

英語検定教科書と学習者のライティングに 出現する語彙について

内田 諭

九州大学言語文化研究院

uchida@flc.kyushu-u.ac.jp



自己紹介

内田 諭 (うちだ さとる)

九州大学大学院言語文化研究院・地球社会統合科学府・共創学部：准教授
専門：認知言語学、コーパス言語学、英語教育学

【近著論文】

Uchida, S. and Negishi, M. (2025). Assigning CEFR-J levels to English learners' writing: An approach using lexical metrics and generative AI. *Research Methods in Applied Linguistics*, 4(2), 100199. <https://doi.org/10.1016/j.rmal.2025.100199>

 CWLA2 Evaluation Results

 Assessment Summary

CEFR-J Level	Total Score
B1.2	55.6

Word Count: 74 | Avg. Sentence Length: 14.8 words/sentence

 Analyzed Text

I agree that parent should limit the amount of time children spend online. First, we may get bad from using the Internet for long time, for example, lack of sleep, bad eyes. Second, we have less communication with family and friends. If you spend a lot on Internet, you will see decrease in your relationship with your friends and your social skills. For these reasons, I am for the limiting time children spend online.

A1 A2 B1 B2 Other content words Misspelled words

調査の目的

【課題】

中高生の**発信語彙**（スピーキング、ライティングで使える単語）を増強したい

【前提】

英語学習において**インプット**（リーディング、リスニング）の重要性は広く認識されており、発信語彙の習得もそれを通して行われると考えられる

【何を見るか】

- 中高生のインプット（教科書）において、**英単語の習得が十分におきるだけの出現**があるか
- 中高生のアウトプット（ライティング）において、**どのような英単語が使われているか**

調査の方法

- 中高の検定教科書における**英単語の種類・頻度**をコーパスから分析する
→ 低頻度の語の割合はどの程度か
- 中高生の英語ライティングにおける**英単語の種類・頻度**をコーパスから分析する
→ 中高生はどのような単語を使っているか

教科書のデータ

【中学校】

- 2025年使用開始の検定教科書6種類 × 3学年
- 各レッスンの本文のデータ（数字・固有名詞を除く34,462語）

【高校】

- 2016年使用開始の英語コミュニケーション6種類 × 3学年
- 各レッスンの本文データ（数字・固有名詞を除く90,192語）

ライティングのデータ

【中学校】

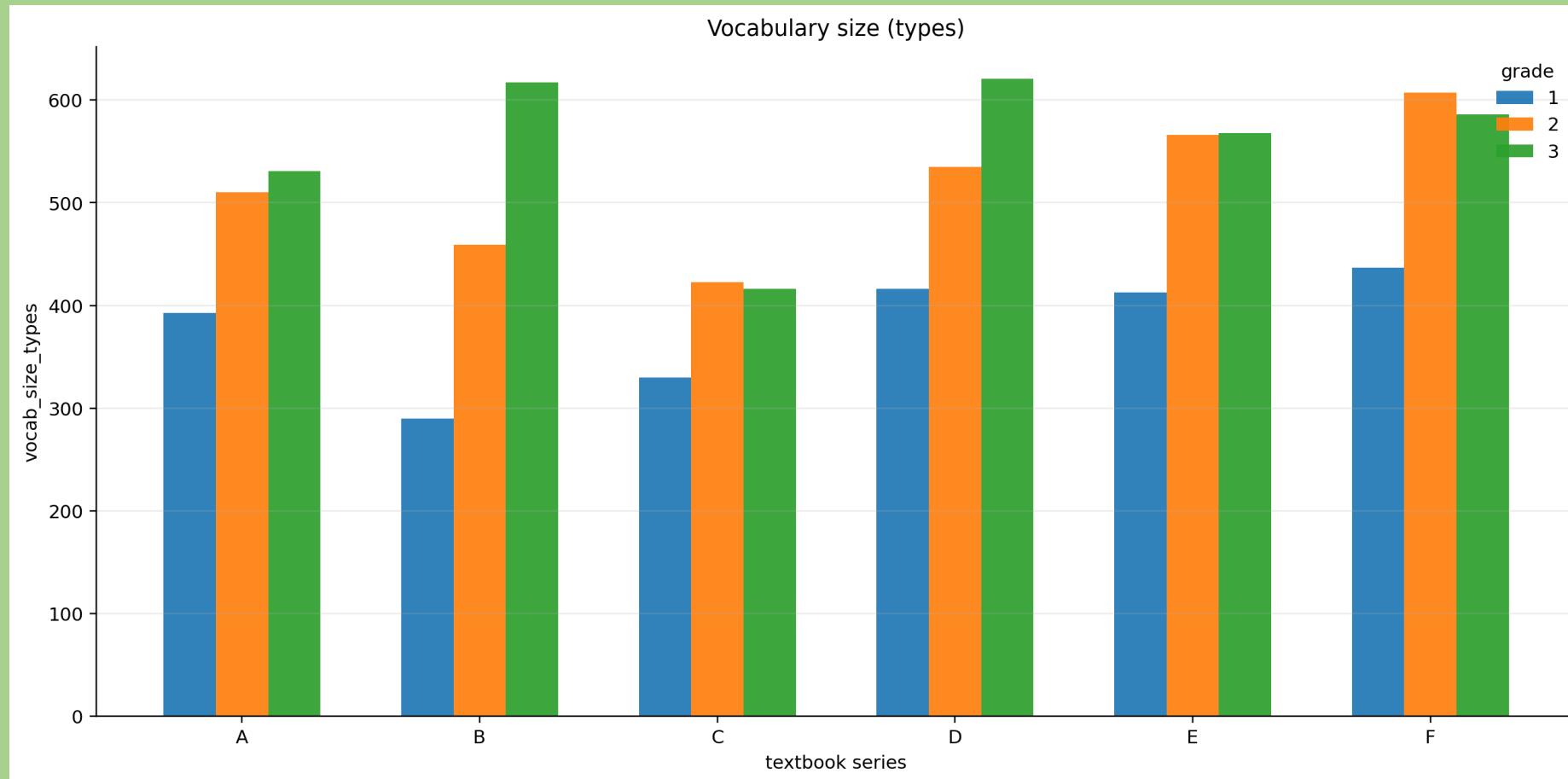
- 2025年7月～12月、トピックは教員が指示
- A中学校3年生 (数字・固有名詞を除く26,722語：75人平均1.9回提出)

