

科学研究費助成事業 研究成果報告書

平成 29 年 6 月 20 日現在

機関番号：32630

研究種目：若手研究(B)

研究期間：2014～2016

課題番号：26770134

研究課題名(和文)日本語文処理における眼球運動計測を用いた実験手法の確立

研究課題名(英文) Establishing the experimental method for eye-tracking research on Japanese sentence processing

研究代表者

新井 学 (Arai, Manabu)

成城大学・経済学部・専任講師

研究者番号：20568860

交付決定額(研究期間全体)：(直接経費) 2,200,000円

研究成果の概要(和文)：本研究では眼球運動測定による日本語理解調査を行い、「予測に基づく文処理モデル」の動詞末位型言語における妥当性を検証すること、そしてその研究成果をもとに眼球運動測定を用いた日本語理解研究の手法を体系的にまとめることを目標とした。研究成果として、日本語では格情報を用いた漸進的処理の中で構造的出現頻度と同時に曖昧性が解消されるまでの時間の長さによって予測的処理が行われるという貴重な知見が得られた。

研究成果の概要(英文)：This research examined the validity of the expectation-based model in the processing of Japanese sentences by using eye-tracking technique. The results from multiple experiments revealed that Japanese readers process sentences incrementally just as English readers but importantly, the frequencies of structural configurations accompanied with case markers and the timing for the arrival of disambiguating information influences the estimation of probabilities for possible structural options. I also published a paper that discusses standard experimental design and reasonable statistical analysis for eye-tracking research on Japanese sentence processing.

研究分野：言語心理学

キーワード：眼球運動 日本語理解 関係節文 予測 動詞末位型言語 統計分析

1. 研究開始当初の背景

眼球運動計測を用いて主部後置型言語である日本語を調査した研究は未だ少なく、その一例として、Mazuka et al. (1997) は日本語においても英語と同様、一時的曖昧性を伴う文を読む際に曖昧性が解消されるリージョンで眼球の停留時間の増加が観測され、誤分析から再解釈を構築する際にかかる処理負荷を量的データによって実証した。

これら過去研究によって文理解における特定の単語の停留時間は前出の情報に影響を受けることが示されており、これは、文理解において個々の語を処理することに次に起こりうる可能な構造的分析の確率を計算するという、文処理研究において近年注目を浴びている「予測に基づく文処理モデル」(Levy, 2006; Hale, 2001)の考え方と一致する。

この文処理モデルではひとつの語を処理する毎に次に来る情報の可能性が制限されると考えられており、その計算された確率に沿って処理の容易さが算出される。これによれば、日本語のような主部後置型言語で「犬が男を噛んだ」という文を理解する際には、動詞情報を得る前に「男」という名詞句を処理するため、"The dog bit the man"のような主部先行型である英語での動詞部での処理負荷とは異なると予想され、その負荷の差が注視の停留時間に反映されると考えられる。しかしこのような仮説を検証した研究は未だ少なく、眼球運動測定を用いた研究はほとんど存在していなかった。

また、眼球運動測定研究は機器の低価格化と共に急速に普及し、現在では世界の多くの研究者・研究室で眼球運動測定研究が行われている。しかしこのように量的データの収集が容易になったのとは対照的に、そこで得られるデータ量は機器の性能向上と共に増大し、それらデータの分析手法は統計解析理論の前進と、様々な分析ツールの開発と共に困難さを増していた。それゆえ眼球運動データの分析に関して研究者間で共有されるスタンダードな手法を確立する必要性が本研究の背景としてあった。

2. 研究の目的

本研究の主な目的の一つは眼球運動測定を用いて実時間上の日本語文理解プロセスを研究し、上記の「予測に基づく文処理モデル」の動詞末位型言語における妥当性を検証することである。

そしてもう一つは日本語文理解における眼球運動測定の実験手法を確立することである。具体的には、様々な構造の日本語文理解における眼球運動を測定し、実験手法・解析方法の解説を含めた論文としてまとめ、今後の日本語文処理研究における眼球運動測定調査の普及に貢献することを目標とした。

3. 研究の方法

本研究で行った眼球運動計測実験の基本的な方法は以下の通りである。

被験者は画面上に提示された文、たとえば(1)のような関係節文を自然なペースで読む際の文字情報上の眼球運動を計測した。

(1) 刑事を呼び出した担当者はとても慌てていた。

測定機から得られる眼球運動データは、各データポイントに記録された時間情報(計測機器のサンプリングレートに依存)と画面解像度に従ったX-Y座標値の位置情報の二つである。これを分析するため、提示文を任意の位置で区切り、複数のリージョンに分割し(この例では「刑事を/呼び出した/担当者は/とても/慌てていた。」のように句ごとに分ける)、各リージョン内で記録された停留(fixation)を元に first pass 等眼球運動計測データ解析に用いられる複数の指標の読み時間の分析を行う。統計検定では観測された停留時間を従属変数とし、「線形混合効果モデル」と呼ばれる被験者における個体差と言語刺激における項目差の二つのランダム要因を同時に説明する統計モデルを用いた分析を行った。

4. 研究成果

4.1.

