社団法人　電子情報通信学会
THE INSTITUTE OF ELECTRONICS,
INFORMATION AND COMMUNICATION ENGINEERS

信学技報
TECHNICAL REPORT OF IEICE.

学習補助のための視点情報に基づく文書アノテーション

大社　綾乃† Kai Kunze†† OlivierAugereau† 黄瀬　浩一†

† 大阪府立大学大学院工学研究科　〒 559–8531 大阪府堺市中区学園町 1–1
†† 慶應義塾大学大学院メディアデザイン研究科　〒 223–8526 神奈川県横浜市港北区日吉 4–1–1
E-mail: †okoso@m.cs.osakafu-u.ac.jp, ††{kai.kunze,augereau.o}@gmail.com, †††kise@cs.osakafu-u.ac.jp

あらまし　本稿では，アイトラッカから取得した視点情報を用いて，文書を読んだ際に難しいと感じた部分を推定し，その情報を視覚化するシステムを提案する．文書を読むという行為から，学習者の理解できていない部分を推定することが出来れば，次の２点で大変有用であると考えられる．まず，学習者は普段の学習において，難しいと感じた部分の全てに印をつける手間を省くことが出来る．そして，指導者はその情報をもとに，学習者に合った指導が可能になる．一般的に人は文書を読む際，難しいと感じると読むスピードが落ちたり，頻繁に読み返す，同じ場所を注視し続けるといった傾向がある．本システムでは，このような特性から読者が難しいと感じた部分を推定し，その結果をもとに，段落ごとおよび単語ごとに文書画像上にアノテーションを付与する．被験者８人，１０文書を用いて，読書中の視点情報から，難しいと感じた部分を推定し，アノテーションを作成したところ，(１）読む速さや注視回数は被験者間の理解度の違いの識別に有効であり，(2）注視時間が難しいと感じた部分の識別に有効であることが分かった．

キーワード　アイトラッカ，視点情報，英語習熟度，アノテーション，学習補助

1. まえがき

日常生活における読む行為をデジタルデータとして記録するリーディングライフログ [1] が注目を集めている．リーディングライフログは，人がどのようなものを，どのくらい，どのように読んだかを記録するものである．例えば，どのようなものかの記録として，文書の種類やジャンルといったものや，どのくらいの記録として，文書数や単語数といったものが挙げられる．また，どのように読んだかの記録としては，目の動きや脳波，表情などを用いて，興味や理解度といったものを記録することが挙げられる．リーディングライフログを解析することにより，ユーザが普段どのような単語や文書に注目しているかを知ることができ，各ユーザの趣味嗜好に合った情報やサービスを提供することが可能になる．また，読んだ単語や文書を記録することで，学習進度の管理や補助にも利用できる．

そこで本稿では，リーディングライフログを利用し，文書を読むだけで学習成果を知るためのデータを視覚的に提供するシステムを提案する．語学学習において，自分の学習成果を正しく知ることは大変重要な要素の１つである．自分の学習成果を知ることで，理解できていない部分を把握したり，今後の学習計画を立てるのに有用である．学習成果を測る方法としては，定期的に試験を受けたり，日々の学習で難しいと感じた単語や文章に印をつけておくことなどが挙げられる．しかし，試験は，その時の体調や問題との相性に影響を受ける場合がある．加えて，出題箇所と関係のない部分については，理解できているかどうかを知ることが出来ないという欠点がある．また，日々の学習で難しいと感じた単語や文章の全てに１つ１つ印をつけておくことは，大変手間がかかり，容易ではない．もし，文書を読むだけで自分の学習成果を知ることができれば大変有用であると考えられる．その結果，学習者は頻繁に試験を受けたり，難しいと感じた部分を文書中に細かく印をつける手間を省くことができる．また指導者は，学習者の学習成果を簡単に知ることができ，個人に合った指導が可能となる．

提案システムでは，学習者が文書を読む際の視点情報の解析から，難しいと感じた部分を推定し，その情報をアノテーションとして文書画像上に印をつける．一般的に人は文書を読む際，難しいと感じると，読む速度が遅くなったり，何度も読み返す，同じ場所を注視し続けるといった傾向がある．そのような特性を視点情報から得ることで，学習者が難しいと感じた部分を推定する．まず視点情報を得るために，アイトラッカを用い，学習者の眼球運動を記録する．次に視点情報から特徴量を抽出し，難しいと感じた部分を推定する．その結果をもとに，段落ごとおよび単語ごとにアノテーションを作成し，文書画像上に表示する．

