

# Web情報探索における視線データの 予備的分析

○高久雅生(情報・システム研究機構)

寺井仁(東京電機大学)

江草由佳(教育政策研究所)

齋藤ひとみ(愛知教育大学)

三輪眞木子(メディア教育開発センター)

神門典子(国立情報学研究所 / 総合研究大学院大学)



# はじめに

- Web上の情報爆発、提供サービスの多様化
- 利用者がどのような行動をとっているか？
- 情報探索行動研究
  - 研究室実験、ログ分析、サンプリング調査、参与観察
- 研究室実験における収集データ
  - ブラウザ行動ログ
  - プロトコル発話
  - インタビュー
  - 事前・事後アンケート
  - 眼球運動

# 眼球運動測定による視線データ分析

- 既存研究における利用例：
  - 認知科学：読みの研究、問題解決
  - ユーザインタフェース、ユーザビリティ
  - Web情報検索研究
    - サーチエンジンなどの検索結果ランキングの改善
    - 上位10件での正解文書数・平均精度
- Web情報探索行動研究における利用例はまだ少なく、分析手法もまだ確立していない
- 予備的な分析結果を報告

# 本研究の位置づけ

- 目的
  - Web利用者の情報探索行動の理解
- 研究手法
  - 研究室実験
  - 分析対象：多様なデータを収集し、分析
    - 眼球運動、ブラウザログ、プロトコル発話、事後インタビュー
- Web探索実験
  - 被験者：大学生11名・ひとり2課題

# 実験の概要

- 被験者11名

- 首都圏地域の大学生（所属大学、専攻分野は限定せず）
- 19～21歳、男性5名・女性6名

- Web情報探索の課題

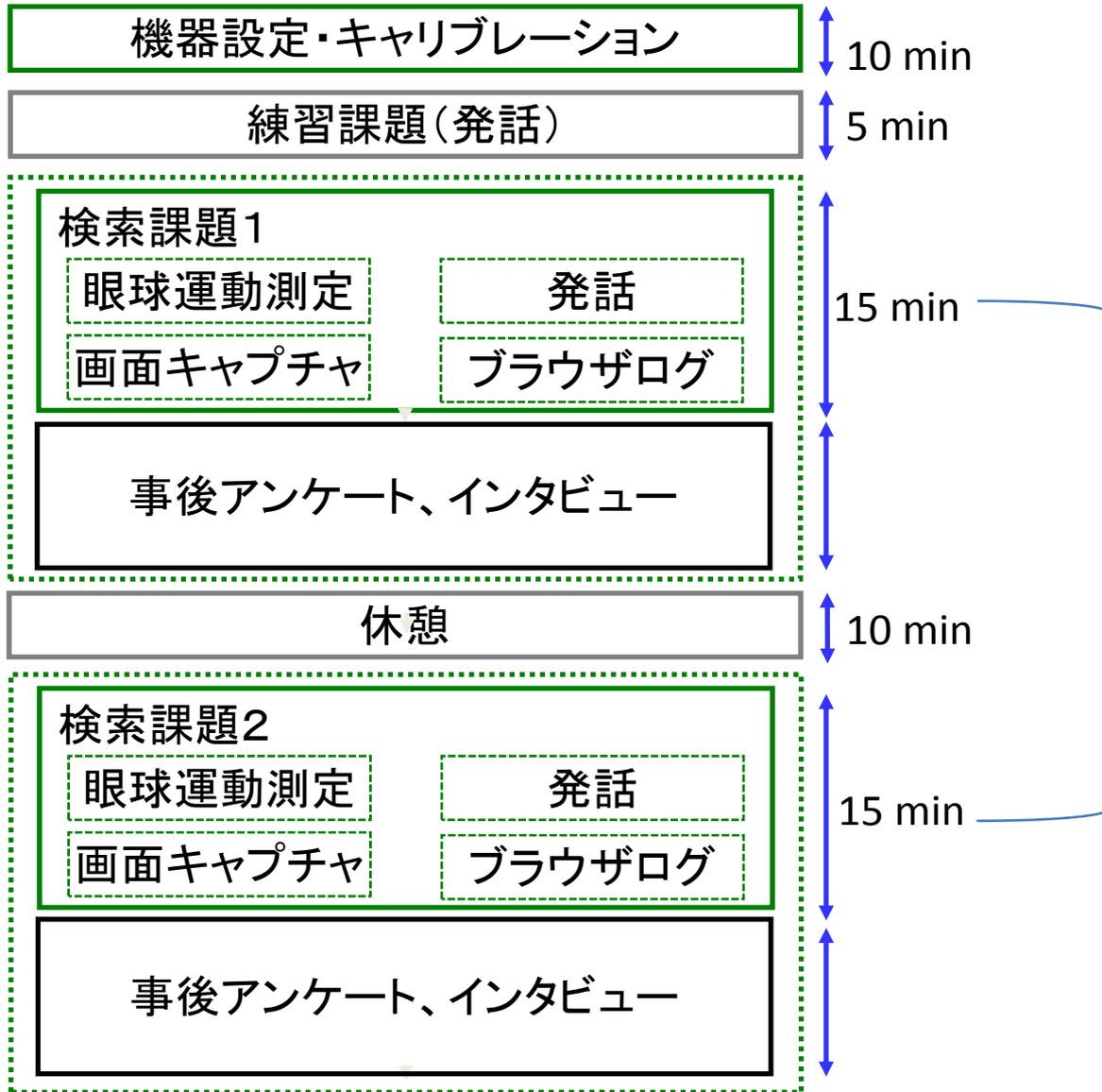
- 約15分間（各課題遂行の制限時間）
- 「世界史」レポートの情報収集
- 国内旅行のための情報収集

