	３	３３	３３歳。
2	25	442	442	2179	2620	歳	歳	３３歳。
2	26	146	311	662	807	。	。	３３歳。
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

りの単語や文節を確認していることがあることがわかる。このような現象をどのように扱うかが問題となる。

表 4 単語単位にデータを集計したものを示す。WIDX が単語順を表す。CIDX 6 と CIDX 5 のように連続して同じ単語を見ている場合には合算されるが、CIDX 12 のように先の方を読んでから戻ってきたと考えられるものについては排除される。

3.2.2　全ての視線走査を表示する形式（Whole Pass）

二つ目の方法は、全ての視線走査を視線走査順に表示する方法である。以下 “Whole Pass” と呼ぶ。この場合、被験者の視線走査順序に沿って視線走査記録を示すことができる一方、原文の文字列情報の可読性がなくなる。

表 5 に単語単位の日本語の視線走査情報（Whole Pass）の例を示す。FIDX が視線走査順、FCNT が注視回数である。FIDX 3 までは前画面から復帰するための注視であると考える。

4.　ChaKi.NET 上での読み時間の可視化

本節ではどのようにして読み時間を可視化するかについて示す。具体的には以下の 3 種類の可視化を検討している。

表 4　日本語の視線走査情報（First Pass）の単語単位の集計

WIDX	FDUR	DSUM	FSTT	FEND	SUW	BPH
1	.	0	.	.	子育て	子育ての
2	.	0	.	.	の	子育ての
3	348	734	857	1222	かたわら	かたわら
4	.	0	.	.	テレビ	テレビ、
5	105	105	3120	3224	、	テレビ、
6	.	0	.	.	ラジオ	ラジオなどで
7	187	187	1280	1466	など	ラジオなどで
8	.	0	.	.	で	ラジオなどで
9	167	167	1490	1656	活躍	活躍中。
10	186	186	1674	1859	中	活躍中。
11	.	0	.	.	。	活躍中。
12	.	0	.	.	３３	３３歳。
13	442	442	2179	2620	歳	３３歳。
14	146	311	662	807	。	３３歳。
⋮	⋮	⋮	⋮	⋮	⋮	⋮

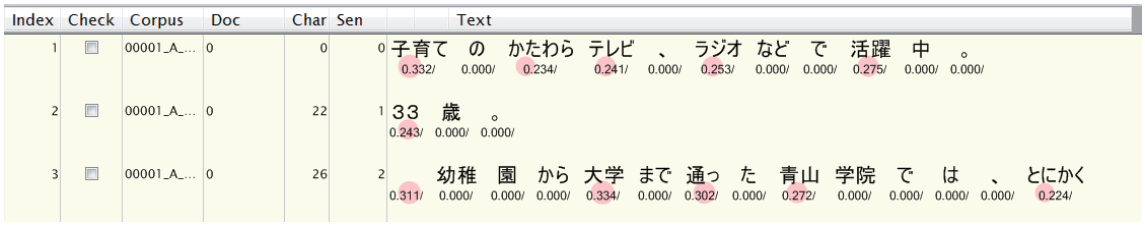
表 5　日本語の視線走査情報（Whole Pass）

LINE	FIDX	FDUR	DSUM	FCNT	FSTT	FEND	CHAR	SUW	BPH
5	1	366	366	1	7	372	年	年	二、三年時に
-99	2	111	111	1	528	638			
2	3	146	311	2	662	807	歳	歳	３３歳。
1	4	180	180	1	857	1036	た	かたわら	かたわら
1	5	168	320	2	1055	1222	か	かたわら	かたわら
1	6	187	187	1	1280	1466	な	など	ラジオなどで
1	7	167	167	1	1490	1656	活	活躍	活躍中。
1	8	186	186	1	1674	1859	中	中	活躍中。
-99	9	120	.	.	2032	2151			
2	10	442	442	1	2179	2620	歳	歳	３３歳。
1	11	234	234	1	2663	2896	か	かたわら	かたわら
1	12	152	320	2	2914	3065	か	かたわら	かたわら
1	13	105	105	1	3120	3224	、	、	テレビ、
2	14	165	311	2	3685	3849	歳	歳	３３歳。
3	15	318	318	1	3889	4206	稚	幼稚	幼稚園から
3	16	167	167	1	4234	4400	大	大学	大学まで
3	17	302	302	1	4425	4726	通	通っ	通った
3	18	259	259	1	4752	5010	学	学院	青山学院では、
3	19	239	239	1	5031	5269	院	学院	青山学院では、
3	20	184	184	1	5312	5495	く	とにかく	とにかく
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