本システムの有効性を検証するために，英語力別に被験者間のアノテーションを比較する．また，難しいと感じた部分がどの程度抽出可能であるかを評価する．なお，本研究は大阪府立大学工学研究科の倫理委員会で承認済みであることを付記しておく．

2. 関連研究

本章では，読書時の眼球運動やリーディングライフログ，デジタルアノテーションに関する関連研究について述べる．

読書中の眼球運動と文書理解度および言語習熟度の関係性を

調査した研究として，Gomez らの研究が挙げられる [2]．Gomez らは，アイトラッカから取得した視点情報のうち，文書理解度に影響を及ぼす特徴量として，文書を読むのにかかる時間や saccade（注視間の素早い目の動き）の平均長などを示している．また，言語習熟度に影響を及ぼす特徴量としては，平均注視時間や saccade の平均長を挙げている．しかし，Gomez らは文書全体の理解度を対象として検証しており，文書の部分的な理解度については言及していない．また，Biedert らは，文書の書き方が読者にとって分かりやすいものかどうかについて視点情報を用いて評価する方法を提案している [3]．視点情報を用いた文書の部分的な評価としては本研究と共通する部分があるが，Biedert らは著者へのフィードバックを目的としている点で，本研究とは目的が異なっている．また，以前我々は，文書中の部分的な理解度を推定するのに有効な特徴量を段落ごと，時間による segment ごと，単語ごとの 3 つの視点から調査した [4]．その結果，segment ごとおよび単語ごとの理解度については，注視回数や注視時間，saccade の平均長が有効であることが分かった．また，段落ごとの理解度については，その段落を読むのにかかった時間および段落に含まれる単語数が有効であることが分かった．

視点情報およびリーディングライフログを利用した学習に関するシステムとして，ユーザが読んだ単語数に着目した Wordometer が挙げられる [5]．Wordometer とは，アイトラッカから取得した視点情報を用いて，ユーザが読んだ単語数を推定し，記録するシステムである．ユーザが日々読んだ単語数を記録することで，学習や読書に対するモチベーションの維持・向上に繋がると考えられる．しかし，現時点では，単語数を推定するのみで，読んだ単語の特定にまでは至っていない．吉村らの研究では，アイトラッカから取得した視点情報を用いて，英語習熟度を推定する手法を提案している [5]．吉村らは，TOEIC スコアに基づく 3 段階の英語習熟度（上級・中級・初級）を高い識別率で推定することに成功している．吉村らの手法は英語習熟度の推定に有効であるが，文書中のどの部分を理解していないかについては推定の対象としていない．文書中のどの部分を難しいと感じているかについて着目した研究として，Kunze らの研究がある [6]．Kunze らの研究ではアイトラッカから取得した視点情報を用いてユーザが難しいと感じた単語の推定を目指しており，本研究と類似する．しかし，定量的な評価には至っておらず，どの単語を読んでいるかについては手動で推定している．また，ユーザに対するビジュアル化について言及していない点も，本研究とは異なっている．

以前我々は，ユーザが難しいと感じた部分を文書画像上にアノテーション付与するシステムを提案した [7]．従来システムでは，読書中の視点情報から，ユーザが難しいと感じた部分を推定し，その情報を文書画像上にアノテーションを付与する．また，段落ごとのアノテーションにのみ焦点を置き，視点情報を手動で文書画像上の段落ごとに分けている．提案システムでは，視点情報の位置修正の自動化および精度向上を行い，段落に対するアノテーションだけでなく，より細かい部分でのアノテーションとして，単語に対するアノテーションを付け加えている．

