



# FUTURE TENSE CONSTRUCTION PREFERENCE IN L2 ENGLISH LEARNERS

DANIEL CRAWFORD

18 APRIL 2024

LING2020 – DATA SCIENCE FOR RESEARCH IN LINGUISTICS

# MOTIVATION

- L2 English Speakers
- Instruction order of future construction (Bardovi-Harlig 2004)
- “going to walk” vs. “will walk”



## RESEARCH QUESTION

IS THERE A CORRELATION BETWEEN FUTURE TENSE  
CONSTRUCTION PREFERENCE AND ENGLISH  
PROFICIENCY FOR L2 ENGLISH LEARNERS?

# METHODOLOGY - CORPUS

- University of Pittsburgh English Language Institute Corpus (PELIC)
  - Juffs, Han, Naismith (2020)
- 7 years of English for Academic Purposes (EAP) (2005-2012)
- Intensive English Program (IEP).
- 1177 students
- 46,230 texts
- 4,250,703 tokens
- Token, Lemma, POS -> determine construction

# METHODOLOGY - PROFICIENCY

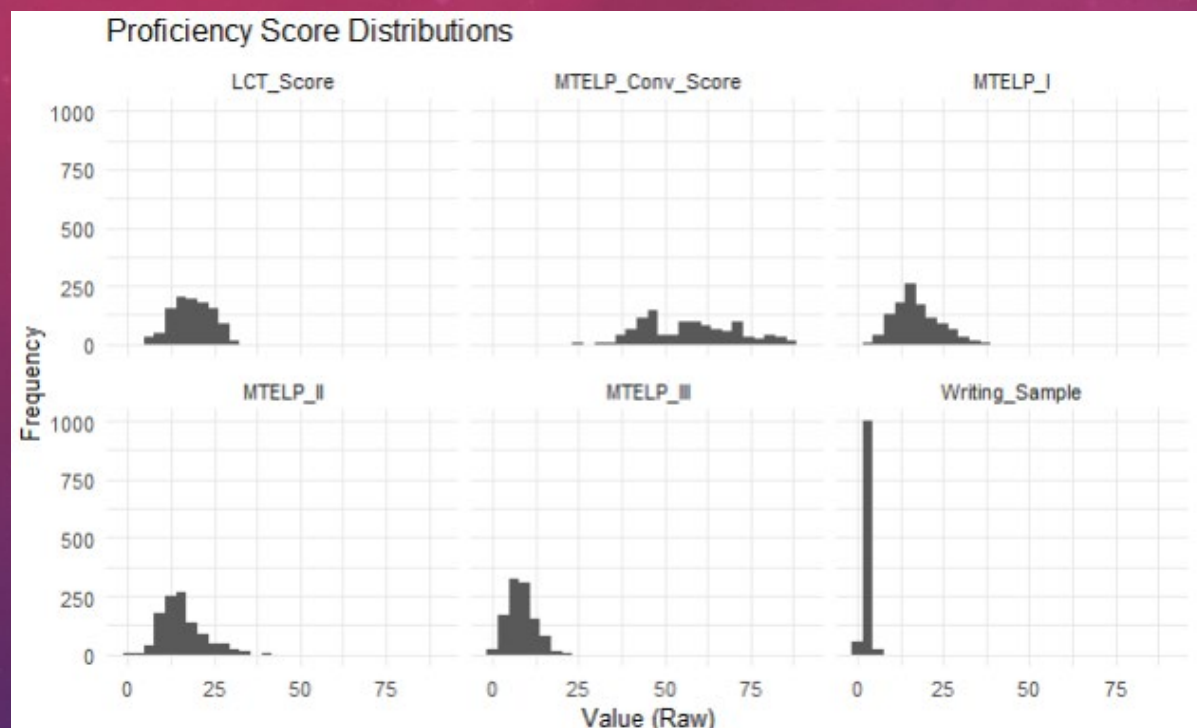
- Michigan Test for English Language Proficiency (MTLEP)
  - I = Grammar, II = Reading, III = Listening, Conv = Combined Score
- Writing Sample: In-house test score
- In House Listening Test (LCT)



