Building on the insights from the English Grammar Profile

From *really good* to *painfully obvious*

Geraldine Mark

L2P workshop Gothenburg 20-21 April 2023

Overview

- Background

- Data

- Methodology

- Insights

- What else?

The background

Describing learner language at CEFR levels.

The background: our task

- To empirically describe what learners can do with grammar across the CEFR proficiency levels.

- To create a resource for English language teaching.

  http://www.englishprofile.org/english-grammar-profile/egp-online

The background: the challenge

Analysis and description of learner grammar competence

- ○ across *six* proficiency levels within the CEFR.
- ○ of *all* aspects of grammar for the ELT community.
- ○ across all L1 backgrounds in the data.

The background: what the EGP is not

Not a contrastive analysis

Not based on one single L1 cohort

Not focussed on just one interlanguage feature

Not a profile of errors

Not prescriptive

EnglishProfile

The CEFR for English

English Grammar Profile   EGP Online   Terms of Use   Grammatical Categories   Glos

English Grammar Profile Onli

Search...   C1   C2   **Search**   **Add Filters**   **Clear Results**

SuperCateg   **SubCategory**

- Select -   **More filters**

Download XLS

Results 1 - 20 of 1222   Sort by: SuperCategory   Ascending   Display #   20

SuperCategory	SubCategory	Level	Can-do statement	Example	Details
ADJECTIVES	combining	A1	**FORM:** COMBINING TWO ADJECTIVES WITH 'AND' Can use 'and' to join a limited range of common adjectives.	Example	Details
ADJECTIVES	combining	A2	**FORM:** COMBINING TWO ADJECTIVES WITH 'BUT' Can use 'but' to join a limited range of common adjectives, after 'be'.	Example	Details
ADJECTIVES	combining	B1	**FORM:** BEFORE THE NOUN Can use a comma to combine two adjectives used before the noun, following the usual order of adjective types.	Example	Details

The data: Cambridge Learner Corpus

Cambridge exams   Written data

aligned to CEFR

55 million+ words

32 million words error-coded

200,000+ exam scripts

Spanning 17 years
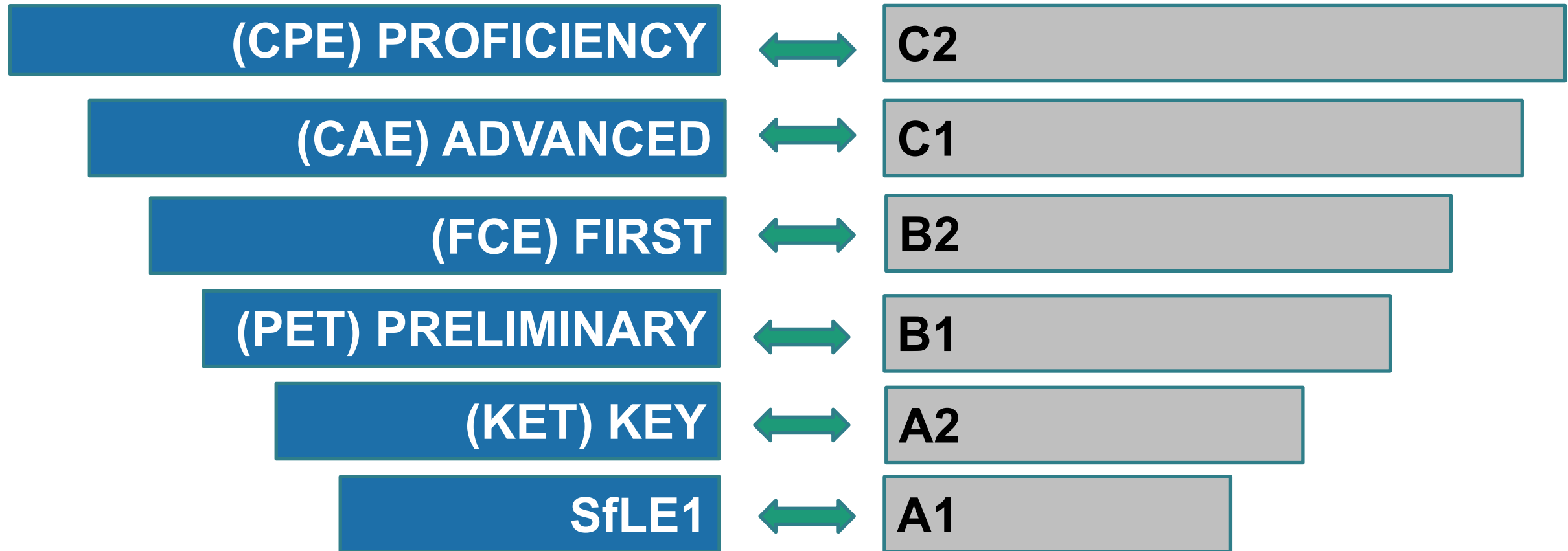
215 countries

143 first languages

Metadata
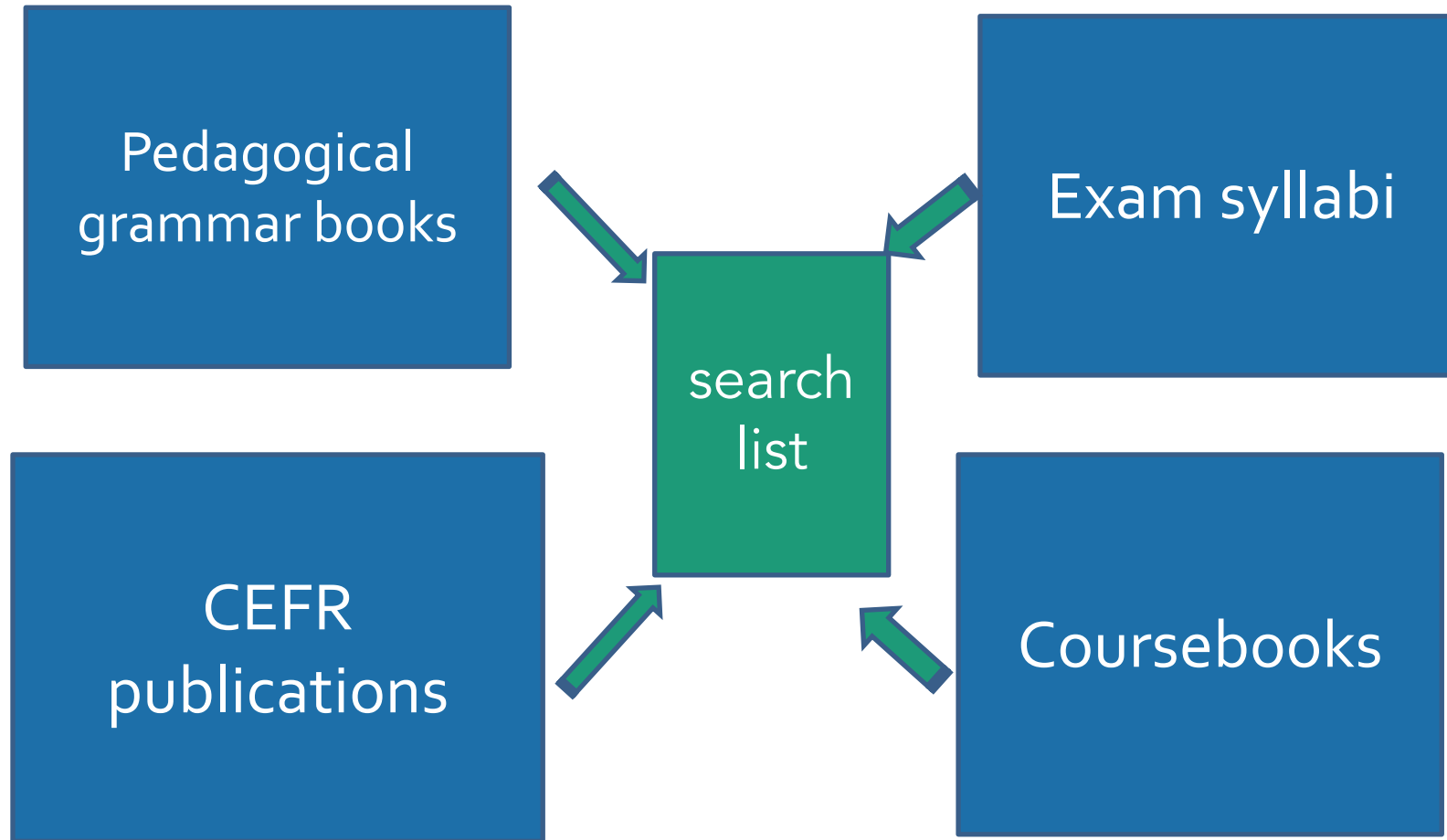Candidates: L1, nationality, age, gender, education
Exam: CEFR level of exam taken and exam performance
Task: question number, task style/format/register
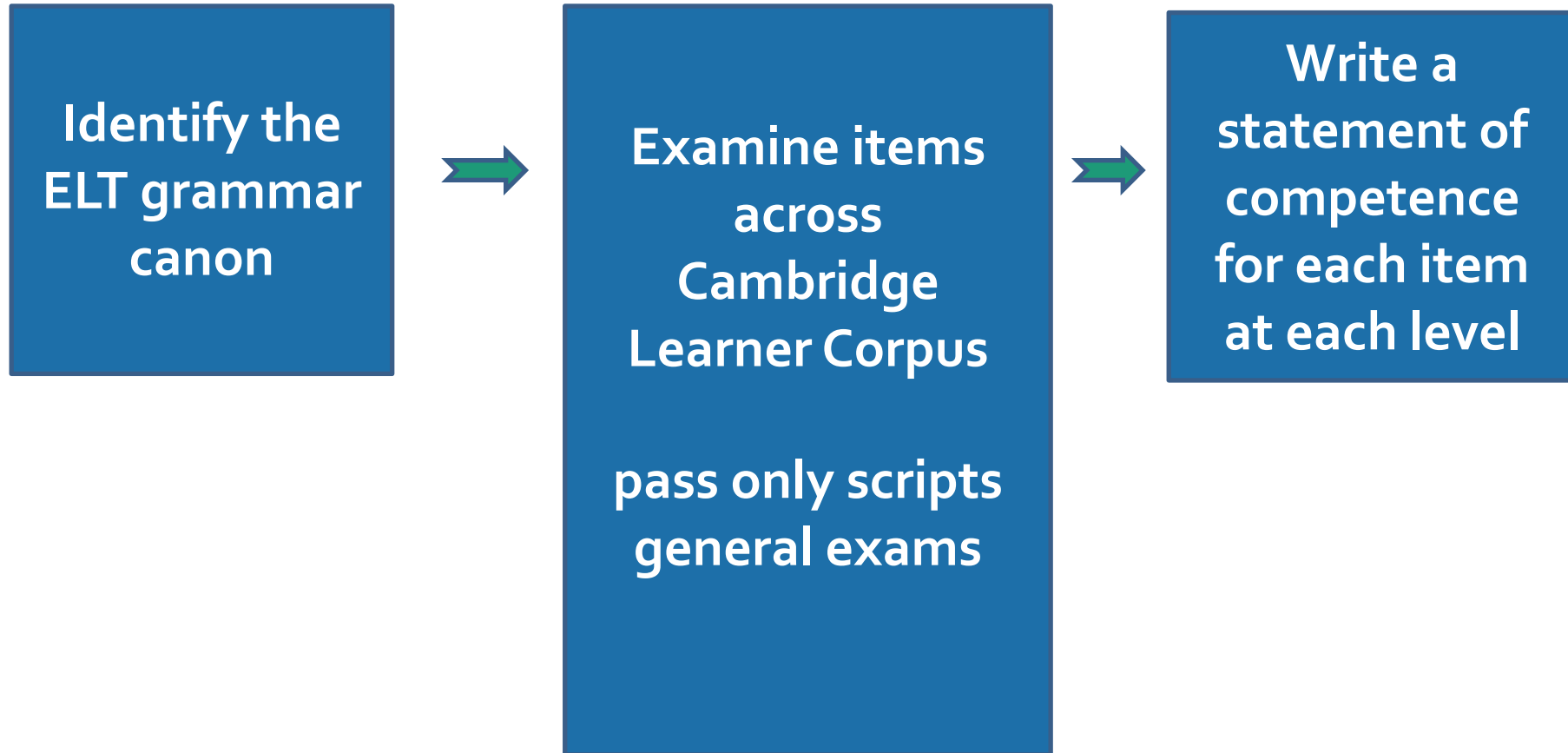
The data: Cambridge exams & the CEFR

(CPE) PROFICIENCY	⟷	C2
(CAE) ADVANCED	⟷	C1
(FCE) FIRST	⟷	B2
(PET) PRELIMINARY	⟷	B1
(KET) KEY	⟷	A2
SfLE1	⟷	A1

The methodology: starting points

The ELT 'grammar canon'

adjectives

adverbs

clauses

conjunctions

determiners

discourse markers

future time

modality

negation

nouns

search list

passives

past

prepositions

present

pronouns

punctuation

questions

reported speech

verbs

focus

A linear methodology

Identify the ELT grammar canon

Examine items across Cambridge Learner Corpus

pass only scripts general exams

Write a statement of competence for each item at each level

A linear methodology: the first fall

Establish the ELT grammar canon ➡ Examine items ➡ Write a statement of competence for each item at each level

