The Effects of Long-term Extensive Reading on Productive Knowledge of High-frequency Vocabulary

Katsuhiro Chiba, Magda Kitano
Bunkyo University, Japan

Background

- Japanese university students lack high frequency vocabulary

Kitano & Chiba, 2018:

- self-reporting yes/no test
- avg. 33 unknown of first 1000 NGSL words (n=84)

Kitano & Chiba, 2020:

- identified 139 high-frequency words commonly unknown by Japanese university students
- 66 of these do not appear in Japanese junior or senior high school textbooks

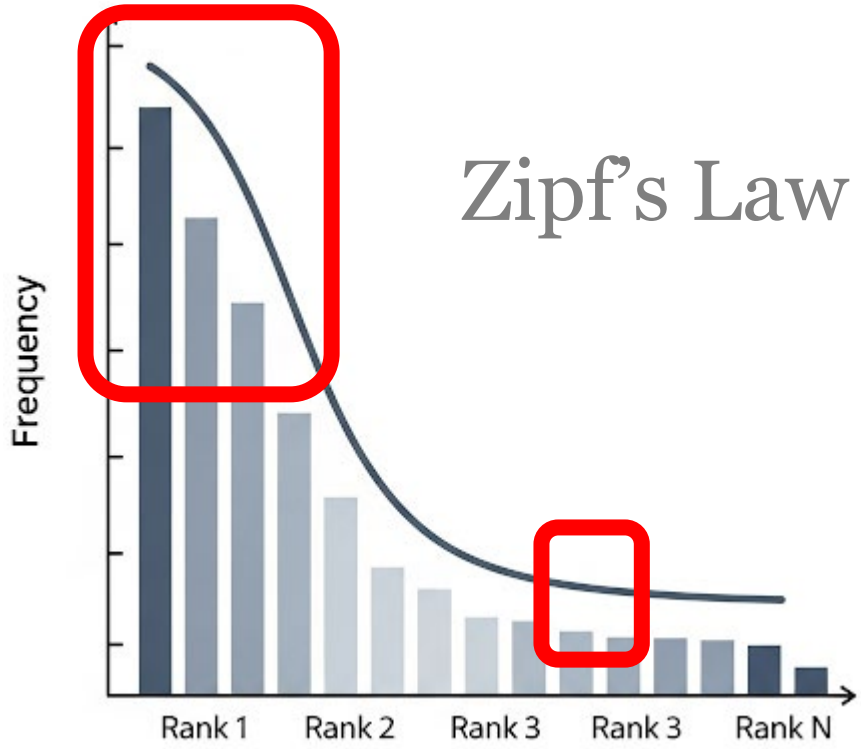
Background

$y \propto \frac{1}{x}$

- Japanese university students lack high frequency vocabulary

High-frequency words:

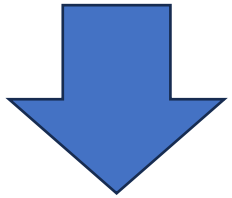
- ✓Make up a large proportion of any text
- ✓Are used primarily in speaking



Zipf's Law

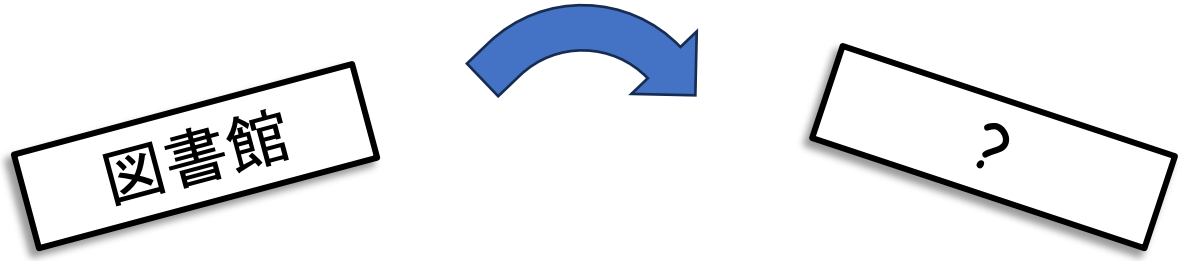
ER for High-frequency Words

- By definition, high-frequency words would be met often through high volumes of ER
- Varied retrieval as well as spaced retrieval is expected if ER is continued over time



ER may be a solution to gaps in high frequency words that vary by learner

Focus on Productive Knowledge



- Productive knowledge of high-frequency words is necessary for basic speaking skills (Nation, 2022)
- Knowledge of various aspects of high-frequency words increases with incidental learning (Webb, 2025), so a productive knowledge test may indicate depth of learning.

Research Question

Does a large volume of extensive reading contribute to productive knowledge of high frequency vocabulary knowledge?

Literature: ER Vocabulary Studies

Single-text studies: after reading a text, students are tested to see if they learned particular words in the text

Horst, Cobb, and Meara (1998)

Zahar, Cobb, and Spada (2001)

Waring and Takaki (2003)

Brown, Waring, and Donkaewbua (2008)

Pellicer-Sanchez and Schmitt (2010)

Alsaif and Masrai (2018)

Literature: ER Vocabulary Studies

Program-wide studies: test vocabulary knowledge of students within a particular ER program. They include pre- and post-testing, and control and treatment groups.

Lee (2007)

Kweon and Kim (2008)

Suk (2017), (2021)

Yamamoto (2011)

Aka (2018)

Webb & Chang (2020)

Nakanishi (2015): meta-analysis

Current Study

High volume ER readers
(400,000 to 2,000,000 tokens)

