

1. 背景と目的

- ◆ 日本語指導が必要な児童生徒が増加 (2018年に5万人超, 10年で1.5倍)
- ◆ 児童向け日本語教育・教材作成が急務
- ◆ 小学校低学年の教室活動の話し言葉で用いられる語彙は？
 - ・ 教育現場でのデータ収集は困難

作業仮説: 「教育番組はその対象学年の日本語母語話者が理解できる語や表現で構成されている」



小学生低学年向け教育番組の語彙を調査

2. 対象と方法

対象

- ・ NHK for School (<https://www.nhk.or.jp/school/>) にて公開されている過去にEテレで放送された教育番組
- ・ 小学3年生以下を対象とする番組
- ・ 2022年度上半期の段階で公開されている番組, 放送回すべて

表1: 分析対象の番組と放送回数, 1話あたりの放送時間

番組名	対象学年	科目	放送回数	時間 (分)
えいごでがんこちゃん	小学1~2年	特活	20	5
おはなしのくに	幼児・小学1~3年	国語	17	10
おぼけの学校たんけんたん	幼児・小学1~2年	生活	16	10
銀河銭湯パンタくん	小学1~2年	道徳	38	10
ことばドリル	小学1~2年	国語	20	10
こどもにんぎょう劇場セレクション	幼児・小学1~2年	その他	10	15
コノマチ☆リサーチ	小学3年	社会	20	10
さんすう犬ワン	小学1~3年	算数	19	10
しぜんとあそぼセレクション	幼児・小学1年	その他	25	15
新・ざわざわ森のがんこちゃん	幼児・小学1~2年	道徳	28	10
すたあと	幼児・小学1年	生活	20	5
で〜きた	幼児・小学1年	特活	20	10
ふしぎエンドレス理科3年	小学3年	理科	20	10
ふしぎがいっぱい3年	小学3年	理科	20	10
理科3年ふしぎだいすき	小学3年	理科	19	15

※番組数に科目間の偏りあり

データの総時間数: 53時間10分

書き起こし作業

- ・ 番組動画を見て日本語を母語とする大学生が手作業にて書き起こし
- ・ できるだけ漢字を用いて書くよう指示
- ・ 著者の1人が全データを点検し, 修正

単語出現頻度の計測

- ・ MeCab+IPA辞書にて形態素解析
 - ・ キャラクタ名, 話し言葉特有の語をユーザ辞書に追加
 - ・ ユーザ辞書に追加した語の半数以上は感動詞
- ・ 表記上の変化や音声変化は「変形」として扱い, 対応する見出し語に集約
 - ・ 例: 「アイデア」・「アイディア」, 「言う」・「ゆう」
- ・ フィラー, 固有名詞, 数詞, 単一の番組にしか出現しない語を除外し, 単語出現頻度を計測

番組ごとの放送時間の差の補正

- ・ 時間正規化単語出現頻度にてランキング
- ・ 旧日本語能力試験の出題基準と比較

$$f'_{ij} = f_{ij} \times \frac{d_i}{D}$$

番組iの放送時間
総放送時間
番組iにおける語jの出現頻度

3. 結果

抽出された語の数

- ・ 延べ語数: 228,012語
- ・ 異なり語数: 3,463語

出題基準との比較

- ・ N4 (最も易しい) に含まれる語のほとんどが含まれている
- ・ 本リストに含まれないN4の語の例
 - ・ 「月曜」, 「おととい」, 「階段」, 「午前」, 「曇り」など

表2: 品詞の異なり語数と延べ語数

品詞	異なり語数	延べ語数
名詞	1,913	57,478
動詞	752	35,532
副詞	287	8,964
感動詞	198	12,940
形容詞	136	6,219
助詞	79	71,922
接続詞	48	1,819
助動詞	19	28,062
連体詞	18	2,827
接頭詞	18	1,374