- （それぞれ具体的なテーマは被験者の興味に応じたものとした）

たとえば、第2次世界大戦、東インド会社の設立から解散まで、アメリカ合衆国の成り立ちなど

たとえば、友達5人と冬休みに沖縄へ、友達2人で春休みに瀬戸内海へ、友達3・4人で冬のスキー場へなど。

# 実験の流れ



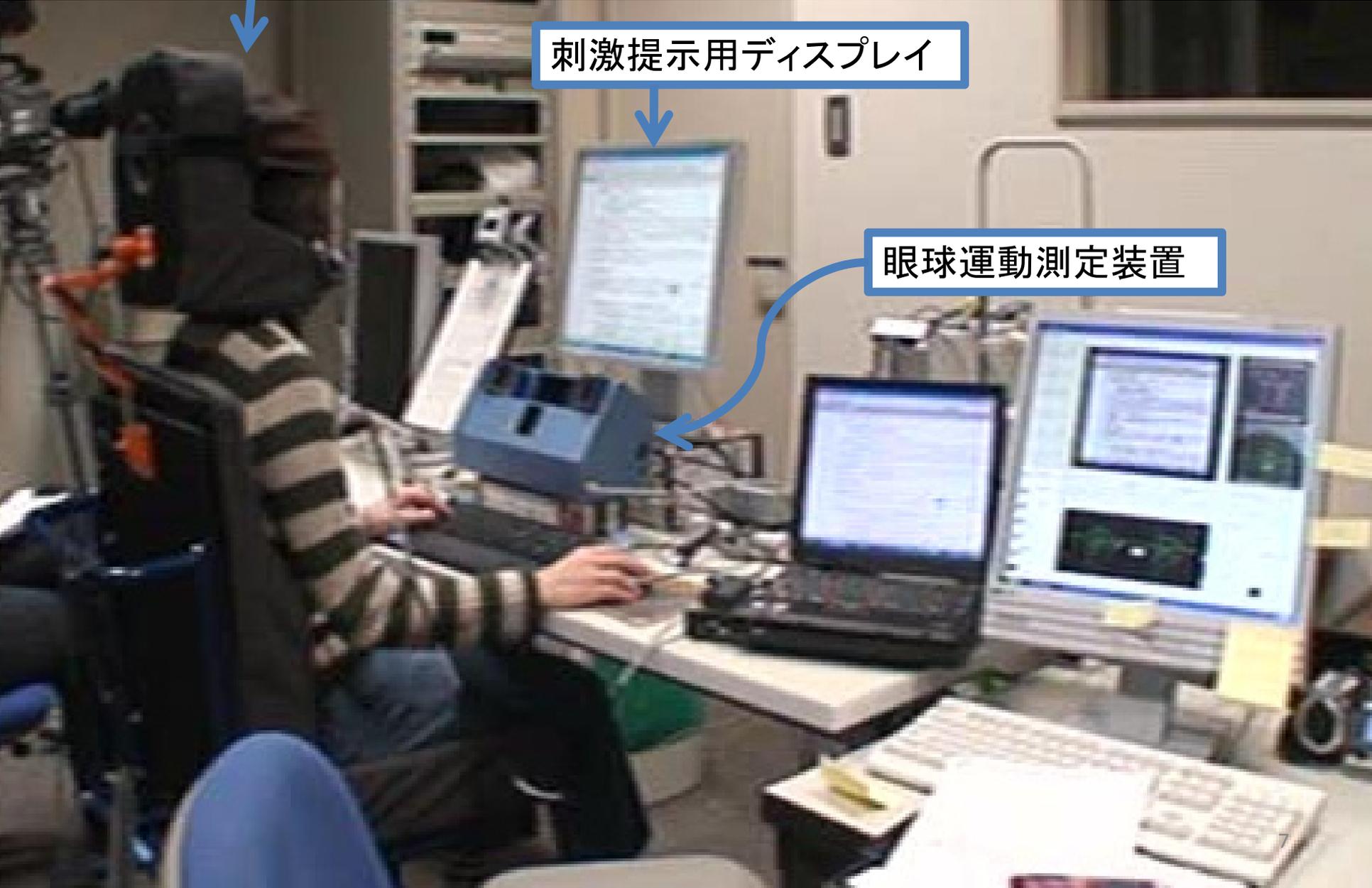
2課題の順序は被験者間で  
カウンターバランス

被験者

# 実験風景

刺激提示用ディスプレイ