No (or little) ER experience

CEFR B1

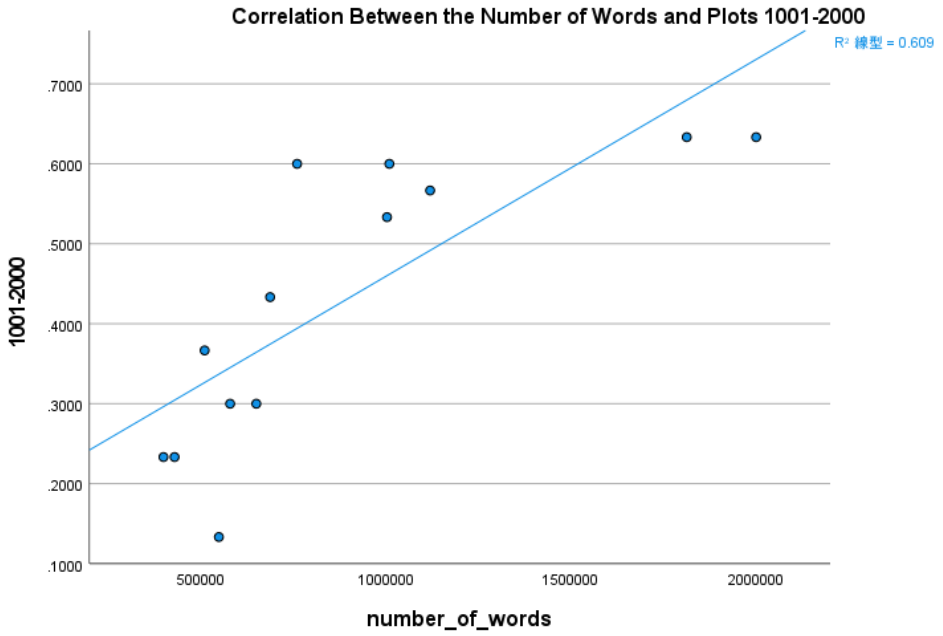
Tested: high-frequency (rank 1-2000) productive knowledge

Current Study

Preliminary findings were presented at Vocab@Vic, 2023:

No significant difference at that point (n=25)

But scores increased with amount of ER read



Methods: Subjects

ER Group
(n=22, 13 Male, 9 Female)

Average Words Read: 1,011,093

SD=534,493.61 , Max=2.600,000, Min=430,000

- university student in Japan
- have read > 400,000 words of ER
- submit reading log

non-ER Group
(n=23, 10 Male, 13 Female)

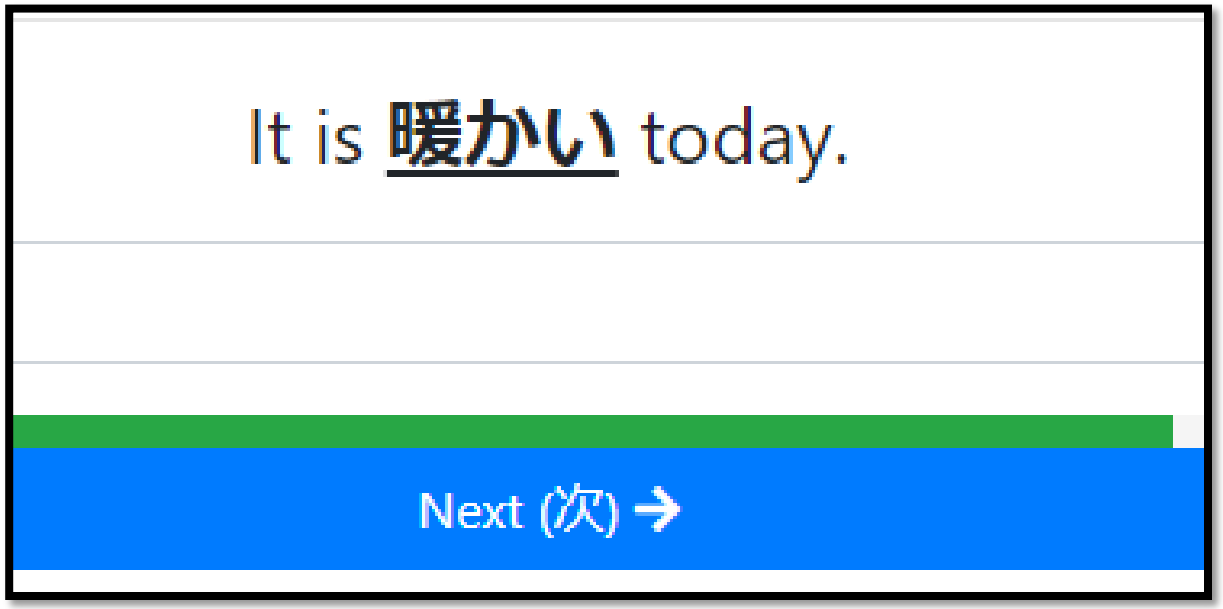
Average Words Read: 2,579

- university student in Japan
- have not done ER (or <30,000 words)

Methods: Vocabulary Test

Vocab Level Test (vlt.carleton.ca)

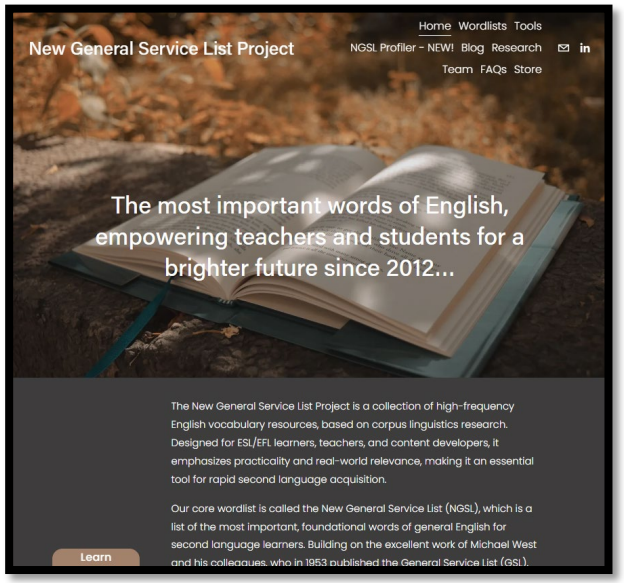
- meaning-recall test
- form-recall test



McLean, S. & Raine, P. (2019). VocabLevelTest.Org [Web application]. Retrieved from https://www.vocableveltest.org