WRONG WAY
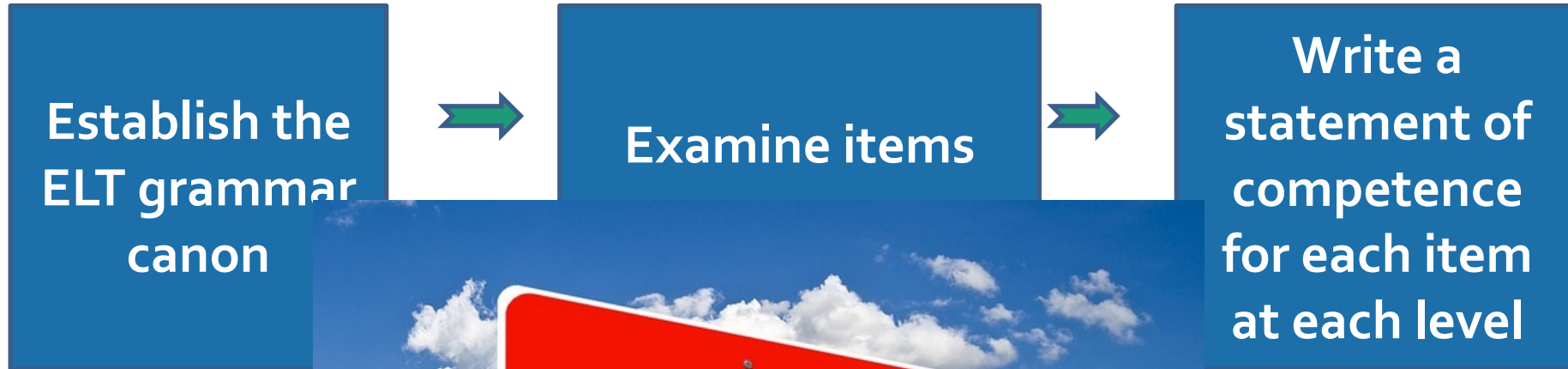
PMW modal verb *would* across pass only data
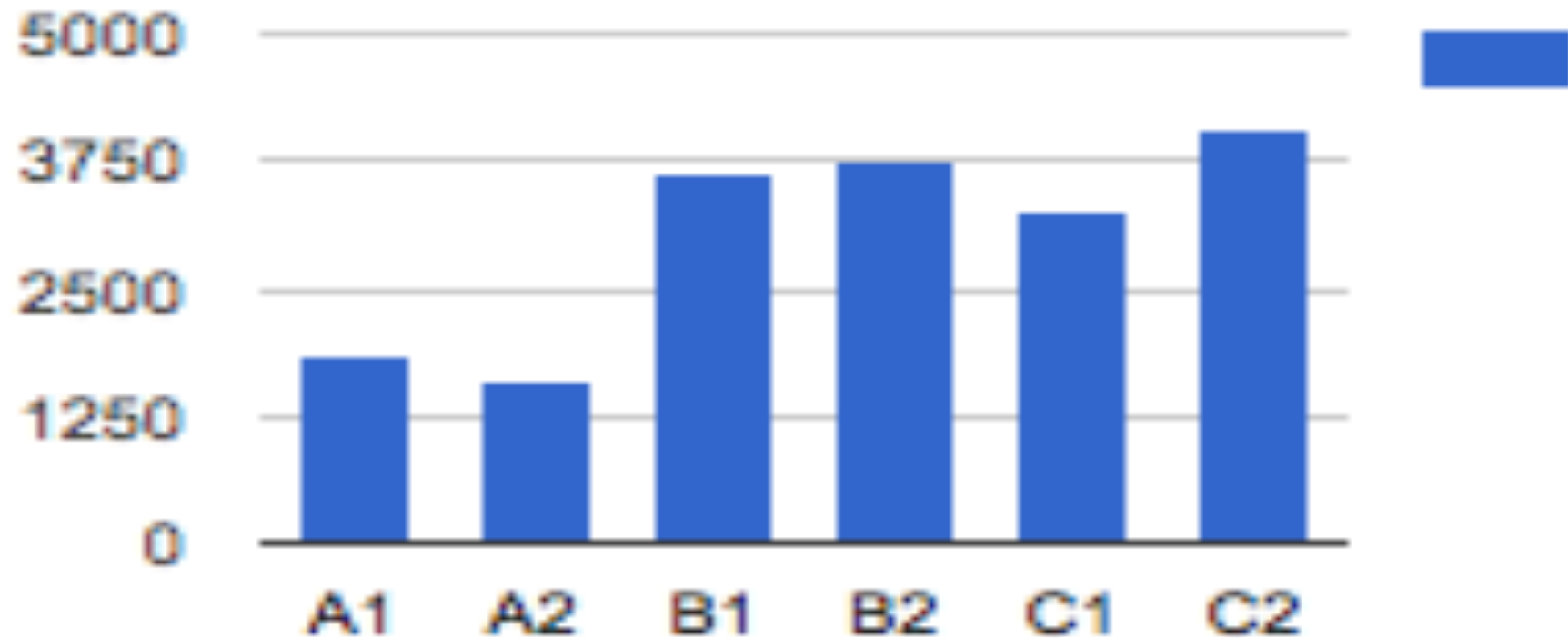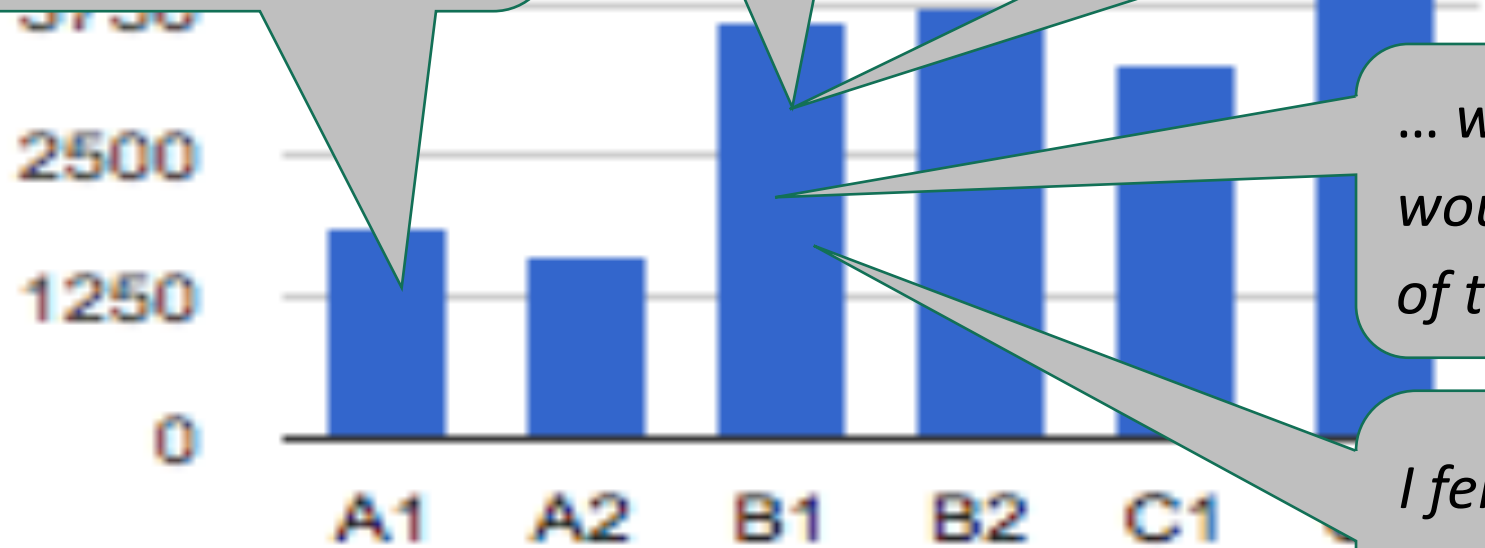
PMW modal verbs ... ss pass only data

I would like to eat some Spanish food (Italian, 2006)

It's your decision but if I had the chance, I would go with my friends. (German 2008)

I would definitely choose to live by the seaside! (Chinese, 2010)

... we thought that the film would be horrible because of the title (Portuguese, 2002)

I felt better when he said that he would be in Fenerbahce. (Turkish, 2009)

A1   A2   B1   B2   C1

0   1250   2500   3750

Waystage: CEFR

A2   Pronouns

2.1   demonstrative (3.1.1, 4.8.1.1, 4.8.1.2)
this, that, these, those

2.2   personal (3.1.1, 4.8.1.1, 4.8.1.2)

2.2.1   subject forms
I, you, he, she, it, we, they

2.2.2   non-subject forms
me, you, him, her, it, us, them

2.3   possessive (4.7.5.1, 4.8.1.1, 4.8.1.2)
mine, yours, his, hers, ours, theirs

KET Handbook

Pronouns

Personal (subject, object, possessive)

Impersonal: it, there

Demonstrative: this, that, these, those

Quantitative: one, something, everybody, etc.

Indefinite: some, any, something, one, etc.

Relative: who, which, that

Van Ek, J A and Trim, J L M (1991a) (1991b) *Waystage 1990,* Cambridge: Council of Europe/Cambridge University Press.
Cambridge ESOL (2009) *Key English Test Handbook for Teachers.*

Waystage: CEFR

A2 Pronouns

2.1 demonstrative (3.1.1, 4.8.1.1, 4.8.1.2)
this, that, these, those

2.2 personal (3.1.1, 4.8.1.1, 4.8.1.2)

2.2.1 subject forms
I, you, he, she, it, we, they

2.2.2 non-subject forms
me, you, him, her, it, us, them

2.3 possessive (4.7.5.1, 4.8.1.1, 4.8.1.2)
mine, yours, his, hers, ours, theirs

KET Handbook

Pronouns

Personal (subject, object, possessive)

Impersonal: it, there

Demonstrative: this, that, these, those

Quantitative: one, something, everybody, etc.

Indefinite: some, any, something, one, etc.

Relative: who, which, that

Van Ek, J A and Trim, J L M (1991a) (1991b) *Waystage 1990,* Cambridge: Council of Europe/Cambridge University Press.
Cambridge ESOL (2009) *Key English Test Handbook for Teachers.*

Possessive pronouns A2

Node form	raw freq		Node form	raw freq
his	20196		theirs	168
Yours	16700		hers	152
yours	3248		Mine	45
mine	1859		Ours	15
His	1595		Hers	3
ours	519		Theirs	2

Mine/mine   A2 examples

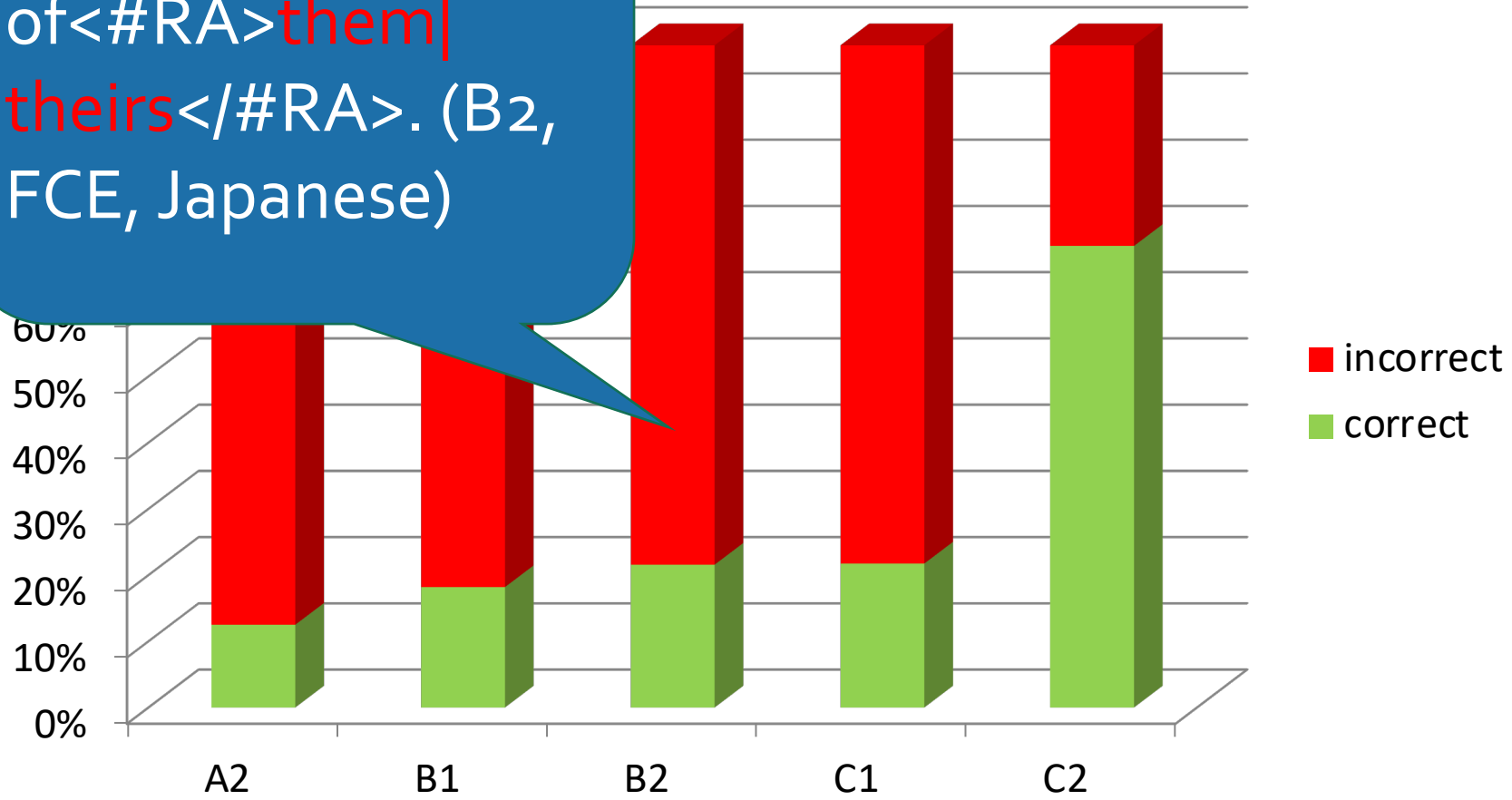- I've just watched a football match between our teams and, of course, mine won … (Portuguese)

