

*BAAL 2013*

# **Researching metadiscourse markers in candidates' writing at Cambridge FCE, CAE and CPE levels**

*Stephen Bax (CRELLA)*

*Daniel Waller (UCLAN)*

*Fumiyo Nakatsuhara (CRELLA)*

**CRELLA**

Centre for Research in English  
Language Learning and Assessment



University of  
Bedfordshire



*This paper reports on research funded through the  
Cambridge ESOL Funded Research Programme, Round  
Three, 2012. Thanks!*

# Plan of the talk

---

- # *Metadiscourse – why study it?*
- # *Our research project*
- # *Implications*

# Metadiscourse

---

- # “the language used to express the author’s awareness and management of the discourse-as-process: which includes management of the organisation of the text, of the participants of the discourse process and...[expression of] the author’s attitude towards the discourse process”

(Burneikaite, 2008:39).

# Examples

---

*e.g.            therefore (logical connective)*  
*essential (emphatic)*  
*so (topic shift)*  
*finally (frame marker)*

# Why Metadiscourse?

---

- Metadiscourse provides a method for examining writer awareness of audience & communicative purpose
- Possible role of discourse in distinguishing higher levels (B2, C1 & C2).
  - B2+ (Council of Europe, 2000: p.35) level defined by “a new focus on discourse skills” esp. with regard to “argument and social discourse”.
- Useful key for language testers?

# Why Metadiscourse?

---

*“[The CEFR] makes specific predictions about the use of such discourse connectives in learner language .....*

*The CEFR predicts that the range of different connectives **expands across proficiency levels**, that more advanced learners make use of less frequent connectives than learners at lower levels, and that learners gain increased control of connectives as they progress.”*

*(Carlsen 2010:191-2, emphasis added)*

# Why Metadiscourse?

---

Shaw and Weir noted with regard to discourse features:

“this parameter might contribute to further grounding of distinctions between levels FCE, CAE and CPE” (Shaw & Weir, 2007: 251).



# Risky conjectures

---

“It is predicted that student narratives that received higher ratings on the TOWL-3 [test] will contain more metadiscourse elements using the classification developed by Hyland (2005), and that fewer metadiscourse markers will be found in narratives that received lower ratings”

(Sandford 2012:13)

**Table 6.1** Functions of metadiscourse in academic texts

Category	Function	Examples
<i>Textual metadiscourse</i>		
Logical connectives	Express semantic relation between main clauses	in addition / but / thus / and
Frame markers	Explicitly refer to discourse acts or text stages	finally / to repeat / here we try to
Endophoric markers	Refer to information in other parts of the text	noted above / see Fig / in section 2
Evidentials	Refer to source of information from other texts	according to X / (1990) / Z states
Code glosses	Help readers grasp meanings of ideational material	namely / e.g. / such as / i.e.
<i>Interpersonal metadiscourse</i>		
Hedges	Withhold writer's full commitment to statements	might / perhaps / possible / about
Boosters	Emphasise force or writer's certainty in message	in fact / definitely / It is clear
Attitude markers	Express writer's attitude to propositional content	Unfortunately / I agree / X claims
Relational markers	Explicitly refer to or build relationship with reader	frankly / note that / you can see
Person markers	Explicit reference to author(s)	I / we / my / mine / our

Following further work by Waller, the following categories and items were identified for analysis:

	<b>Metadiscourse category</b>	<b>No. of lexical items included</b>	<b>Examples of lexical items included in each category</b>
1	FM - Announce goals (AG)	18	I would like to, I will focus on, the aim
2	Attitude markers (AM)	25	even, have to, hopefully, important
3	Code glosses (CG)	16	for example, in other words, such as, that is to say
4	Emphatics (EM)	38	actually, certainly, in fact, must
5	Endophorics (EN)	13	discussed above, example, section
6	Evidentials (EV)	24	according to, argue, claim, show
7	Hedges (H)	47	about, could, possible, would
8	FM - Label stages (LC)	12	all in all, in conclusion, overall, to sum up
9	Logical connectives (LC)	39	also, although, in addition, on the other hand
10	Person markers (PM)	4	I, me, mine, my
11	Relational markers (EM)	21	consider, find, let's, you
12	FM - Sequencing (S)	16	finally, firstly, last, to start with
13	FM - Topic shifts (TS)	8	In regard to, now, to come back to, well
	Total	281	

# How did the analysis operate?

---

1. Text Inspector automated analysis for 900 Cambridge expository essays
2. Manual analysis of 200 texts
  - 3 independent raters
3. Revision of statistics accordingly

