

# Investigating the lexical bundles used in an email writing task in an English proficiency test

Li Zhi & Alex Volkov  
Paragon Testing Enterprises, Inc.

# Agenda

- Background
- Methodology
- Results
- Conclusions and Implications

# About Paragon Testing Enterprises

- A subsidiary of the University of British Columbia
- The developer of
  - The Canadian English Language Proficiency Index Program (CELPIP) Test
  - The Language Proficiency Index (LPI)
  - The Canadian Academic English Language (CAEL) Assessment (starting 2015)

# Current project

## Motivation

Provide information for LINC and ESL classroom in terms of the use of lexical bundles across performance levels in writing.

Do test-takers of various proficiency levels use lexical bundles differently **in terms of their functions/ syntactic structure/use throughout the emails?**

Research  
questions

Corpus

Selection  
criteria

Cleaning  
the list

Coding



TESTING ENTERPRISES  
**Paragon**

# Writing an email

## Writing Task 1: Writing an Email

You recently made reservations for dinner at a very famous and expensive restaurant in town. However, the meal and the service were terrible. The restaurant manager was not available to solve the problem, so you left without a resolution.

**Write an email to the restaurant's manager in about 150-200 words. Your email should do the following things:**

State what problems you had with the food you ordered.

Complain about the service.

Describe how you want the restaurant to resolve the problem to your satisfaction.



# Lexical Bundles

*Lexical phrases, lexical bundles, word strings, word sequences, formulaic sequences.*

Should the units be meaningful? Should they be structurally complete?

Biber and Barbieri (2006): “most frequently occurring sequences of words, usually not structurally complete and not idiomatic in meaning.”

# Target Corpus

Proficiency level	# of texts	# of words	# of tokens	Average length
CELPIP 4	2500	392,625	13,505	157
CELPIP 7	2500	482,605	13,773	193
CELPIP 10	2500	482,681	14,989	193

Frequency and range

Biber and Barbieri (2006): 40 times per million, in 20 texts

Challenges of our corpus

Antconc 3.4.4w

Research  
questions

Corpus

Selection  
criteria

Cleaning  
the list

Coding



TESTING ENTERPRISES  
**Paragon**

# Cleaning the list

- **Combined Bundles**

*from you kind regards (I look forward to hearing from you. Kind Regards);*

*may concern i am (To whom it may concern. I am...)*

- **Prompt Specific Units**

*daughter's birthday party; dear mr smith I*

- **Length of Bundles**

*you in advance for (Thank you in advance for);*

*hesitate to contact me (Do not hesitate to contact me)*

- **Bundles as Whole Clauses**

*To whom it may concern; Thank you for taking the time to read this email*

Research  
questions

Corpus

Selection  
criteria

Cleaning  
the list

Coding



TESTING ENTERPRISES  
**Paragon**



# Qualitative and Quantitative Steps

- Functions (Biber et al., 2004; Cortes 2004)
  - Stance: *I am unable to, I am sure you,*
  - Referential: *in regard to the, at the same time*
  - Discourse organizer: *I am writing to, to let you know*
  - Other: *thank you very much, to whom it may concern*
- Syntactic structures (Chen & Baker, 2014)
  - Verb phrase-based: *if you need any*
  - Noun phrase-based: *the end of the*
  - Preposition phrase-based: *for a long time*