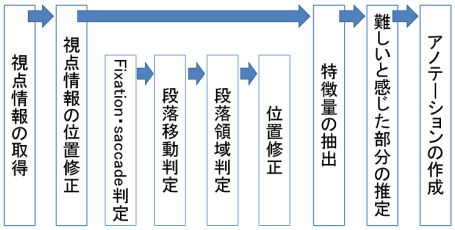


図 1　提案システムの流れ

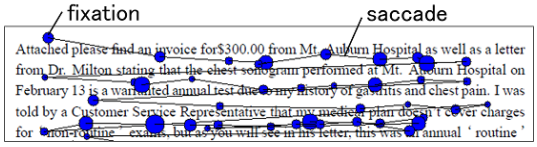


図 2　視点情報に fixation-saccade 判定を施したもの

3.　提案システム

3.1　提案システムの流れ

図 1 に提案システムの流れを示す．まず，アイトラッカを用いてユーザの眼球運動を記録し，視点情報を取得する．アイトラッカから得られる視点の位置には誤差が含まれるため、位置を修正する．その際，段落ごとに位置を修正するため，各視点情報がどの段落に属するかを求める．次に，段落ごとおよび単語ごとに視点情報を複数の特徴量で表す．このようにして得られた特徴量を用いて，難しいと感じた部分を推定する．その結果に基づいてアノテーションを作成し，文書画像上に表示する．

3.2　視点情報の取得

アイトラッカを用いて視点情報を取得する．アイトラッカとは，眼球運動を測定する装置である．本システムでは，据置型のアイトラッカを用いて，ディスプレイ上に表示された文書を読む際の視点情報を取得する．

3.3　視点情報の位置修正

視点位置の修正は 4 つのプロセスに分類される．まず，アイトラッカから得られた視点情報を，注視を意味する fixation と注視間の素早い目の動きを意味する saccade に分ける．そして得られた fixation 情報から、段落移動を検出する．これは視点の位置修正を段落ごとにおこなうためである．次に各 fixation がどの段落に属するかを求める．最後に，段落ごとに平行移動および拡大縮小をおこなうことで，視点の位置を修正する．以下に各プロセスについて詳しく述べる．

3.3.1　fixation-saccade 判定

アイトラッカから得られた視点情報を，fixation と saccade に分ける．fixation は視点がある範囲内に一定時間停留することであり，saccade は fixation 間の素早い目の動きのことである．人は fixation と saccade を繰り返すことで、ものを見たり文書を読む．本システムでは，Buscher らの手法 [8] を用いて，図 2 のように視点情報を fixation と saccade に分ける．

3.3.2　段落移動判定

段落ごとに視点の位置を修正するため，fixation 情報を用いて，図 3 のようにどこで段落移動をしたかを推定する．ある

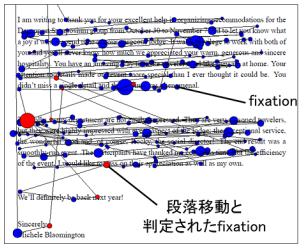


図 3　段落移動判定を施した fixation

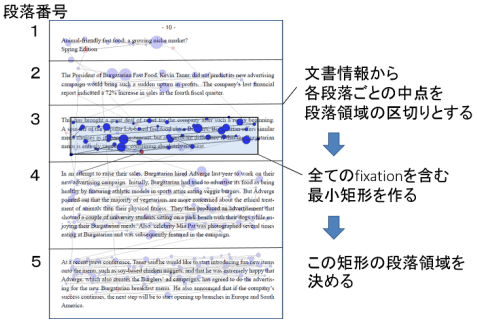
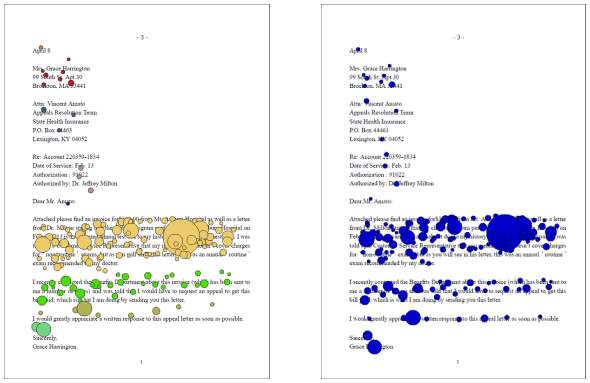