# Texts

---

1. 1200 Cambridge expository essays, 400 at each level – FCE (B2), CAE (C1), CPE (C2) - mixed nationalities/L1 groups
2. Reviewed, then 900 selected
3. Analysed with Text Inspector
4. Manual analysis of 200 texts
  - 3 independent raters
5. Revision of statistics accordingly

# Note

---

1. Different kinds of texts (L2)
2. Shorter
3. Not fully academic



A tool for teachers, editors and linguists

## Analysing metadiscourse markers

### How to use this webpage

Paste your text into the box on the right and click 'submit' at the bottom. You will then see each word in your text coloured and numbered according to Hyland's categorisation of metadiscourse markers.

You can then edit the list - in particular choose IGNORE if the system has analysed a word or phrase incorrectly.

Then press RECALCULATE and you will get an analysis of terms used in your text, along with calculations concerning Type-Token ratio and so on.

At the bottom of that final page the analysis is set out in a form for easy copying, for pasting into a spreadsheet.

This version of Text Inspector is limited to 2000000 characters.

You have  characters left.

Teachers are important figures throughout our young lives, and the impression that give us influence us not just during these years but often for along time after we have left academic learning.

However, it is necessary for teachers to keep students engaged and motivated in lessons. If they do not have the capability to do so, the students are likely to become uninterested and in many cases, their behaviour becomes unmanageable and raucous.

Submit Query



A tool for teachers, editors and linguists

Recalculate\_Scores

teachers are **important** [2 attitude marker] figures throughout **our** [11 relational marker] young lives **and** [9 logical connective] the impression that give **us** [11 relational marker] influence **us** [11 relational marker] not just during these years **but** [9 logical connective] **often** [7 hedge] for along time after **we** [11 relational marker] have left academic learning **however** [9 logical connective] it is necessary for teachers to keep students engaged **and** [9 logical connective] motivated in lessons if they do not have the capability to do **so** [13 topic shift] the students are **likely** [7 hedge] to become uninterested **and** [9 logical connective] in many cases their behaviour becomes unmanageable **and** [9 logical connective] raucous

Recalculate\_Scores



# Text Inspector

## Summary analysis of your text

Text elements	
Tokens (total items)	73
Types (unique items)	57
Type-token ratio	0.78
Tokens per type	1.28
Corrections	
Corrections as a percentage of original tokens	

Metadiscourse elements (summary)						
	Tokens	Types	Type-token ratio	Tokens per type	Tokens per 1000 words	Types per 1000 words
1. Announce goals:			0.00	0.00	0.0000	0.0000
2. Attitude markers:	1	1	1.00	1.00	13.6986	13.6986
3. Code glosses:			0.00	0.00	0.0000	0.0000
4. Emphatics:			0.00	0.00	0.0000	0.0000
5. Endophorics:			0.00	0.00	0.0000	0.0000
6. Evidentials:			0.00	0.00	0.0000	0.0000
7. Hedges:	2	2	1.00	1.00	27.3973	27.3973
8. Label stages:			0.00	0.00	0.0000	0.0000
9. Logical connectives:	6	3	0.50	2.00	82.1918	41.0959
10. Person markers:			0.00	0.00	0.0000	0.0000
11. Relational markers:	4	3	0.75	1.33	54.7945	41.0959
12. Sequencing:			0.00	0.00	0.0000	0.0000
13. Topic shifts:	1	1	1.00	1.00	13.6986	13.6986
Totals:	14	10	Average: 0.33	Average: 0.49	191.7808	136.9863

## Detailed analysis

||

<i>Metadiscourse markers</i>	<i>1. Announce Goals</i>	<i>2. Attitude markers</i>	<i>3. Code glosses</i>	<i>4. Emphatics</i>	<i>5. Endophorics</i>	<i>6. Evidentials</i>	<i>7. Hedges</i>	<i>8. label stages</i>	<i>9. Logical connectives</i>
<i>Quantity Tokens/Types</i>	/	1/1	/	/	/	/	2/2	/	6/3
<i>Items</i>		<i>important : 1</i>					<i>likely : 1 often : 1</i>		<i>and : 4 but : 1 however : 1</i>
<i>Metadiscourse markers</i>	<i>10. Person markers</i>	<i>11. Relational markers</i>	<i>12. Sequencing</i>	<i>13. Topic shifts</i>					
<i>Quantity Tokens / Types</i>	/	4/3	/	1/1					
<i>Items</i>		<i>our : 1 us : 2 we : 1</i>		<i>so : 1</i>					