- Mine is red. (Spanish – European)

- I bought a new jacket too, because mine was very old … (Italian)

- I bought a pullover and trousers because mine are too small for me. (Korean)

- I enjoyed mine very much. (Tagalog)

- A friend of mine bought a very nice sweater for me. (Turkish)

theirs
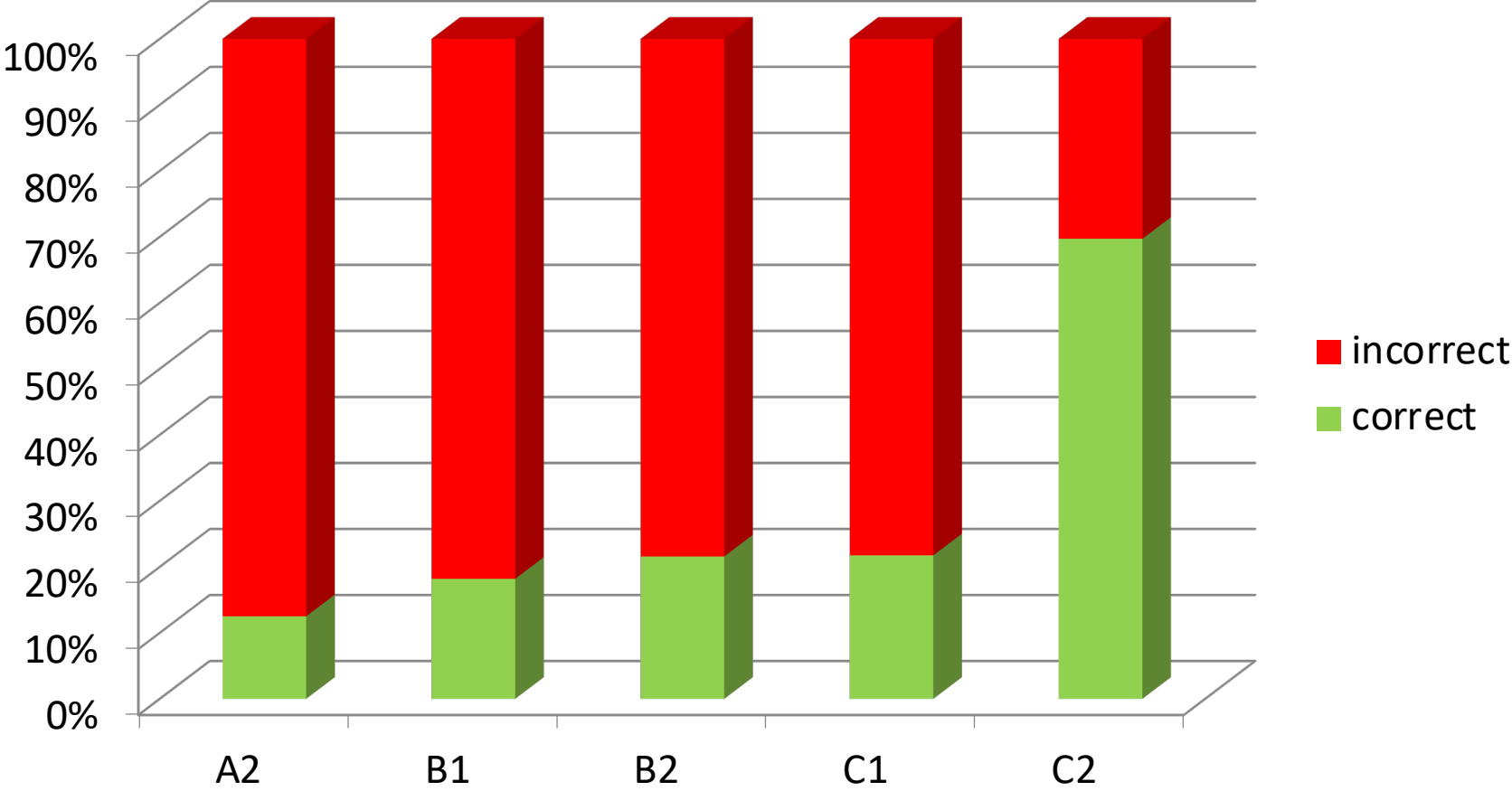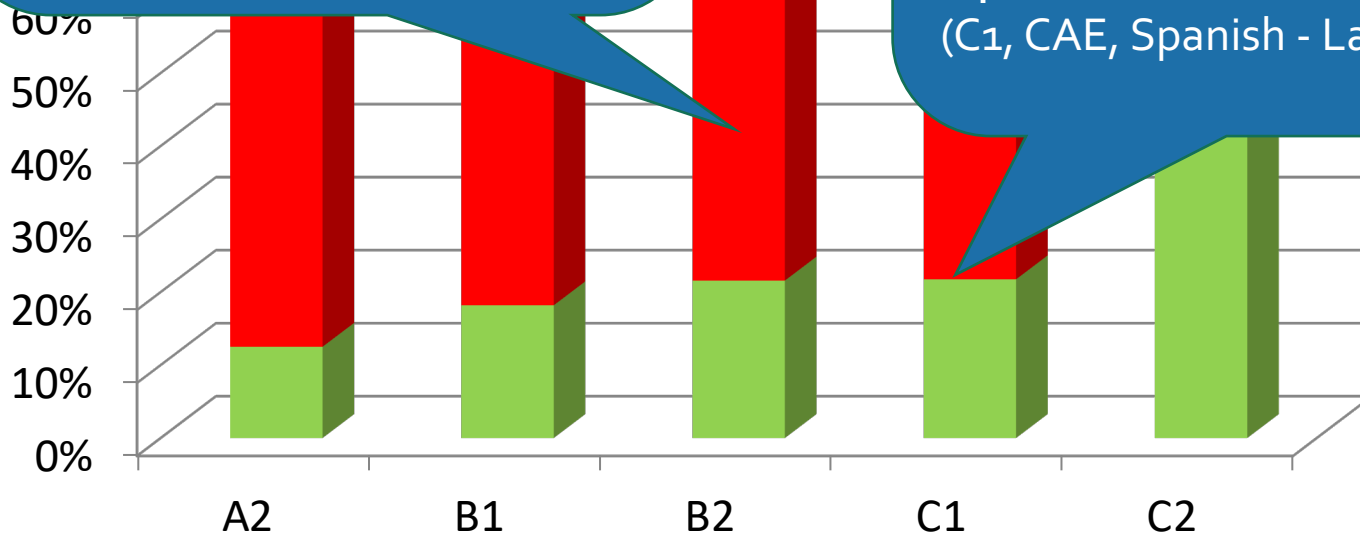
A linear methodology: back to the drawing board
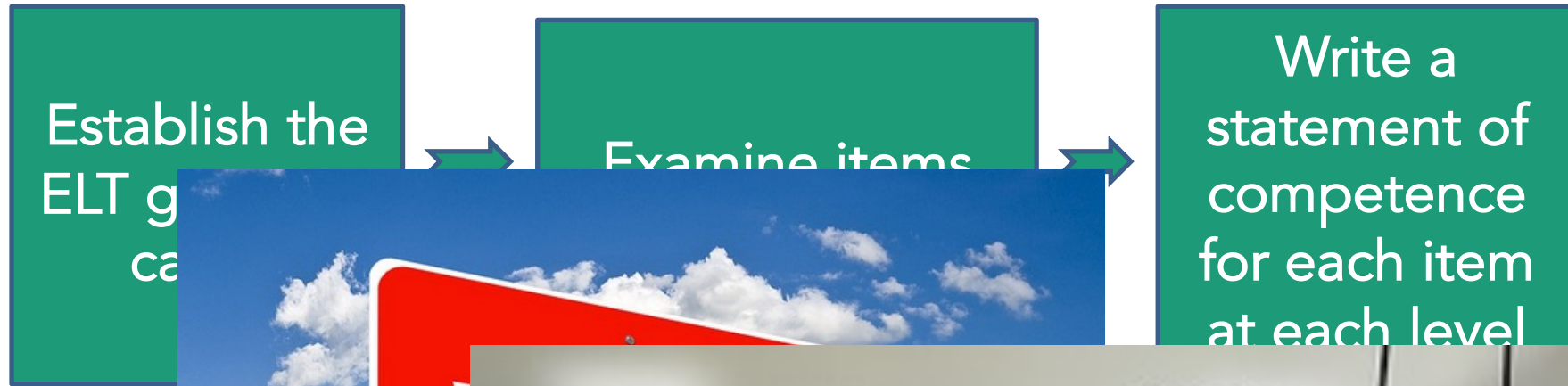
How did we really get there?
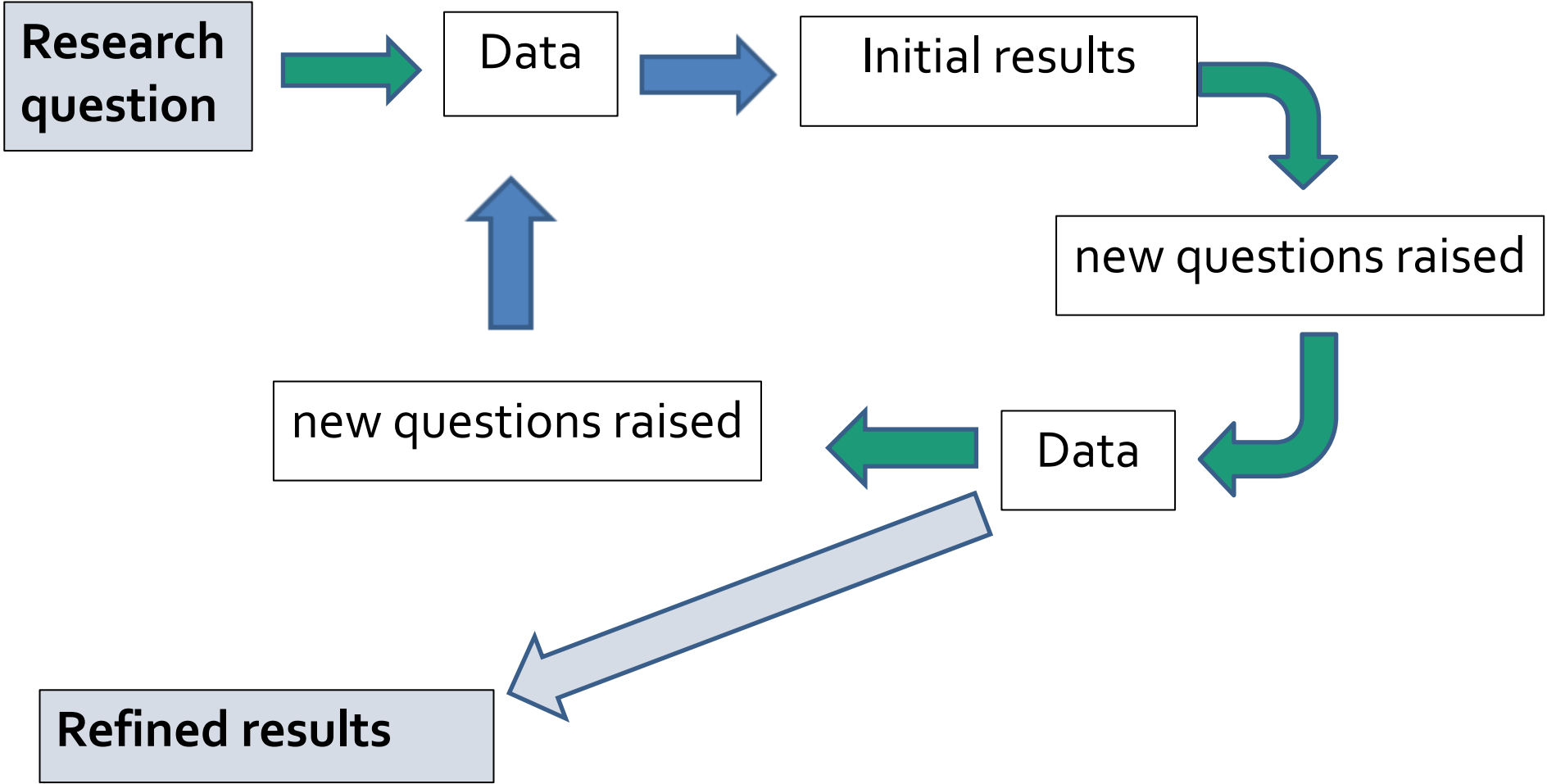
The methodology

