

Exploring the order of acquisition of **lexical collocations** among Chinese learners of English: A cross-sectional study

Jing Xu

Cambridge Assessment English

BAAL TEA SIG, 29 March 2019





Collocates of 'Conference'

attend

annual

participants

at

hold

regional

conference

venue

in

well-attended

sponsor

during

proceedings

Collocates of 'Brexit'

endless Brexit negotiations

continue

Construct Definition of ‘Lexical Collocation’

Co-occurrence (e.g. Firth, 1957; Sinclair, 1991)

Cause (v.)

- confusion
- rift
- anger
- frustration
- a kidney stone

Semantic prosody (e.g. Carter, 1998; Firth, 1957)

Holistic storage & processing (e.g. Conklin & Schmitt, 2012)

Conveying precise meaning (e.g. Moon, 2008, p. 243)

“[Words] are interconnected, not isolates ... meaning is derived from context, and ... collocation is key.”

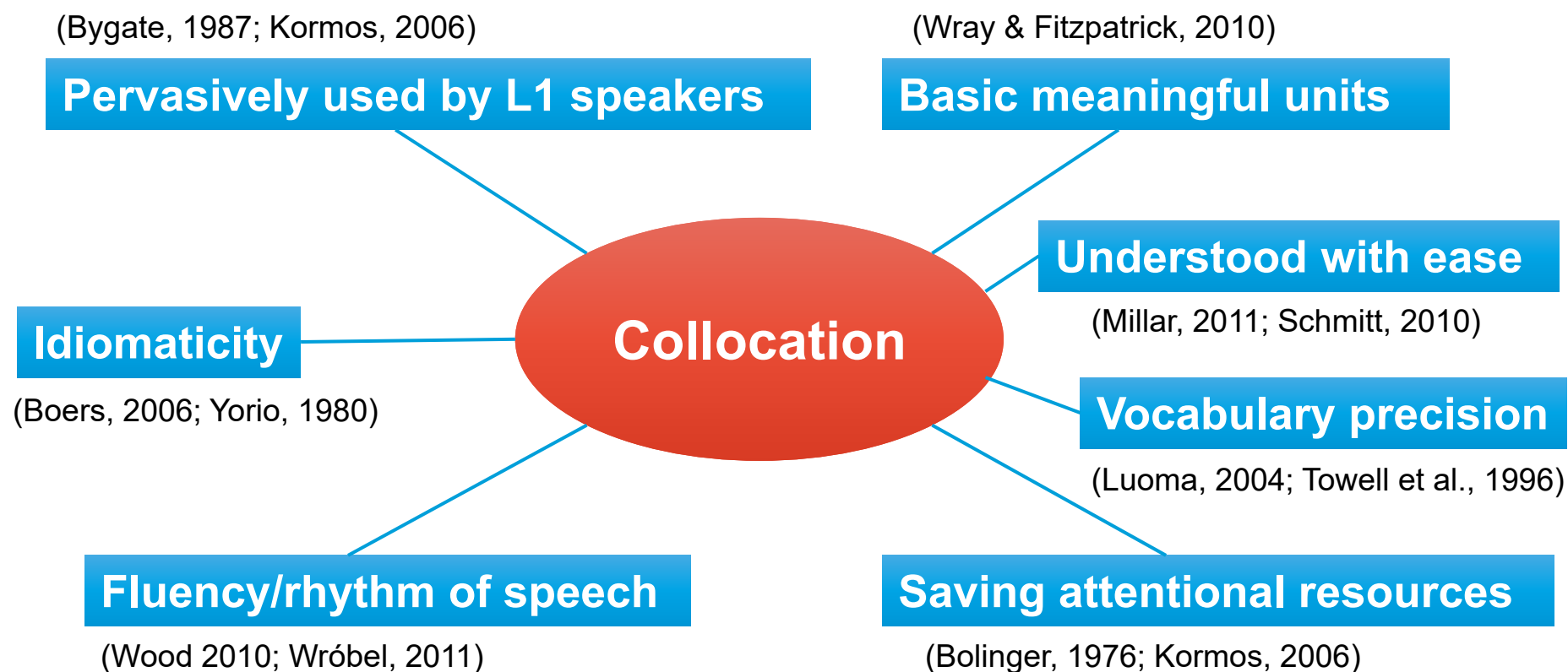
Examples of L2 Collocational Errors

| Native speakers | L2 learners |
|-----------------------|------------------------|
| Fight fiercely | Fight severely |
| Eating habits | Food habits |
| Gain knowledge | Study knowledge |
| Make plans | Gather plans |

Previous Research on L2 Collocation Acquisition

- A positive relationship between L2 collocational knowledge and general L2 proficiency (e.g. Al-Zahrani, 1998; Bonk, 2000; Gitsaki, 1999; Hsu, 2007; Keshavarz & Salimi, 2007)
- Collocation patterns vary in difficulty
 - Lexical more difficult than grammatical (Gitsaki, 1999)
 - Adverb-adjective the easiest (Martyńska, 2004)
- Limitation: No spoken data