眼球運動測定装置



19インチ液晶ディスプレイ

画面キャプチャ、ブラウザログ、視線データを計測

Windows XP

Mozilla Firefox(全画面モード)

眼球運動測定装置  
NAC社製 EMR-AT-VOXER



# 分析手法

探索過程全体  
被験者9名分

- 分析1

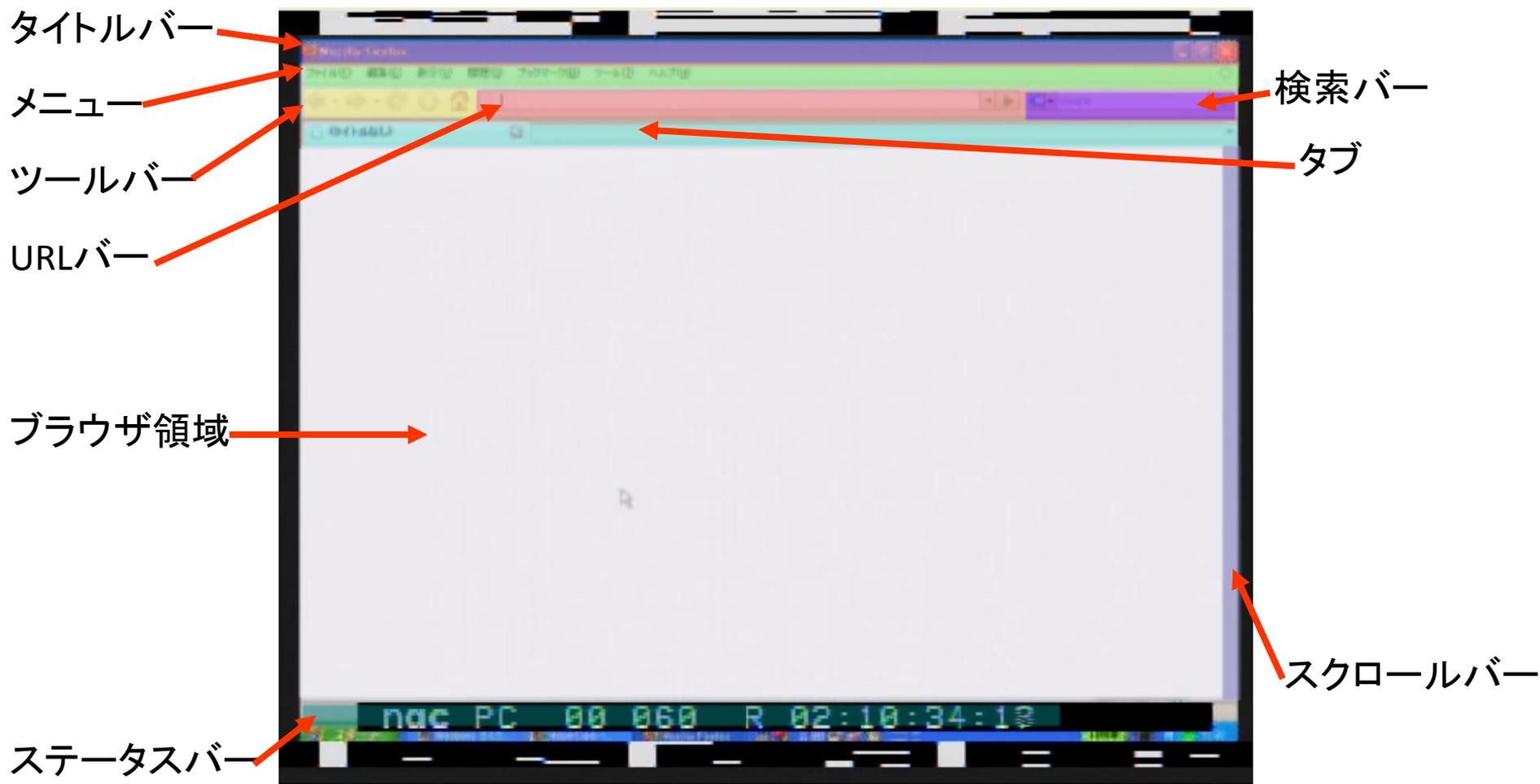
- 探索行動全体における画面中の注視領域の分析
- 眼球運動分析ツールによる自動解析結果