- **Frequency**
- **Correct uses**
- Distribution
- Spread of language families
- Spread of contexts of use
- Task effect

- Opportunity of use

Example: [word="had|'d"][tag="RB"]{0,2}[tag="V.N"]

Level	Raw frequencies	PMW	Errors	%
BNC	117, 116	61.3		
A1	1	11		
A2	194	188	96	49.5
B1	2,720	712	532	19.5
B2				
C1				
C2				

*I <#TV> **had seen** | saw </#TV> Akropolis yesterday.*
[A2, Dutch, 1999]

Example: [word="hadl'd"][tag="RB"]{0,2}[tag="V.N"]

Level	Raw frequencies	PMW	Errors	%
BNC	117, 116	61.3		
A1	1	11		
A2	194	188	96	49.5
B1	2,730	712	532	19.5
B2	8,093	1333		
C1	2,756	460		
C2	13,213	1999		

The methodology

- **Frequency**
- **Correct uses**
- Distribution
- Spread of language families
- Spread of contexts of use
- Task effect: (SkE text visualisation)

- Opportunity of use

Is it frequent?

Is it correct?

Good range of learners?

Different L1 speakers?

Different contexts ?

Not affected by task?

B1 Statements of 'grammar knowledge'
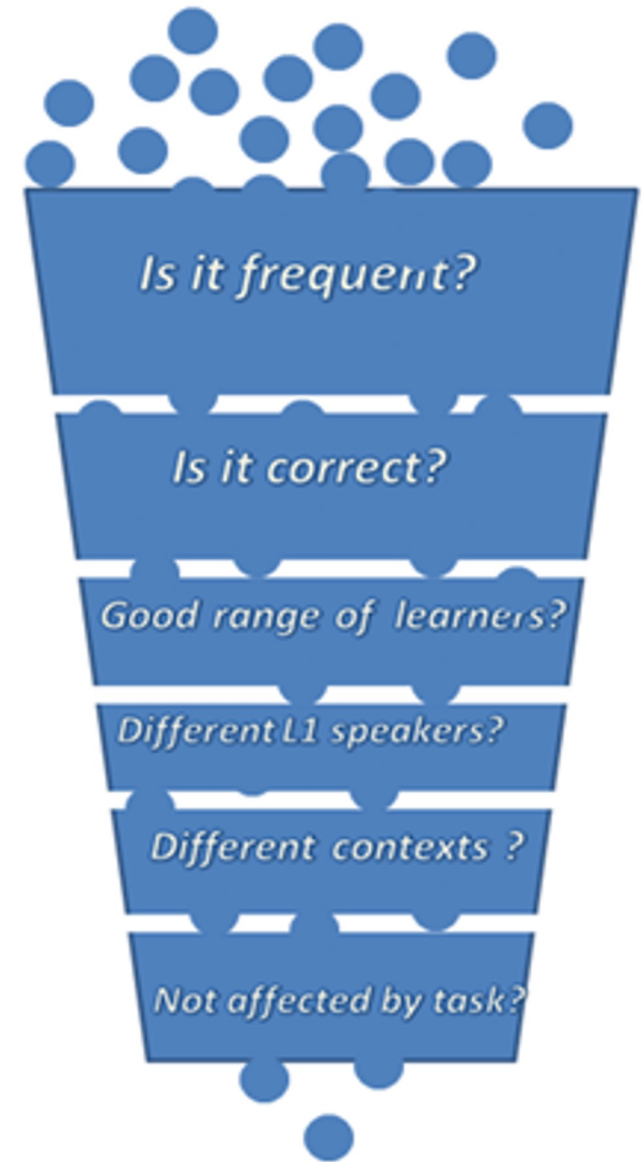
FORM:
Can use the affirmative form

FORM: Can use the past per
of adverbs (including never, ever, just, always, already)
normal mid-position.

USE: Can use the past perfect simple to talk about a time
before another time in the past.

Actually I'd arranged an appointment with my doctor before calling you.
(Arabic – Meghreb, 2008)

The methodology

- **Frequency**
- **Correct uses**
- Distribution
- Spread of language families
- Spread of contexts of use
- Task effect
- Opportunity of use

Is it frequent?

Is it correct?

Good range of learners?

Different L1 speakers?

Different contexts ?

Not affected by task?
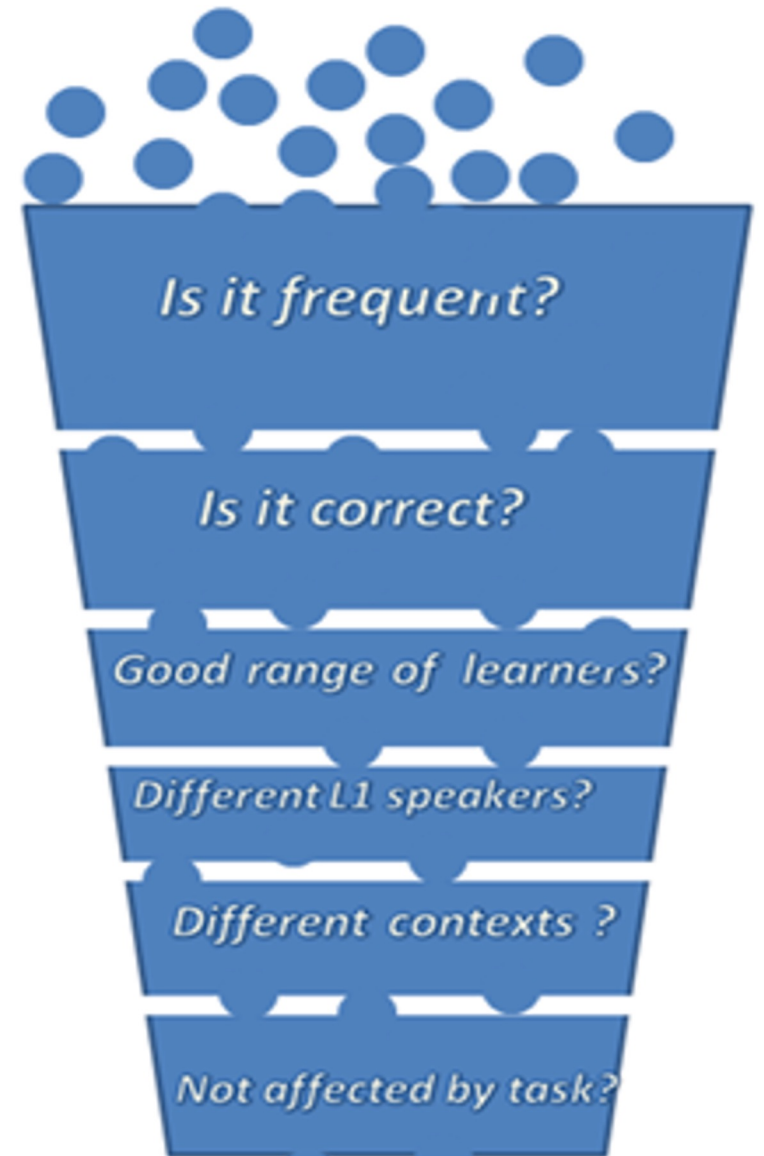
O'Keeffe, A., & Mark, G. (2017). The English Grammar Profile of learner competence: Methodology and key findings. *International Journal of Corpus Linguistics*, 22(4), 457-489.

B2 Statements of 'competence'

FORM/USE: Can use the past perfect simple after *because* to give explanatory information.

USE: Can use the past perfect simple in relative clauses to give background information.

USE: Can use the past perfect sim... changed.

USE: Can use the past perfect si... verb in the reporting clause is in t...

USE: Can use the past perfect sim... verb in the reporting clause is in the past simple.

> *He went to the Hotel because he wanted to find out who was the person that had sent him a mysterious letter ...*
>
> (Chinese, 1997)

B1 [tag="JJ"][word="enough"]

old enough	84		quiet enough	4
big enough	69		much enough	3
good enough	49		comfortable enough	3
large enough	9		strong enough	2
warm enough	6		safe enough	2
responsible enough	5		mature enough	2
brave enough	5		hot enough	2

old/big enough

- So, tell them that you're **old enough** to go on holidays with your friends and that you will be very careful. (B1, PET, Greek)

- So, if I were you I would try to explain to my parents that I'm **old enough** to go on a trip with my friends. (B1, PET, Basque)

- My bed is too small and my cupboard isn't **big enough** to put all my shirts and trousers into. (B1, PET, German)

- I moved to a new home because my mother is going to have a baby and our previous home wasn't **big enough**! (B1, PET, Italian)

Is it good enough?

- At the moment I'm not **good enough**. (B1, PET, German)
- A standard room is **good enough**. (B1, BECP, Chinese)
- I decided to join three lessons each week because my English grammar is not **good enough**. (B1, PET, Hungarian)
- I think that your English skills are **enough good** and you can manage. (B1, SfLE3, Polish)

responsible/comfortable enough

- My bed is a small size so I'm not **enough comfortable**. (B1, PET, Turkish)
- My bedroom is not very big but it's **enough confortable** for me. (B1, PET, Spanish - European)
- you can also make an appointment between your parents and your friend's parents so that they get convinced that you are **enough reponsable** to go with your mates on vacations. (B1, PET, Spanish - European)
- He was worried about dissappoint his mom for not been **enough responsible** … (B1, PET, Spanish - Latin American)

http://www.englishprofile.org/component/grammar/content/37

Element	**FORM:** WITH 'ENOUGH'
SuperCat	ADJECTIVES
SubCat	modifying
Lexical Range	N/A
Level	**B1**
Cando	Can use 'enough' to modify adjectives.
Corrected Learner Example	My room is big enough, there is a lot of space, and for this reason I will buy a stereo. (Greece; B1 THRESHOLD; 2001; Greek; Pass) I thought they had the same problem as me: we don't think our English is good enough. (China; B1 THRESHOLD; 2002; Chinese; Pass)
Uncorrected Learner Example	My room is big enough and there is a lot of space and for this reason I will buy a stereo. (Greece; B1 THRESHOLD; 2001; Greek; Pass) I thought they got the same problem as me, we don't think our English is good enough. (China; B1 THRESHOLD; 2002; Chinese; Pass)
Comments	There are lots of errors with word order for this feature, e.g. … 'you are enough responsible to stay for some days without your parents' - but at B1 the structure is there.

Some key insights

- Lexico-grammatical development

- Competence emerges both syntactically and functionally

- Error and competence are not mutually exclusive

- Mismatch between ELT syllabi and competence

Lexico-grammatical development

Past simple: Student performance		
	Raw frequency	**PMW**
A1	63,508	27,151
A2	244,043	28,601
B1	209,997	24,391
B2	288,339	25,715
C1	235,808	25,760
C2	222,208	27,795

Past simple: A1

*She **was** a teacher in Iraq.* (**A1**, Kurdish)

*It **was** a pretty bag. I **liked** it very much* (A1 , Chinese)

*Last night I **went** to the disco.* (**A1**, Spanish)

Can use common regular and irregular verbs in the past simple, to talk about some everyday events or states in the past.

Past simple: B1

*He always **seemed** so confident to me!*

(Brazil; B1 THRESHOLD; 2002; Portuguese; Pass)

Can use the past simple to talk about habitual states or actions.

Past simple: B2

I wanted to know if the rooms are single or double, if they have showers and if there is room service. …

(**B2**, FCE, Spanish – European)

Can use *I wondered* and *I wanted*, as politeness structures, when making polite requests and thanking.

Past simple: B2

[letter of complaint]
*It would be better if you **provided** a lunch of higher quality in future.*   (C1, CAE, German)

Can use *if* + past simple as a politeness structure (especially in the context of letters)

Past simple: C1

*It is true that it **did** rain during the day but it was not a "sudden downpour", as you wrote. It was more a small shower …*

(C1, CAE, Swedish)

Can use emphatic **did** to defend arguments and refute inaccuracies.