The Role of Collocation in Speaking



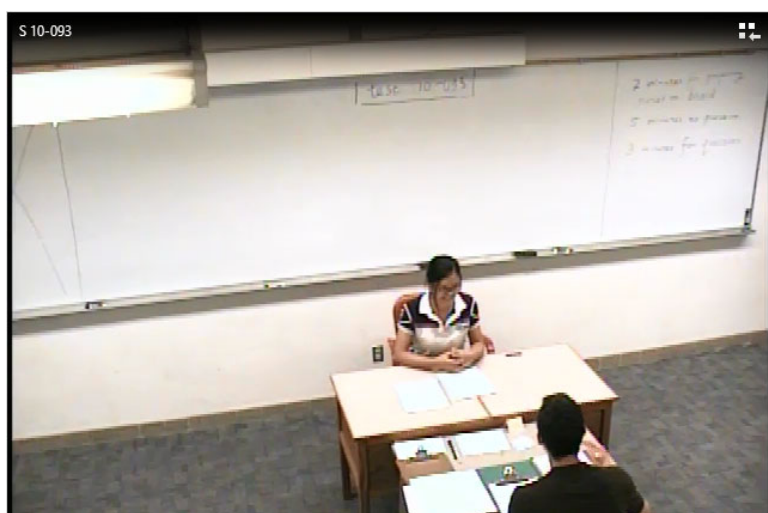
Research Questions (RQs)

1. Do L2 learners of higher oral proficiency demonstrate better lexical collocation performance in free speaking than those of lower oral proficiency?
2. Do L2 learners' lexical collocation performance in free speaking vary across different speaking contexts?
3. What is likely the order of lexical collocation acquisition among Chinese learners of English?

Methodology

Spoken Data

- SPEAK and TEACH exams (Iowa State University)
- Stratified sampling from 4 oral proficiency levels
- 60 Chinese candidates
- Double transcription



Methodology

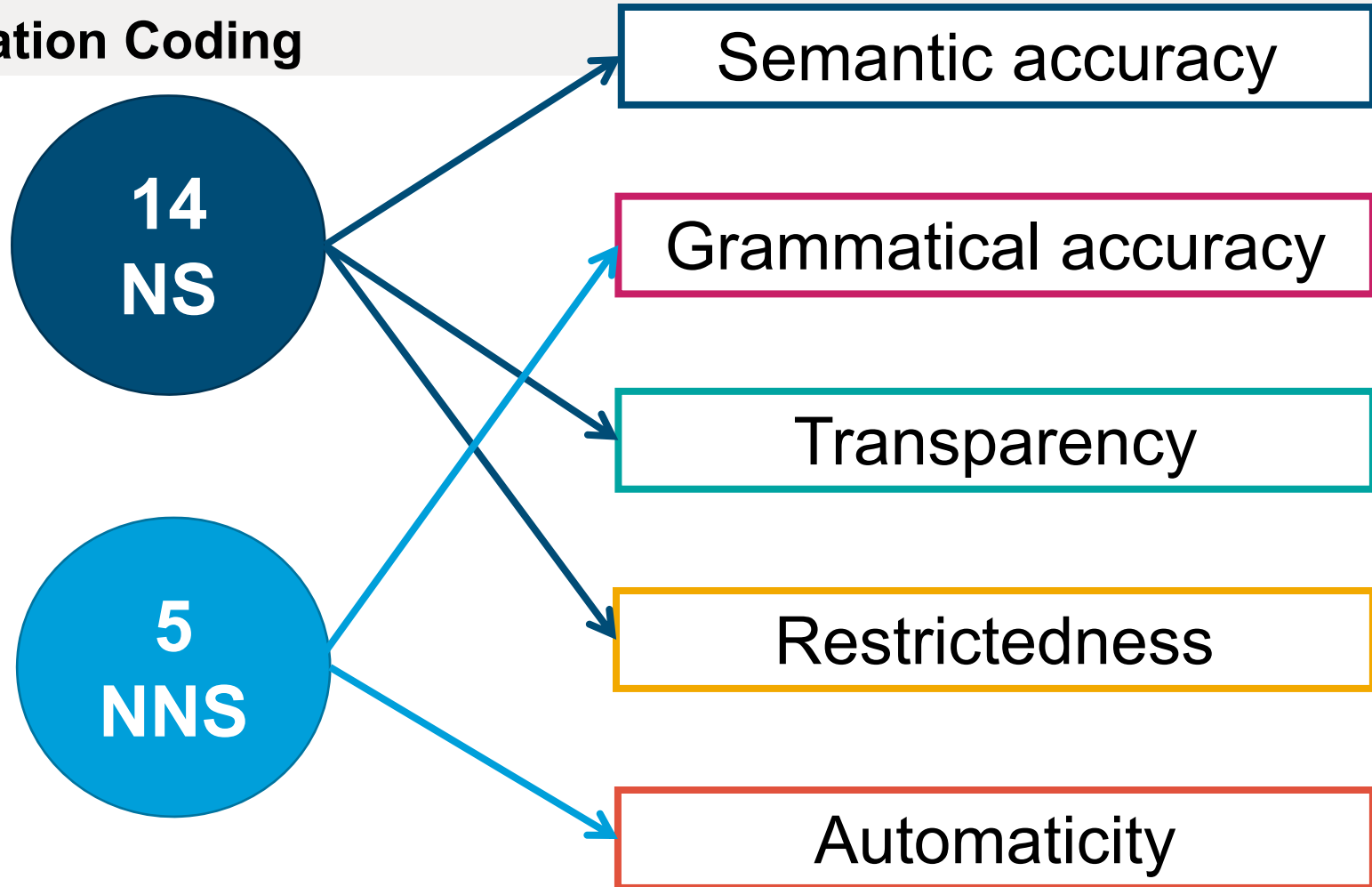
Collocation Identification

| POS Pattern | SPEAK (general) | TEACH (academic) |
|--------------------|------------------------|-------------------------|
| ADJ-N | social network | logical connection |
| ADV-ADJ | equally important | closely related |
| ADV-V | privately owned | randomly selected |
| N-N | college education | energy intake |
| N-of-N | tail of a kite | steepness of a line |
| N-V | a plane lands | demand increases |
| V-N | fly a kite | consume oxygen |
| PHV-N | put out a fire | cut down cost |