【高校】

- 2025年10月～12月、トピックは教員が指示
- B高校1年生 (同96,623語：316人平均3.3回提出)
- B高校2年生 (同143,539語：375人平均3.8回提出)
→スペルミスは除外

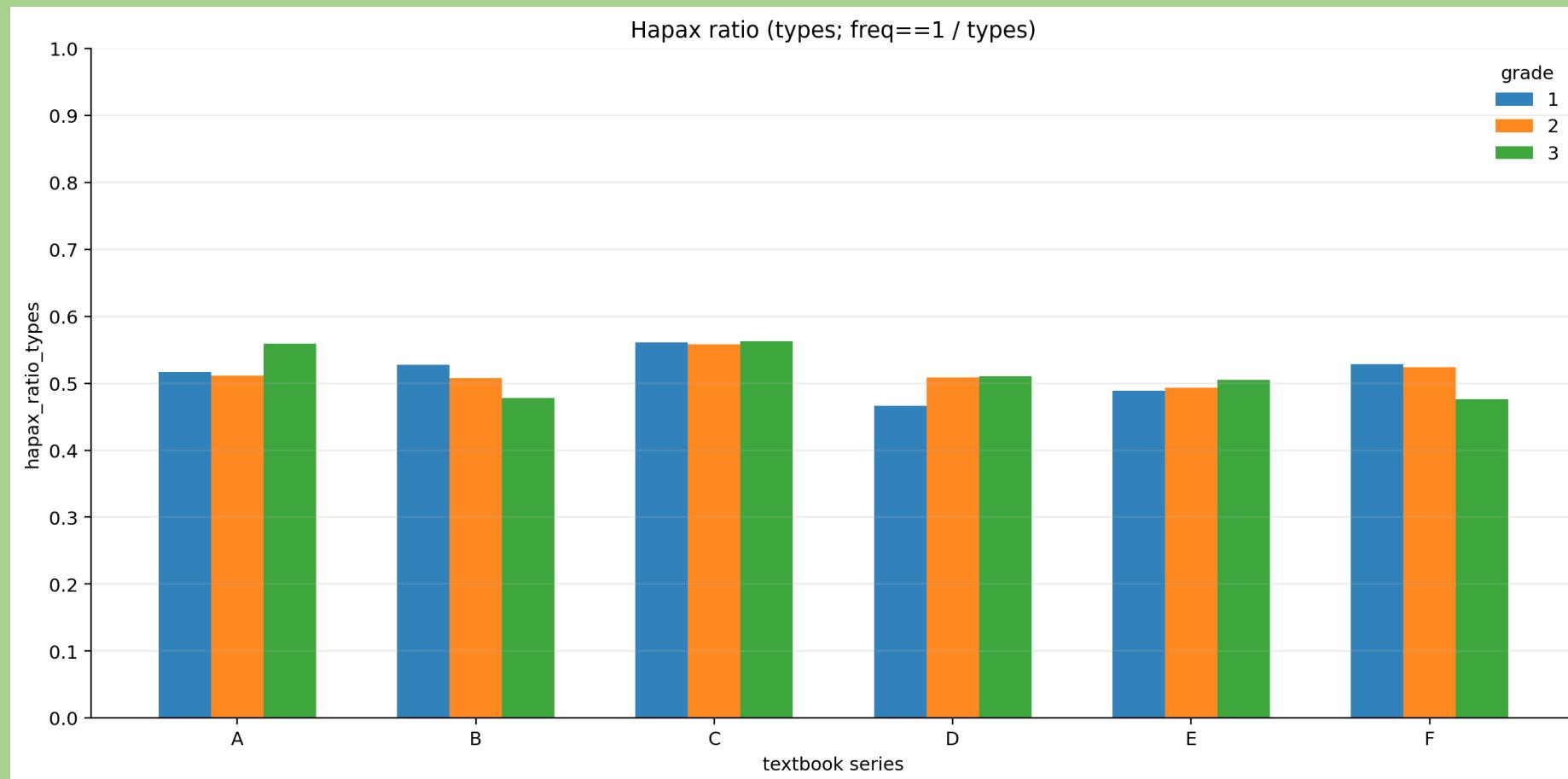
beause -> because (freq=1853) ↓
becaase -> because (freq=1853) ↓
becaouse -> because (freq=1853)
becase -> because (freq=1853) ↓
becasue -> because (freq=1853) ↓
becaouce -> because (freq=1853) ↓
becaue -> because (freq=1853) ↓
becaus -> because (freq=1853) ↓
becouse -> because (freq=1853) ↓
becuse -> because (freq=1853) ↓