- 分析2

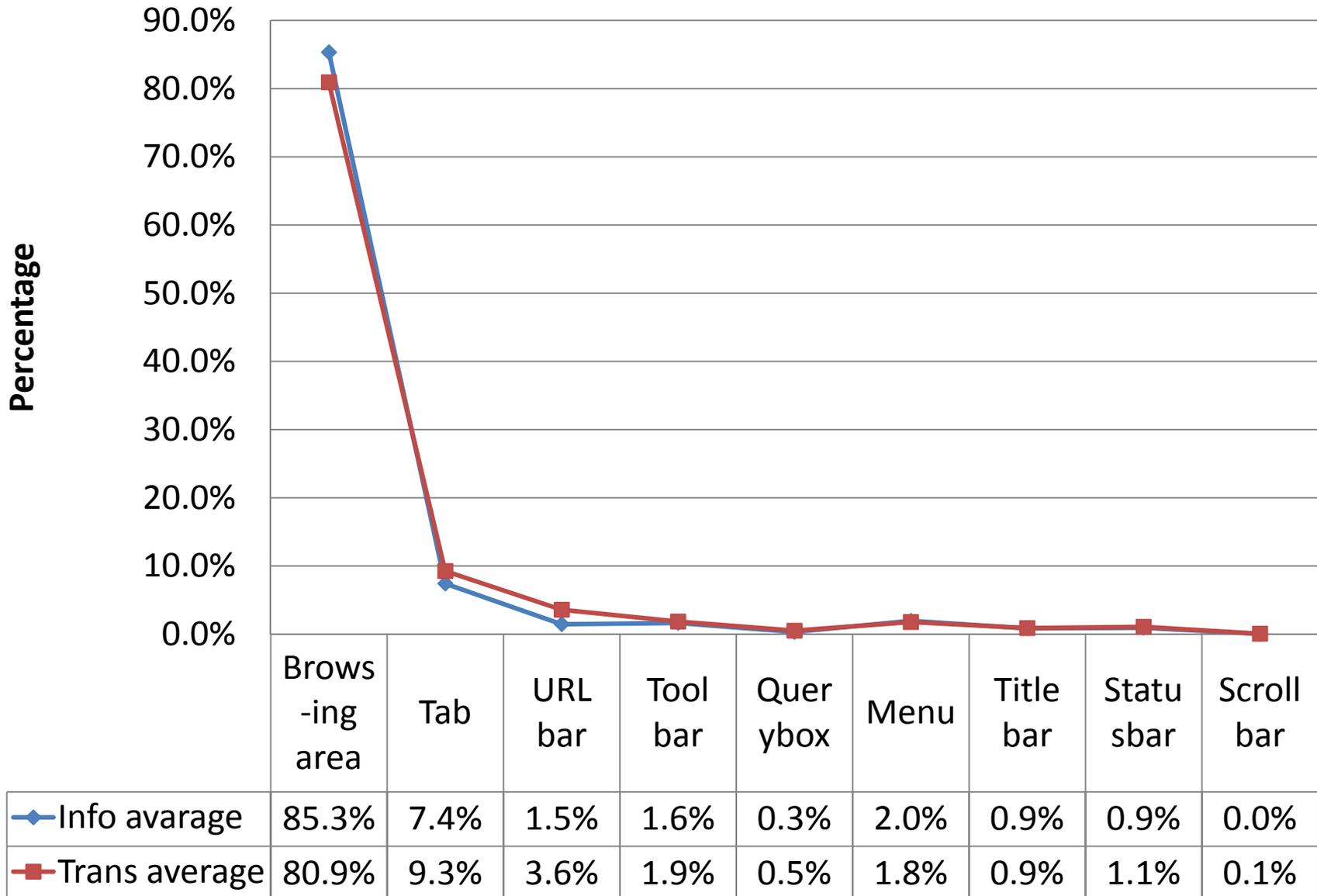
- サーチエンジンにおける検索結果ページのみを対象として、視線位置の推移を分析
- 人手によりページ内要素をタグ付け