Methodology

Collocation Coding



Semantic Accuracy/Prosody

| Coding | Examples |
|------------------|---|
| 1 = Unacceptable | I <u>applied for doctor</u> to study Biology. Correction: I applied for admission to a doctoral program in Biology |
| 2 = Substandard | I <u>applied for her permission</u> to date her. Correction: I tried to ask her out. |
| 3 = Native-like | The preserve today maintains a limited access status, which means an individual must <u>apply for permission</u> to enter the preserve. |

Grammatical Accuracy

Coding

1 = Erroneous

Examples

Determiner error:
Sometimes I think, maybe
made mistake.

Word-form error: China's
ongoing experimentals with
election are successful.

Function-word error: My
father is a more heavy
smoker than me.

0 = Error free

Transparency

Coding

1 = Partially figurative

0 = Literal

Examples

Blow a chance

Blow a trumpet

Restrictedness/Precision

Coding

Examples

Highly restricted

It is almost impossible to find a substitute for the modifier in the collocation

Blow a chance
an act of violence
Bees buzz

Moderately restricted

The modifier may be substituted with a small number of words

Deeply (closely, heavily,
intimately, very much)
involved

Automaticity

Coding

0 = Uninterrupted utterance

1 = Interrupted utterance

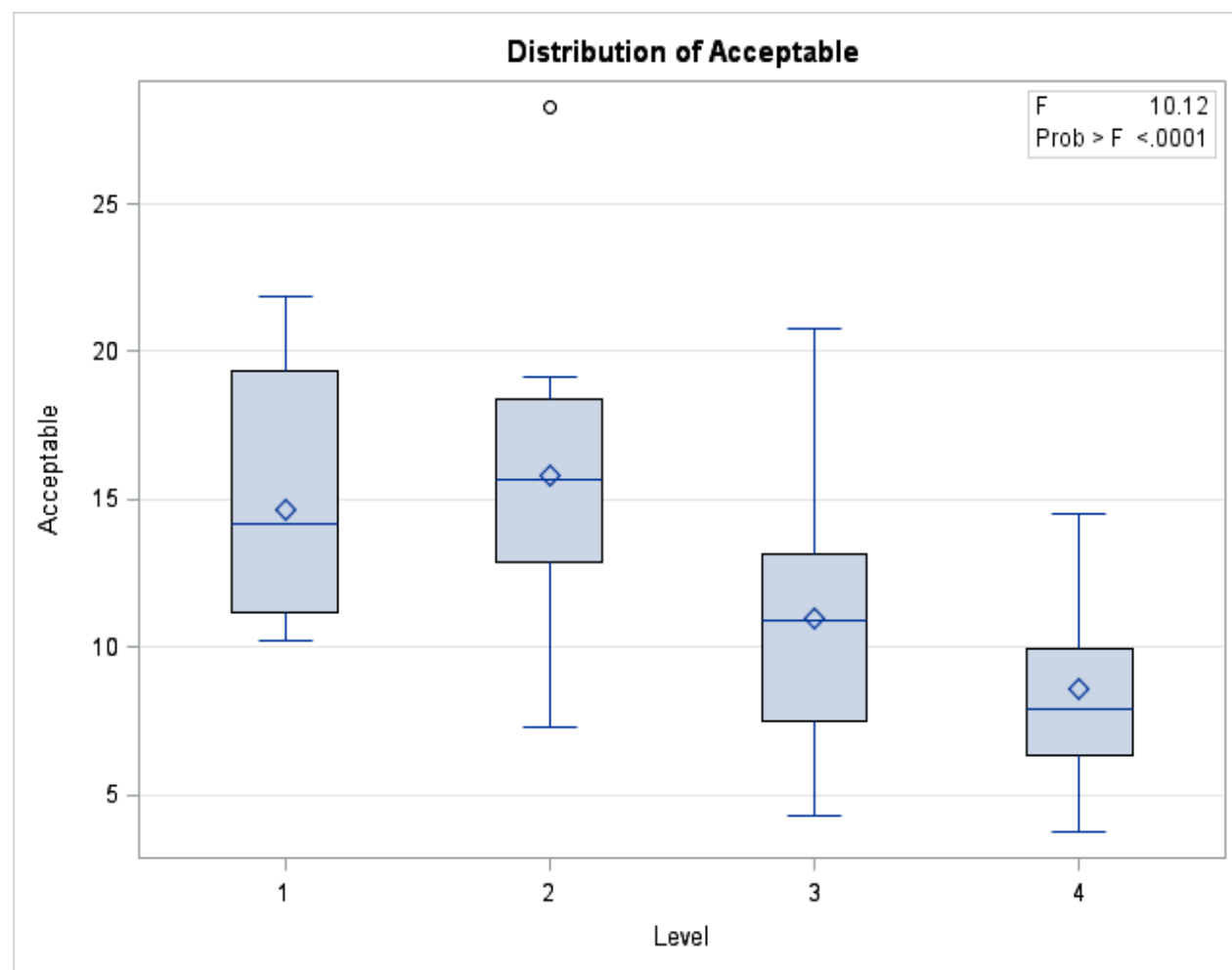
Examples

So you should pay attention to that especially during the midterm exam.