中学教科書の分析結果：語彙サイズ



【シリーズ累積】 A社: 979, B社: 920, C社: 779, D社: 1044, E社: 1019, F社: 1057

中学教科書における頻度1の単語の割合



※A~Fはシリーズ名（出版社）を表す

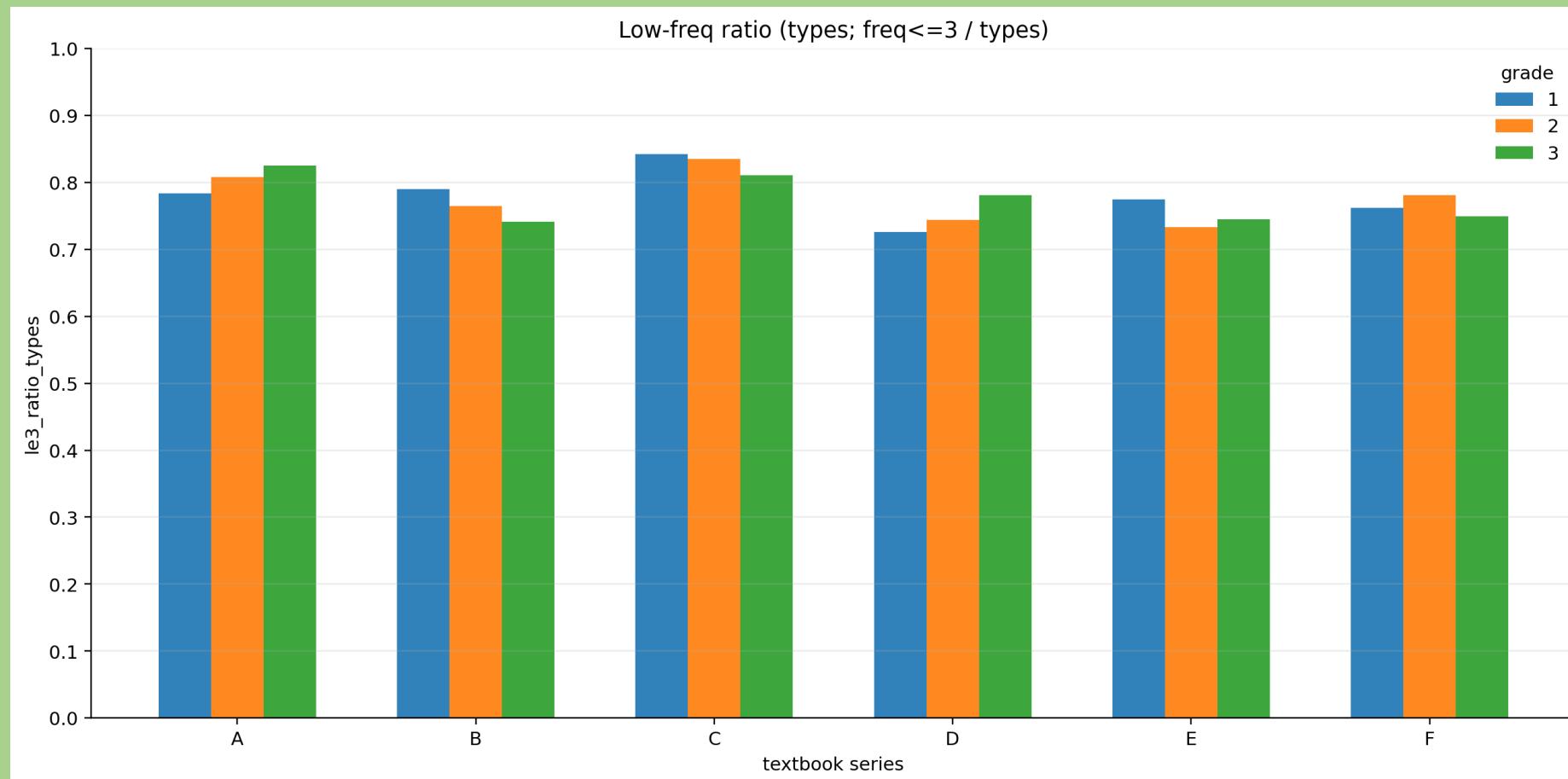
中学教科書における頻度1の単語のCEFRレベル

	A1	A2	B1	B2	#N/A	語数
A	30%	29%	21%	6%	15%	452
B	37%	27%	19%	6%	11%	422
C	45%	28%	15%	3%	10%	368
D	37%	26%	18%	5%	15%	483
E	34%	29%	19%	6%	12%	465
F	33%	28%	16%	7%	15%	477
平均	36%	28%	18%	5%	13%	2667