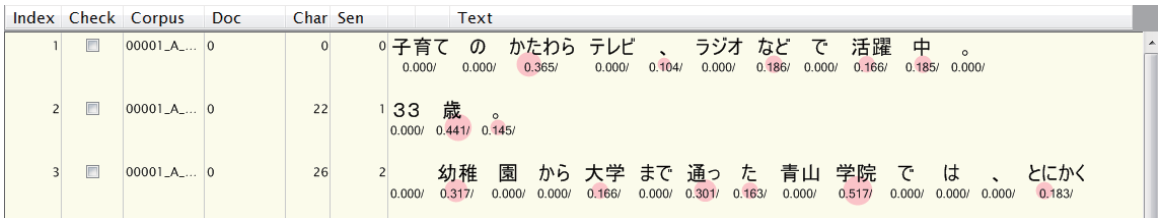
- 自己ペース読文実験で得られた読み時間の可視化
- 視線走査実験で得られた読み時間の可視化 (First Pass:コーパス出現順単語呈示)
- 視線走査実験で得られた読み時間の可視化 (Whole Pass:読み時間順単語呈示)

一つ目の自己ペース読文実験で得られた読み時間の可視化を図 4 上段に示す。自己ペース読文実験で得られた読み時間は文節単位のものである。読み時間は文節単位に割り当てられているが、ChaKi.NET において時間情報の割り当て単位は形態素単位になっている。検索の簡便

自己ペース読文実験の ChaKi.NET 上での可視化（文節頭に注視時間情報付与）



日本語視線走査実験 (First Pass) の ChaKi.NET 上での可視化（文字単位の注視時間情報を単語単位に集積）



日本語視線走査実験 (Whole Pass) の ChaKi.NET 上での可視化（文字単位の注視時間情報を単語単位に表示。集積は行わない。）

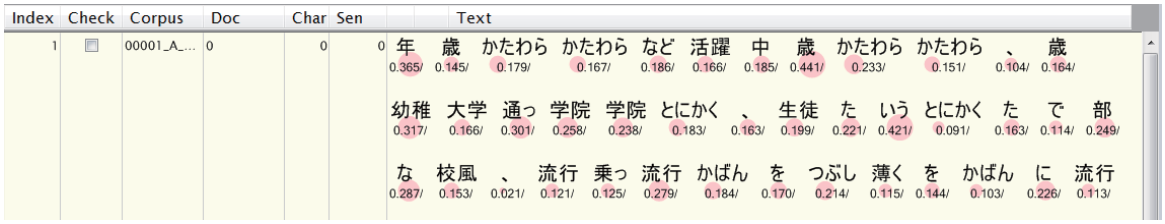


図 4　ChaKi.NET 上での文の読み時間の可視化

のために文節頭の形態素に文節全体の時間情報を付与している。読み時間の計測においては後戻りや文節飛ばしを許さないために全ての文節頭の形態素に時間情報が付与される。また視線走査装置とは異なり、注視している時間だけでなく視線を移動する時間（サッケード）も各形態素に割り当てられる。元データに付与されている形態論情報や係り受け情報との親和性がよく ChaKi.NET 上で統合して検索しやすい。

二つ目の 3.2.1 節で示したデータ表現形式 (First Pass) の可視化を図 4 中段に示す。一つ目の可視化との違いとして、全ての文節に読み時間が付与されない、文節中で真に注視された単語に読み時間が付与される、必ずしも文字列順に読まれているわけではないという点がある。この三点から時間情報に基づいて有用な検索をすることは難しい。さらに同一単語内で連続する注視点についてはその注視点間を移動する時間は読み時間として含まれているが、それ以外の注視点間を移動する時間は読み時間として含まれない点がある。また自動処理により処理されるデータでは、前の画面から復帰する途中の移動で注視されたものについては一回目の注視点 First Pass として認定されてしまうという点がある。

三つ目の 3.2.2 節で示したデータ表現形式 (Whole Pass) の可視化を図 4 下段に示す。上記二つと異なり文字列の線形順序の情報が失われており、そのままでは元の言語現象を問い合わせることが難しい一方、注視順に呈示されている。言語分析には不適かもしれないが、読み戻しや時間をかけて読んでいるところや読み飛ばしているところを重点的に観察することで要約補助用途などの利用が考えられる。