検索結果ページ限定  
被験者4名分

# 分析1: 探索行動全体の注視点領域



# 分析1: 結果



## 分析2:

# 検索結果ページにおける視線移動

- 人手で視線(アイマーク)位置をタグ付け
  - 0.5秒間隔フレームで検索結果ページ内のページ要素をタグ付け
  - 22領域を設定
    - タイトルバー、メニュー、ブックマーク、ツールバー、URLバー、検索バー、検索バーボタン、タブ、サービスリンク、クエリボックス、検索ボタン、スクロールバー、ヒット件数、スペルチェック、スポンサーリンク、タイトル、スニペット、URL、関連検索、次へ、ページ内検索、ステータスバー、その他、エラー

1 amazon - Google 検索 - Mozilla Firefox

2 ファイル(F) 編集(E) 表示(V) 履歴(H) 3 ブックマーク(B) ツール(T) ヘルプ(H)

4 5 http://www.google.co.jp/search?q=amazon&lr=lang\_ja&ie=utf-8&oe=utf-8&aq: 6 Google 7

8 (タイトルなし) amazon - Google 検索

9 ウェブ 画像 地図 ニュース グループ Gmail more ログイン

10 amazon 11 検索 検索オプション 表示設定

12

13 amazon に一致する日本語のページ 約 25,600 件中 1 - 10 件目 (0.06 秒)

14 Amazon スポンサーリンク  
www.amazon.co.jp Amazonはこちら 1500円以上国内配送無料。代引OK

15 もしかして: [アマゾン](#)

16 Amazon.co.jp: 通販サイト 本、DVD、CD、ヘルス&ビューティー ...

17 Photoshop Elements その場で割引。4月30日までのキャンペーン期間中、対象の『Photoshop Elements』が、その場で1000~4000円割引。さらに対象商品を購入された方の中から、「アマゾンギフト券1万円分」を抽選で20名様にプレゼント。

18 www.amazon.co.jp/-/113k - キャッシュ - 関連ページ

19 関連検索: [青いケシ 写真](#), [青海省 青いケシ](#), [青いケシ ホリデユラ](#), [青いケシ 植物園](#), [ヒマラヤ 青いケシ 登山家](#), [血ガ嶺 青いケシ](#)

20 1 2 3 4 5 6 7 8 9 10 次へ

21 検索: 完了

22

# 分析ツール

行動発話プロトコル分析

ファイル(F) ヘルプ(H)

Yahoo! JAPAN 検索

ウェブ検索結果 (検索条件の見直し)

1. 三国志 三国志に関する本、雑誌をチェック。 - Yahoo!オークション  
 絵巻光輝の「三国志」を愛む、無料での立ち読みも可能。 - Yahoo!コミック  
 三国志のゲームソフト情報をチェック。 - Yahoo!ゲーム

2. 三国志 三国志 三国志に関する本、雑誌をチェック。 - Yahoo!オークション  
 絵巻光輝の「三国志」を愛む、無料での立ち読みも可能。 - Yahoo!コミック  
 三国志のゲームソフト情報をチェック。 - Yahoo!ゲーム