# Manual analysis (sample)

---

1. Manual review of data, 3 coders
2. Aim: to check that markers are what they claim to be (e.g. so)
3. Use of AntConc concordancer
4. Revised statistics

# Key research issues

---

## *Quantity and variety of markers*

- # There will be little difference in the **overall quantity of metadiscourse markers** used by lower and higher level writers (from Burneikaite, 2008)  
**or**
- # Higher levels will use **more** metadiscourse markers (Sandford 2012)

# Types of markers used

---

- # Higher level writing will show significantly lower use of certain **common Logical Connectives** (Burneikaite 2008, Hawkey and Barker 2004, Carslen 2010).

# Types of markers used

---

- # Higher level writing will show significantly higher use of *endophoric markers*, referring to information in other parts of the text (e.g. *noted above*) (from Burneikaite 2008)
- # Higher level writing will show significantly higher use of *evaluative markers*, in particular *emphatics* (e.g. *definitely*) (Burneikaite 2008)

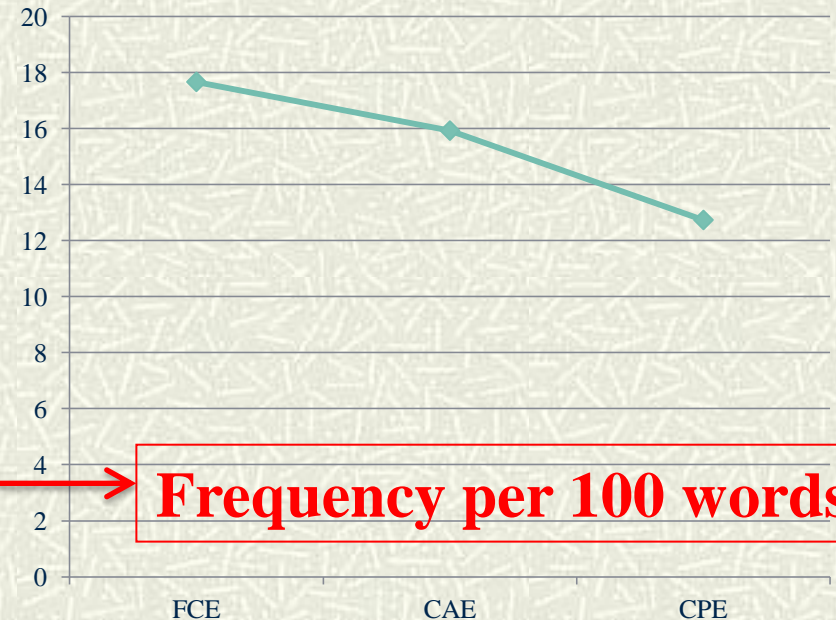
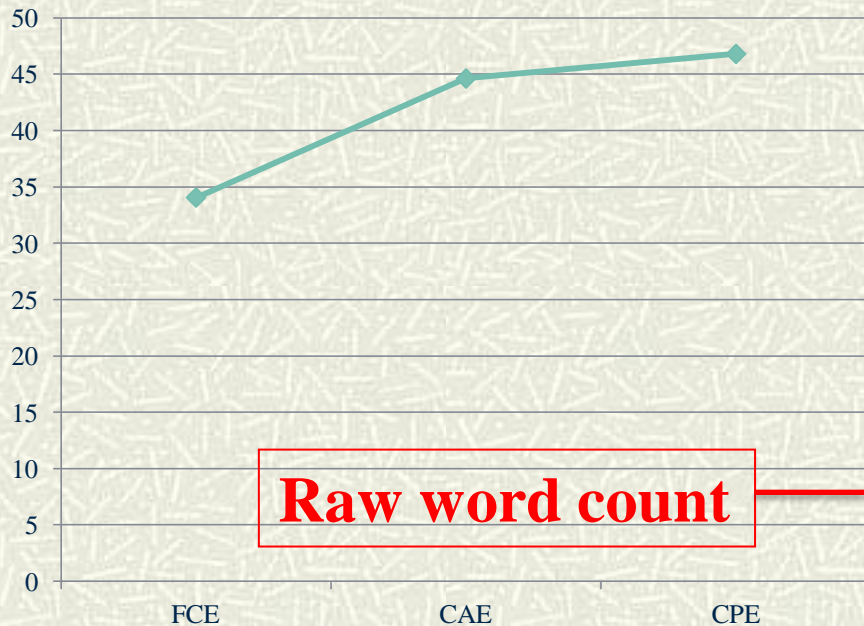
# Metadiscourse categories and lexical items