Ok so, first I think the government should set up a very strict rule **to uh to- to-ban** the air pollution

RQ1: Differences in Collocations across Levels

Normalised Frequency of **Native-like Collocations** by Levels



$F(3, 56) = 10.12, p < .01,$
 $\eta^2 = .36$

Post-hoc Scheffé:

Level 1 > Level 4, $p < .01$

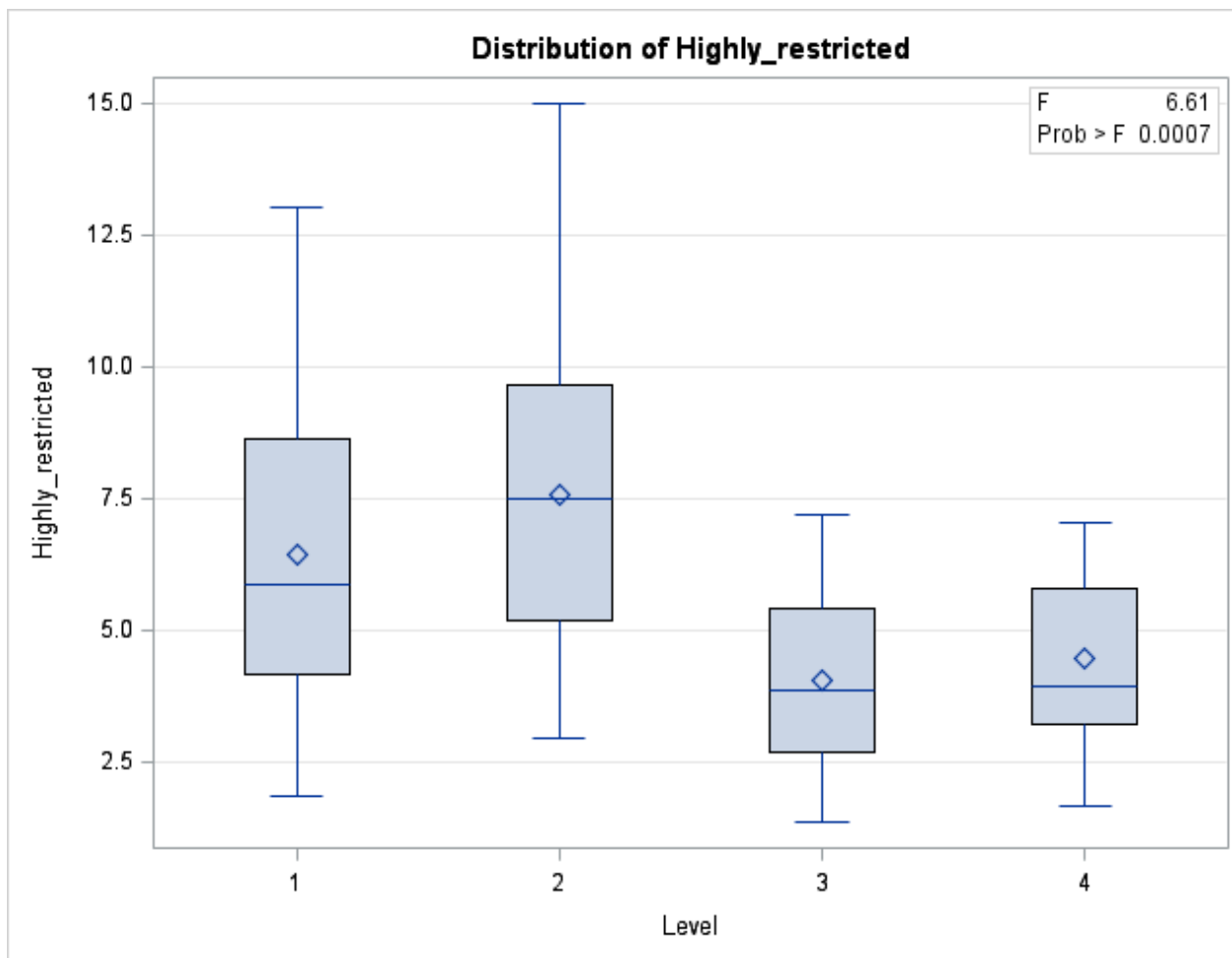
Level 2 > Level 3, $p < .05$

Level 2 > Level 4, $p < .01$



RQ1: Differences in Collocations across Levels

Normalised Frequency of **Restricted Collocations** by Levels



$$F(3, 56) = 6.61, p < .01, \eta^2 = .27$$

Post-hoc Scheffé:

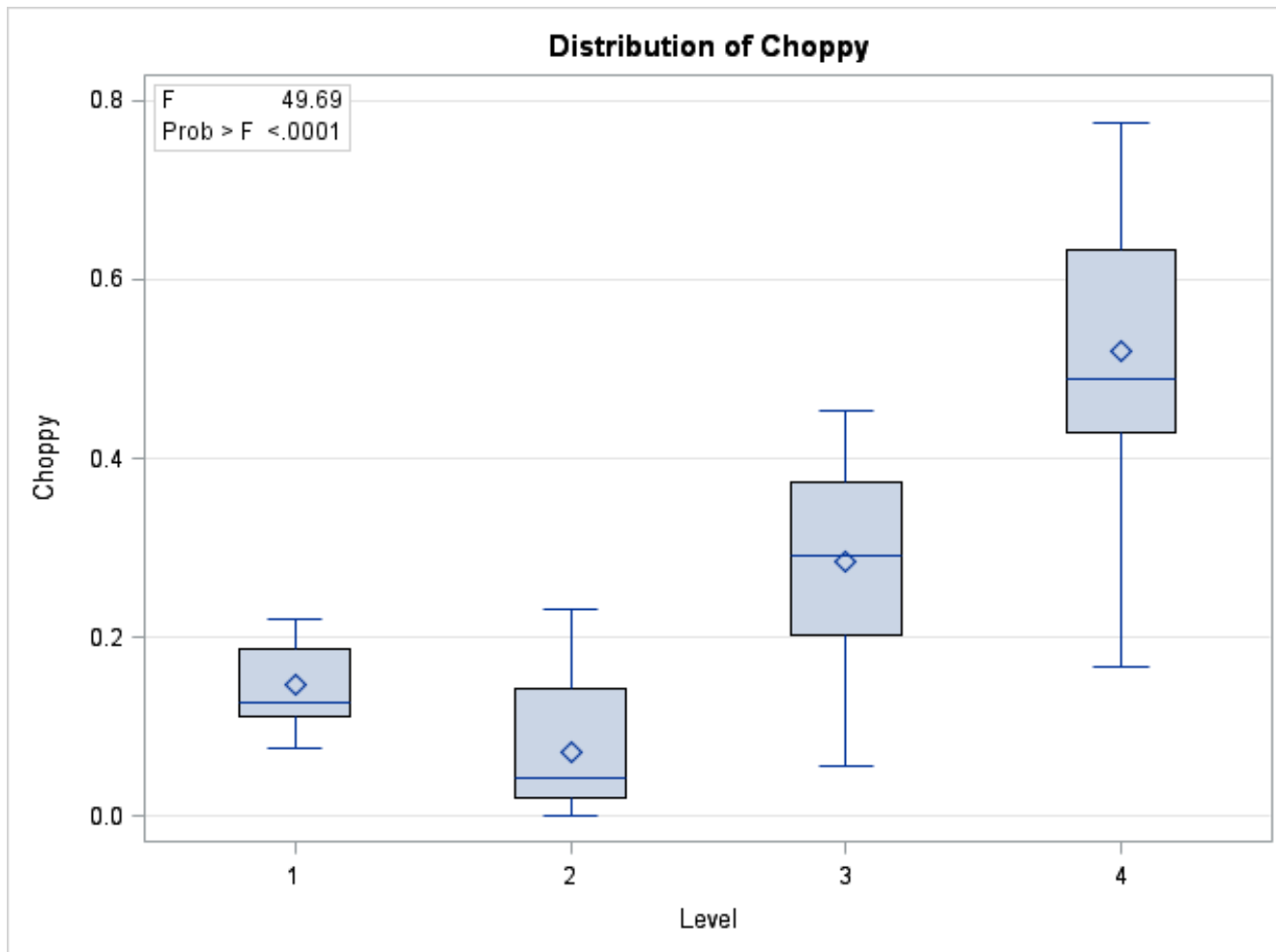
Level 2 > Level 3, $p < .05$

Level 2 > Level 4, $p < .05$



RQ1: Differences in Collocations across Levels

Normalised Frequency of **Choppy Collocations** by Levels



$F(3, 56) = 49.69, p < .01,$
 $\eta^2 = .69$

Post-hoc Scheffé:

Level 1 < Level 3, $p < .01$

Level 1 < Level 4, $p < .01$

Level 2 < Level 3, $p < .01$

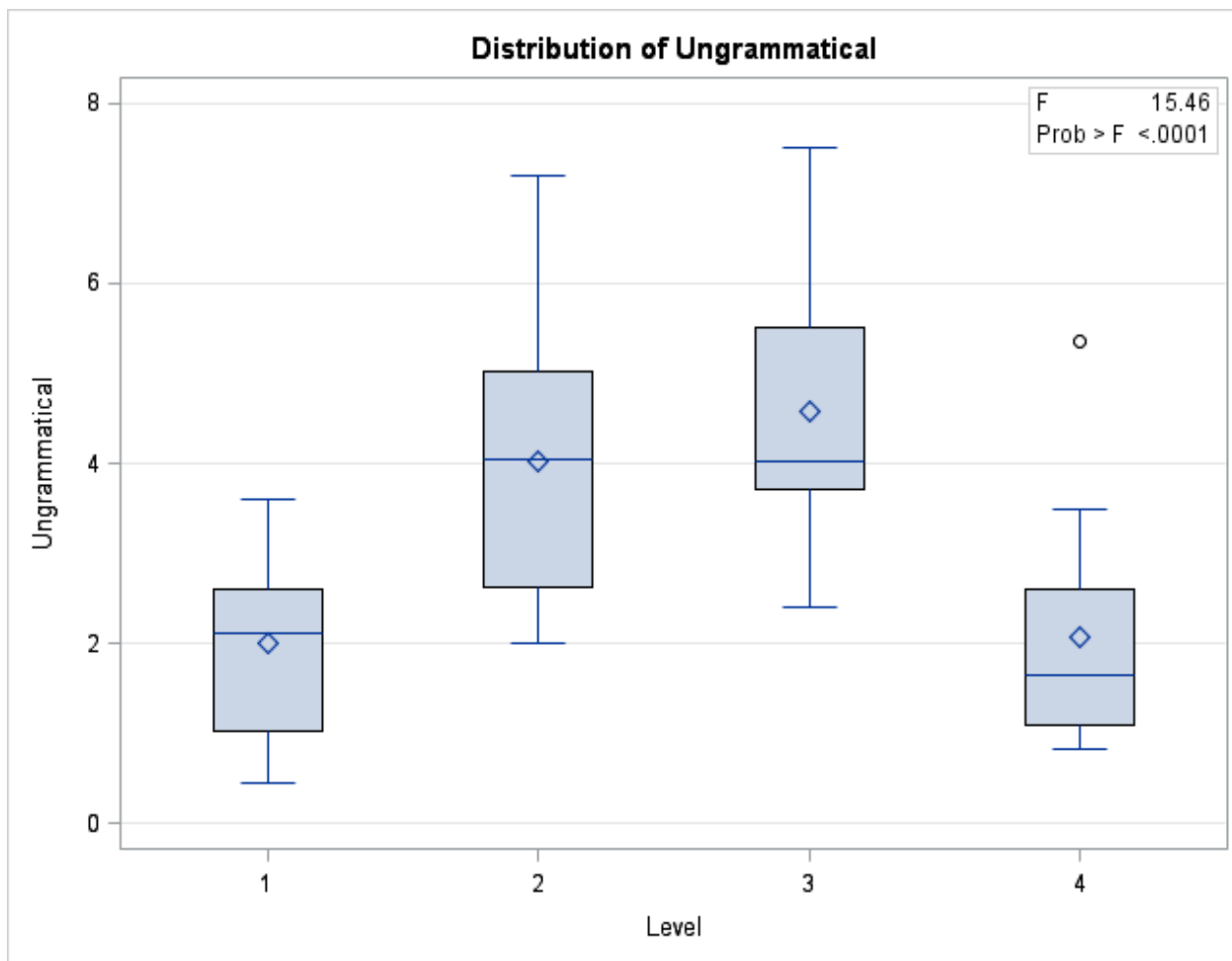
Level 2 < Level 4, $p < .01$

Level 3 < Level 4, $p < .01$



RQ1: Differences in Collocations across Levels

Normalised Frequency of **Ungrammatical Collocations** by Levels