3. 三国志 三国志に関する本、雑誌をチェック。 - Yahoo!オークション  
 絵巻光輝の「三国志」を愛む、無料での立ち読みも可能。 - Yahoo!コミック  
 三国志のゲームソフト情報をチェック。 - Yahoo!ゲーム

4. 三国志 三国志に関する本、雑誌をチェック。 - Yahoo!オークション  
 絵巻光輝の「三国志」を愛む、無料での立ち読みも可能。 - Yahoo!コミック  
 三国志のゲームソフト情報をチェック。 - Yahoo!ゲーム

一時停止 01:07

X:\cores2007\_exp\sub\sub\_11\sub\_11\_0003.avi 26.9246575 full

対象

一覧ページ  
 特定のページ  
 一覧ページの視線

リセット

行動

search  close  
 link  jump  
 next  change  
 return  browse  
 system  submit  
 focus  bookmark  end

リセット

注視領域(データは備考1に記録されます)

タイトルバー  検索バー  検索ボタン  タイトル  ページ内検索  
 メニュー  検索バーボタン  スクロールバー  スニペット  ステータスバー  
 ブックマーク  タブ  ヒット件数  URL  その他  
 ツールバー  サービスリンク  スペルチェック  関連検索  \*\*\*\*\*  
 URLバー  クエリボックス  スポンサリング  次へ

リセット

ビデオ同期  関連情報表示

Position	対象	対象_詳細1	対象_詳細2	対象_詳細3	対象_詳細4	対象_詳細5	対象_詳細6	行動	行動_詳細1	行動_詳細2	行動_詳細3	行動_詳細4	行動_詳細5
1	28.405948							focus					
2	32.805948							jump					book.mark
3	34.305948	特定のページ			Yahoo! JAPAN	http://yahoo.co.jp	-1						book.mark
4	50.2909796							focus					search
5	66.102948							search	Yahoo	三国志			
7	67.102949	一覧ページの視線	1	8,150,000	Yahoo!検索 - 三国志	http://search.yal	1	0	サービスリンク	Yカテゴリ	75,162		
	67.602949	一覧ページの視線							サービスリンク	Yオークション	120,183		
9	68.102949	一覧ページの視線							その他		322,100		
10	68.602949	一覧ページの視線							関連検索		277,88		
11	69.102949	一覧ページの視線							その他		332,108		