本研究の理論的骨組みである「予測に基づく文処理モデル」の動詞末位型言語における妥当性を検証するための眼球運動測定研究の成果は以下の通りである。

4.1.1. 関係節構造を使った Digging-in 効果の検証

Digging-in 効果と呼ばれる構造的曖昧性をもつ関係節の長さによる処理コストが、頻度情報に基づく予測に起因するか、あるいは構造分析における統語的結びつきの時間的変化によるのか眼球運動測定を用いて検討した。具体的には今までの研究においては、この digging-in 効果を反映しているとされる処理負荷が、漸進的分析へのコミットメントの掘り下げに起因するのではなく、実際には文中の曖昧箇所を追加された語句による余分な処理コストが原因かもしれないという代替可能性を否定することができなかった。この問題は文の中で付加詞の位置が自由ではない英語を対象とする研究では解決することが不可能であった。そのためこの研究では、付加詞の位置が比較的自由にこの検証が可能である日本語を用い眼球運動測定実験を行った。その結果、この掘り下げ効果は付加詞の追加処理コストによって生じるのではなく、特定の分析に対してコミットした時間的(また距離的)長さによって生じていることが実証された(つまり「掘り下げ効果」の解釈を支持した)。この研究のもう一つの重要な点として、比較された異なる位置に付

加詞を持つ文の間には出現頻度の差が見られなかったことが挙げられる。これは頻度情報に基づいて特定の言語情報の出現予測確率を計算し、その確率と実際のインプットのズレ(「予測エラー」と呼ばれる)から処理負荷を見積もる「予測に基づく文処理モデル」の予測が当てはまらないことを意味し、代わりに漸進的分析に対するコミットメントの度合いは動的な自己組織化プロセス(dynamical, self-organizing processes)によって、構造的曖昧性が継続する時間と共に強化されるとするモデルを支持した。眼球運動測定実験の結果を報告した論文を執筆・投稿し、国際学術雑誌『PLOS ONE』に投稿し、2016年に採択・出版された。

4.1.2. 主語関係節文と目的語関係節文における処理負荷のアシメトリーの研究

日本語関係節文の理解について眼球運動測定を用いた実験を複数行い、動詞末位型言語である日本語では、主語関係節の処理的容易さを作業記憶の負荷ではなく頻度情報に基づく統語構造の確率的予測によって説明できることを明らかにした。

具体的には(2a)のような主語関係節は(2b)のような目的語関係節よりも処理が楽であることが過去の研究からわかっているが、その理由として(2a)のように目的格が文頭に現れ次に動詞が続くような文では主文構造を取る確率が低く、関係節構造である確率が高い。それゆえ、関係節の主部が現れる前に関係節の予測が起こっている可能性が考えられ、(2b)のような通常主節構文として理解され関係節主部(「担当者は」)が現れた時点で初めて関係節であることが明らかになる文に比べて処理コストが低いことがわかった。また、(2a,b)で関係節主部にかかる修飾句を操作することで実際に主部の予測が起きているか検証した。この結果主部の予測が起こる目的語関係節文(2a)で無生名詞の所有格(「指紋鑑定の」)を見た時のみ処理困難が観測された。

(2a) 特捜部の刑事を携帯電話でこっそりと呼び出した(物静かな/指紋鑑定の)担当者はとても慌てていた。

(2b) 特捜部の刑事が携帯電話でこっそりと呼び出した(物静かな/指紋鑑定の)担当者はとても慌てていた。

この結果によって今まで議論されてきた構造的依存関係を構築する情報間の距離及び記憶の影響とは別の、頻度に基づく予測処理によって主語関係節の選好性を説明できるという重要な研究結果が得られた(現在この研究成果をまとめた論文を執筆しており、今年の夏頃に言語心理学の分野で最も権威ある国際学術雑誌『Journal of Memory and Language』に投稿する予定である)。

4.2. 眼球運動測定を用いた日本語理解研究の実験およびデータ分析手法の体系化

この目標のため、2015年9月に国立国語研究所にて行われたシンポジウム「統計的言語研究の現在」にて講演し、その後その内容をもとに統計数理研究所が発行する『統計数理』に学術論文「言語理解研究における眼球運動データ及び読み時間データの統計分析」(Douglas Roland 博士との共著)を投稿し、査読者から頂いた多くの有益なコメントをもとに大幅な修正を加えた後無事採択され、2016年12月に出版された。