頻度1のA1の単語の例 (Aより) : add, ask, between, bird, clean, culture

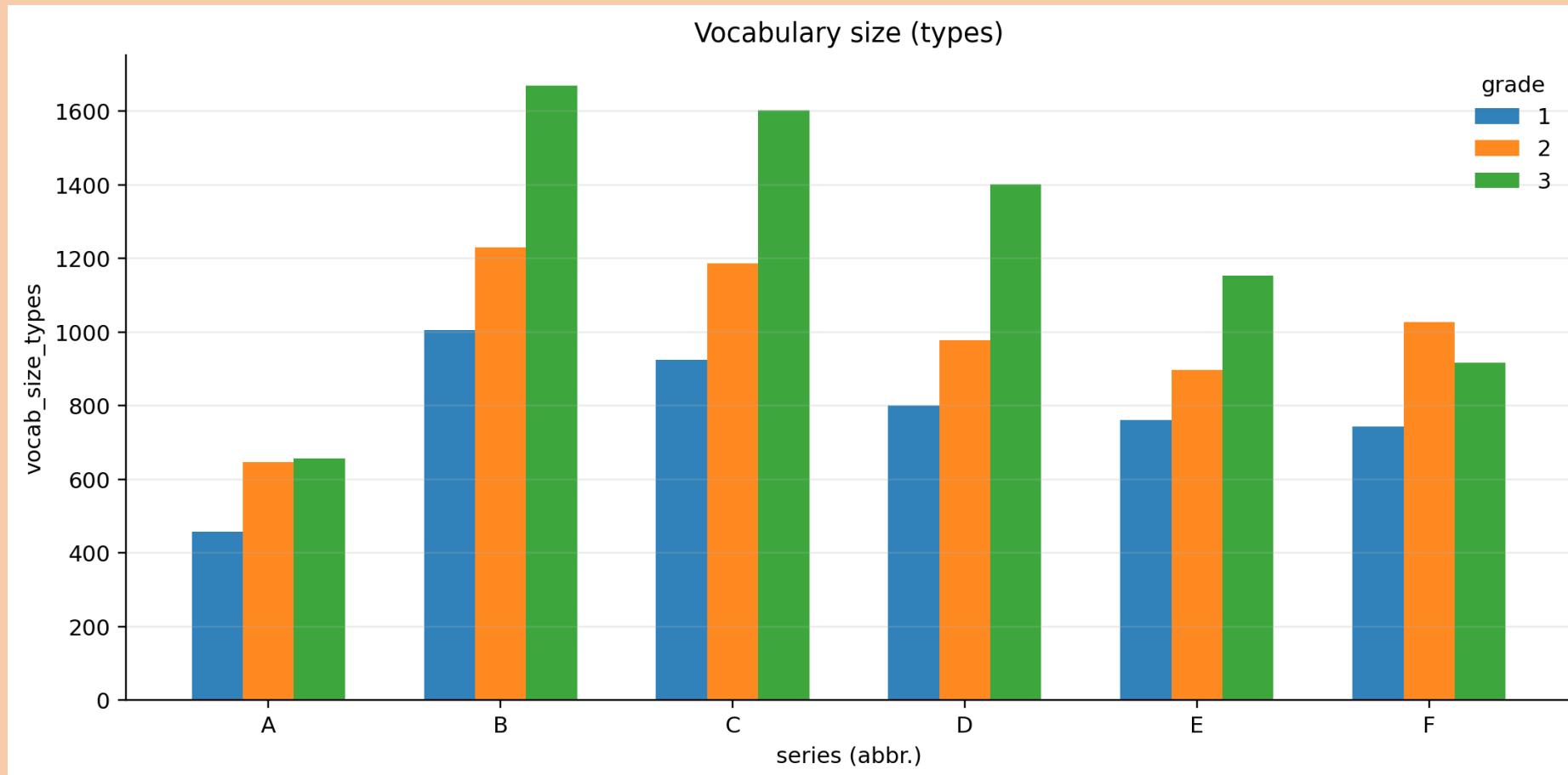