## Coding Agreement:

Level 10	93%
Level 7	88%
Level 4	85%
Average	89%

Research  
questions

Corpus

Selection  
criteria

Cleaning  
the list

Coding



TESTING ENTERPRISES  
**Paragon**

# Results – Overview

	Total types	Total tokens
CELP 4	82	3957
CELP 7	99	6257
CELP 10	95	6350

Results  
Overview

Results  
RQ1

Results  
RQ2

Results  
RQ3

Results  
RQ4



TESTING ENTERPRISES  
**Paragon**

# Results – Functions of Lexical Bundles

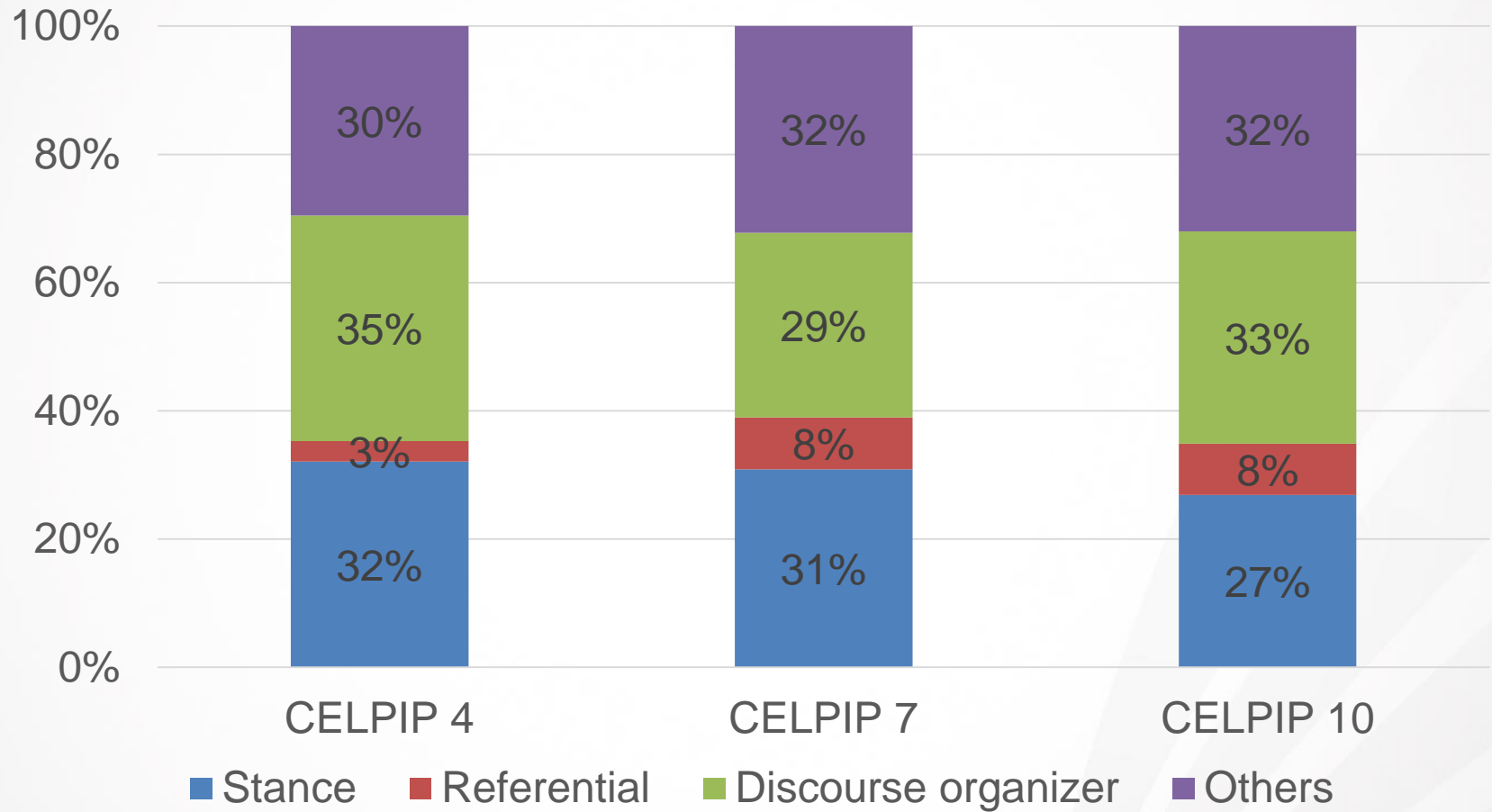
- Functions of lexical bundles across proficiency levels (token)