Methods: Vocabulary Test

- New General Service List
- Two bands: 1-1000, 1001-2000
- 30 items tested per band
- students sat the test with a researcher present



www.newgeneralservicelist.com

Results

Descriptive Statistics

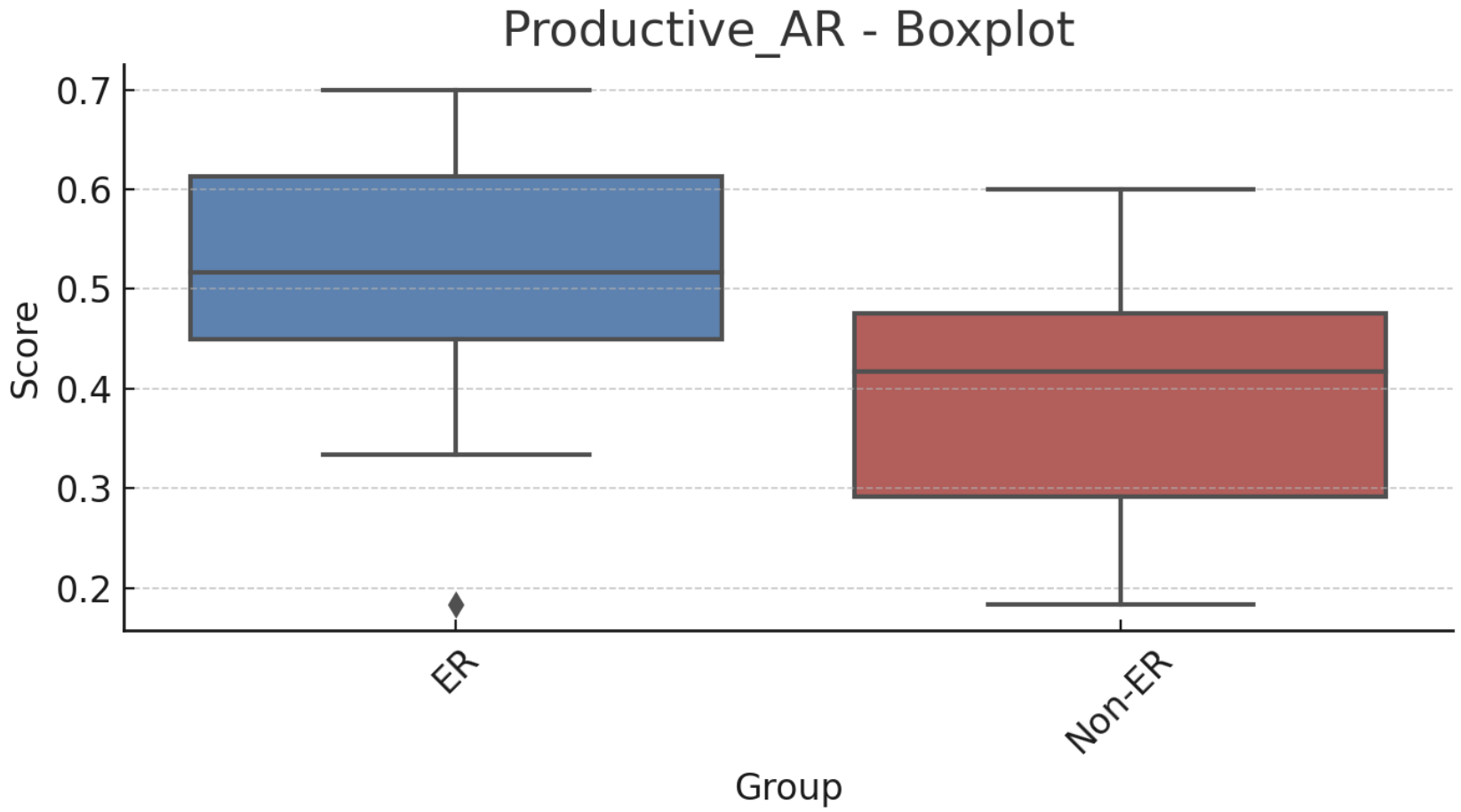
Participant Proficiency: CASEC

Group	Mean	SE
ER	612.955	23.744
Non-ER	504.174	17.889

Accuracy Rate Results

Group	Mean	SE
ER	0.513	0.027
Non-ER	0.399	0.024

Accuracy Rate Results: Overall



Is it really ER?

Isn't the high accuracy rate on vocabulary tests due to original English proficiency rather than the effect of extensive reading?

