

# 12

## あなたの語彙数測ります

### 令和版語彙数推定テスト

#### どんな研究

単語親密度とは、語のなじみ深さを評定実験によって数値化したものです。NTTが20年以上前に調査した約7万7千語の単語親密度は、語彙数推定テストをはじめ、今でも多くの分野で活用されています。我々は経年変化の調査や、新語追加のため、**令和版単語親密度調査**に取り組みました。

#### どこが凄い

単語親密度の調査を**16万3千語**という類を見ない規模で実施しました。クラウドソーシングを用いながら、スクリーニングを丁寧に実施することで、信頼度の高い調査結果を得ることに成功しました。これにより、20年前との比較が可能になりました。これを基に、令和版語彙数推定テストを作成しました。

#### めざす未来

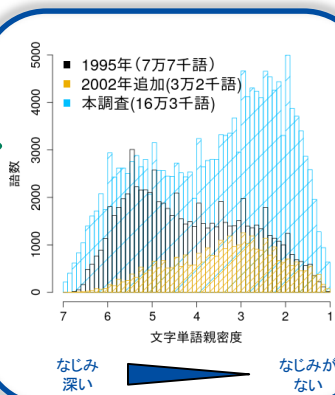
本調査結果を用いることで、**現代に適した語彙数の推定**が簡単に行えるようになります。小学生～高校生を含む幅広い年代の語彙数調査に取り組んでいます。今後、語彙数と読解力や学力全般との関係の調査・分析を実施、**効果的な学習支援の実現**をめざします。

#### 単語親密度クラウド調査

良好な評価者に  
繰り返し評定を依頼

#### 令和版 単語親密度データベース

過去データとの度数分布比較



評価者

事前スクリーニング

漢検2級以上  
or  
漢字の読みテストの  
成績上位

親密度評定

1セット1000～2000語  
単位で評価  
(5%の語は重複評価)

評価者

事後スクリーニング

重複評価の結果、  
相関係数が 0.5 以上

ポイント①

漢字の知識があり、  
評価が一貫している  
評価者に繰り返し依頼

ポイント②

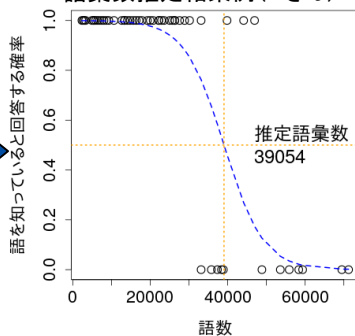
新語を含め、16万3千語調査  
例) スマートフォン、コンビニ、  
ヘテロダイマー、臘八、…

#### 語彙数推定方法

知っている語をチェックしてください。

- 銀行
- 担当
- ...
- 親疎
- 改鑄

語彙数推定結果例(Aさん)



ポイント③

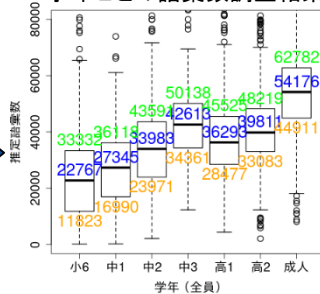
50語程度の簡単チェック

語彙数推定テストはこちらから:

<http://www.kecl.ntt.co.jp/icl/lirg/resources/goitokusei/>



学年ごとの語彙数調査結果



ポイント④

小・中・高・成人 約4600人で  
語彙数調査を実施

#### 関連文献

- [1] S. Fujita, T. Kobayashi, “単語親密度の再調査と過去のデータとの比較,” 言語処理学会第26回年次大会, pp.1037–1040, 2020.
- [2] S. Fujita, T. Kobayashi, T. Yamada, S. Sugawara, T. Arai, N. Arai, “小・中・高校生の語彙数調査および単語親密度との関係分析,” 言語処理学会第26回年次大会, pp.355–358, 2020.

#### 連絡先

藤田 早苗 (Sanae Fujita) 協創情報研究部 言語知能研究グループ  
Email: cs-openhouse-ml@hco.ntt.co.jp