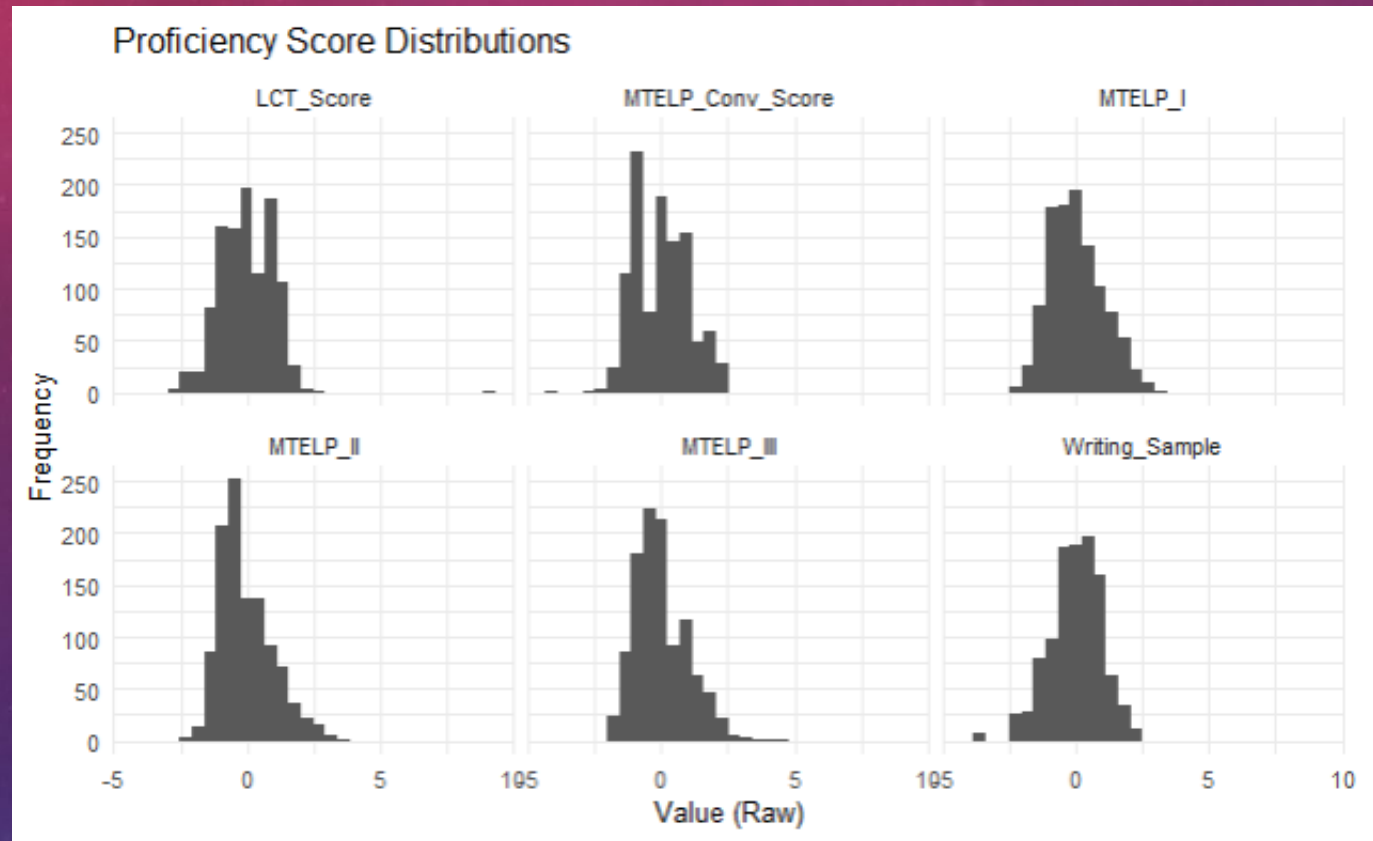
# SUMMARY STATISTICS

- 1078 students
- 29 different native languages
- Will Construction: 15,210
- Going To Construction: 1,728