図 4　段落領域判定の手順



位置修正前　　位置修正後

図 5　位置修正を施す前後の fixation

表 1　アノテーション作成に利用する特徴量

グループ	特徴量	記号	詳細
段落	読む速さ	S	単語数/fixation 合計持続時間
	読み返し回数	R	-
	fixation 平均持続時間	Q_a	fixation 合計持続時間/fixation 回数
単語 イディオム	fixation 持続時間	Q_t	単語領域に含まれる fixation 合計時間

fixation を F_n とし，その座標を (X_n, Y_n) とする．H_1 を閾値とすると，

$$|Y_n - Y_{(n+1)}| > H_1 \tag{1}$$

を満たす場合，F_n と F_{n+1} の間を段落移動とみなす．

3.3.3　段落領域判定

段落移動を判定した後，各 fixation がどの段落領域に含まれるかを求める．その手順を図 4 に示す．まず，予め文書情報を抽出しておき，段落間の二等分線を段落領域の区切りとしておく．次に，段落移動と判定されてから次に段落移動と判定されるまでを 1 つの segment とする．その segment 内に含まれる全ての fixation を含む最小矩形を作る．その矩形が最も広範囲に含まれる段落領域を，その segment の属する段落と決める．

3.3.4　位 置 修 正

段落ごとに視点情報の位置を修正する．まず，ある段落に含まれる全ての fixation を包含する最小矩形を作る．また，文書情報を用いて，その矩形が属する段落のテキストの重心を包含する最小矩形を作る．そして平行移動により，fixation の矩形とテキストの矩形の中心を合わる．また縦方向において，テキストの矩形に合わせて fixation の矩形に拡大縮小を施し，視点情報の位置を修正する．横方向においては，fixation の矩形の両端に位置する fixation がどの単語を読んでいるのか特定できないため，縮小拡大は施していない．この場合でも，文書のサイズとアイトラッカの特性上，横方向のスケーリングは比較的誤差が小さいため，拡大縮小の処理を施さなくても問題は生じにくい．このようにして修正された fixation を図 5 に示す．図 5 の左図は段落領域判定後の様子を表しており，右図は段落ごとに位置修正を施した様子を表している．この位置修正後の fixation の位置をもとに，アノテーションを付与する．

3.4　特徴量の抽出

人は文書を読む際，難しいと感じると，読む速度が遅くなったり，何度も読み返す，同じ場所を注視し続けるといった傾向がある．そのような特性を利用し，fixation の回数や持続時間，読み返しの回数などを求め，特徴量とする．本システムでは，難しいと感じた部分の推定を段落ごとと単語ごとにおこなう．そのため特徴量は，段落に対するものと単語に対するものに分けて考案する．表 1 にアノテーション作成に使用する特徴量を示す．まず段落に対する特徴量として，読む速さ S，読み返しの回数 R，fixation の平均持続時間 Q_a を用いる．単語に対する特徴量としては，fixation の合計持続時間 Q_t を用いる．Q_t は，図 6 のように単語領域を定めたとき，各単語領域に含まれる fixation の合計持続時間を指す．読み返しの判定方法について，以下に詳細を述べる．

読み返しの判定は，視点情報の位置修正の際に同時に行う．図 7 のように，ある fixation を F_n とし，その座標を (X_n, Y_n) とする．また，その前後の fixation とのなす角度を A_n とする．段落移動の際は，縦方向に大きく視点が移動する傾向がある．また，改行の際，視点は読んでいた行の行末から次の行の行頭へ向かう動きをする．その動きは縦方向には小さいが，横方向に大きく，それまで読んでいた時の視点の移動方向とは逆向きになる．そして，読み返しは，場面によって，段落をまたぐものや，単語，イディオムの読み返しなど様々である．本システムでは，単語，イディオム程度の部分を繰り返し読むことを読み返しとする．その動きは横方向にも縦方向にも小さく，それまで読んでいた時の視点の移動方向とは逆向きになる．これらの特性を利用し，図 8 のフローチャートにそって，段落移動・改行・読み返しを判定する．ここで H，W，K は閾値を指す．以上の方法で求めた読み返しの回数を特徴量として用いる．