Category	No. of items	e.g.
1. Announcing goals	18	I would like to, I will focus on, the aim
2. Attitude markers	25	even, have to, hopefully, important
3. Code glosses	16	for example, in other words, such as
4. Emphatics	38	actually, certainly, in fact, must
5. Endophorics	13	discussed above, example, section
6. Evidentials	24	according to, argue, claim, show
7. Hedges	47	about, could, possible, would
8. Label stages	12	all in all, in conclusion, overall, to sum up
9. Logical connectives	39	also, although, in addition, on the other hand
10. Person markers	4	I, me, mine, my
11. Relational markers	21	consider, find, let's, you
12. Sequencing	16	finally, firstly, last, to start with
13. Topic shift	8	In regard to, now, to come back to, well
<b>Total</b>	<b>281</b>	

# Use of Metadiscourse Markers

Length of writing and use of MMs across 3 levels

	FCE(B2)	CAE(C1)	CPE(C2)
Mean No. of lexical tokens	192.69	279.92	365.98
Mean No. of MMs (word count)	34.05	44.61	46.79
Mean No. of MMs (freq. per 100 wds)	17.67	15.93	12.74



**Raw word count**

**Frequency per 100 words**



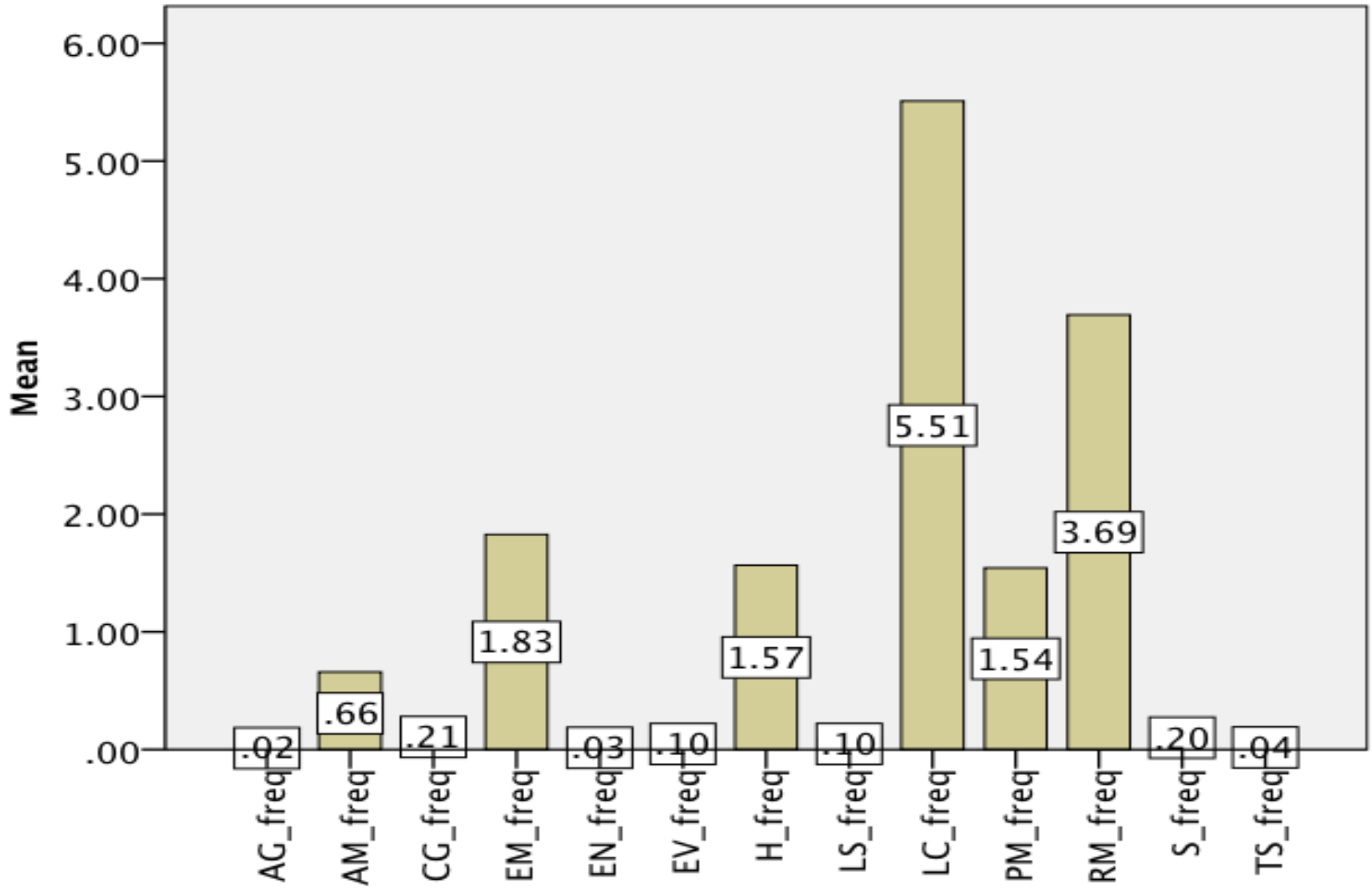
# Key research issues

---

## *Quantity and variety of markers*

- # There will be little difference in the **overall quantity of metadiscourse markers** used by lower and higher level writers (from Burneikaite, 2008)  
**or**
- # Higher levels will use **more** metadiscourse markers (Sandford 2012) **No**

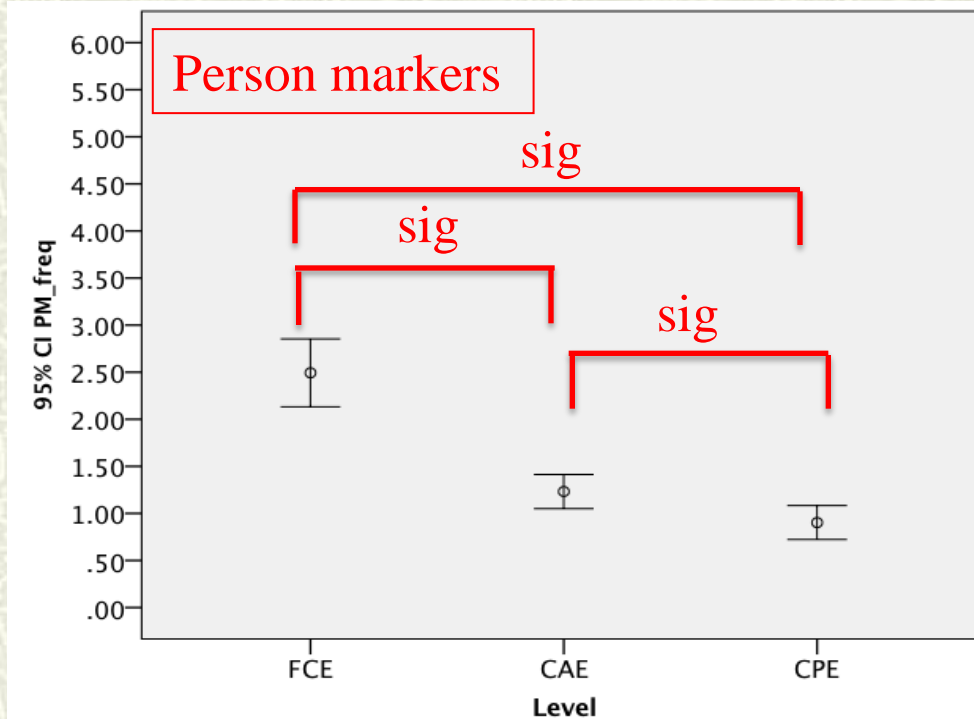
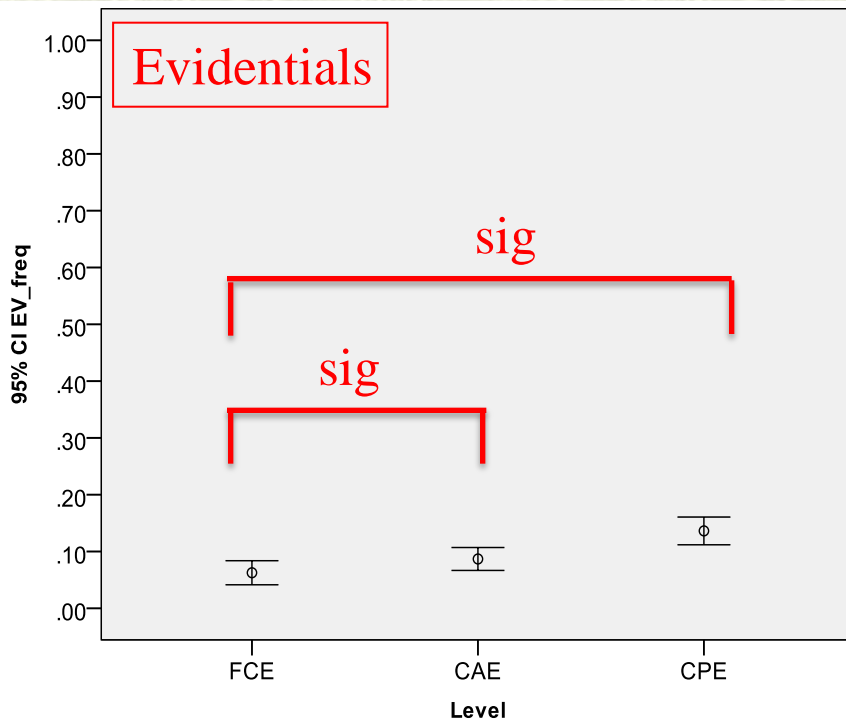
# Average use of each of the 13 metadiscourse categories (per 100 words) by all candidates