	<b>Stance</b>	<b>Referential</b>	<b>Discourse organizer</b>	<b>Others</b>
CELPPIP 4	1269 (32%)	126 (3%)	1393 (35%)	1169 (30%)
CELPPIP 7	1932 (31%)	506 (8%)	1802 (29%)	2017 (32%)
CELPPIP 10	1706 (27%)	507 (8%)	2105 (33%)	2032 (32%)



# Results – Functions of Lexical Bundles

(tokens in percentage)



Results  
Overview

Results  
RQ1

Results  
RQ2

Results  
RQ3

Results  
RQ4



TESTING ENTERPRISES  
**Paragon**

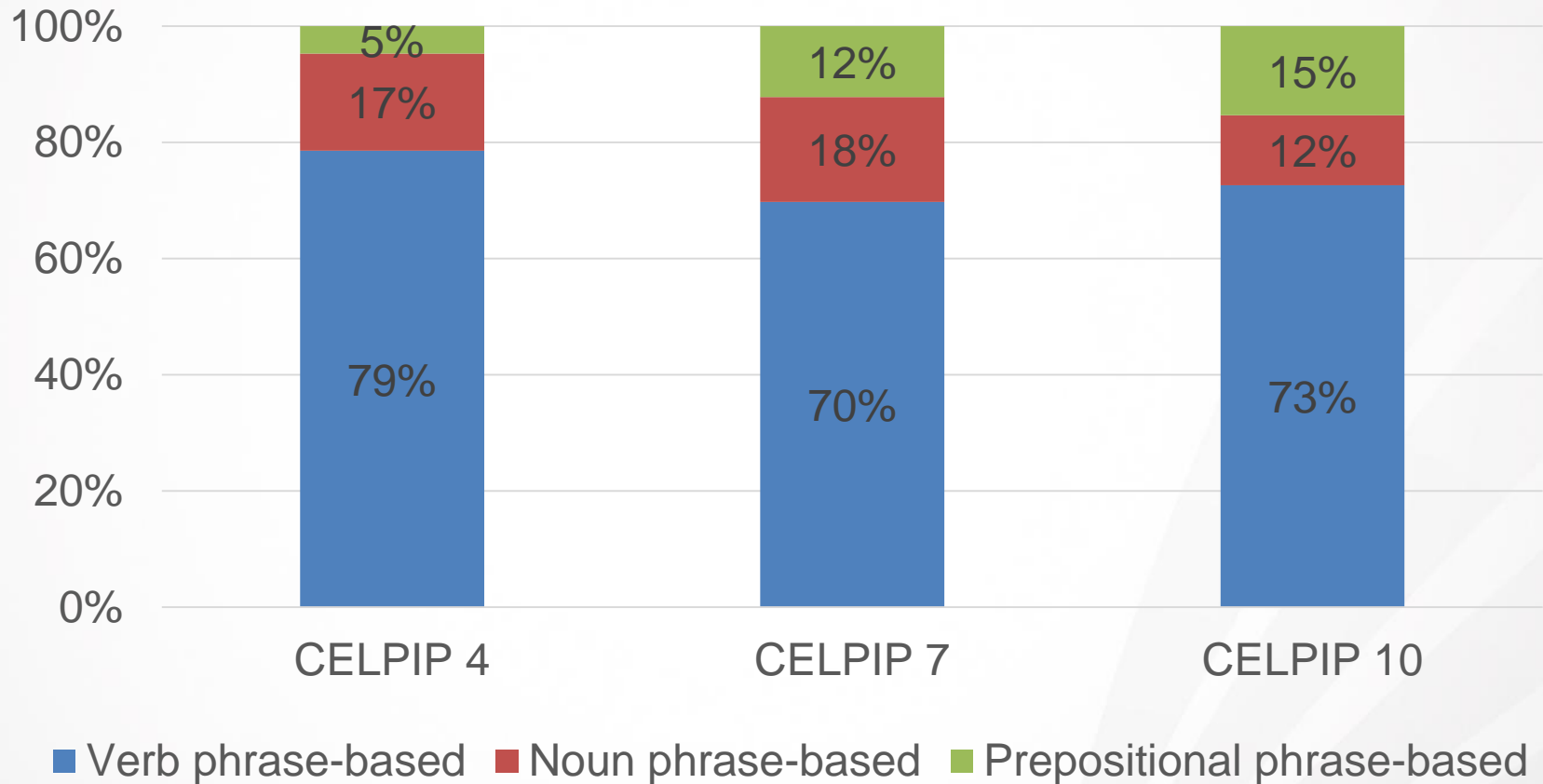
# Results – Syntactic Structures of Lexical Bundles