3.5　ユーザが難しいと感じた部分の推定

難しいと感じた部分の推定は段落ごとと単語ごとに行う．ど

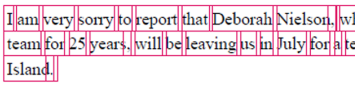


図 6　単語領域の例

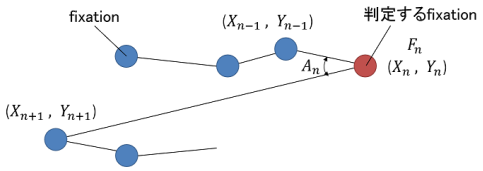


図 7　段落移動・改行・読み返し判定の際の fixation の様子

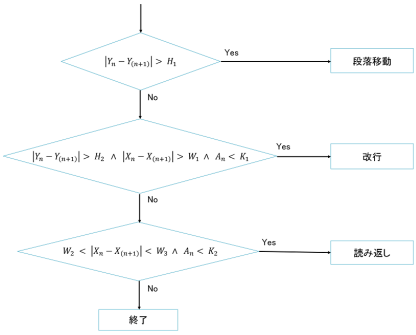


図 8　段落移動・改行・読み返し判定のフローチャート

の段落を難しいと感じたかを推定するために，読む速さ，読み返しの回数，fixation の平均持続時間を用いる．読む速さが遅いほど，読み返しの回数が多いほど，fixation の平均持続時間が長いほど，その段落を難しいと感じた可能性が高いと考えられる．また，単語やイディオムなど，より細かい部分を難しいと感じたか推定するために，単語領域内の fixation の合計持続時間を用いる．ある単語に含まれる fixation の合計持続時間が長いほど，その単語を長く注視しているため，その単語を難しいと感じた可能性が高いと考えられる．本システムでは，これらの情報を用いて難しいと感じた部分を推定する．

3.6　アノテーションの作成

段落ごとのアノテーションと単語ごとのアノテーションを作成する．その例を図 9 に示す．

各段落の左右に位置するマークは，段落ごとの特徴量をもとに作成するアノテーションである．亀のマークは読む速さを表し，読む速さが遅いほど亀の数が多くなる．矢印のマークは読み返しの回数を表し，読み返しの回数が多いほど矢印の数が増える．そして，虫眼鏡のマークは fixation の平均持続時間を表し，虫眼鏡の数が多いほど fixation の平均持続時間が長いことを表す．いずれも 3 段階アノテーションで表すが，アノテーションがつかない段落もある.（以後，4 段階のアノテーションと表現する.）アノテーションの段階を決めるために，閾値をもうける．まず全ユーザの全文書数の特徴量（S，R，Q_{ave}）に対して正規化をおこなう．これは英語力別に被験者間でアノテーションの比較をおこうことを目的としているためである．そして，それぞれの特徴量に対し適当な閾値を設定し，4 段階にわけ，アノテーションに対応させる．

図 9　提案システムにより得られるアノテーションの例

表 2　被験者の分布

クラス	TOEIC リーディングスコア	人数
A	400〜495	1
B	300〜399	4
C	200〜299	3

テキスト上に記された矩形のマークは，単語ごとの特徴量をもとに作成するアノテーションである．D_1 を閾値とすると，

$$Q_t > D_1 \tag{2}$$

を満たすとき，文書画像上に表示する．またそれぞれの閾値は，各被験者の文書ごとに適当な値を設定する．これは被験者によって fixation の持続時間に大きく差があったため，また，文書によって難易度が異なっているためである．

4.　実　　験

本章では提案システムの検証について述べる。被験者間の英語能力の違いがアノテーションに及ぼす影響と、難しいと感じた部分の推定精度の 2 点を検証した。

4.1　実 験 概 要

名の被験者に TOEIC 問題集より抜粋した英文 10 文書を読んでもらい，その際の視点情報をアイトラッカにより記録した．被験者は大学生および大学院生 10 名（男性 6 名，女性 4 名）である．そのうち 2 名においては，アイトラッカによる視点情報の取得に失敗していたため，本実験のデータには用いなかった．被験者の英語力の指標として TOEIC のリーディングスコアを用いた．被験者の TOEIC スコアを 3 クラスに分類したものを表 2 に示す．リーディングスコアは 495 点満点で，最低点は 200 点，最高点は 450 点であった．また，本実験では，SMI 社の据置型アイトラッカ RED 250 を用いた．