※A~Fはシリーズ名 (出版社) を表す

中学教科書における低頻度単語の割合



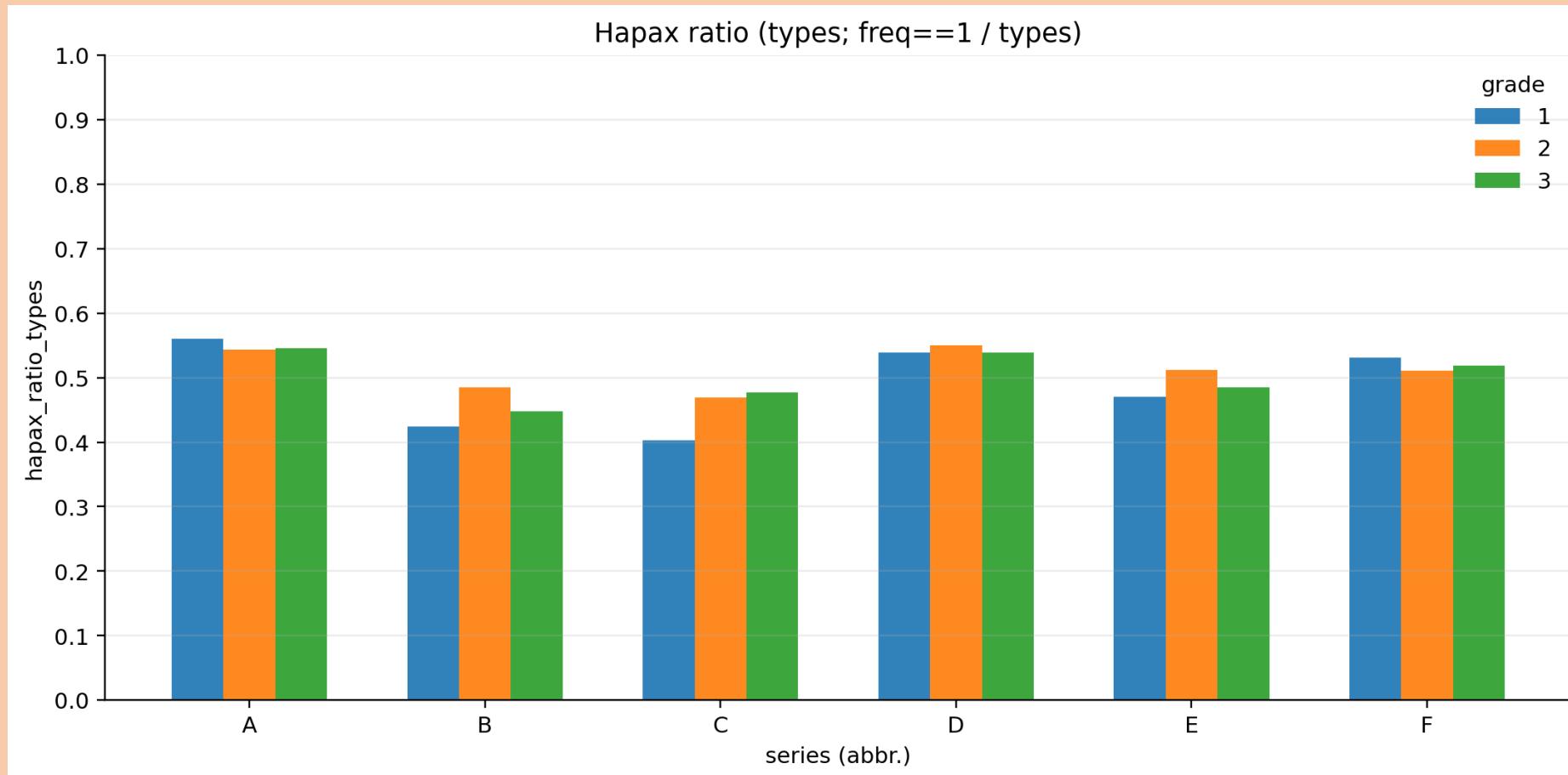
※A~Fはシリーズ名（出版社）を表す

高校教科書の分析結果：語彙サイズ