# TEST SCORES

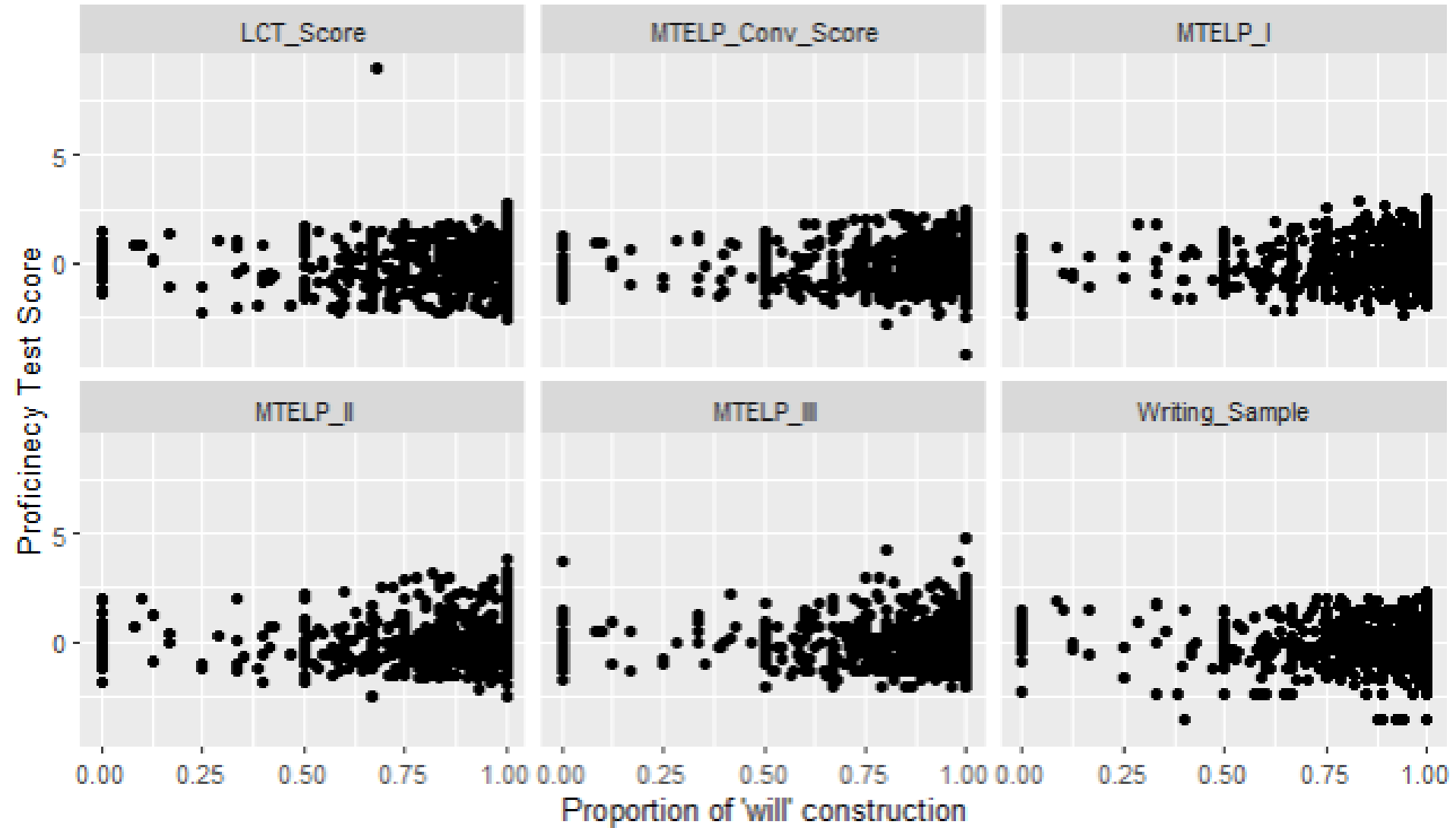


# TEST SCORE NORMALIZED

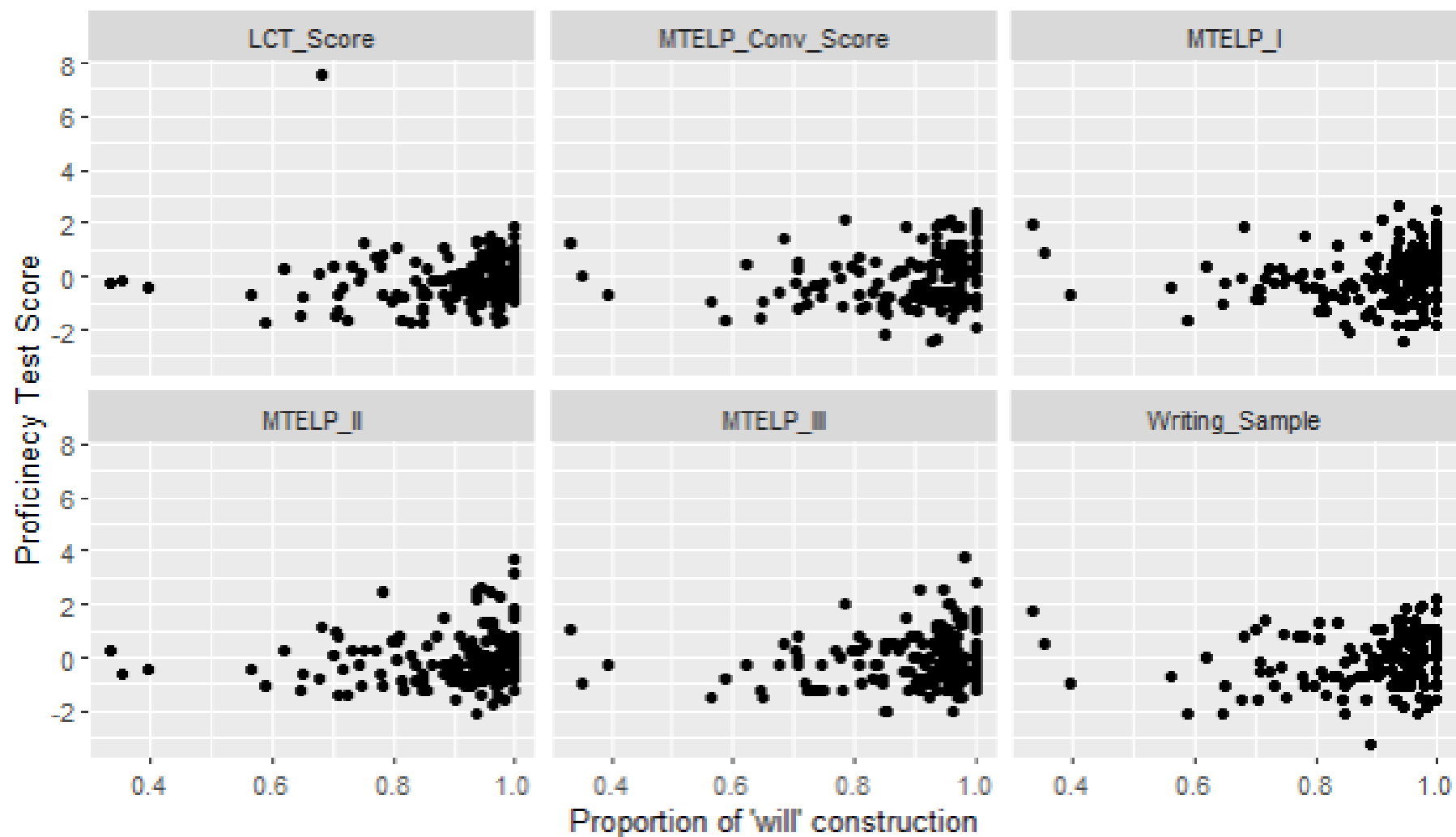




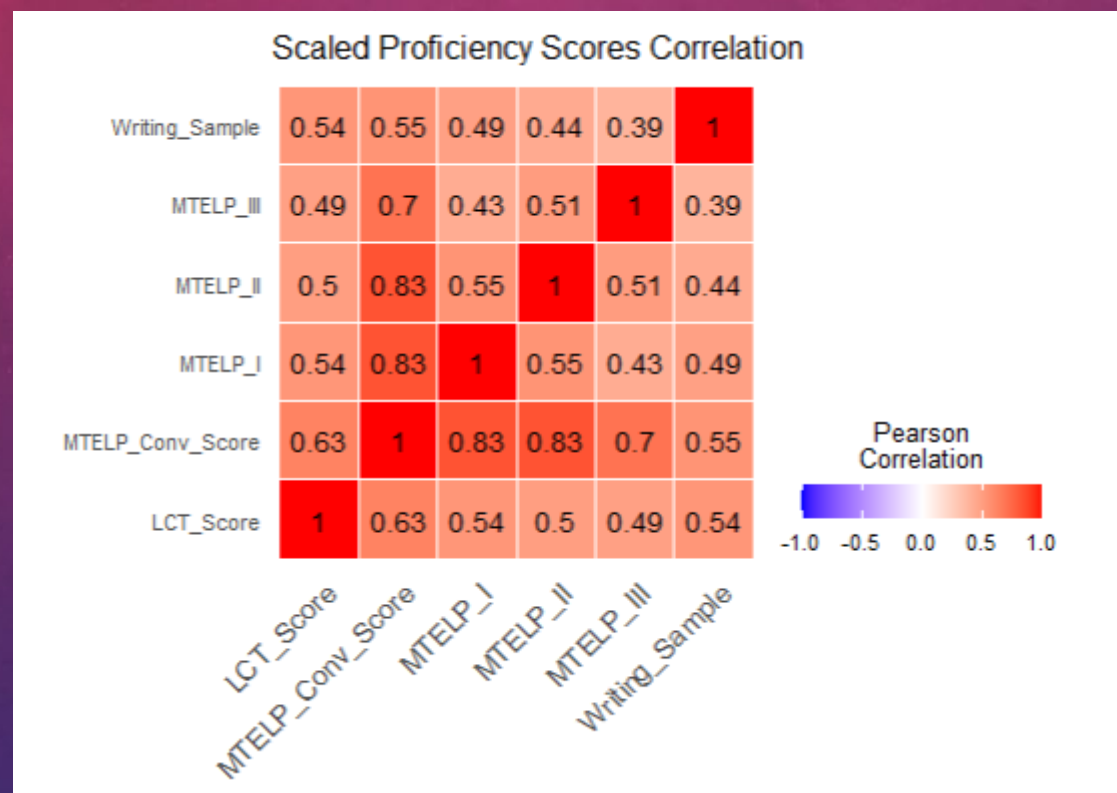
## Proficiency Score Distributions



## Proficiency Score Distributions (more than 30 uses of future tense)

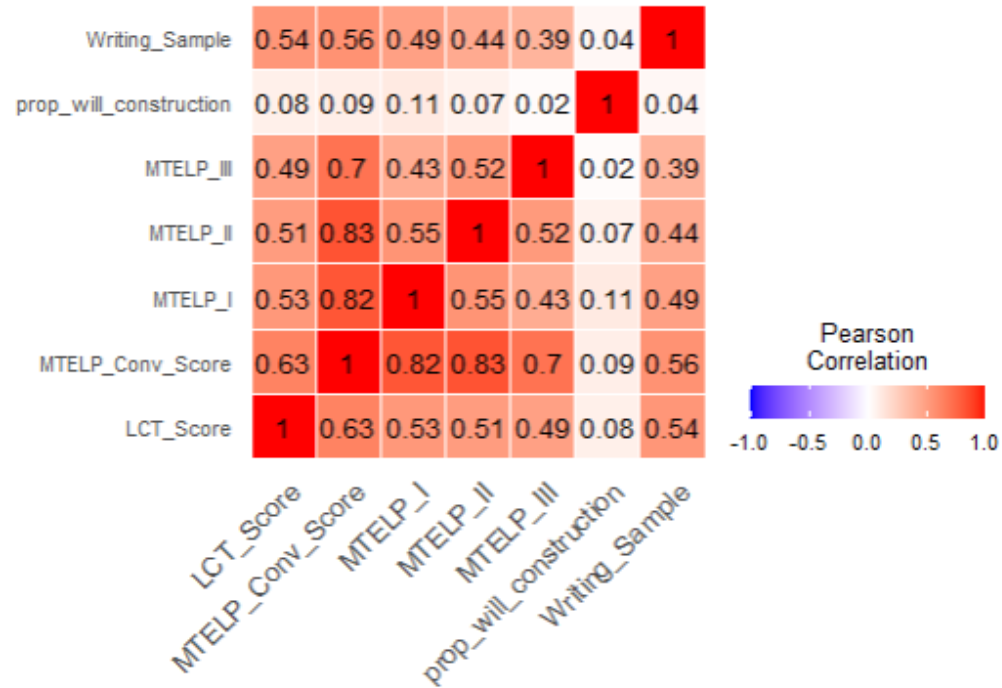


# CORRELATION AMONG PROFICIENCY SCORES

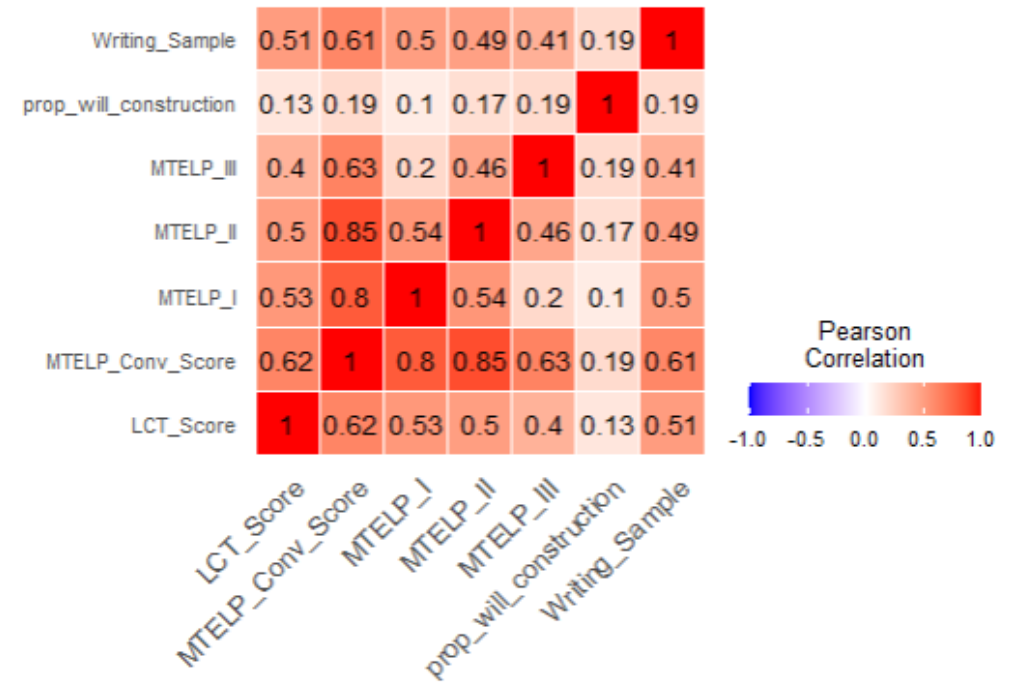




Scaled Proficiency Scores Correlation



Scaled Proficiency Scores Correlation (30 plus)



# LINEAR MODELS – PROPORTION OF ‘WILL’

## Residuals:

Min	1Q	Median	3Q	Max
-0.91352	-0.04394	0.08376	0.12017	0.19061

## Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	0.877726	0.006548	134.035	<2e-16 ***