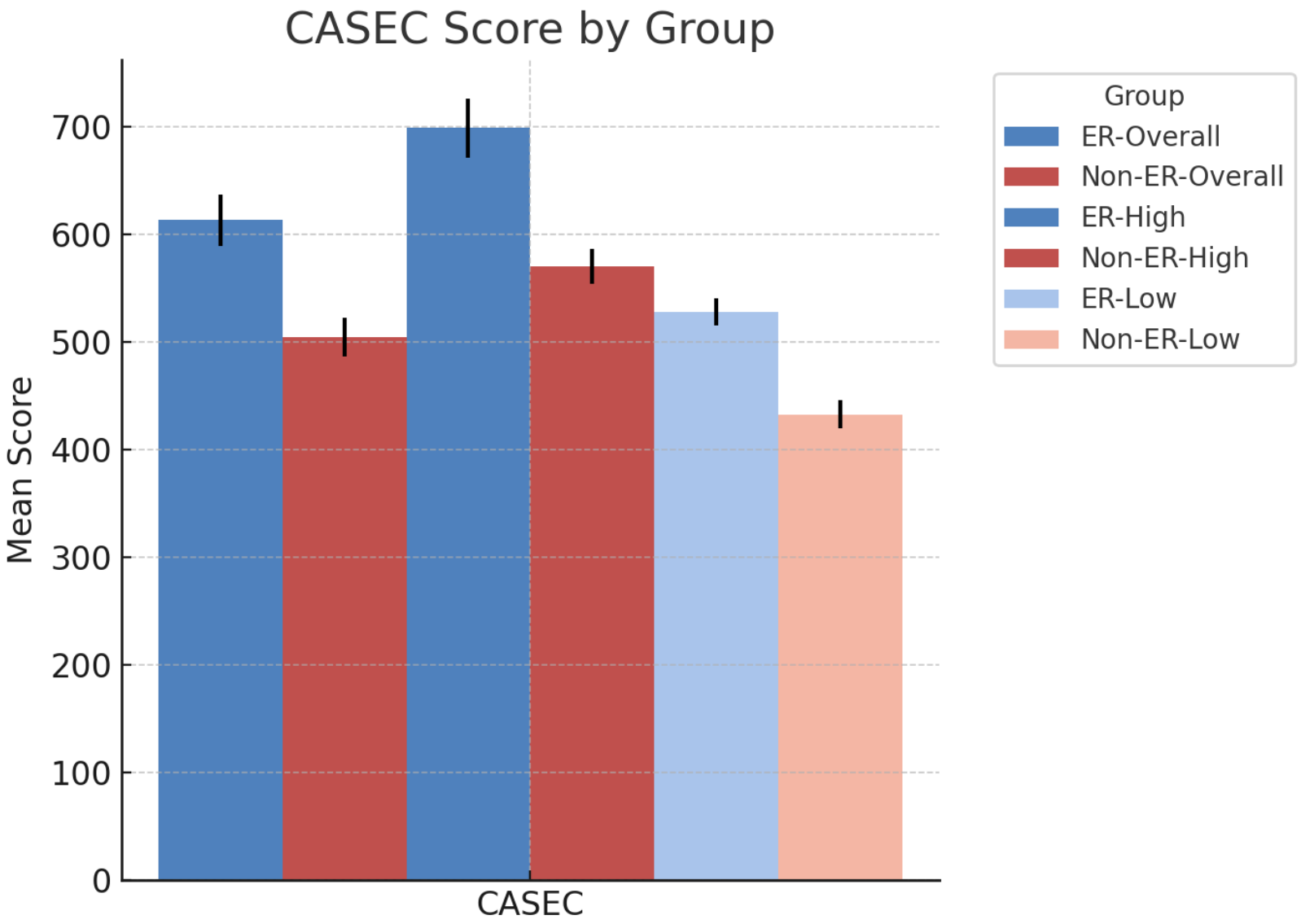
Grouping

We divided the subjects into two proficiency groups, high and low, by CASEC scores.

Proficiency Groups: CASEC Scores

Proficiency	Group	Mean	SE
Overall	ER	612.955	23.744
	Non-ER	504.174	17.889
High Proficiency	ER	698.273	27.422
	Non-ER	569.917	16.151
Low Proficiency	ER	527.636	12.651
	Non-ER	432.455	13.158

Proficiency Groups: CASEC Scores



Accuracy Rate by Proficiency Groups

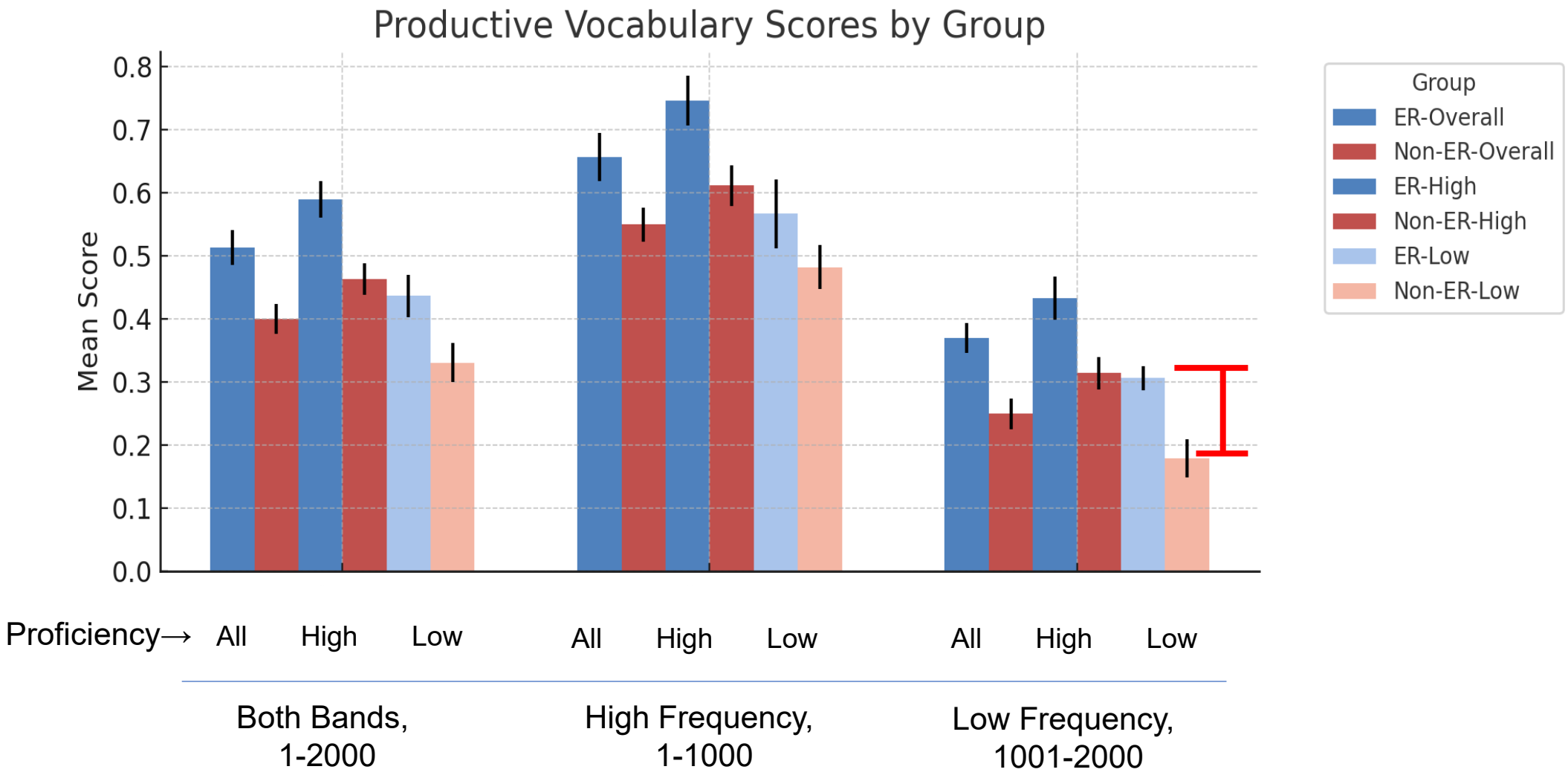
Proficiency	Group	Mean	SE
Overall	ER	0.513	0.027
	Non-ER	0.399	0.024
High Proficiency	ER	0.589	0.029
	Non-ER	0.462	0.025
Low Proficiency	ER	0.436	0.034
	Non-ER	0.33	0.031

