

山内 暉

211X122X

2021年6月13日

どのひらがなが強く, 弱いのか

1.Introduction

日頃から, 友達としりとりをして時間を潰すことが多い.そこで友達は「り」で始まる言葉で攻めてくるのだが, その根拠というものはあるのだろうか?と感じた.データを実際に分析してみると「ら」で攻めるというのは非合理的かもしれないし, 実際にはもっと良い方法があるのかもしれない.また, なるべくしりとりを高速で長続きさせるゲームをすることもあるのだが, その場合では一般的に弱いとされるひらがなで攻めるのがラリーを続けるコツとなる.したがって, しりとりにおいてひらがなの強弱というものを考察するのは面白いトピックだと感じた.最終レポートではそれを解き明かすべく, データの可視化に取り組もうと考えた.

2.Method

まず, 前処理として, 辞書に含まれている名詞から「どのひらがなが何個頭文字になっているか」と「どのひらがなが何個最後尾になっているか」をカウントする.そのために『分類語彙表増補改訂版データベース』[1]を利用する.そのデータをpythonで読み込み, 分類が名詞になっているもののみを抽出し, それらの全てから頭文字と最後尾をカウントした.なお, 普段濁点を取っても良いルールでやっている為, 濁点の有無は関係なくデータを取る.そのデータを書き出し, 前処理は終了となる.それらのデータをバーチャートで表示し, 各ひらがなごとに頭文字と最後尾になった数をビジュアライズすることで, どのひらがなが強く, どのひらがなが弱いのかを考察する.また, しりとりということで比較的軽いテーマであるので, ワードクラウドを表示す