一時停止 01:08

X:\cores2007\_exp\sub\sub\_11\m2u00210.avi 20.5757935 full

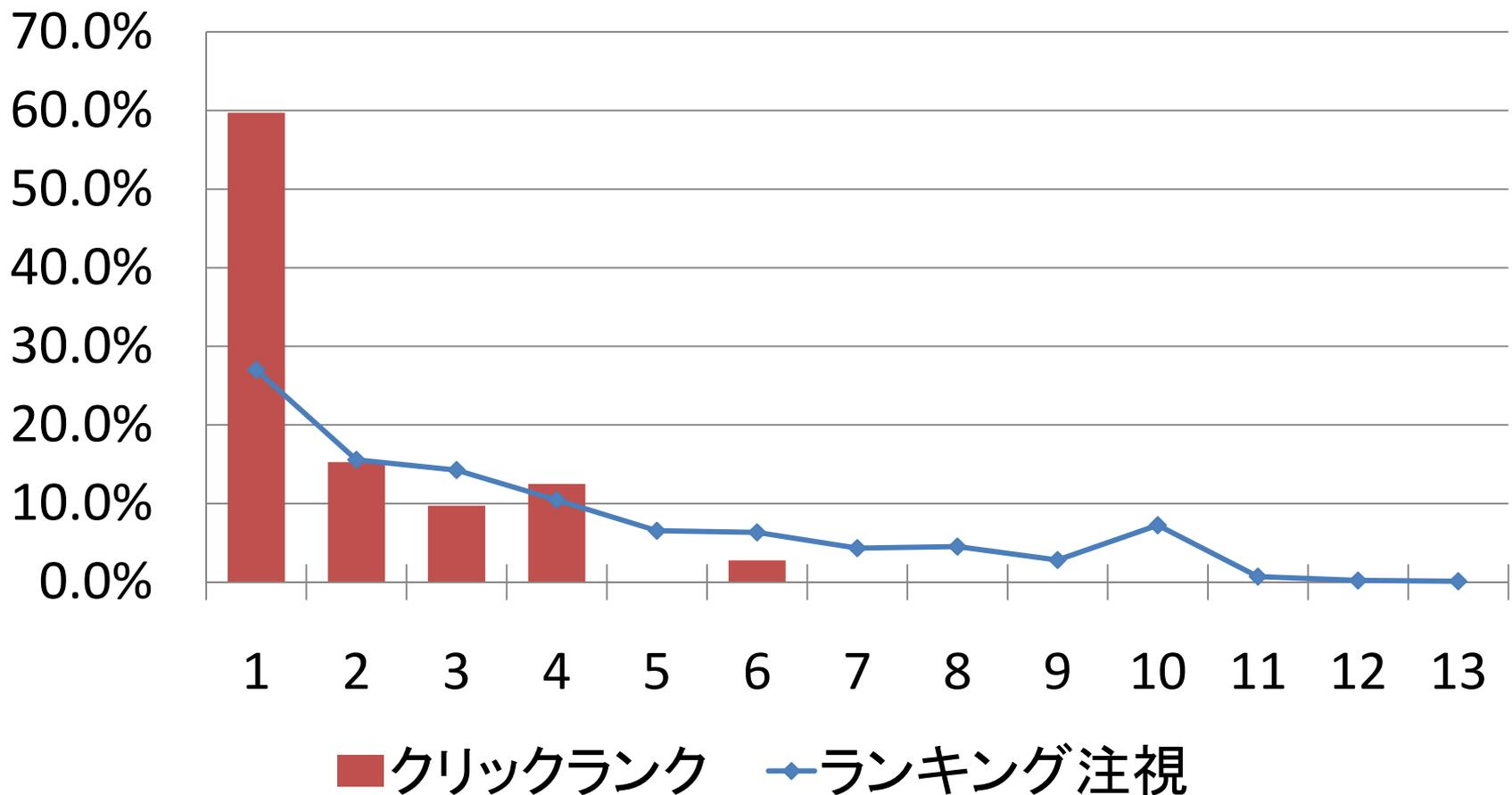
**タグ付け結果**

注視対象	Sub1		Sub7		Sub10		Sub11	
	Info	Trans	Info	Trans	Info	Trans	Info	Trans
その他	18	44	1	12	17	28	72	24
クエリボックス	6	17	0	13	13	20	58	23
スクロールバー	1	0	0	0	0	0	0	0
サービスリンク	0	0	0	0	7	4	18	29
ステータスバー	1	0	0	0	0	0	1	0
タイトルバー	2	5	0	0	0	0	0	0
スポンサーリンク	0	9	0	6	12	28	1	7
タブ	5	9	0	3	31	59	6	5
ヒット件数	0	0	0	0	3	0	1	0
ツールバーボタン	1	4	0	2	4	0	0	0
メニュー	0	1	0	0	0	0	0	0
関連検索	1	12	0	8	5	1	10	1
*****	44	348	1	34	33	25	23	38
次へ	0	1	0	0	0	1	0	2
URLバー	2	0	0	1	2	0	3	0
検索ボタン	0	2	0	0	0	0	2	2
スニペット	70	37	0	45	98	96	80	23
タイトル	36	30	3	56	43	82	61	61
URL	25	21	1	17	27	41	25	18

ヒット件数など、ほとんど注視を集めない領域も

ランキング部分に集中

# ランキング部分の注視とクリックランク



# Scanpath

- ランキング部のみに着目
- 各ランク間の遷移を分析

– (図2の場合)