次に実験手順について述べる．実験の様子を図 10 に示す．まずアイトラッカのキャリブレーションをおこない，キャリブレーションが十分に作動しているか確認する．次に，被験者はディスプレイ上に表示される英文書を読む．この際の視点情報を記録する．その後，英文書に関する問題が書かれた紙を渡し，被験者はその問題に答える．この際，被験者は英文書を見ることはできない．問題を問いてもらう目的は英文書を真剣に読んでもらうためであり，問題の正答数は実験には用いない．最後に、被験者は英文書を印刷した紙を受け取り、難しいと感じた部分に印をつける。被験者の書いたアノテーションの例を図 11 に示す。この手順を 10 文書分繰り返す．

4.2　実験 1 ：英語力別の被験者間の比較検証

英語力別の被験者間において，提案システムから得られたア

図 10　実験の様子

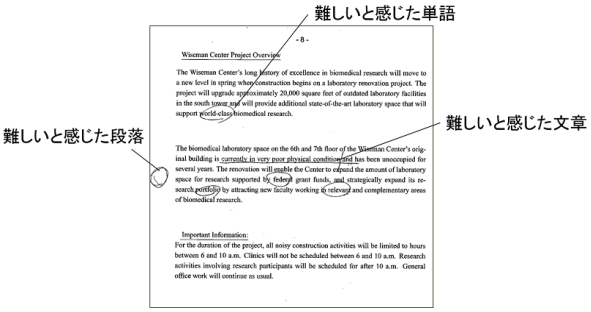
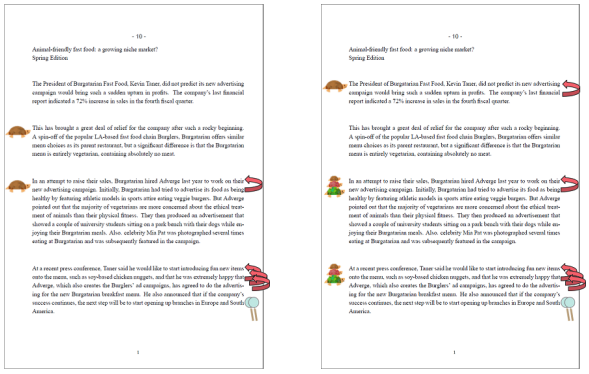


図 11　被験者によるアノテーションの例



クラス A　　クラス C

図 12　英語力によるクラス別被験者のアノテーション例

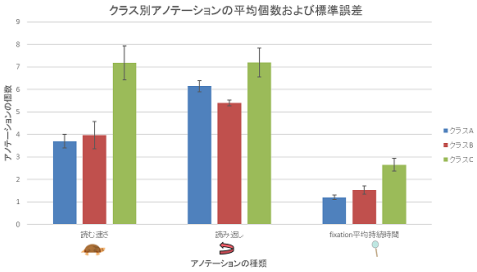


図 13　英語力別クラスごとのアノテーションの数の比較

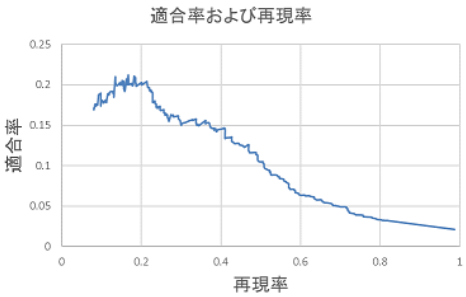


図 14　アノテーションの閾値を変化させた時の適合率および再現率

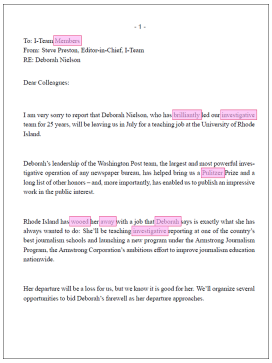
ノテーションにどのような差が見られるかを比較し，評価する．実験で記録した視点情報を元に提案システムを用いてアノテーションを作成した．英語力でクラス分けした被験者間で，段落に対するアノテーションを比較する．ここで段落に対するアノテーションのみを比較するのは，段落に対するアノテーションは全被験者の全文書を用いて閾値を設定しており，被験者間での比較が可能だからである．