Productive Vocabulary

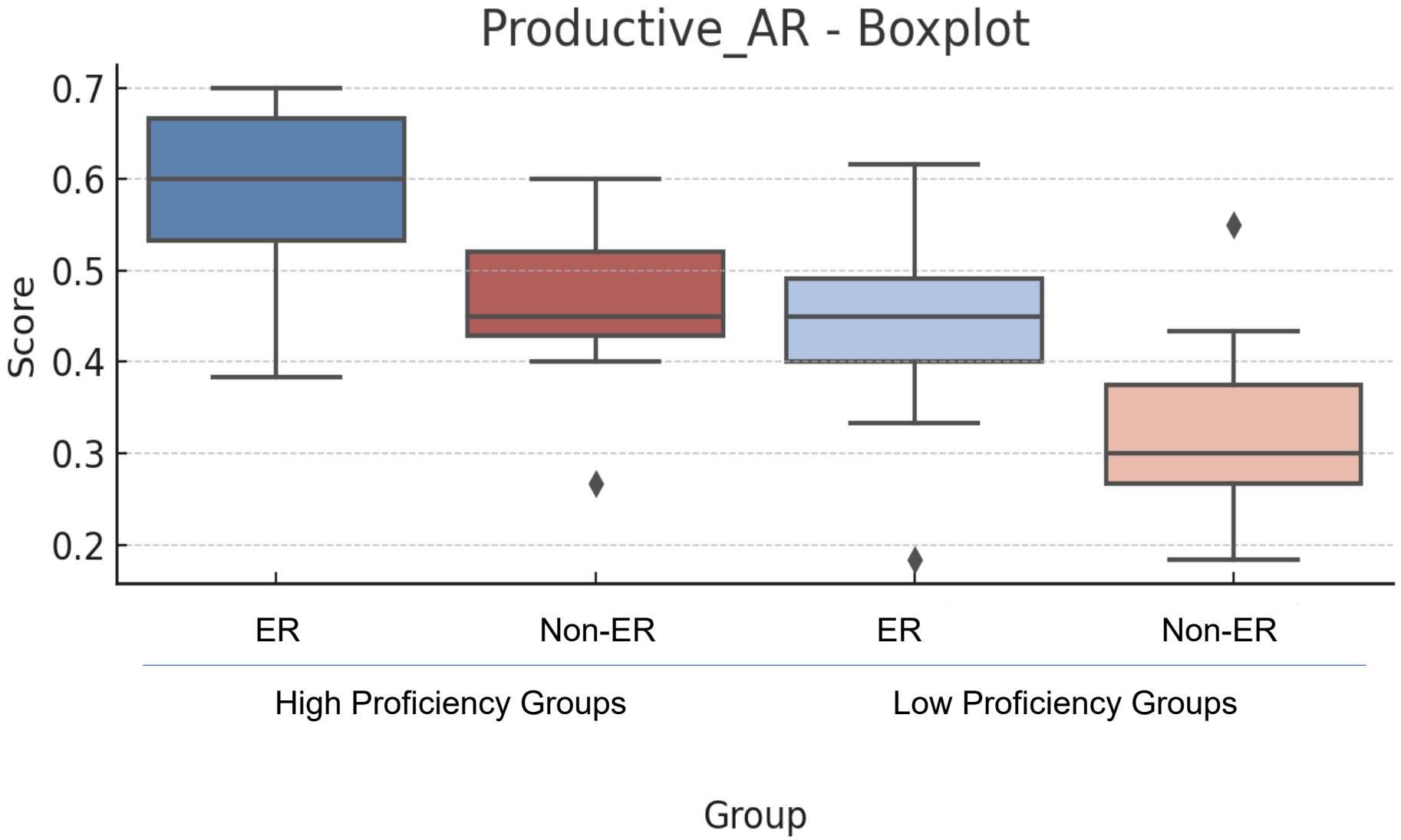
by Proficiency Groups and Frequency Band

Metric	Proficiency	Group	Mean	SE
High Frequency Band: Rank 1 - 1000	Overall	ER	0.656	0.038
		Non-ER	0.549	0.027
	High Proficiency Groups	ER	0.745	0.039
		Non-ER	0.611	0.032
	Low Proficiency Groups	ER	0.567	0.055
		Non-ER	0.482	0.035
High Frequency Band: Rank 1001 - 2000	Overall	ER	0.37	0.024
		Non-ER	0.249	0.024
	High Proficiency Groups	ER	0.433	0.034
		Non-ER	0.314	0.026
	Low Proficiency Groups	ER	0.306	0.019
		Non-ER	0.179	0.031

