

視線情報による母語・第二言語の推定—多言語表示を用いて

王 琛[†] 浅井 洋樹[†]
[†]早稲田大学大学院基幹理工学研究科

山名 早人[‡]
[‡]早稲田大学理工学術院

1. はじめに

Google Glassをはじめとしたインテリジェント技術の進化、普及とともに、視線情報を利用した研究が数多く行われるようになってきた[1]。2020年の東京オリンピック開催を控え、空港や街中でデジタルサイネージに代表される多言語表示が増えている。しかし、こうしたデバイスでは、ユーザの明示的な表示言語選択無しに、ユーザ自身の母語を表示することはできない。ユーザが当該画面を瞬間的に見るだけでユーザの母語を自動検知できれば、非侵襲での自動言語選択を実現することができ、その応用は格段に広がると思われる。本稿では、上記を実現するにあたり、視線測定を用いた母語・第二言語の判定可能性についての検討結果を報告する。

2. 視線測定

本研究では、被験者に対する負担を無くすため、ウェアラブルな視線測定装置ではなく、設置式の装置 Eye Tribe¹を使用する。実験では、ディスプレイ上に複数言語で記述された文字を「全文表示」、「スクロール表示」、「言語毎の表示位置変更」など異なる表示方法で複数回表示し、被験者の視線を時系列データとして取得する。「国旗の表示」、「一単語のみの表示」についても検証する。また、Juddら[2]による「人間は画面の中央を見ることが多い」という知見も考慮し、画面中央への配置が言語推定に影響を及ぼすかについても調査する。

3. 理解できる言語の推定

人はどこかに注視する時、たまにちらりと注視点外に視線を移すことがあるが、このような不安定な視線データは採用しない。次に、被験者がセンテンスを読んでいるかどうかを判定する。センテンスの途中までしか読まない場合、被験者は当該センテンスの記述言語を理解しているわけではなく、好奇心で見ている可能性が高い。以上を実現するため、時系列の視線データから各単語に対する停留時間を求める(図1)。そして、注視単語が予め用意されている「各センテンス内の重要単語(キーワード)」と一致するかどうかにより、被験者の当該言語の理解度を推定する。一致度合いが高いほど、理解度が高いと推定する。

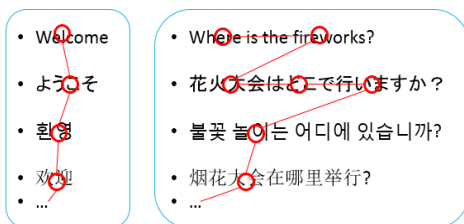


図1 停留注視点の例(赤丸:注視点)

具体的には、視線情報としての時系列視点座標を一定時間間隔で切り分け、言語別に注視の時間割合を算出する。図2は注視時間のタイムラインの例(被験者1名)を表し、図3は注視時間割合の例を示している。画面全体をスキャンする視線を含め、注視時間割合の第一ピーク、すなわち最初に一番長く停留する言語が母国語、第二ピークが第二言語、第三ピークが第三言語になる可能性が高いことがわかる。この方針に基づき、被験者3人(国籍:日本, 韓国, 中国)の視線情報から表示方法ごとに推定したところ図4の結果が得られた。

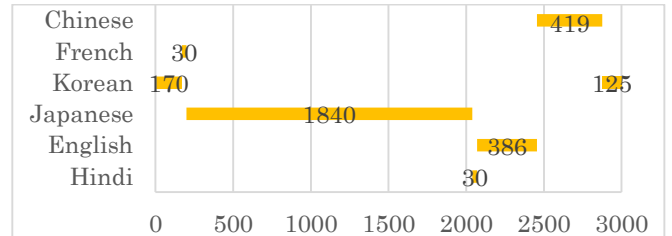


図2 注視時間タイムラインの例 [ms] (日本人)

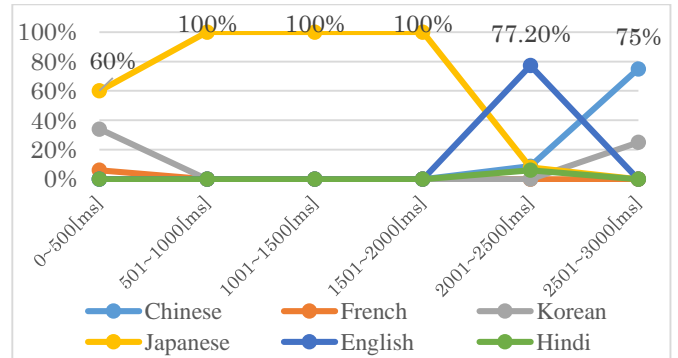


図3 注視時間割合の例 (日本人)

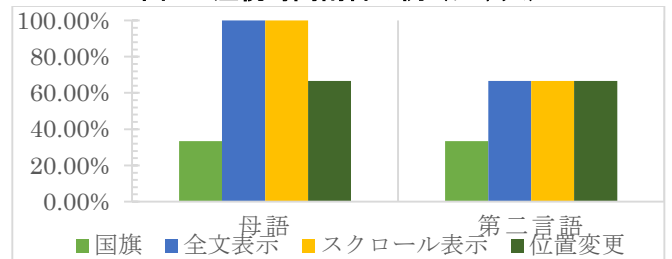


図4 表示方法ごとの推定正解率

4. 今後の課題

特徴量として、注視時間割合以外に、他の特徴量を用いて精度向上を行う予定である。

参考文献

[1] 大野健彦. “視線から何がわかるか 視線測定に基づく高次認知処理の解明.” *認知科学*, Vol. 9, No.4, pp.565-579, 2002.
 [2] Judd, Tilke, et al. “Learning to predict where humans look.” *12th Int’l Conf. on Computer Vision*, 2009.

¹ <https://theyetribe.com/>