LCT_Score	0.010937	0.008851	1.236	0.217
MTELP_I	0.020962	0.014941	1.403	0.161
MTELP_II	0.004578	0.014347	0.319	0.750
MTELP_III	-0.012149	0.010744	-1.131	0.258
MTELP_Conv_Score	0.003907	0.026926	0.145	0.885
Writing_Sample	-0.007985	0.008211	-0.973	0.331

# LINEAR MODELS – PROPORTION OF ‘WILL’ – 30PLUS

Residuals:

Min	1Q	Median	3Q	Max
-0.60370	-0.02259	0.03388	0.07260	0.14293

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	0.9048205	0.0091370	99.028	<2e-16 ***
LCT_Score	0.0001043	0.0120123	0.009	0.993
MTELP_I	-0.0148313	0.0278126	-0.533	0.595
MTELP_II	-0.0041690	0.0246332	-0.169	0.866
MTELP_III	0.0049430	0.0197463	0.250	0.803
MTELP_Conv_Score	0.0274367	0.0510778	0.537	0.592
Writing_Sample	0.0130082	0.0118877	1.094	0.275



# LINEAR MODEL – MTELPI, LCT, WRITING SAMPLE

## Residuals:

Min	1Q	Median	3Q	Max
-0.91222	-0.04219	0.08376	0.12000	0.17888

## Coefficients:

	Estimate	Std. Error	t value	Pr(> t )	
(Intercept)	0.877670	0.006544	134.123	< 2e-16	***
LCT_Score	0.009151	0.008331	1.099	0.27225	
Writing_Sample	-0.008407	0.008070	-1.042	0.29781	
MTELP_I	0.022657	0.008153	2.779	0.00556	**

# CONCLUSION

1. There is an overwhelming preference for 'will' construction over 'going to' construction
2. This investigation found no distinct correlation in proficiency exam scores and construction preference
3. Naturally, there was high correlation among Proficiency Exam scores
4. The grammar section of the MTELP was more of a predictor than any other section, as well as combined score.

# FUTURE DIRECTION

- Longitudinal (3% of students had proficiency measured more than once)
- Considerations for question/prompt topics – What was the need for future tense
- Constructions in speech



# REFERENCES

- Juffs, A., Han, N-R., & Naismith, B. (2020). The University of Pittsburgh English Language Corpus (PELIC) [Data set]. <http://doi.org/10.5281/zenodo.3991977>
- Bardovi-Harlig, K. (2005). Proceedings of the 7th Generative Approaches to Second Language Acquisition Conference (GASLA 2004) , ed. Laurent Dekydtspotter et al., 1-12. Somerville, MA: Cascadilla Proceedings Project.
- Ben Kubacki