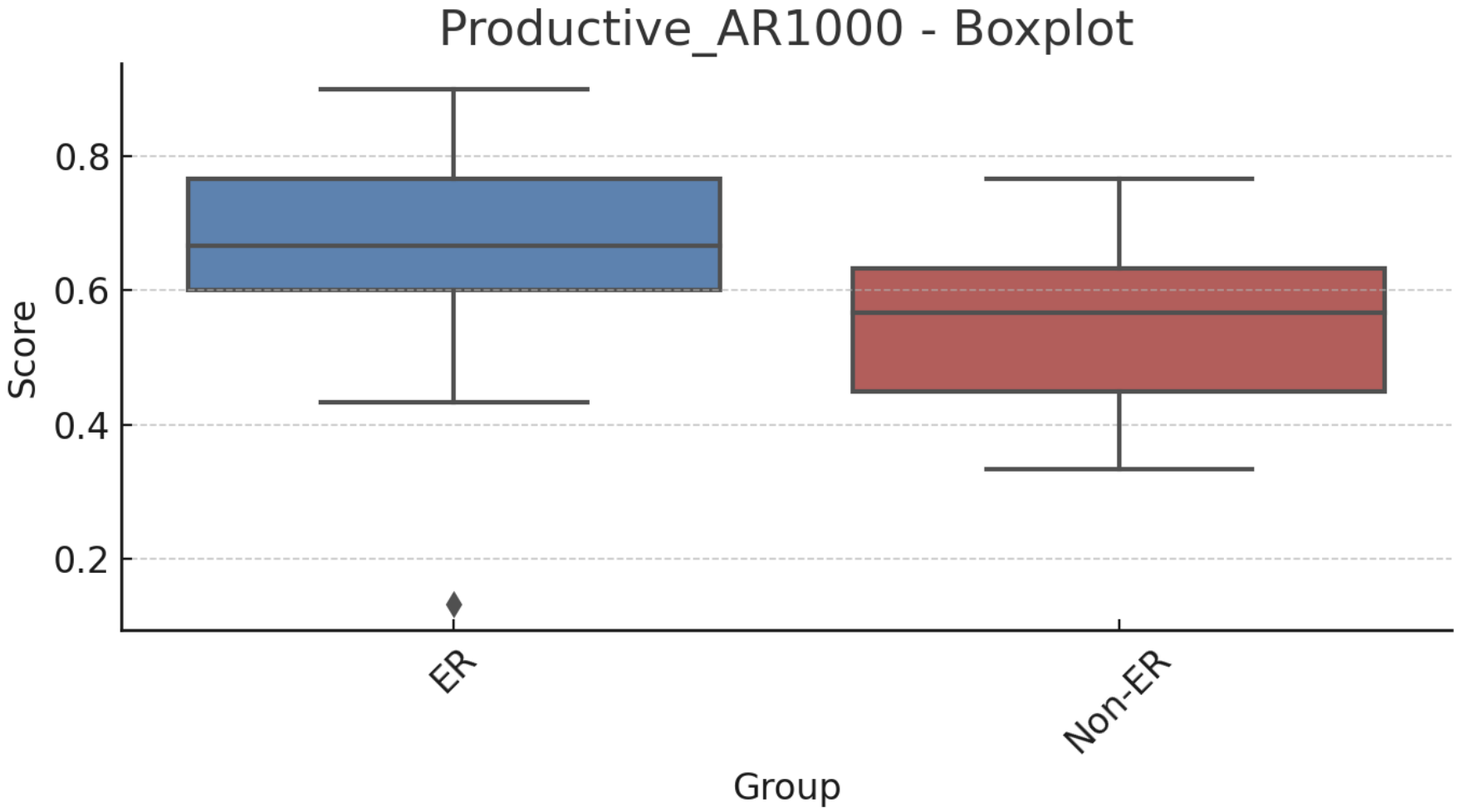
Accuracy Rate by Proficiency Groups and Frequency Band



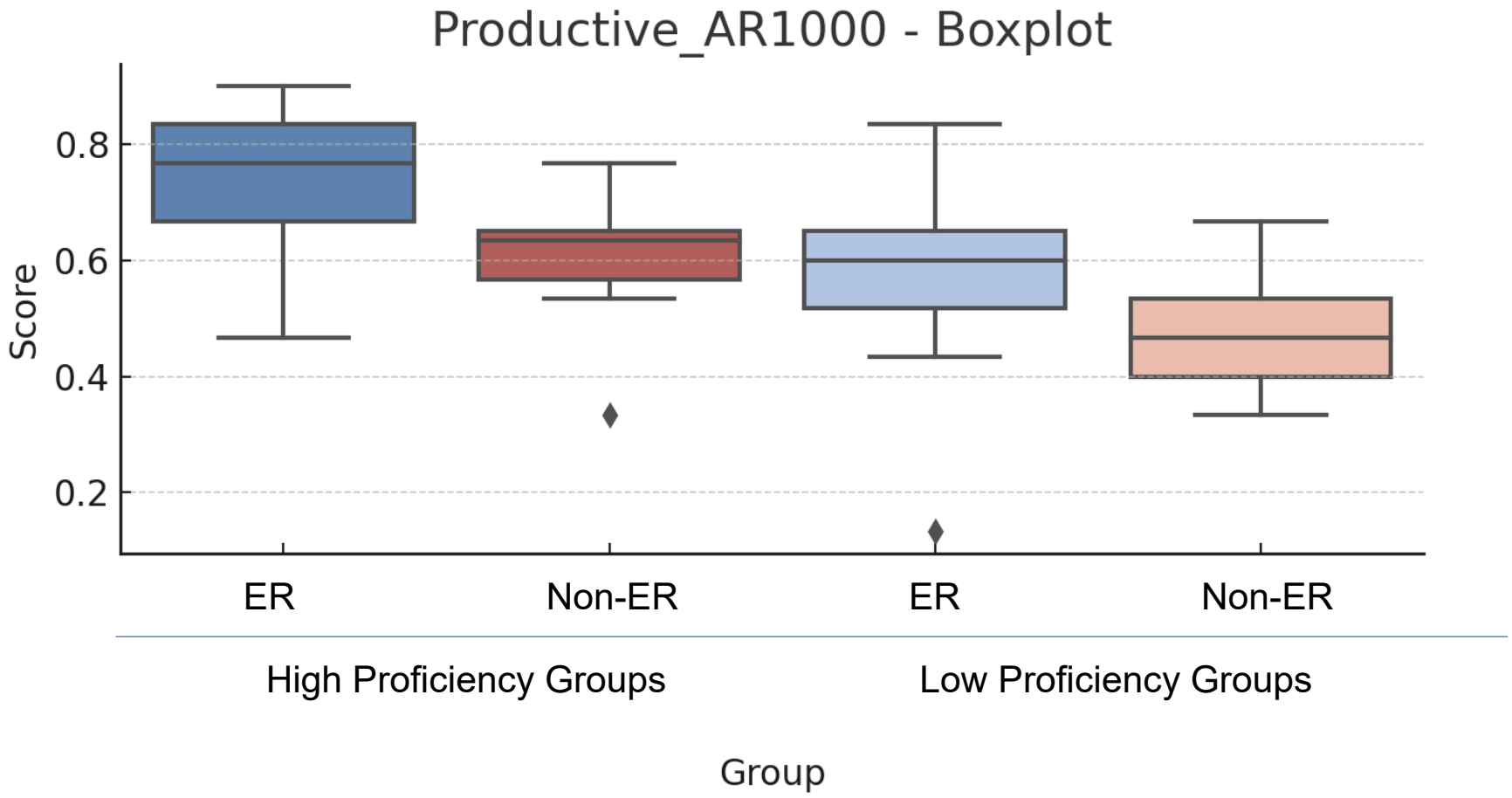
Accuracy Rate: High-Low Comparison (Both Bands, 1 – 2000)