英語力の高いクラスと低いクラスのアノテーションを図 12 に示す．また，クラス別のアノテーションを定量的に評価するため，各文書のアノテーションに含まれるマーク数を比較したグラフを図 13 に示す．今回用いた文書に含まれる段落数は平均 6.7 であり，アノテーションに含まれる亀や矢印，虫眼鏡の数は 0 ～ 3 である．例えば図 12 のクラス C のアノテーションの場合，亀の数が 7，矢印の数が 5，虫眼鏡の数が 2 と数える．図 13 より，クラス A とクラス B にはあまり違いが見られないが，クラス A・B とクラス C とは顕著な差が見られる．また読み返しのアノテーション数は各クラスであまり大きな違いが見られない．原因として，被験者によってどの程度理解できるまで読んだかに違いがあることが挙げられる．一部の被験者，特に上級者は，一通り読み終えた後に繰り返し部分的に文書を読み直し，文書の内容の記憶に努めていた様子が見られた．その結果，このようにクラスごとに読み返しのアノテーション数に差が見られなかったのだと考えられる．読み返しのアノテーション数は英語力別クラスを特定する要素としては不十分であるが，各文書のどの段落を難しいと感じたかにの推定に寄与すると考えられる．以上より，読む速さや fixation の平均持続時間のアノテーション数により，英語力の中上級者と下級者を判別することが可能である．下級者を判別可能であるという結果から，指導者がどの学習者が理解できていないのかを知ることが可能であるという提案システムの有効性を示していると言える．しかし，上級者と中級者を判別できるようなアノテーションおよび特徴量の考案が必要である．

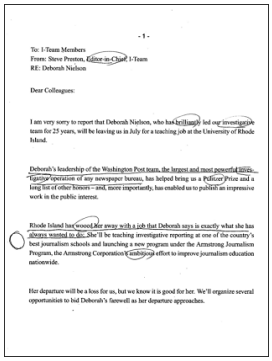
4.3　実験 2：ユーザが難しいと感じた部分の推定精度検証

ユーザが難しいと感じた部分の推定精度について検証する．実験で記録した視点情報を用いて，提案システムからアノテーションを作成した．被験者ごとの単語に対するアノテーションに着目し，実験の際に被験者に記入してもらったアノテーションと比較する．提案システムにより得られたアノテーションと被験者に記入してもらったアノテーションを図 15 に示す．また，定量的に評価するため，適合率および再現率を求めた．本検証では，被験者が難しいと感じた単語およびイディオムに着目した．各単語領域に含まれる fixation の合計持続時間 Q_t に対し，閾値を T とすると，

$$Q_t > T \tag{3}$$



提案システムによる
アノテーション



被験者による
アノテーション

図 15　単語に対するアノテーションの比較

を満たす場合に，正しく検出できたとして評価した．また，この評価は単語単位でおこなうが，被験者がイディオムを難しい部分として印付けした場合，イディオムを一単語として評価した．図 14 に閾値 T を変化させた時の適合率および再現率を表すグラフを示す．その結果，F 値が最大の時，適合率は 15.6%，再現率は 36.2% となった．この時の F 値は，21.8% であった．適合率，再現率ともに高い精度を得ることはできなかった．その原因について考察する．

まず，提案システムでは難しいと感じた部分として出力されたにもかかわらず，被験者が難しいと感じていない例として，人名や地名などの固有名詞が挙げられる．固有名詞は初見の際には注視が続くが，２度目以降は文脈を理解するのに必要ではないと判断され，読み飛ばす場合があったと考えられる．また，本検証では評価の対象としなかったが，被験者には難しいと感じた文章にもアノテーションを記入してもらっている．難しいと感じた文章に含まれる単語が，提案システムによって難しいと感じた部分として出力される例があった点も原因の１つであると考えられる．この点について，評価方法の改善および，難しいと感じた文章と単語での特徴量の違いをより細かく分析する必要がある．次に，被験者が難しいと感じたにもかかわらず，提案システムによって検出されなかった例として，複数回出現する単語が挙げられる．難しいと感じた単語が複数回出現した場合，二度目以降は検出されないことが多かった．予め文書から単語情報を抽出しておき，一度難しいと感じた部分として出力された単語は，二度目以降自動で補完することが有用であると考えられる．その他にも，視点情報の位置修正の失敗も適合率および再現率の低下の原因であると考えられる．