5. 課題

本節ではデータ表現形式と可視化の課題について検証する。

検証の前提として、可視化の目的について示す。我々の研究目的は心理言語研究と言語処理の二つの目的がある。一つ目の目的は例文が統制されている心理言語実験報告が、コーパス中の実際の用例でも観察できるかどうかを検証することである。コーパスで対象となる言語現象を形態論情報と係り受け構造により絞り込み検索をし、そこで観察される読み時間のふるまい俯瞰してみることが必要になる。二つ目の目的は言語処理応用である。既存の言語解析開発においてはコーパスに対して専門家がアノテーションを行い、それを機械学習に基づいて再現することで解析器を構成することが多い。一方、かな漢字変換などにおいては、系列学習のような技術が開発される一方で、予測変換といった実際の人間のふるまいを再現するようなモデルも提案されている。かな漢字変換は「書く」行為のモデル化だが、「読む」行為においても同様なモデル化が可能になると考える。今後ラップトップ PC やスマートフォンなどのカメラを利用して、大量の視線走査情報情報が得られることが考えられるが、これらを用いた自動要約・難読箇所検出・テキスト簡単化などが可能になるのではないかと考える。

このような目的のもと、現状の課題について示す。

- 周辺視野の可視化
  現在は何らかの基準で代表する形態素に読み時間情報を付与している。しかし、周辺視野で隣接している形態素や文節なども同時に見ている可能性があり、読み飛ばされているのか、周辺視野で確認しているのかを識別するような情報が必要になる。
- 複数人被験者のデータの可視化
  心理言語実験においては複数人の被験者データの分析が必要になる。現在、一つのテキストに対して一つの読み時間情報しか格納できない。
- 文字数やモーラ数による正規化
  心理言語実験においては、平均時間を評価したり、呈示順・文字列長・モーラ長に基づいて回帰分析を行いその残差を評価したりすることが行われる。読み時間そのものでは個人差などがあり何らかの正規化機構や統計論的モデル化が求められる。
- 文字列の線形順序と視線走査順序の相互関係
  自己ペース読文実験結果のように文字列の線形順序と視線走査順序が一致する場合については問題が起こらないが、視線走査実験のように文字列の線形順序と視線走査順序が必ずしも一致せず、読み戻しや読み飛ばしが起きるよう場合は、その相互作用の可視化が重要になる。
- 読み戻しの可視化
  自動要約や難読箇所検出においては読み戻しや読み飛ばしの可視化が重要になる。読み飛ばしは文字列の線形順序で First Pass の表現方法でも十分読み飛ばされていることが確認できる一方、読み戻しについては可視化できているとは言い難い。現在 Whole Pass などを用いて ChaKi.NET の WordList 機能を用いることで、仮想的に読み戻し回数の可視化が可能であるが、十分とは言えない。

このような課題に対し、ChaKi.NET のセグメント・リンク・グループの利用を検討している。ChaKi.NET では、形態論情報や係り受け構造に追加して、任意の構造としてセグメント・リンク・グループを表現することができる。例えば、周辺視野の情報をセグメントで表現したり、First Pass の表現上で視線の推移をリンクで表現したり、Whole Pass の表現上で同一形態素をグループで表現したりすることが考えられる。

6. おわりに

本稿では、コーパス管理システム ChaKi.NET の連続値データ型を用いて、読み時間を格納し可視化する手法について紹介した。現在作成中の BCCWJ に対する読み時間アノテーションは他の用途に利用されることも歓迎である。それぞれの用途に基づいて、必要な情報や利用者系は異なってくると考える。ポスター発表ではコーパスやツールに対する要望などを伺いたい。

謝辞

本研究の一部は科研費基盤 (B)「言語コーパスに対する読文時間付与とその利用」、国語研基幹型共同研究プロジェクト「コーパスアノテーションの基礎研究」および国語研「超大規模コーパス構築プロジェクト」によるものです。

参考文献

浅原正幸・森田敏生 (2013).「コーパスコンコーダンサ『ChaKi.NET』の連続値データ型」 第 4 回コーパス日本語学ワークショップ予稿集, pp. 249–256.

Kennedy, A., and J. Pynte (2005). “Parafoveal-on-foveal effecits in normal reading.” *Vision Research*, 45, pp. 153–168.

Matsumoto, Yuji, Masayuki Asahara, Kou Kawabe, Yurika Takahashi, Yukio Tono, Akira Ohtani, and Toshio Morita (2005). “Chaki: An annotated corpora management and search system.” *Proc. of the Corpus Linguistics Conference Series (Corpus Linguistics 2005)*.

関連 URL

- 「ChaKi.NET」 Web ページ：`http://sourceforge.jp/projects/chaki/releases/`
- 「Linger」Web ページ：`http://tedlab.mit.edu/~dr/Linger/`