- Syntactic structures of lexical bundles across proficiency levels (token)

	<b>Verb phrase-based (VP)</b>	<b>Noun phrase-based (NP)</b>	<b>Preposition phrase-based (PP)</b>
CELP 4	3110 (79%)	660 (17%)	187 (5%)
CELP 7	4366 (70%)	1129 (18%)	762 (12%)
CELP 10	4613 (73%)	764 (12%)	973 (15%)



# Results – Syntactic Structure of Lexical Bundles (token in percentage)



# Results – Lexical bundles in email sections

- Opening
  - *Dear sir or madam, Hope this email finds you well, Good day*
- Stating the purposes
  - *I am writing to, to bring to your attention*
- Body of message
  - *I would appreciate it if you could, may I suggest that, I need your help*
- Closing
  - *Thank you for your time, Kind Regards, I look forward to*



# Results – Lexical bundles in email sections

	CELP 4	CELP 7	CELP 10
<b>Opening</b>			
<b>Addressing the reader &amp; introduce self</b>	Dear sir or madam	Dear Sir or Madam	Dear sir or madam
	To whom it may concern	To whom it may concern	To whom it may concern
	Good day		Hope this email finds you well
	How are you		
	My name is	My name is	My name is
<b>Starting Purposes</b>			
<b>Introducing the purpose</b>	First of all I	First of all I	I am writing this letter/email
	I am writing this letter because	I am writing to VP	I am writing to (you today)
	I am writing to you	I am writing in order to	It has come to my attention that
	I am writing to Verb	I am writing this email/letter	I would like to take this opportunity to
		I am writing to you to	
	To inform you Prep	The reason why I	To inform you of/that
	To let you know that/about	To bring to your notice/attention	To let you know that
		To let you know that	To thank you for
		To tell you that	To bring to your attention
		To complain about the	To express my concern
		To inform you about	In regards to the
		To thank you for	I am writing to you regarding
		have noticed that	have noticed that
<b>Body of Message</b>			
<b>Requesting action</b>	I want you to Verb	I would like to ask/request you about	I would like you to
	I would like you to Verb	I think you should	I would appreciate your
	I need your help	I would really appreciate if you could	I would really/greatly appreciate if
	You can help me	I would like you to	I would be grateful
		I would request you to	(Please) let me know if/what
		I want you to	I was wondering if
<b>Offering information/ action</b>		I request you to	
	Please let me know	Please let me know (if you)	If you need any
			If you would like
<b>Expressing desire</b>			I am/would be happy to
	I would like to VP (suggest)	I would like/love to VP	I would like/love to
	I want to Verb	I want to	I would like NP
<b>Making suggestions</b>		It would be great	It would be great
		If it is possible	Would be a great
			May I suggest that
			I would suggest that
			I think it would be
			Would it be possible
<b>Expressing hope</b>	I hope you can/will	I hope you Verb	If you would be
	I hope you Verb (understand my situation)	I hope that you Verb	I hope you can/will (consider)
		I am hoping for	I hope that you
<b>Other bundles</b>	At the same time	I am looking for	I am hoping that
	As soon as possible	I am very/really sorry	I am sure that
	A lot of people	As soon as possible	As you are aware
	Is very important to	For a long time	For the past years
	I am very happy	On the other hand	In the (near) future
	I am sorry	As soon as I	At the same time
		One of the best	As soon as possible
		In the middle of	As a result of
		I am (pretty) sure that	As a result I
		Is very important to	At your earliest convenience
		At the same time	In the last few months
		Most of the time	For the first time
			In the last few months
<b>Closing</b>			
<b>Showing gratitude</b>	Thank you for reading	Thank you very much/so much	Thank you again for
	Thank you for your Noun	Thank you in advance	Thank you in advance
	Thank you so much/very much	Thank you for your time/kind consideration	Thank you so/very much
	Thank you for taking the time to read this letter/email		Thank you for your time/consideration/understanding