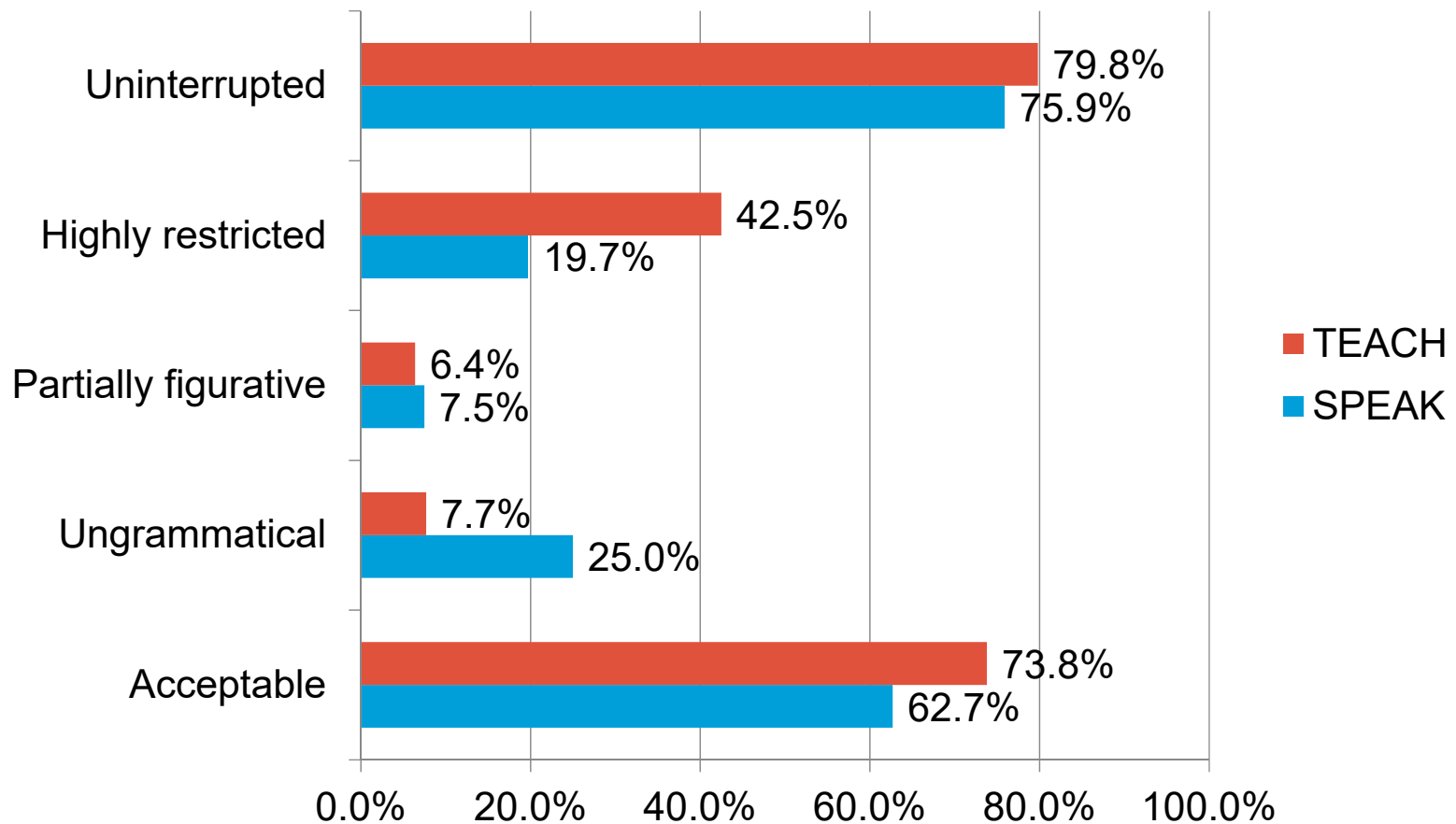


$F(3, 56) = 15.46, p < .01,$
 $\eta^2 = .46$

Post-hoc Scheffé:

Level 1 < Level 2, $p < .01$
Level 1 < Level 3; $p < .01$
Level 2 > Level 4, $p < .01$
Level 3 > Level 4, $p < .01$

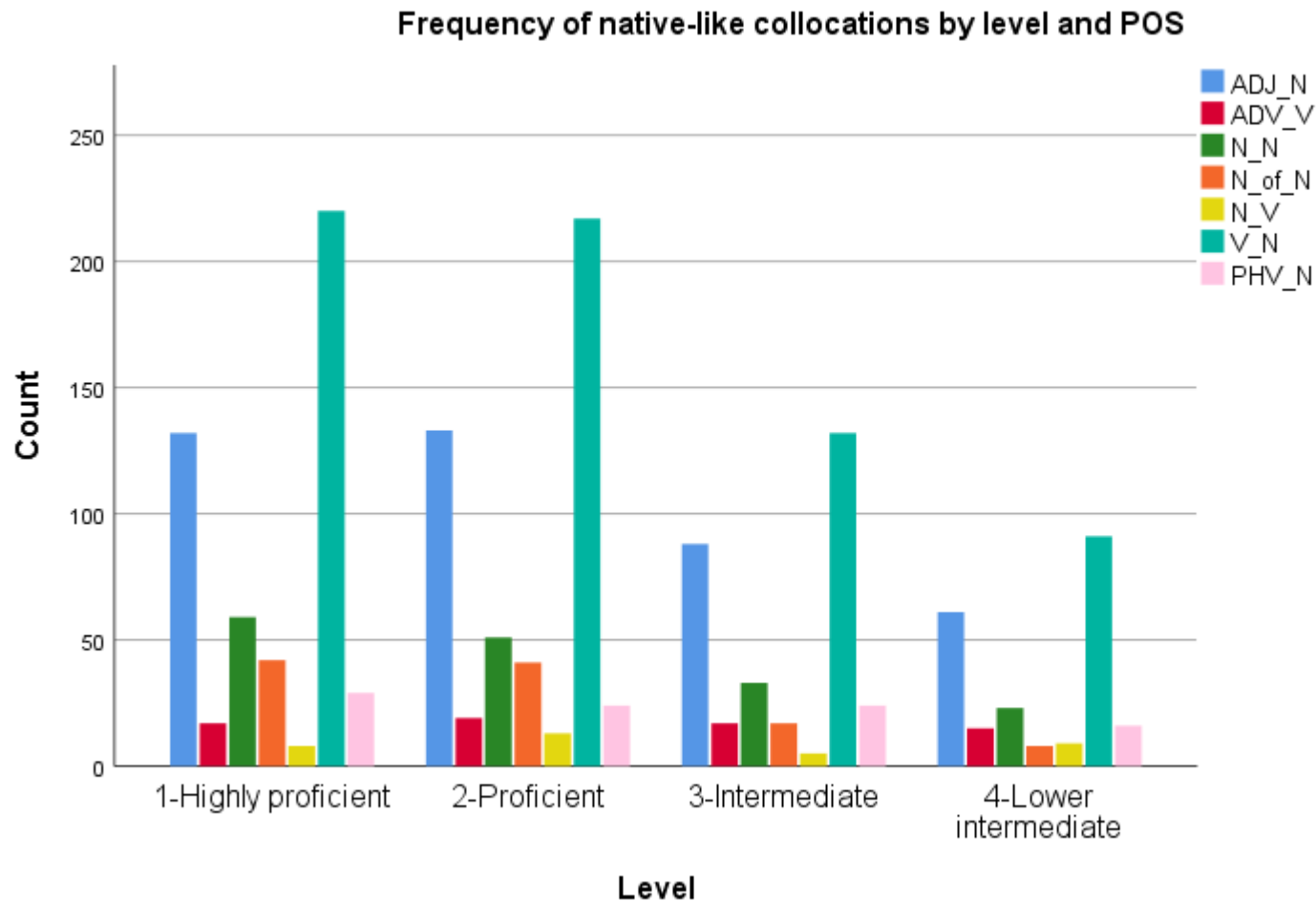
RQ2: Differences in Collocations across Speaking Contexts