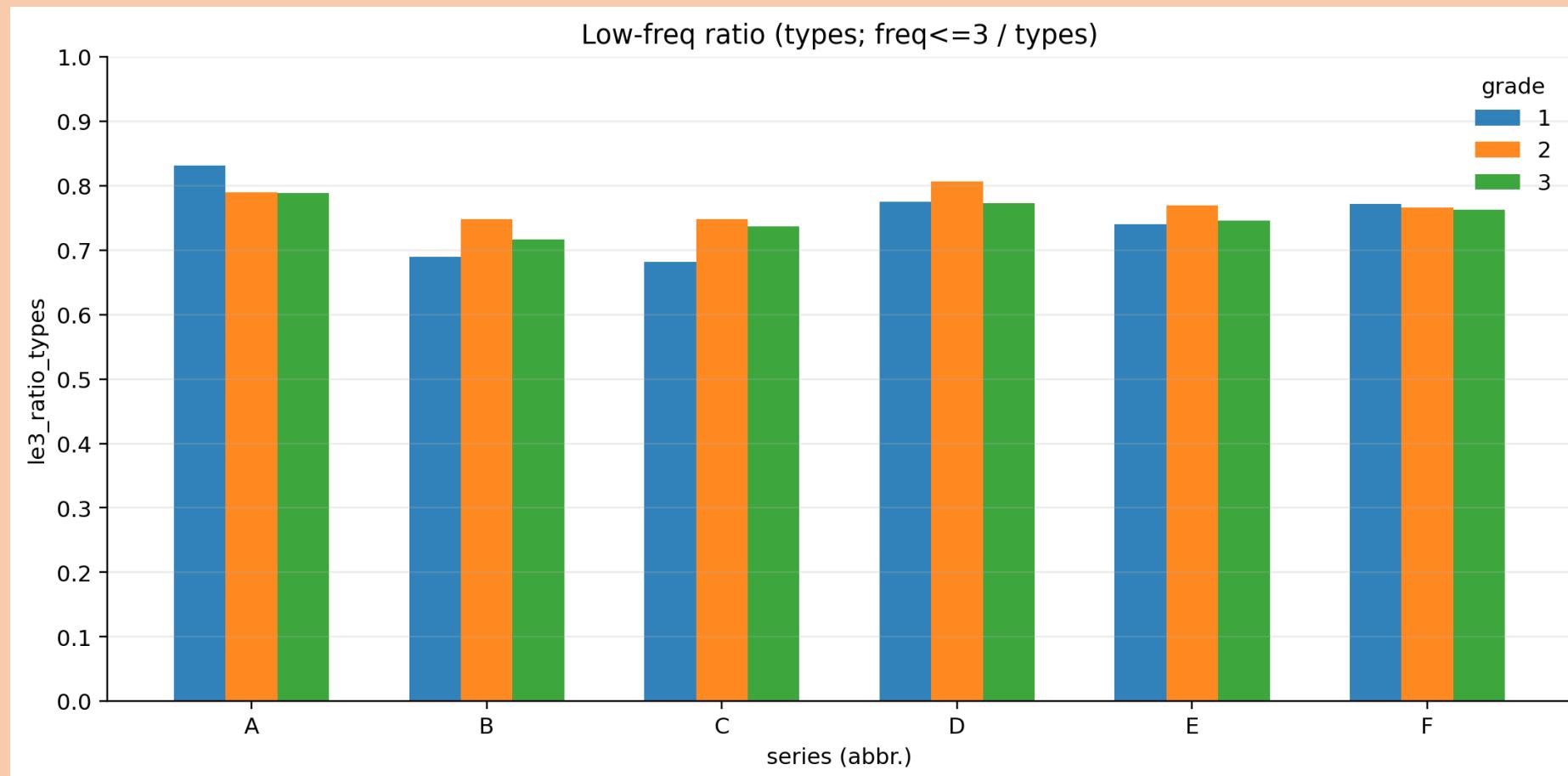
【シリーズ累積】 A社:1149, B社: 2422, C社: 2333, D社: 2042, E社: 1808, F社: 1747