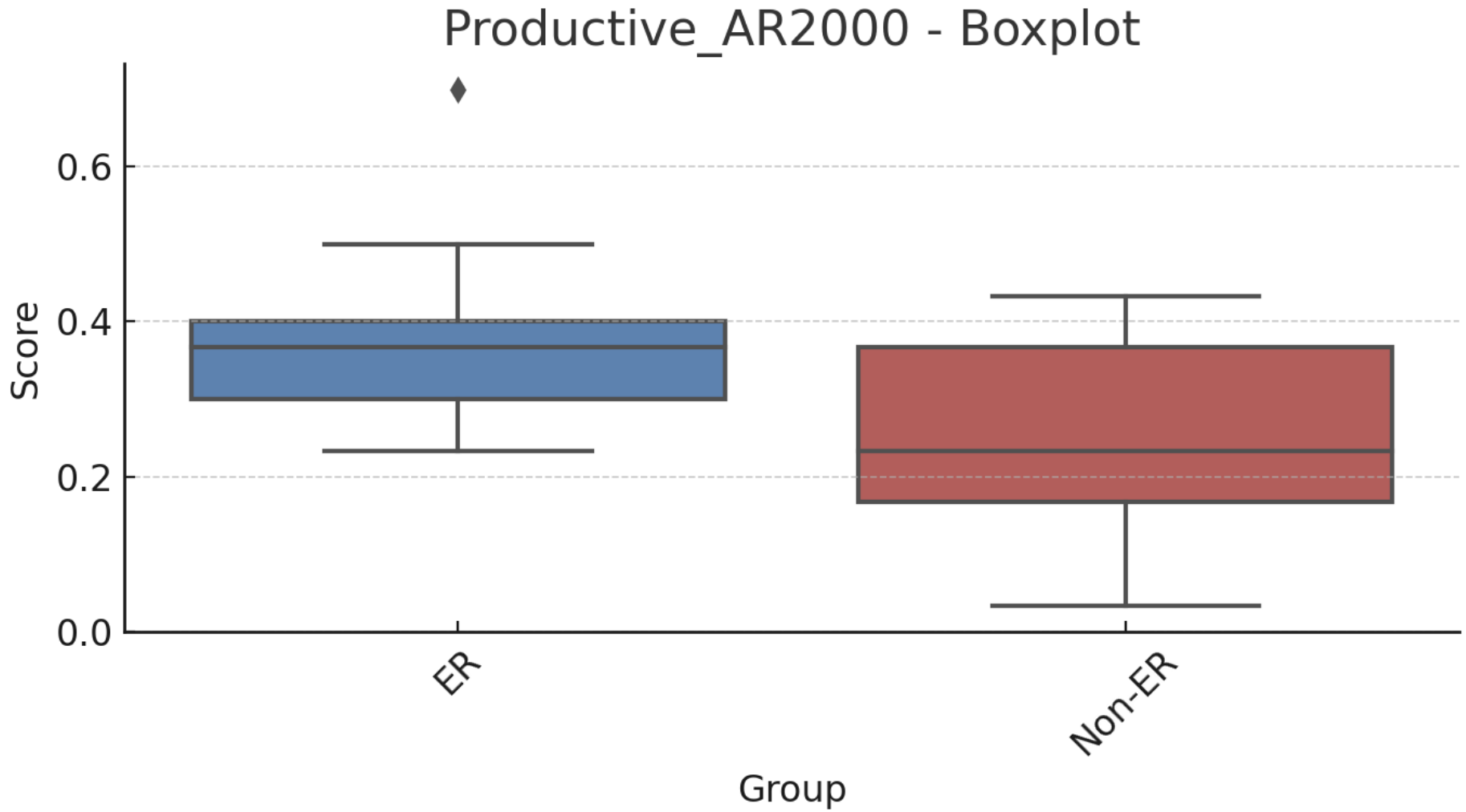
Accuracy Rate: All Participants, High Frequency Band (1-1000)