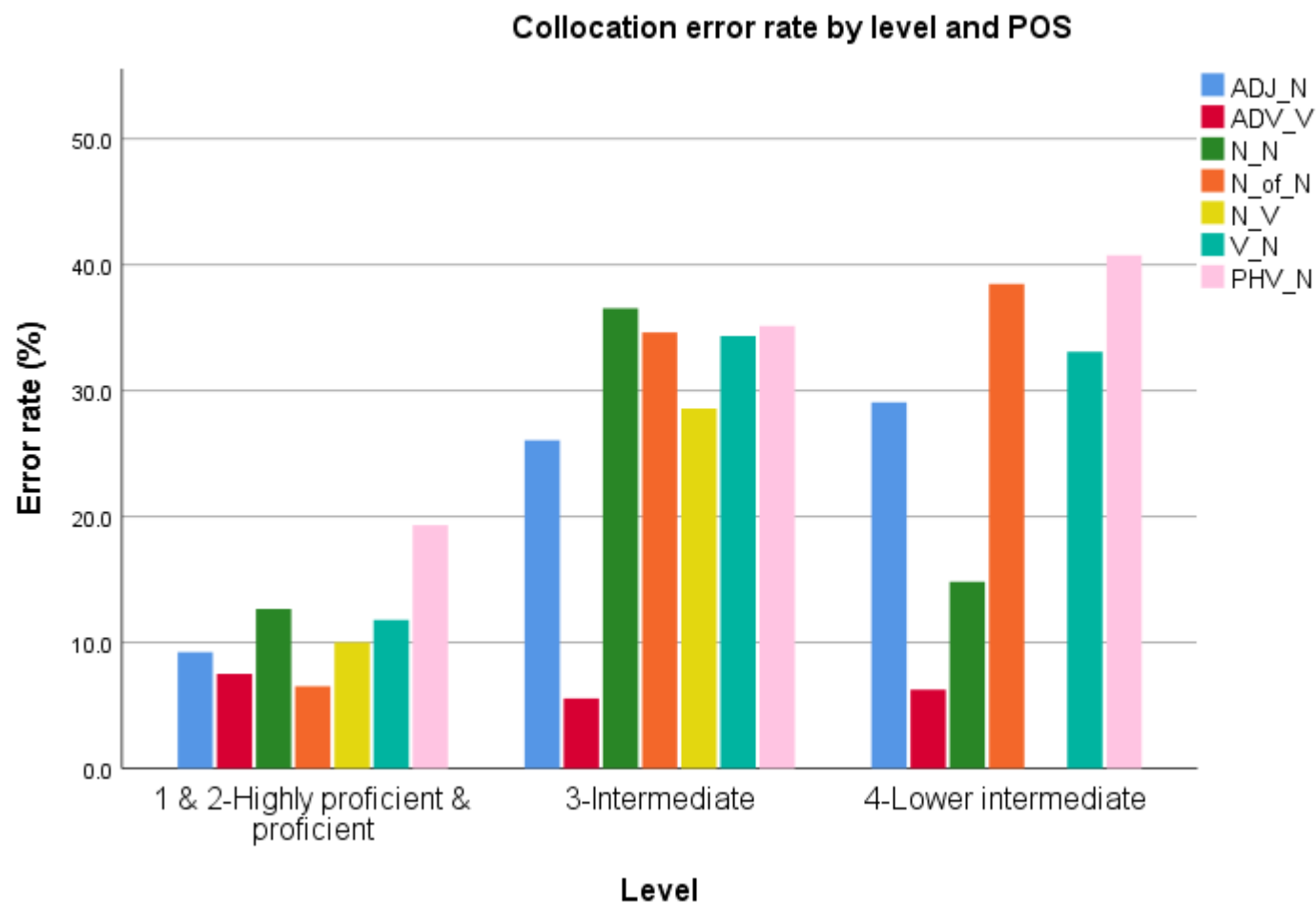
RQ3: Order of Acquisition

Distribution of Native-like Collocations by Level and POS



RQ3: Order of Acquisition

Error Rate by Level and POS



Conclusion

1. Evidence of varied oral collocation performance across proficiency levels
2. Evidence of varied oral collocation performance across speaking contexts
3. Order of acquisition
 - PHV-N not acquired even at the advanced level
 - ADJ-N, V-N, N-N, N-of-N, N-V likely acquired at the advanced level
 - ADV-V acquired first or the easiest

Implications

1. L2 speaking theory



2. Automated speaking assessment
 - Construct expansion
 - Malpractice detection

3. Automated feedback for learning

Directions for Future Research

1. Cross-validation with native spoken corpora
2. Speakers of other L1s

Acknowledgements

-  Small Grants for Doctoral Research
-  Academic Communication Program
- Transcribers and annotators

Alicia Christy

Audrey Risius

William Lewis

Erik Voss

Jim Ranalli

Erin Todey

Jenny Anderson

Stephanie Link

Todd Paben

Zhi Li

Mandy Qian

Mo Chen

Hong Ma

Ivon Katz

Dan Douglas

Felicity Douglas

Kimberly Levelle

Jesse Gleason

Relevant Articles

- Xu, J. (2015). *Predicting ESL learners' oral proficiency by measuring the collocations in their spontaneous speech*. Unpublished doctoral dissertation. Iowa State University. Ames, IA.
- Xu, J. (2018). Measuring “spoken collocational competence” in communicative speaking assessment. *Language Assessment Quarterly*, 15(3), 255-272.