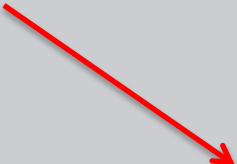

# Use of MMs across the three levels of proficiency (FCE, CAE, CPE: N=300 each)

The Kruskal-Wallis Tests → Post hoc tests (Mann-Whitney Tests):

1. **No sig difference:** Announce goals, Attitude markers, Code glosses, Logical connectives, Sequencing
2. **Increase as the level goes up:** Endophorics, Evidentials
3. **Decrease as the level goes up:** Emphatics, Hedges, Label stages, Person markers, Relational markers, Topic shifts



# Logical connectives: “No sig difference”

Trend	Item
No diff 	also, although, besides, in addition, on the contrary, on the other hand, so
Up 	as a result, consequently, even though, furthermore, however, moreover, nevertheless, since, therefore, though, thus, whereas, while, yet
Down 	because, but
Down-Up 	and

# Types of markers used

---

# Higher level writing will show significantly lower use of certain common Logical Connectives (from Burneikaite, 2008, Hawkey and Barker 2004, Carslen 2010).

# In part

# Types of markers used

---

- # Higher level writing will show significantly higher use of *endophoric markers*, owing to writers' experience in structuring longer texts (from Burneikaite, 2008)
- # Not conclusive – few examples (short texts)



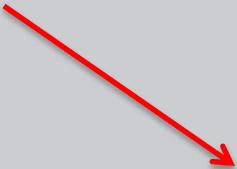

# Evaluative markers (e.g. 'clearly')

---

# Higher level writing will show significantly higher use of *evaluative markers*, in particular *emphatics* (Burneikaite, 2008)



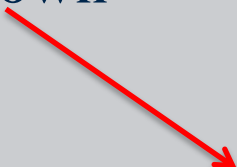
# **Partly**

# Emphatics: “Decreased use as the level goes up”

Trend	Item
No diff 	clearly, definitely, even if, I believe, in fact, must, never, obviously, of course, sure, won't
Up 	actually, certainly, indeed, essential, should, the fact that, undoubtedly
Down 	always
Up-Down 	know, will



# Hedges: “Decreased use as the level goes up”

Trend	Item
No diff 	about, almost, could, couldn't, generally, in general, mainly, mostly, often, possibly, usually, wouldn't
Up 	likely, may, perhaps, probably, possible, seems
Down 	sometimes
Up-Down 	maybe, might, would

# What does this tell us?

---

- # Value in analysing metadiscourse in learner writing
- # BUT: we must look beyond the categories at individual items
- # Analyse individual items more closely, considering cognitive complexity and lexical frequency
- # A more varied picture than is sometimes presented
- # Exams – should elicit longer texts so as to allow for more ‘advanced’ metadiscourse use

# Risky conjectures revisited

---

“It is predicted that student narratives that received higher ratings on the TOWL-3 [test] will contain more metadiscourse elements using the classification developed by Hyland (2005), and that fewer metadiscourse markers will be found in narratives that received lower ratings”

(Sandford 2012)

Too risky

*BAAL 2013*

**Thank you**

*Stephen Bax (CRELLA)*

*Daniel Waller (UCLAN)*

*Fumiyo Nakatsuhara (CRELLA)*

**CRELLA**

Centre for Research in English  
Language Learning and Assessment



University of  
Bedfordshire