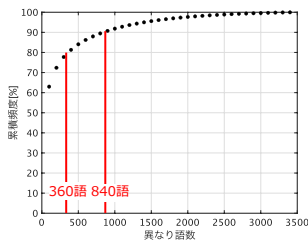


図1: 時間正規化単語出現頻度のカバー率

表4: 上位1,000語の中で能力試験N1および級外の語

品詞	単語
名詞	葉っぱ, お前, おいら, あたし, 幼虫, ドリル, 産む, 俺, 雌, 人造, バッタ, 蝶, 明かり, モンシロチョウ, 奴, ホウセンカ, お宝, 阿婆, ウサギ, お家, 本日, オクラ, 僕ら, あんた, カマキリ, とこ, お湯, お化け, お客, フラッパ, 変身, アリ, おら, カブトムシ, わし, ヒマワリ, 子象, 挑戦, 昆虫, 正解, 尻尾, 茎, 銀河, トノサマバッタ, さなぎ, 美味しい, タイム, 木の葉, ホームページ, モール, カエル, 和尚, どんぐり, クリップ, キャベツ, リス, 鶴, 作戦, 勝手, タスキ, 折り紙, 同士, 草むら, セミ, 砂鉄, ロボット, 給食, ナイス, 作業, 通り道, 粘土, 隊員
動詞	やってくる, 知れる, 気が付く
副詞	なんだか, もぐもぐ, もしも, ごしごし, わくわく, 本当は
接続詞	まずは
連体詞	そういう
形容詞	格好いい, だかい, 気持ちいい
感動詞	こりゃ, オッケー, やあ, ふーん, ガーン, ようこそ
助詞	によって

日本語を母語とする児童であれば知っている語が多い

表3: 時間正規化単語出現頻度の上位30語

No.	見出し語	変形	品詞1	品詞2	多義語	出現頻度
1	の		助詞	格助詞/終助詞/連体化		638.473
2	て	てーっ	助詞	格助詞/接続助詞		622.911
3	だ	じゃろ/じゃん/だあ だーだあーだーい だーだろおだあー だーいだーだろお	助動詞			573.879
4	た	たあーたーたーたり	助動詞			551.149
5	に		助詞	格助詞/副詞化		498.028
6	は		助詞	係助詞		496.908
7	が		助詞	格助詞/接続助詞		441.583
8	を		助詞	格助詞		385.870
9	ない	なーい/せん/ないっ なきや	助動詞			265.519
10	よ	よお/よおー/よーん	助詞	終助詞		261.958
11	する		動詞	自立		243.801
12	と		助詞	格助詞/接続助詞 副詞化/並立助詞		223.834
13	ね	ねえ/ねえ/ねー/ねえー ねん	助詞	終助詞 終助詞		223.235 223.235
14	か		助詞	副助詞/並立助詞/終助詞		214.412
15	ます	まーす	助動詞			211.635
16	ん		名詞	非自立		209.591
17	も		助詞	係助詞		184.514
18	で		助詞	終助詞/接続助詞		182.903
19	いる		動詞	自立	○	180.533
20	です	でーす	助動詞			156.544
21	な		助詞	終助詞		153.212
22	う		助動詞			122.343
23	から	からあ	助詞	終助詞		118.270
24	なる		動詞	自立/非自立	○	103.197
25	何	なあに/なに/なーん	名詞	代名詞		101.080
26	てる		助動詞			98.100
27	事		名詞	一般/接尾/非自立		92.074
28	これ	これっ	名詞	代名詞		88.113
29	お		接頭詞	名詞接続		82.558
30	できる		動詞	自立/非自立		79.870

4. まとめと課題

Eテレの小学校低学年向け番組にて使用される語を調査

- ・ 小学校生活に必要な語が多く含まれる
- ・ 能力試験のレベルとは必ずしも対応しない
- ・ 一部の番組で役割語が多用されているので, 日本語を母語としない児童への配慮が必要

課題

- ・ 対象のデータの量が不十分
- ・ 実際の教室活動の音声調べているわけではない
- ・ 得られたデータをどのようにして実際の学習支援につなげていくのか

リストは以下よりダウンロードできます

<https://www.konan-u.ac.jp/hp/kitlab/etv/>