高校教科書における頻度1の単語の割合



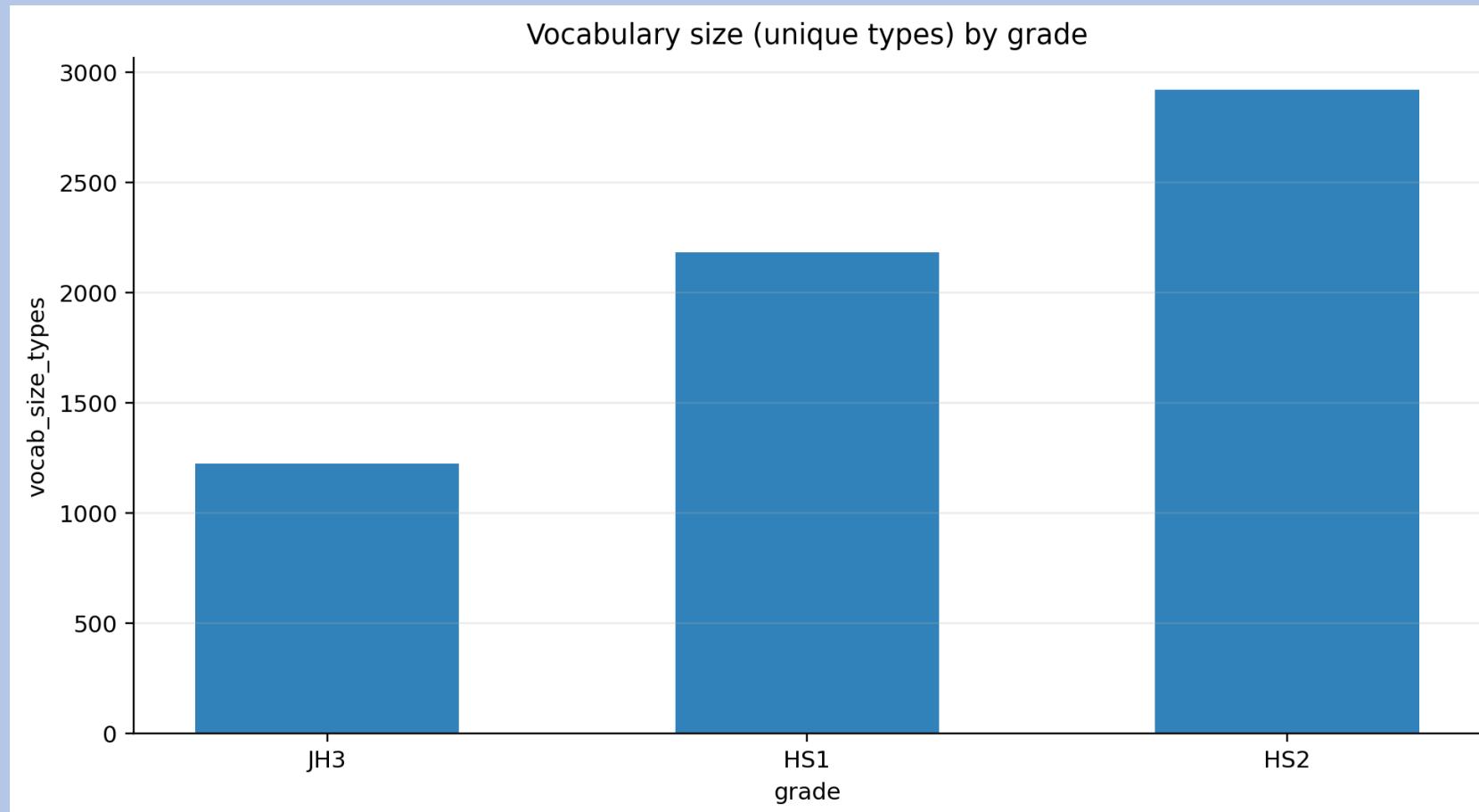
※A~Fはシリーズ名（出版社）を表す

高校教科書における低頻度単語の割合



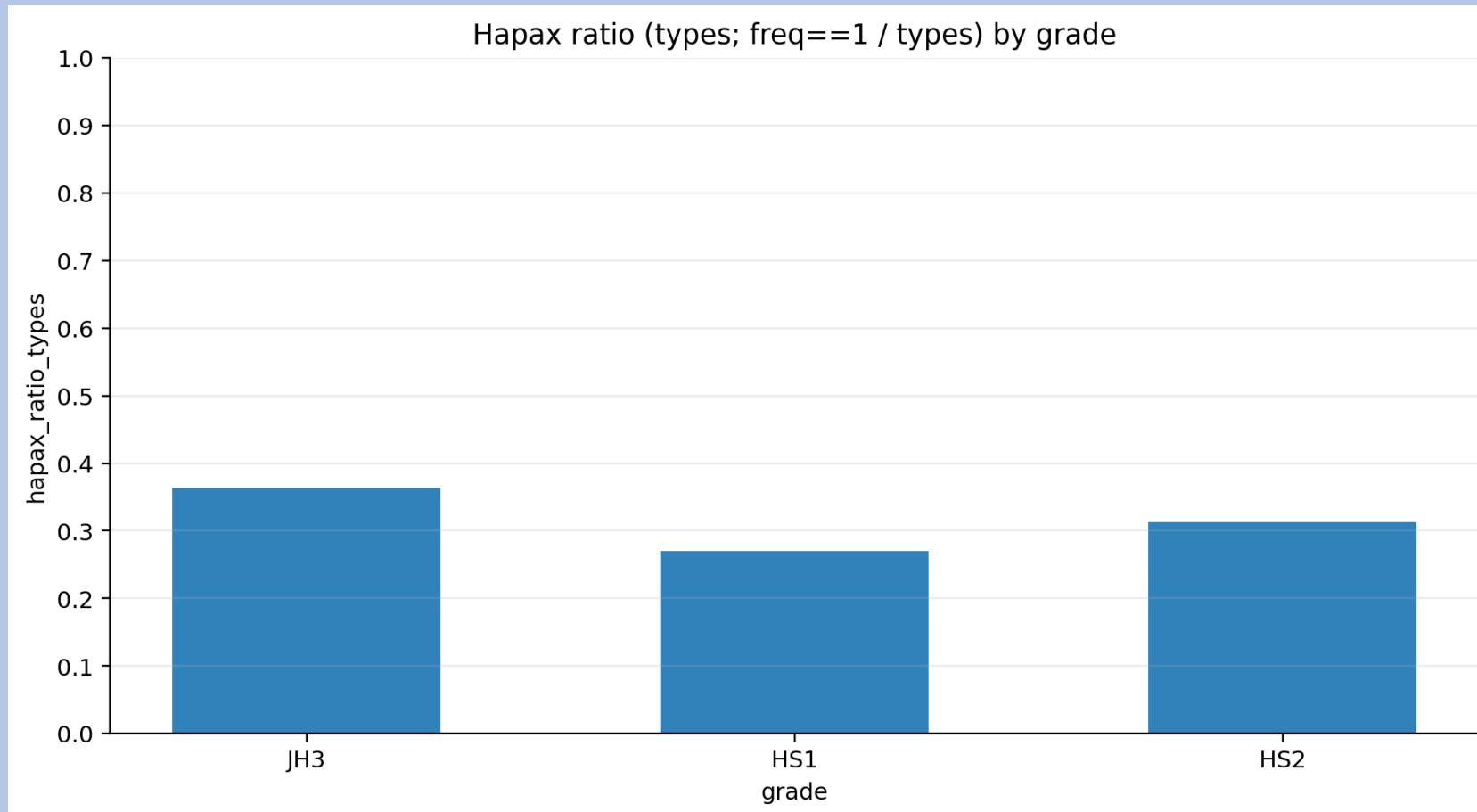
※A~Fはシリーズ名（出版社）を表す

ライティングの学年別の語彙サイズ



※それぞれの学年全体のコーパスで算出

ライティングの頻度1の単語の割合



※それぞれの学年全体のコーパスで算出

ライティングで高頻度の教科書低頻度語

【抽出条件】**中学教科書**（6社合算）で頻度3以下、**中学生ライティング**で頻度10以上、**CEFR-J Wordlist**でAレベルの単語

cafe (A1), deep (A2), grandfather (A1),
hobby (A1), math (A1), round (A2),
skill (A1), smile (A1), society (A2), subject (A1)

まとめ

- 中学・高校ともに**頻度1の単語**の割合は**50%程度**
→半数以上がAレベル (A1、A2) の単語
- 中学・高校ともに**頻度3以下の単語**は**70%以上**
→発信語彙の習得のために十分な頻度がないと考えられる (cf. Morita et al., 2025)
→Sato (2025)では高校英語教科書によるばらつきが大きく、新出単語の頻度が学習のために十分でないことを指摘
- ライティングで頻度1の単語の割合は**約30%程度**
→一定の単語を繰り返し使う傾向がある
→トピックにもよるが学習者が頻繁に使うが教科書には十分に出現しない語がある

Morita, M., Kajiro, T., Uchida, S., Takahashi, Y., & Muto, K. (2025). More words, fewer encounters? A study of Japanese EFL textbooks. SLRF2025.

Sato, T. (2025). Vocabulary analysis of high school English textbooks: Can common vocabulary for high school students be identified? *JASELE Journal*, 36, 65–80.