An iterative approach: past simple development

1. Increase in verb types
2. L1:L2 form distribution convergence the higher the levels
3. Growth of functions (including pragmatic)
4. Increase in errors at B1

Error pattern with *must* PMW
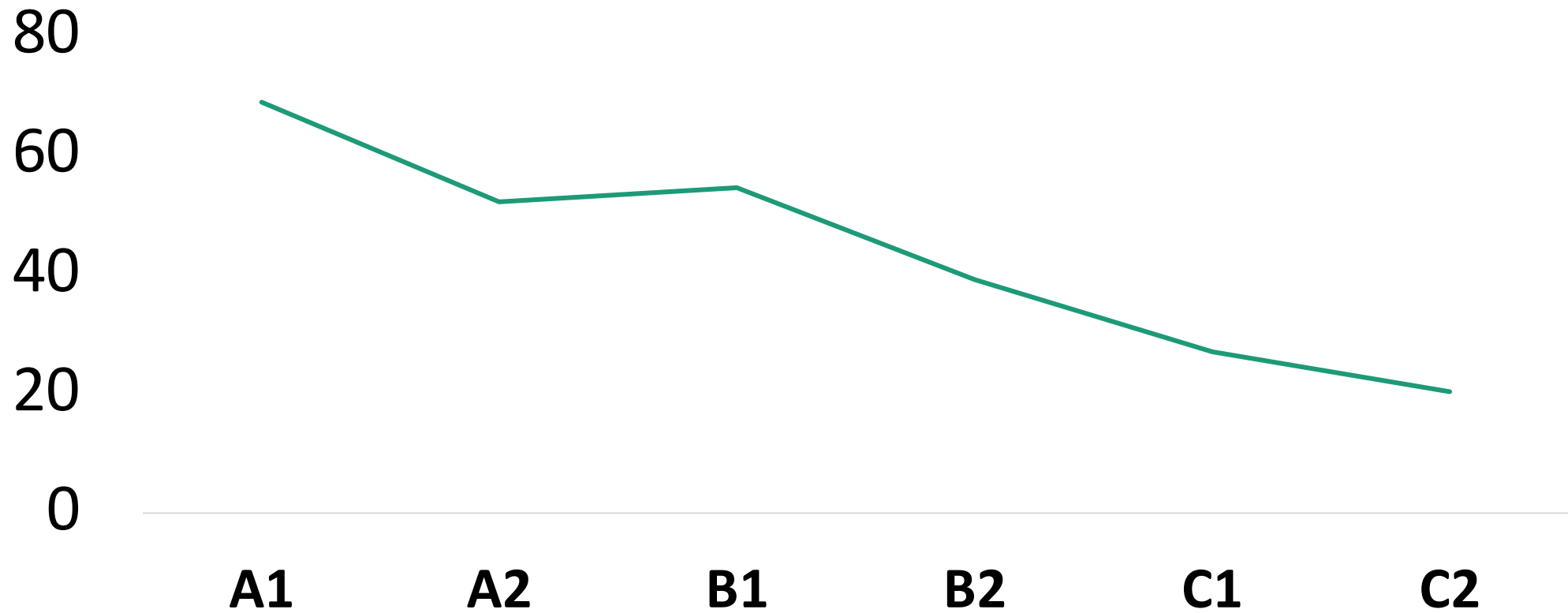
Flat-lining?

Higher level learners appear to demonstrate less progress.
They are taking more risks.
They have more 'error opportunity'.

Accuracy-complexity trade-off / B2-C2 stabilisation line (Thewissen 2013)

Illustrating a developmental pathway: *must*	A1	A2	B1	B2	C1	C2
Form: Can use affirmative and negative		✓	✓	✓	✓	✓
Function: Can use must to talk about obligation and necessity		✓	✓	✓	✓	✓
Form: Can use *must* with adverbs			✓	✓	✓	✓
Form: Can use ellipted forms of *must*			✓	✓	✓	✓
Function: Can use *must* to talk about deductions and conclusions			✓	✓	✓	✓
Function: Can use *must* to make invitations			✓	✓	✓	✓
Function: Can use *must* for strong suggestions			✓	✓	✓	✓
Form: Can use *must have +ed*				✓	✓	✓
Function: Can use *must* to express concession (*I must admit*)				✓	✓	✓
Function: Can use *must* for emphasis				✓	✓	✓
Function: Can use *must* to express rules				✓	✓	✓
Form: Can use ellipted form *must have + -ed* (without subject)					✓	✓
Form: Can use inverted patterns of must after negatives for focus / emphasis					✓	✓
Form: Can use past negative form						✓

Grammatical development	≠	Only through new structures

Grammatical development	=	Same structures New uses

Beyond syntactic stabilisation

"developmental endpoint"
 "should not … be interpreted as a sign of no learning" (Thewissen 2013)

 A *syntactic* developmental endpoint.

  not "linguistic rigor mortis"
  (Larsen-Freeman 2006)

adverbs (69 in EGP)

*Take care. See you **soon**.*
***Sometimes** I go to the beach.* (A1)

*If I were you I'd **certainly** choose not to go with your parents.* (B1)

***Apparently** someone made a mistake writing the details I gave over the phone.* (C1)

*He was quite nice to me, **surprisingly** and he told him that he would do his best but that the music was not too loud.* (C2)

adverb + adjective

	PMW
C2	6649
C1	4514
B2	3330
B1	2919
A2	635
A1	199

It is **painfully obvious** **that it** will rescue our city from pollution and fumes...
It is **highly unlikely** **that** the goods can vanish from your warehouse without being noticed.

I'm **really happy** **to** see you.
It is **really nice** **that** you come to visit me.

My home is **very nice**.
This is a **very nice** neighbourhood.

Grammatical development doesn't look like this …

| present simple | ➕ | past simple | ➕ | modal verbs | ➕ | countable nouns | ➕ |
| conditionals | ➕ | subject pronouns | ➕ | past perfect | ➕ | future perfect |
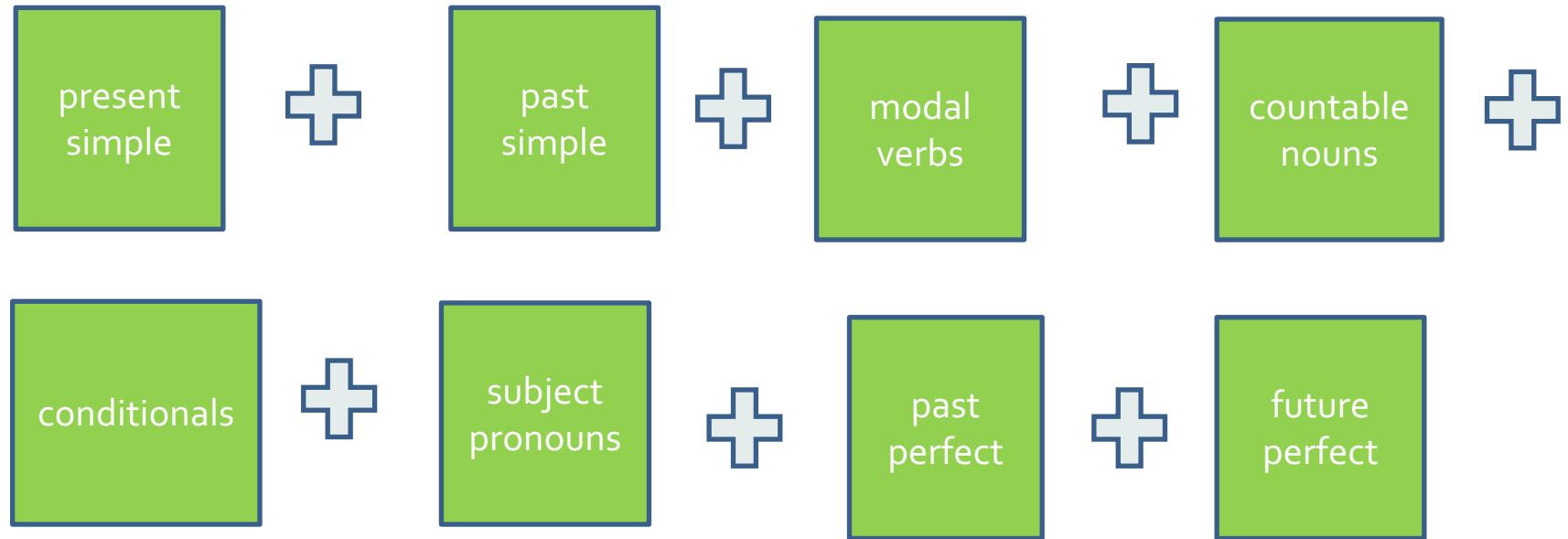
It looks more like this ...

'grammatical polysemy'

use

use

use

form

C2 DISAPPROVAL
*Indeed no one can imagine what children **will** do!*

C1 HABITUAL, TYPICAL
*... a dish that **will** usually contain some kind of sauce*

B1 PREDICTIONS
*... as soon as you finish your school, **it'll** be harder to meet your old friends.*

A1 FUTURE ARRANGEMENTS
*I'**ll** come to see you tomorrow morning.*

pronoun + *will* + main verb
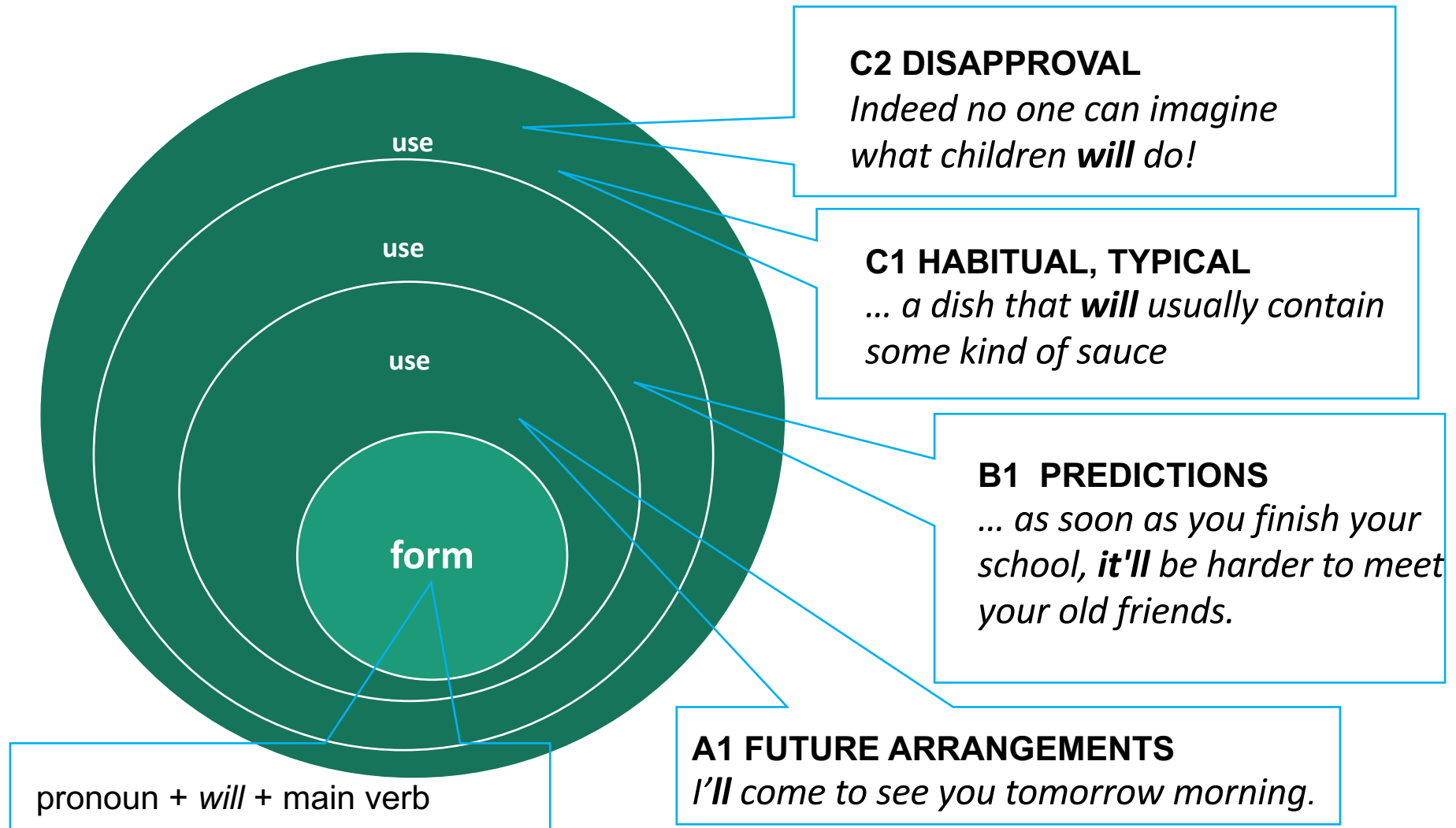
The meaning is in the phrase