ることで、厳密ではないものの興味の引きやすいチャートを表示し、視覚的にこの調査の面白さを訴える。

3.Result

バーチャートの表示結果は図1のようになった。

また、ワードクラウドでの表示結果は頭文字になっている場合が図2, 最後尾の場合は図3のようになった。

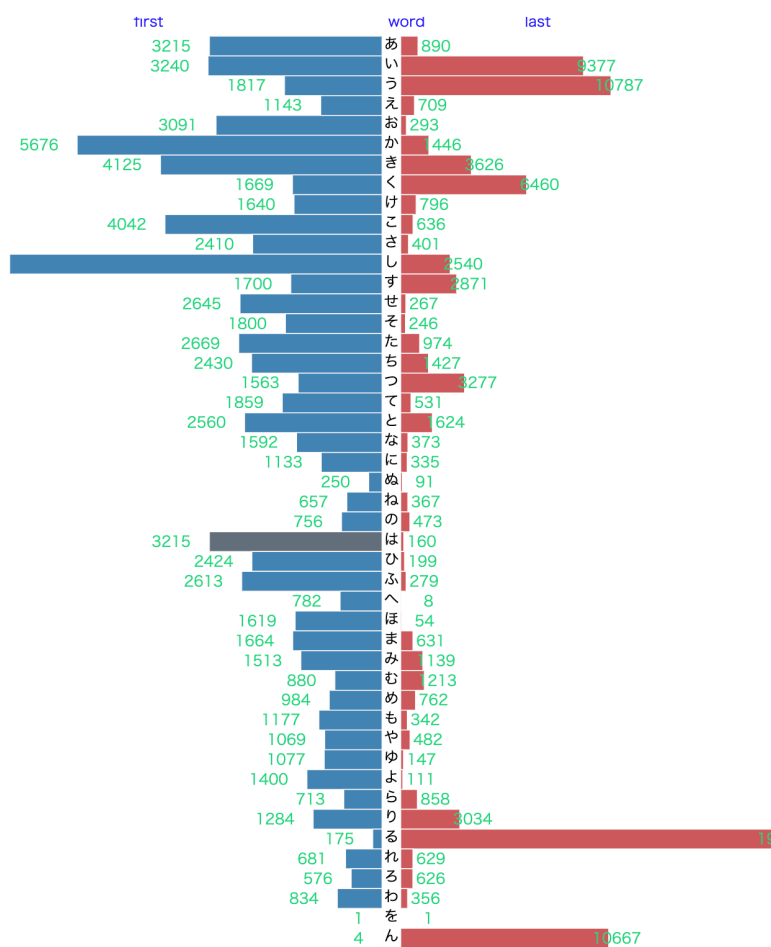


図1: ひらがなが出現した数

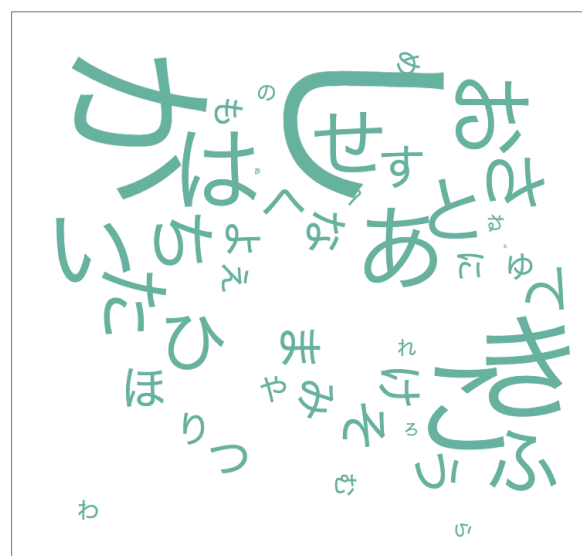


図2: 頭文字のワードクラウド

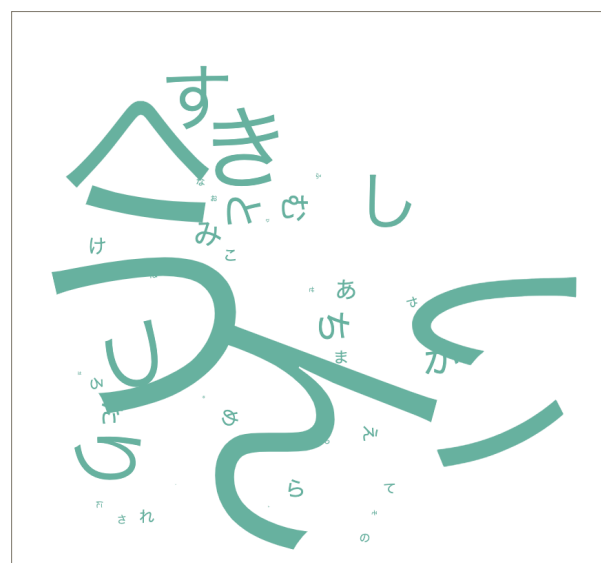


図3: 最後尾のワードクラウド

4. Discussion

まず、強いひらがなとはなにだろうか？しりとりルールは「相手が言葉が出てこなくなったら勝ち」または「相手が最後に”ん”がついたら勝ち」を満たせばいい訳だが、基本的にしり通りの終わりは前者となるため、前者を満たすひらがなを考える。前者を満たすようなひらがなに要求される条件は「頭文字になることが少ない代わりに、最後尾になることが多い」ひらがなであると考えられる。なぜなら自らが平仮名を決めて攻める場合だと自らが最後尾の単語は調整でき、相手はそれを頭文字として受け取るためである。図1を見てみると、この条件を満たすのは「る」であることが確認できる。

逆に弱いひらがなとはなにだろうか？先程の定義の逆が弱いひらがなであるので、それを考えると「頭文字になることが多く、代わりに最後尾になることが少ない」ひらがなであるがそれは「し」であることが確認できる。また、「は」「か」も弱いと言えるだろう。

強弱とは関係なく図1, 2, 3をみて言えることであるが、一般的な語彙では頭文字になるひらがなは比較的ばらけているが、最後尾になる単語は非常に偏りがあり、「ん」で終わる単語は非常に多いということがわかる。したがってしりとりはデータ分析のない時代に生まれたゲームだが、しり通りの「ん」で終わったら負けというルールはデータ的に見ても勝負の良いアクセントになっているということがわかった。しかし、「る」は強い単語であるためここを最後に出すとルールに抵触するようになっているのも面白いかもしれない。

5. Conclusion

しりとりで強いひらがなは「る」であり、弱いひらがなは「し」「は」「か」であることがわかった。また、単語の頭文字はさまざまなひらがながあるが、最後尾にくる単語は非常に偏りがあることがわかり、日本語の面白さを認識することとなった。

5. Reference

[1] 「分類語彙表増補改訂版データベース」, 2021/06/01, <https://ccd.ninjal.ac.jp/goihyo.html>