• Rank2 – Rank1 – Rank1 – Rank2 – Rank3 – Rank4 – Rank5 – Rank1

= 2-1-1-2-3-4-5-1

- 圧縮Scanpath

= 2-1-2-3-4-5-1

- Scanpath長

– 各Scanpathに含まれる注視点の数

– 上の例では、

- Scanpath長 = 8
- 圧縮Scanpath長 = 7

図2

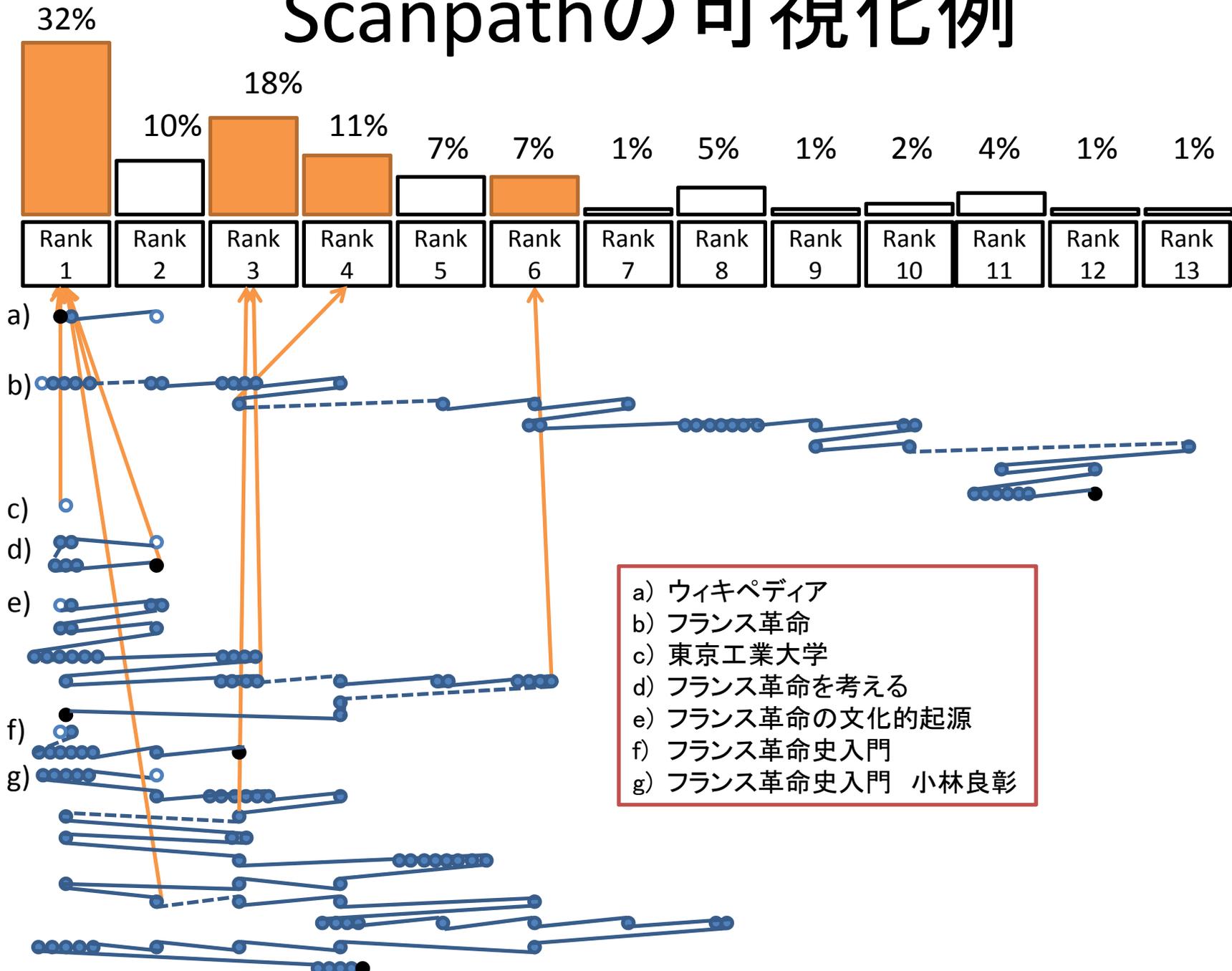


# Scanpath結果

- Scanpath長：
  - 平均16.0, 中央値13.0, 標準偏差12.7
- 圧縮Scanpath長
  - 平均3.6, 中央値2.5, 標準偏差3.7
- 先行研究(Lorigo et al., 2008)
  - Google検索結果におけるScanpath長：
  - 「平均16.0; 圧縮平均3.2」

類似する結果

# Scanpathの可視化例



# まとめ

- Web情報探索実験の視線分析
  - 探索行動全体での注視領域(分析1)
  - サーチエンジン検索結果ページでの注視(分析2)
    - 領域毎の分析、ランク単位Scanpath分析とその可視化
- 今後の課題
  - 検索結果ページ以外のページにおける視線分析
  - 発話データやインタビューデータとの統合的分析
  - 検索結果評価や情報獲得の過程のより精緻な分析