この論文では、言語理解研究における眼球運動測定実験、その中でも視覚世界実験と読み実験によるデータ、さらに自己ペース読み課題を用いた読み時間のデータに対して、今現在広く利用されている分析方法を比較し、その有効性について検討した。また、主に線形混合モデル及び一般化線形混合モデルを用いたデータ解析手法について基礎的な概念から具体的な適用方法まで紹介し、これらのモデルを慎重に且つ論理的な手順に従い適用することは今までのデータ集約を必要とした分散分析などの手法と比べて多くの利点があることを主張した。また、未だコンセンサスの得られていない様々な問題、具体的には外れ値の処理の方法、一般化線形混合モデルを使う際の応答変数の分布ファミリーの選択およびリンク関数の指定、また最適モデルの選択方法など、分析結果を左右する重要な項目についても扱った。この論文はこれから自分のデータ分析に線形混合モデル等の分析を応用したい、または既に使っているがどのように適用すれば良いのか十分に理解していない研究者にとって分析方法のマニュアルとしても利用できる有益なものとなった。

5. 主な発表論文等

〔雑誌論文〕(計 4 件)

Yamada, T., Arai, M., & Hirose, Y. (2017). Unforced revision in processing relative clause association ambiguity in Japanese: Evidence against Revision as Last Resort, *Journal of Psycholinguistic Research*, pp.661-714. (査読有) doi:10.1007/s10936-016-9457-8

新井学, Douglas Roland, (2016). 言語理解研究における眼球運動データ及び読み時間データの統計分析, 「統計数理」, 64巻、特集「統計的言語研究の現在」, 統計数理研究所, pp.201-231. (査読有)

新井学 (2016). 第二言語の理解と学習における誤りの重要性, 成城大学『経

済研究』第214号, 73-96. (査読無)

Arai, M., & Nakamura, C. (2016). It's harder to break a relationship when you commit long. PLOS ONE. <https://doi.org/10.1371/journal.pone.0156482> (査読有・オープンアクセス)

[学会発表](計 7件)

Arai, M. & Kahraman, B. Distance-independent factor for the processing asymmetry of Japanese subject/object relative clauses, 言語科学会第18回国際年次大会(JSLs 2016), 2016年6月4日, 東京大学(東京).

Arai, M. Importance of making (prediction) errors in L2 processing, 22nd Annual Conference on Architectures and Mechanisms for Language Processing, 2016年9月1日-3日, Bizkaia Aretoa, ビルバオ(スペイン).

Arai, M., Chang, F., & Ishikawa, M. Do word-based expectations influence structural priming? 22nd Annual Conference on Architectures and Mechanisms for Language Processing, 2016年9月1日-3日, Bizkaia Aretoa, ビルバオ(スペイン).

Arai, M., Nakamura, C., & Hirose, Y. Priming from a previously predicted but unseen verb, 22nd Annual Conference on Architectures and Mechanisms for Language Processing, 2016年9月1日-3日, Bizkaia Aretoa, ビルバオ(スペイン).

Uchida, S., Arai, M., Miyamoto, E. T., & Hirose, Y. Processing Japanese sentences as a zero-sum game, 言語科学会第17回国際年次大会(JSLs 2015), 2015年7月18日-19日, 別府国際コンベンションセンター(九州).

Uchida, S., Arai, M., Miyamoto, E. T., & Hirose, Y. Prune early or prune late? Surprisal will cost you either way, Annual CUNY conference on Human Sentence Processing, 2015年3月19日-21日, ロサンジェルス(アメリカ).

Nakamura, C., Arai, M., & Hirose, Y. What is helpful for native speakers can be misleading for L2 learners: Evidence for misinterpretation of contrastive prosody, Annual CUNY conference on Human Sentence Processing, 2015年3月

19日-21日, ロサンジェルス(アメリカ).

[図書](計 0件)

[産業財産権]

出願状況(計 0件)

名称:
発明者:
権利者:
種類:
番号:
出願年月日:
国内外の別:

取得状況(計 0件)

名称:
発明者:
権利者:
種類:
番号:
取得年月日:
国内外の別:

[その他]
ホームページ等 なし

6. 研究組織

(1) 研究代表者
新井 学 (Arai, Manabu)
成城大学・経済学部・専任講師
研究者番号: 20568860

(2) 研究分担者
なし

(3) 連携研究者
なし

(4) 研究協力者
なし