- *I'll **see** you tomorrow.*
- ***You will** see me tomorrow.*

- pronoun + modal + main verb

- *You must bring crisps … (A1)*
- *I must admit … (C2)*
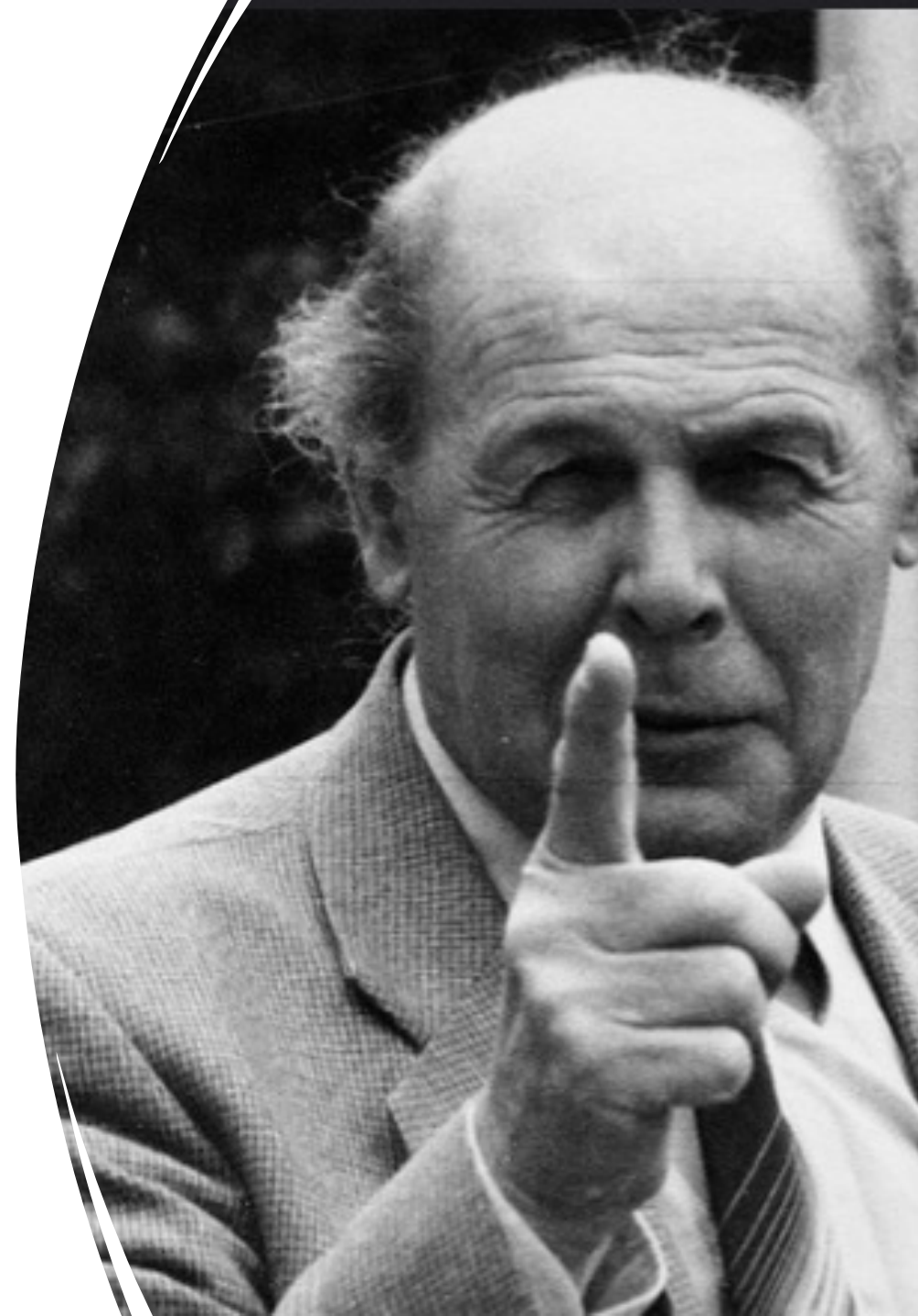
Some key insights

- Lexico-grammatical development

- Competence emerges both syntactically and functionally

- Error and competence are not mutually exclusive

- Mismatch between ELT syllabi and competence

Slots and frames
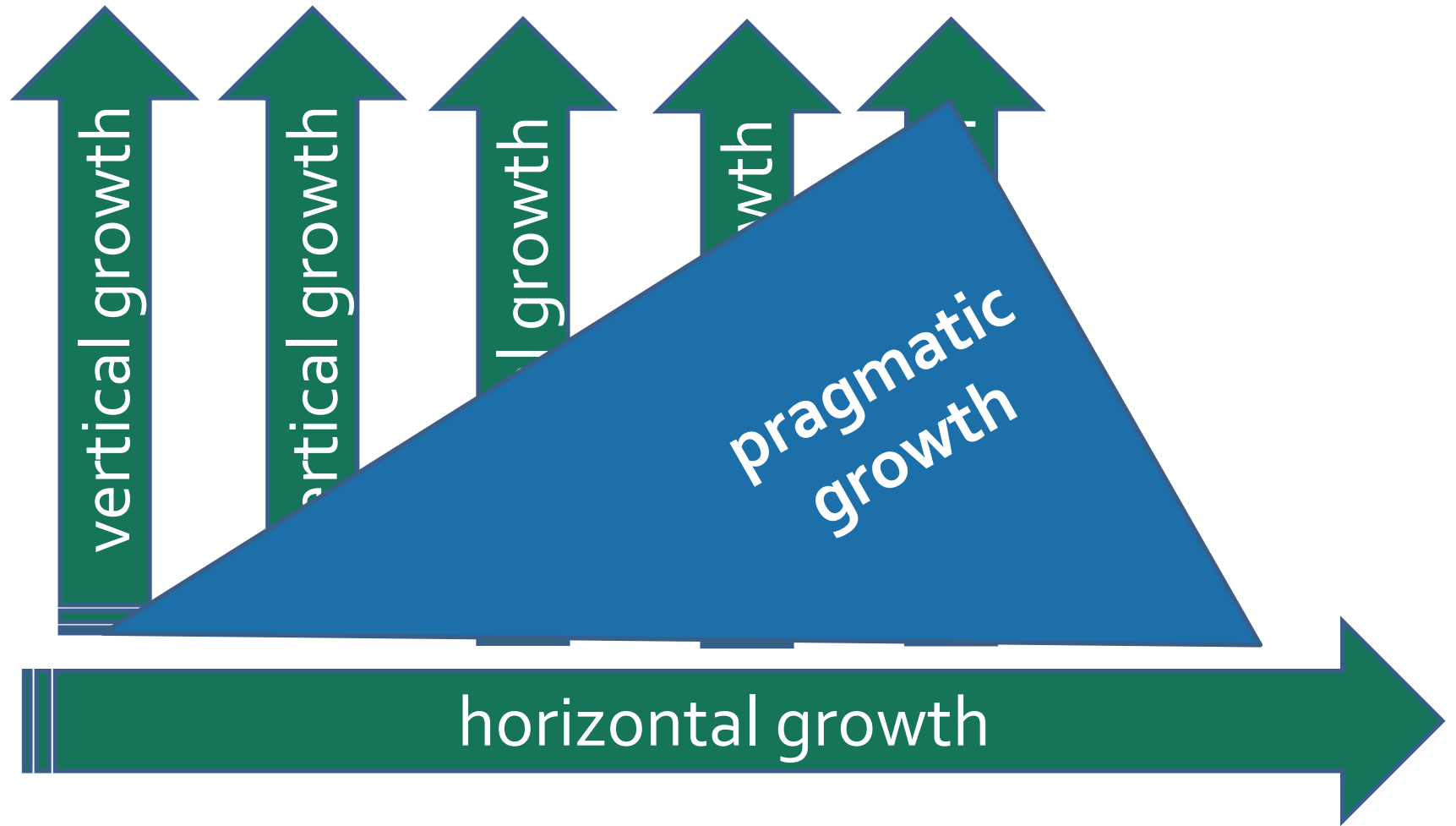
Other ways in

- Top-down

- Bottom-up

Past simple [tag="V.D"] Top 10

BNC	%	A1		A2		B1		B2		C1		C2	
was	20.2	✓	18.1	✓	23.6	✓	23.8	✓	24.6	✓	25.8	✓	25.0
had	10.9	✓	3.4	✓	4.9	✓	7.0	✓	8.8	✓	9.9	✓	11.9
were	5.3	✓	5.1	✓	5.2	✓	6.9	✓	8.1	✓	9.7	✓	7.4
said	4.8			✓	2.3	✓	1.8	✓	1.4	✓	1.3	✓	1.2
did	3.1			✓	2.9	✓	3.3	✓	3.6	✓	4.1	✓	4.1
came	1.3					✓	2.6	✓	1.1	✓	1.1	✓	1.1
went	1.2	✓	11.2	✓	8.1	✓	3.6	✓	3.0	✓	1.8	✓	1.5
looked	1.1												
thought	1.1												
took	1.0										✓	1.3	✓	1.3
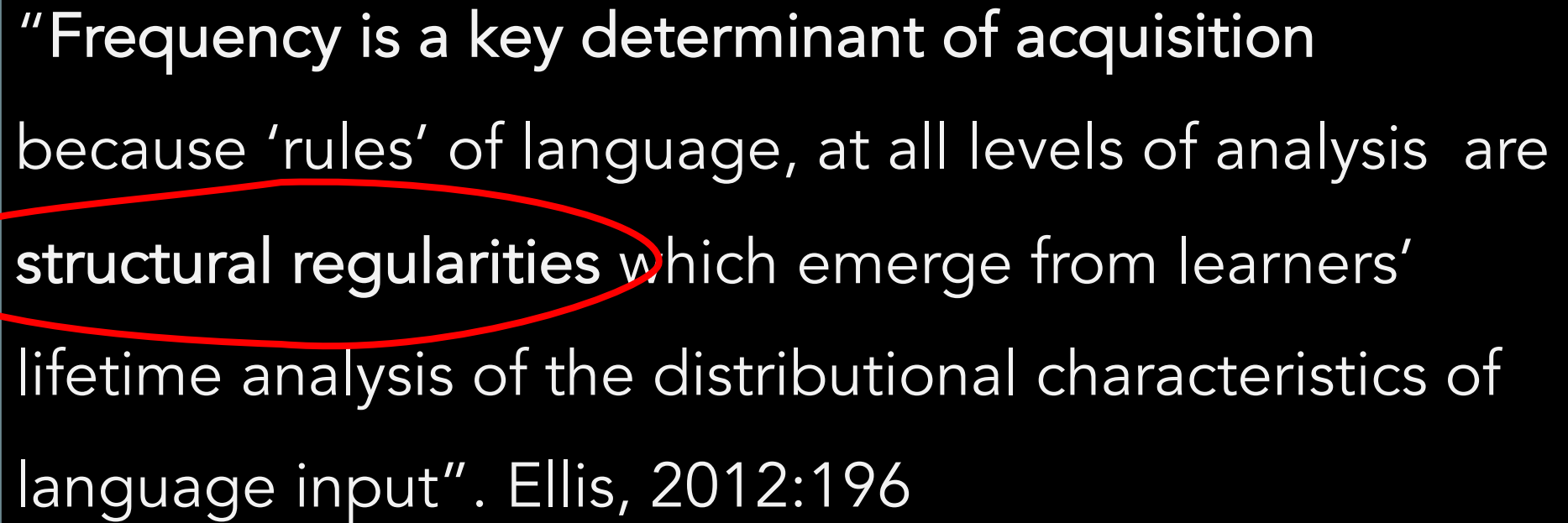
Development in L2 – usage-based

- **A theory of language acquisition** from first language acquisition studies
- Language comes from **experience**
- Our brains are sensitive to the **frequency of forms and meanings**.
- There is a **pattern** of **language development** seen in L1 development.

formula → limited slot and frame pattern → full productive system

Investigating whether/how this occurs and develops in learner language.