以上の点をふまえ，本検証では高い精度を得ることは出来なかったが，視点情報を利用することで難しいと感じた部分を推定することには改善の余地が多くあると言える．

5. ま と め

本稿では，アイトラッカから取得した視点情報を用いて，ユーザが難しいと感じた部分を推定し，文書画像上に段落ごとおよび単語ごとにアノテーションを付与するシステムを提案した．段落ごとの読む速さ，読み返しの回数，fixation の平均持続時間を難しいと感じた部分の推定に用いる特徴量とし，段落ごとにアノテーションを作成した．英語力によるクラス間で，段落ごとのアノテーションを比較したところ，読む速さおよび fixation の平均持続時間のアノテーションにおいて，英語力の上・中級者と下級者の間に大きな差が見られた．提案システムから英語力の低いユーザの判別が可能であると言える．しかし，読み返しの回数によるアノテーションについては，英語力別クラス間に違いが見られなかった．読み返しのアノテーションは英語力別クラスを特定する要素としては不十分であるが，各被験者が各文書のどの段落を難しいと感じたかの推定に寄与すると考えられる．この点についての更なる分析が必要である．また，ユーザが難しいと感じた部分を提案システムにより得られたアノテーションからどの程度推定可能であるかを検証した結果，適合率は 15.6%，再現率は 36.2% であり，十分な精度を得ることが出来なかった．しかし，特徴量の工夫や，より細かい分析により，精度向上の余地があると考えられる．

今後の課題としては，難しいと感じた部分をより高精度に推定できるように，特徴量の再考が必要である．ユーザにとってより分かりやすいよう，アノテーションを工夫することも課題の一つである．また，より多くのデータ収集や分析，実用的な環境での実験も必要である．

謝辞　本研究の一部は, JST CREST および日本学術振興会科学研究費補助金 基盤研究 (A)(25240028) ならびに若手研究 (B)(26730095) の補助による．

文　　献

[1] 川市仁史, K Kunze, 黄瀬浩一. 文書画像検索を用いたリーディングライフログの提案. 電子情報通信学会技術研究報告, Vol. 112, No. 495, pp. 255–260, 2013.

[2] Pascual Martínez-Gómez and Akiko Aizawa. Recognition of understanding level and language skill using measurements of reading behavior. In *Proceedings of the 19th international conference on Intelligent User Interfaces*, pp. 95–104, 2014.

[3] Ralf Biedert, Andreas Dengel, Mostafa Elshamy, and Georg Buscher. Towards robust gaze-based objective quality measures for text. In *Proceedings of the Symposium on Eye Tracking Research and Applications*, pp. 201–204, 2012.

[4] Kai Kunze, Hitoshi Kawaichi, Kazuyo Yoshimura, and Koichi Kise. The wordometer–estimating the number of words read using document image retrieval and mobile eye tracking. In *Document Analysis and Recognition, 2013 12th International Conference on*, pp. 25–29, 2013.

[5] 吉村和代, Kai Kunze, 黄瀬浩一. 読書時の眼球運動を利用した英語習熟度推定法. 電子情報通信学会技術研究報告, pp. 63–68, feb 2015–123.

[6] Kai Kunze, Hitoshi Kawaichi, Kazuyo Yoshimura, and Koichi Kise. Towards inferring language expertise using eye tracking. In *CHI'13 Extended Abstracts on Human Factors in Computing Systems*, pp. 217–222, 2013.

[7] Ayano Okoso, Kai Kunze, and Koichi Kise. Implicit gaze based annotations to support second language learning. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication*, pp. 143–146, 2014.

[8] Georg Buscher, Andreas Dengel, and Ludger van Elst. Eye movements as implicit relevance feedback. In *CHI'08 extended abstracts on Human factors in computing systems*, pp. 2991–2996, 2008.