Results:  
Overview

Results  
RQ1

Results  
RQ2

Results  
RQ3

Results  
RQ4



# Results – Lexical bundles & length of residence at CELPIP 7 (n = 700)

less than 1 year	3-5 years	6-10 years
i would like to	i would like to	i would like to
i am writing this	i am writing this	i am writing this
to whom it may	i am writing to	to whom it may
i am writing to	dear sir madam i	as soon as possible
as soon as possible	i am looking forward	i am writing to
dear sir madam i	as soon as possible	thank you very much
thank you very much	to whom it may	i am looking forward
i am looking forward	please let me know	dear sir madam i
to inform you that	thank you for your	to hear from you
to hearing from you	to hearing from you	thank you for your
good day i am	to hear from you	to hearing from you
thank you for your	i hope you will	please let me know
to hear from you	not be able to	not be able to
i am writing you	dear sir i am	to inform you that
to let you know	i am writing you	dear sir or madam
not be able to	at the same time	i look forward to
please let me know	i am going to	to let you know
dear sir or madam	thank you very much	if you have any
first of all i	to inform you that	at the same time
at the same time	thank you so much	thank you so much
thank you so much	to let you know	for your kind consideration
thank you in advance	i would like you	thank you in advance
dear sir i am	first of all i	i hope you will
i just want to	i am planning to	let me know if
will be able to	i look forward to	will be able to

Results  
Overview

Results  
RQ1

Results  
RQ2

Results  
RQ3

Results  
RQ4



TESTING ENTERPRISES  
**Paragon**

# Conclusions

- Somewhat different patterns in the use of lexical bundles of stance function and referential function
- Differences in the use of PP-based lexical bundles across proficiency levels
- Different amounts of lexical bundles used across sections of emails by test-takers of different proficiency levels
- Minimal impact of length of residence on test-takers' use of lexical bundles at intermediate proficiency level (CELPPI 7)

# Implications

- Use lexical bundles to
  - Raise awareness of importance of appropriate expressions
  - Teach formulaic expression in email writing

# Thank you!

# Investigating the lexical bundles used in an email writing task in an English proficiency test

## Questions?

[research@paragontesting.ca](mailto:research@paragontesting.ca)

# References

- Biber, D. (2006). *University Language: A Corpus-based Study of Spoken and Written Registers*. Amsterdam, the Netherlands: John Benjamins.
- Biber, D., & Barbieri, F. (2007). Lexical bundles in university spoken and written registers. *English for Specific Purposes*, 26, 263–286
- Chen, Y.-H., & Baker, P. (2014). Investigating criterial discourse features across second language development: Lexical bundles in rated learner essays, CEFR B1, B2 and C1. *Applied Linguistics*, amu065–. <http://doi.org/10.1093/applin/amu065>
- Cortes, V. (2013). The purpose of this study is to: Connecting lexical bundles and moves in research article introductions. *Journal of English for Academic Purposes*, 12, 33-43.
- Hyland, K. (2008). As can be seen: lexical bundles and disciplinary variation. *English for Specific Purposes*, 27, 4–21
- Hyland, K. (2012). Bundles in academic discourse. *Annual Review of Applied Linguistics*, 32, 150-169.
- O’Keeffe, A., McCarthy, M. & Carter, R. (2007). *From Corpus to Classroom: Language Use and Language Teaching*. Cambridge University Press.
- Sinclair, J. (1991) *Corpus, Concordance and Collocation*. Oxford: Oxford University Press.