Putting together the building blocks

"Frequency is a key determinant of acquisition because 'rules' of language, at all levels of analysis are structural regularities which emerge from learners' lifetime analysis of the distributional characteristics of language input". Ellis, 2012:196

Building blocks

FORM MEANING MAPPINGS

- morphemes: *-aholic, -s*

- simple words: *nut*

- idioms: *It is driving me nuts*

- abstract syntactic frames: Subject-Verb-Object-Object

  She gave him some flowers TRANSFER VAC

Building blocks of a language that derive from form-meaning pairings that become cognitively **entrenched**
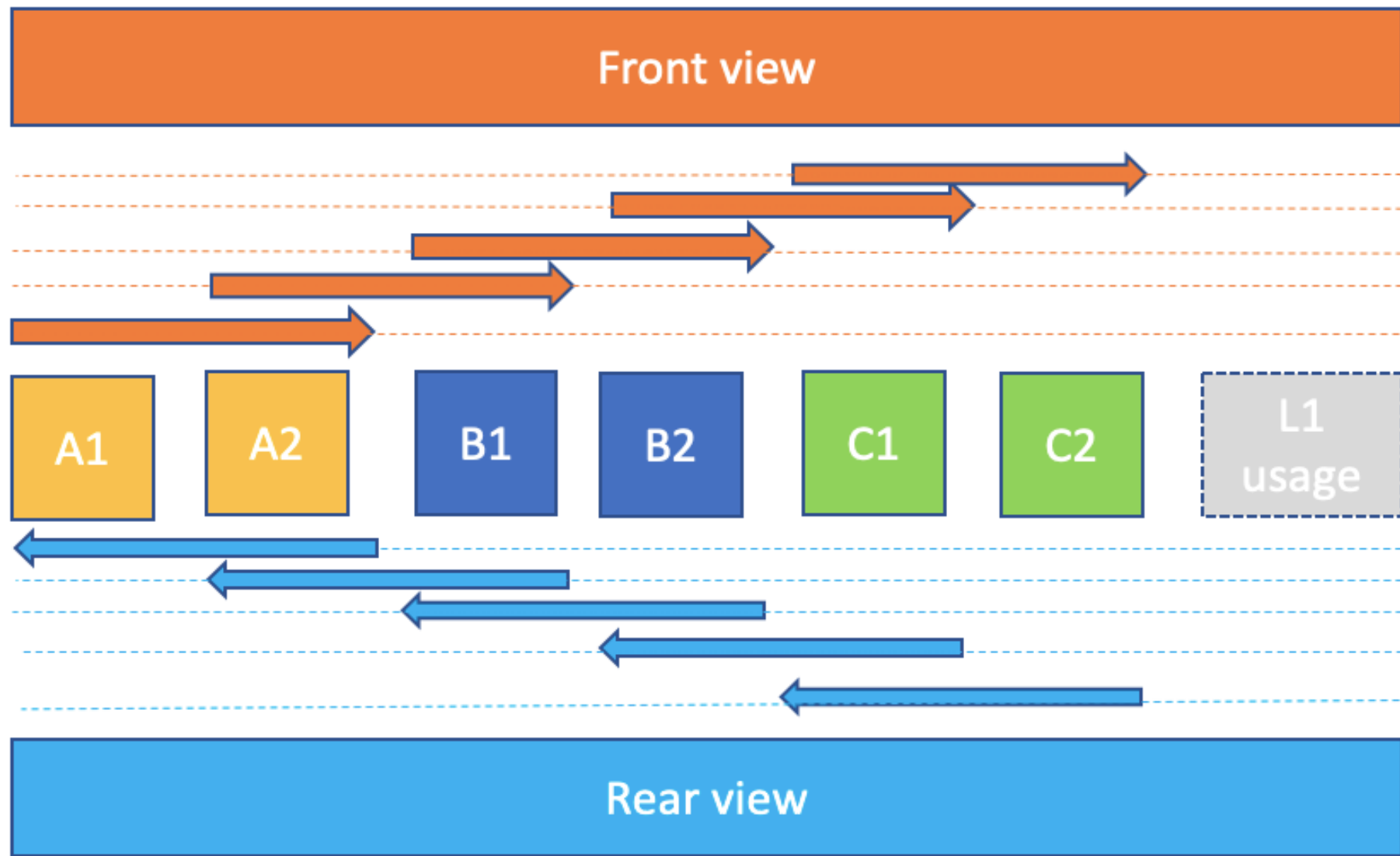
Sequences of 4-grams A2

word + word word + word

How are you ?	to go to the	! How are you	I think you should
are you ? I	I would like to	. I think that	Thank you for your
I "m going to	Would you like to	you ? I "m	. I do n"t
, How are you	in the centre of	. I hope you	are a lot of
. See you soon	. Would you like	to a new shop	you like to go

Sequences of 4 POS tags

tag + tag + tag + tag (open slot sequences)

e.g. noun + preposition + determiner + noun

determiner + adjective + noun + preposition

The yellow house with. (A2)

A wide range of (C1/C2)

A2 sequences of 4 POS tags

NN IN DT NN	PP MD VV IN	NN IN PPZ NN	DT NN IN DT
IN DT JJ NN	SENT PP MD VV	DT NN SENT PP	NP NP , PP
IN DT NN SENT	NN SENT PP VVP	NP NP NP NP	PP MD VV PP
IN DT NN IN	DT JJ NN IN	VV IN DT NN	SENT PP VVP PP
IN PPZ NN SENT	DT JJ NN SENT	PP VVP TO VV	JJ NN SENT PP

NN IN DT NN = noun + preposition + determiner +noun

e.g. *centre of the city, front of the cinema, price of the ticket*

Questions

- Is development in L2 writing observable through the frequency and distribution of POS sequences across proficiency levels?

- How does POS sequence usage develop across proficiency levels?
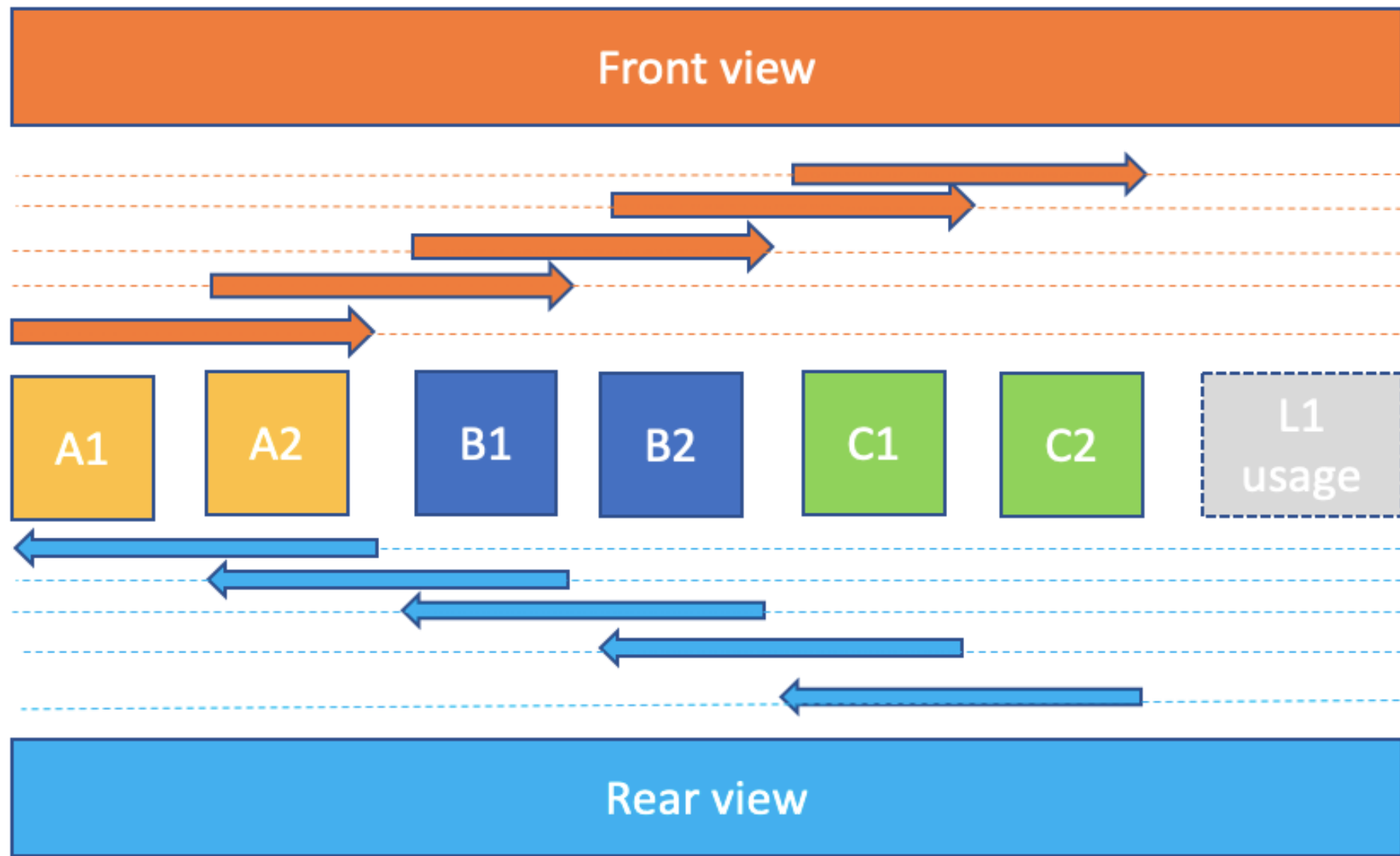
Top 5 A1 POS-tag sequences

POS tag sequence

A1 ranking

difference in ranks at other levels

Rank difference between top A1 sequences and other levels					
A1 4-gram POS tag sequences	**A2**	**B1**	**B2**	**C1**	**C2**
1 .+pronoun+modal+verb SENT PP MD VV *. I would like*	-6	-16	-20	-35	-69
2 pronoun+modal+verb+preposition PP MD VV IN *You can come to*	-4	-10	-34	-58	-93
3 preposition+determiner+noun+. IN DT NN SENT *in the morning.*	0	1	0	-3	-3
4 preposition+posspronoun+noun+. IN PPZ NN SENT *to my house.*	-1	-2	-4	-13	-14
5 noun+.+pronoun+presentsimpleV NN SENT PP VVP *phone. I like*	-3	-10	-19	-38	-85
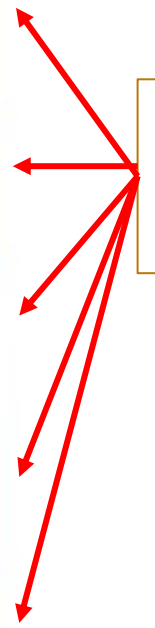
Top 5 C2 POS-tag sequences

POS tag sequence

C2 ranking

difference in ranks at other levels

	Rank difference between top C2 sequences and other levels					
	4-grams POS tag sequences	**A1**	**A2**	**B1**	**B2**	**C1**
1	noun+preposition+determiner+noun NN IN DT NN _aim of this report_	-9	0	0	0	0
2	preposition+determiner+adjective+noun IN DT JJ NN _on the other hand_	-60	0	-1	0	0
3	preposition+determiner+noun+preposition IN DT NN IN _in the middle of_	-12	-1	-2	-1	-1
4	determiner+adjective+noun+preposition DT JJ NN IN _a great deal of_	-105	-5	-5	-2	-1
5	determiner+noun+preposition+determiner DT NN IN DT _the aim of this_	-27	-11	-2	0	2
	preposition+determiner+noun+ IN DT NN SENT					
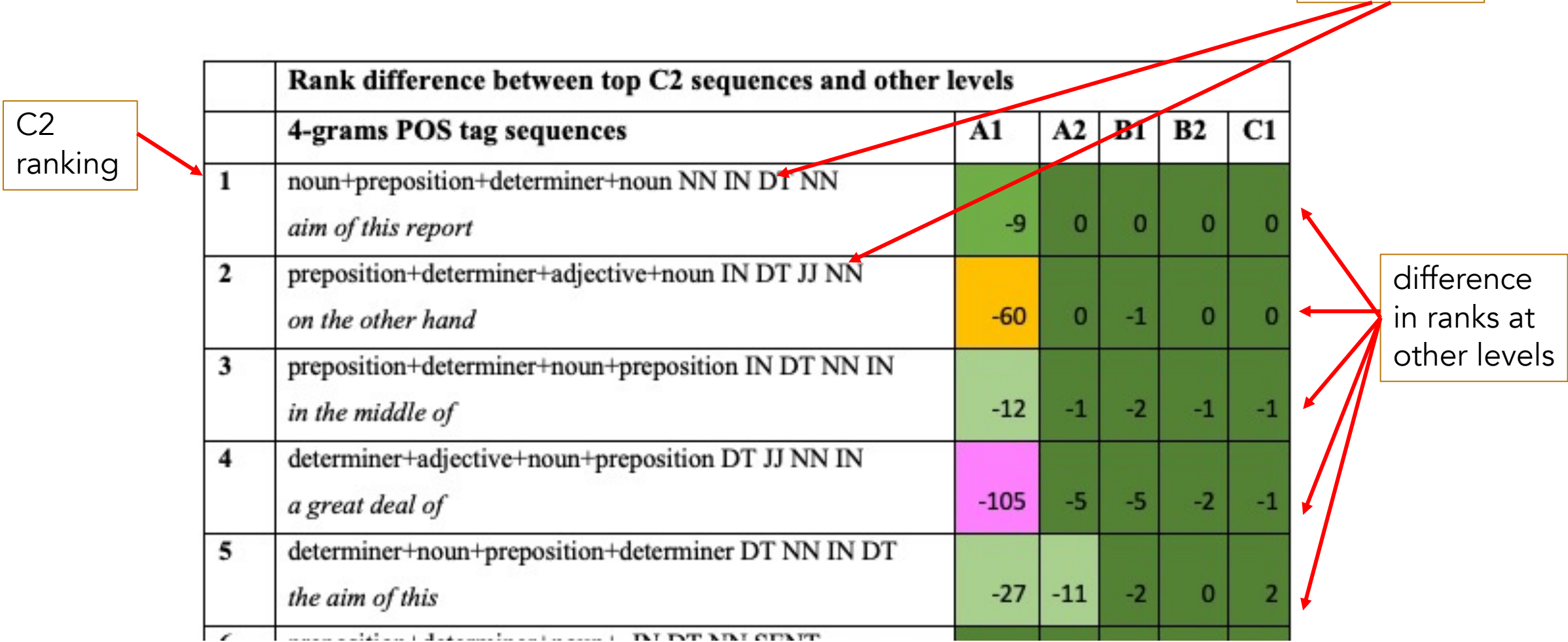
FINDINGS