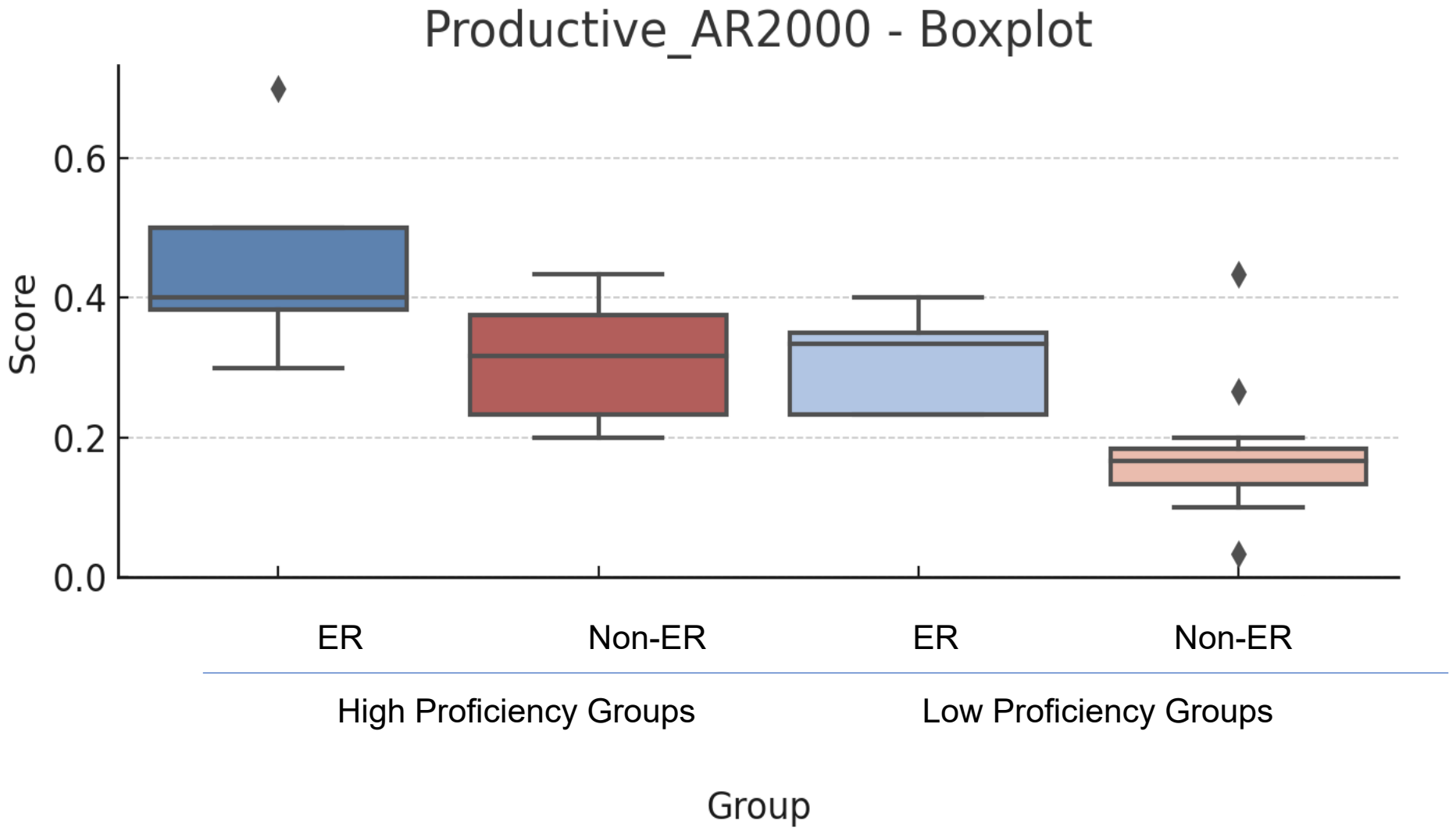
Accuracy Rate: High Frequency Band (1-1000)



Accuracy Rate: All Participants, Low Frequency Band (1001-2000)



Accuracy Rate: Low Frequency Band (1001-2000)



ANOVA

ANOVA Summary: Productive

Frequency	Effect	F	p	FDR p	η^2
Overall	ER Group	15.666	0.000	0.001	0.196
Overall	CASEC Level	23.324	0.000	0.000	0.291
Overall	Interaction	0.125	0.726	0.816	0.002
High (1-1000)	ER Group	7.227	0.010	0.016	0.115
High (1-1000)	CASEC Level	14.044	0.001	0.001	0.224
High (1-1000)	Interaction	0.365	0.549	0.706	0.006
Low (1001-2000)	ER Group	19.386	0.000	0.000	0.235
Low (1001-2000)	CASEC Level	21.983	0.000	0.000	0.267
Low (1001-2000)	Interaction	0.020	0.890	0.890	0.000

Interpretation of ANOVA Results

• Interaction Effect (ER × CASEC): Not significant.

• Main effects:

- ER Group: Significant across all productive vocabulary measures.

- CASEC Level: Also significant across all measures.

▶ Interpretation:

→ Both Extensive Reading (ER) and English proficiency contribute to vocabulary growth.

→ ER has an effect regardless of proficiency level.

Pairwise Comparison of Productive Vocabulary Scores with FDR Correction

Our ANOVA showed that extensive reading had a statistically significant main effect on productive vocabulary scores, meaning that its positive impact was observed regardless of the learner's CASEC score.

However, to determine **which proficiency groups** were specifically benefiting from ER, we proceeded with pairwise comparisons.

Given that multiple comparisons increase the risk of false positives, we applied **FDR correction**, which offers a balanced approach to statistical rigor and discovery.

Pairwise Comparison: ER vs Non-ER (High Proficiency Groups)

Measure	Mean (ER-High)	Mean (Non-ER-High)	FDR p-value	Cohen's d
Overall	0.589	0.463	0.010	1.397
High Frequency (1-1000)	0.745	0.611	0.023	1.109
Lower Frequency (1001-2000)	0.433	0.314	0.023	1.170

Pairwise Comparison: ER vs Non-ER (Low Proficiency Groups)

Measure	Mean (ER-Low)	Mean (Non-ER-Low)	FDR p-value	Cohen's d
Overall	0.436	0.330	0.037	0.993
High Frequency (1-1000)	0.567	0.482	0.210	0.555
Lower Frequency (1001-2000)	0.306	0.179	0.010	1.497

Overall Pairwise Comparison

Proficiency Group	Measure	Mean (ER)	Mean (Non-ER)	FDR p-value	Cohen's d
High	Overall	0.589	0.463	0.010	1.397
Low	Overall	0.436	0.330	0.037	0.993
High	High Frequency	0.745	0.611	0.023	1.109
Low	High Frequency	0.567	0.482	0.210	0.555
High	Lower Frequency	0.433	0.314	0.023	1.170
Low	Lower Frequency	0.306	0.179	0.010	1.497

Key Findings from Pairwise Comparisons

- ER learners consistently outperformed Non-ER learners across all vocabulary measures.
- Significant differences were observed even when controlling for proficiency (High vs High, Low vs Low).
- Cohen's d values suggest large effects of ER, especially in:
  - Low frequency bands of Lower proficiency learners.
- Results confirm that ER contributes to productive vocabulary growth across proficiency levels.

Conclusion

- Extensive Reading (ER) improves High- frequency productive vocabulary knowledge.
- This effect is independent of proficiency level.

▶ Implication:

→ ER is effective for learners at both high and low proficiency levels.

→ Supports the use of long-term ER as a general vocabulary acquisition strategy.

Limitations and Future Goals

- Difficulty of getting high volume readers.
- Fostering life-long readers.
- What research method can we use to prove effects of ER scientifically, overcoming the variables of long term activity?

The Effects of Long-term Extensive Reading on Productive Knowledge of High-frequency Vocabulary

Katsuhiro Chiba, Magda Kitano
Bunkyo University, Japan

references

Thank you!