Sequence usage converges as proficiency increases

C2 View Top 100

	C2:A1	C2:A2	C2:B1	C2:B2	C2:C1
no instances	16	5	0	0	0
+/-500	10	9	0	0	16
+/-100 to 500	34	29	29	15	29
+/-31 to 100	21	23	33	33	18
+/-11 to 30	13	19	16	27	37
+/-10	3	4	6	6	
+/-5	3	11	13	19	

Legend: ■ +/-5 ■ +/-10 ■ +/-11 to 30 ■ +/-31 to 100 ■ +/-100 to 500 ■ +/-500 ■ no instances
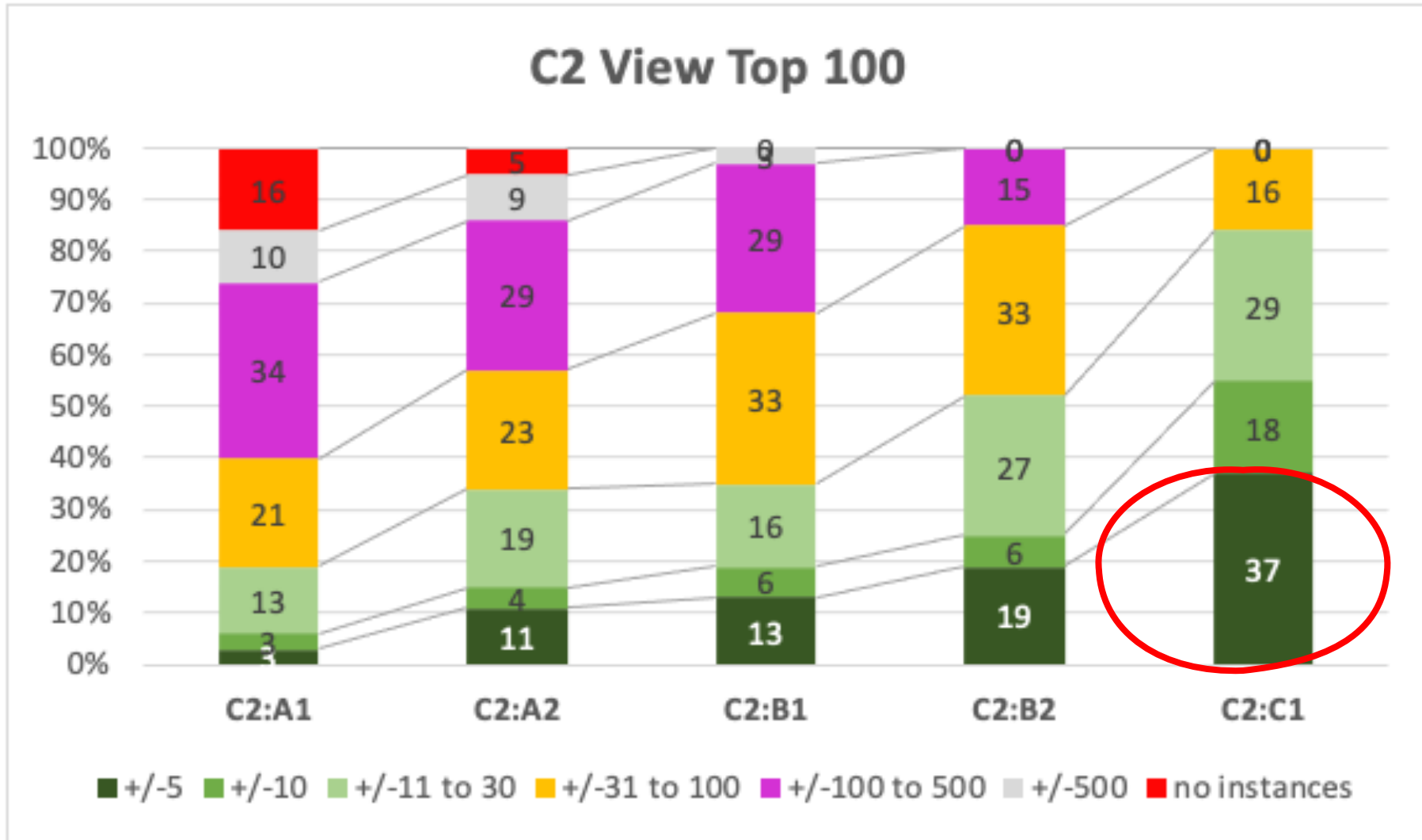
Sequence types

- Core
- Emerging
- Decreasing

Core sequences: A1 to C2

A1 rank		A2
2	PP MD VV IN	-4
11	NN IN PPZ NN	0
19	VV IN DT NN	5
26	PP MD VV DT	2
29	PP VBP VVG TO	-5
37	VBP VVG TO VV	-4

A2 rank		A1	B1
1	NN IN DT NN	-9	0
2	IN DT JJ NN	-60	-1
4	IN DT NN IN	-11	-1
9	DT JJ NN IN	-100	0
11	NN IN PPZ NN	0	-2
18	PP MD VV PP	-28	0
24	PP MD VV DT	-2	-4
31	DT NN IN NN	-53	-1
33	PP VVP PP MD	-62	2
42	IN DT NN CC	-137	2

B1 rank		A2	B2
1	NN IN DT NN	0	1
3	IN DT JJ NN	1	0
5	IN DT NN IN	1	-1
7	DT NN IN DT	-9	2
9	DT JJ NN IN	0	3
10	TO VV DT NN	-18	1
13	NN IN PPZ NN	2	-2
25	VV DT NN IN	-11	5
28	PP MD VV DT	4	1
39	IN DT NN PP	-23	1

B2 rank		B1	C1
1	NN IN DT NN	0	0
2	IN DT JJ NN	-1	0
4	IN DT NN IN	-1	0
5	DT NN IN DT	-2	2
6	DT JJ NN IN	-3	1
9	TO VV DT NN	-1	0
13	PP MD RB VV	-6	1
14	NN IN PPZ NN	1	-1
20	VV DT NN IN	-5	4
31	TO VV IN DT	-14	-1
41	IN DT NN PP	2	4
42	DT NN IN PPZ	7	3
48	VBZ DT JJ NN	-25	1
49	MD VV DT NN	-20	0

C1 rank		B2	C2
1	NN IN DT NN	0	0
2	IN DT JJ NN	0	0
3	DT NN IN DT	-2	-2
4	IN DT NN IN	0	1
5	DT JJ NN IN	-1	1
9	TO VV DT NN	0	-1
10	DT NN IN NN	-8	3
11	JJ NN IN DT	-12	0
12	PP MD RB VV	-1	-3
13	NNS IN DT NN	-13	0
14	NN IN DT JJ	-19	2
15	NN IN PPZ NN	1	1
16	VV DT NN IN	-4	-1
22	IN DT NN CC	-6	3
24	VVN IN DT NN	-26	4
32	TO VV IN DT	1	-3
34	RB IN DT NN	-11	5
42	NN IN DT NNS	-43	-1
47	VBZ DT JJ NN	-1	2

C2 rank		C1
1	NN IN DT NN	0
2	IN DT JJ NN	0
3	IN DT NN IN	-1
4	DT JJ NN IN	-1
5	DT NN IN DT	2
7	DT NN IN NN	-3
10	TO VV DT NN	1
11	JJ NN IN DT	0
12	NN IN DT JJ	-2
13	NNS IN DT NN	0
14	NN IN PPZ NN	-1
15	PP MD RB VV	3
17	VV DT NN IN	1
19	IN DT NN CC	-3
20	VVN IN DT NN	-4
29	RB IN DT NN	-5
35	TO VV IN DT	3
43	NN IN DT NNS	1
45	VBZ DT JJ NN	-2
48	JJ IN DT NN	-3

- Top ranking sequences are identical across all levels.
- Noun phrases dominate and becoming increasingly relevant
- Sequences with adjectives in noun phrases are a marker of transition between A1 and A2
- There is stabilisation between A2 and B1
- There is a leap in development between B1 and B2

Increase in emerging sequences

A1 rank		A2
10	NN IN DT NN	9
15	IN DT NN IN	11
32	DT NN IN DT	16
36	IN DT NN NN	11
40	NP NP NP NP	27
43	TO VV DT NN	15
46	PP MD VV PP	28

A2 rank		A1	B1
13	NP NP NP NP	-27	9
16	DT NN IN DT	-16	9
28	TO VV DT NN	-15	18
29	PP MD RB VV	-177	10
36	VV DT NN IN	-60	11
40	PP VHP TO VV	22	6
44	DT NN IN PPZ	-23	9
46	PP VVD DT NN	16	19

B1 rank		A2	B2
19	PP MD RB VV	-10	6
29	VV IN DT NN	15	7
32	DT NN IN NN	1	14
40	IN DT NN CC	-2	12
41	PP MD VV TO	-19	25
42	MD VV TO VV	-19	27
43	IN DT NN NN	18	13
45	TO VV IN DT	15	14
47	JJ NN IN DT	10	24

B2 rank		B1	C1
18	DT NN IN NN	-14	8
23	JJ NN IN DT	-24	12
26	NNS IN DT NN	-32	13
28	IN DT NN CC	-12	6
30	IN DT NN NN	-13	10
33	NN IN DT JJ	-34	19
39	DT NN IN NNS	-16	8
43	DT NN TO VV	-67	14
45	RB IN DT NN	-57	11
50	VVN IN DT NN	-161	26

C1 rank		B2	C1	C2
29	DT NN TO VV	-14	6	
31	DT NN IN NNS	-8	9	
38	VV DT JJ NN	-15	6	
39	DT NN IN PPZ	-3	14	
45	JJ NN IN NN	-68	17	
46	IN DT JJ NNS	-58	12	

C2 rank		C1
22	DT NN IN NNS	-9
23	DT NN TO VV	-6
25	DT NN IN PPZ	-14
28	JJ NN IN NN	-17
32	VV DT JJ NN	-6
34	IN DT JJ NNS	-12
38	NN IN JJ NNS	-20
40	IN PPZ JJ NN	-12
42	IN NN IN DT	-14
46	JJ NNS IN DT	-15
49	IN NN TO VV	-25
50	DT NN IN JJ	-22

- Noun phrases continue to dominate

- Sequences containing tensed verbs peak at B1

- There is a leap in development between B1 and B2

Decreasing sequences

A1 rank		A2
9	PP VVP TO VV	-6
12	VV IN PPZ NN	-66
16	IN PPZ NN IN	-58
18	PP VHP TO VV	-22
22	CD NN "" NN	-118
24	IN CD : CD	-164
30	PP VVD DT NN	-16
31	PP VBP VVG IN	-55
35	IN CD NN ""	-142
41	NN IN CD NN	-271
48	MD VV IN PPZ	-172
49	PP VVP DT NN	-16
50	DT NN VBZ IN	-341

A2 rank		A1	B1
6	PP MD VV IN	4	-6
14	VV IN DT NN	-5	-15
15	PP VVP TO VV	6	-7
25	IN DT NN NN	-11	-18
30	TO VV IN DT	-31	-15
34	PP VBP VVG TO	5	-16
35	PP VVP RB VV	-184	-25
37	JJ NN IN DT	-337	-10
38	DT NN IN NNS	-43	-17
39	DT NN IN PP	-27	-15
41	VBP VVG TO VV	4	-15
45	TO VV IN PP	-8	-40
49	MD VV IN DT	-61	-76

B1 rank		A2	B2
4	NP NP NP NP	-9	-8
12	PP MD VV IN	6	-24
18	PP MD VV PP	0	-7
22	PP VVP TO VV	7	-12
27	PP VVD DT NN	-19	-33
31	PP VVP PP MD	-2	-35
34	PP VHP TO VV	-6	-38
35	DT NN IN PPZ	-9	-7
36	PP VVD TO VV	-47	-10
38	NN IN PPZ NNS	-346	-197
48	PP VBP VVG IN	-38	-40
50	PP VBP VVG TO	16	-73

B2 rank		B1	C1
12	NP NP NP NP	8	-6
15	MD VV TO VV	-27	-13
16	PP MD VV TO	-25	-17
22	VV IN DT NN	-7	-8
25	PP MD VV PP	7	-25
27	PP MD VV DT	-1	-8
34	PP VVP TO VV	12	-29
36	PP MD VV IN	24	-24
46	PP VVD TO VV	10	-36

C1 rank		B2	C2
18	NP NP NP NP	6	-33
20	IN DT NN NN	-10	-6
28	MD VV TO VV	13	-27
30	VV IN DT NN	8	-6
33	PP MD VV TO	17	-38
35	PP MD VV DT	8	-26
37	IN DT NN PP	-4	-7
49	MD VV DT NN	0	-13
50	PP MD VV PP	25	-86

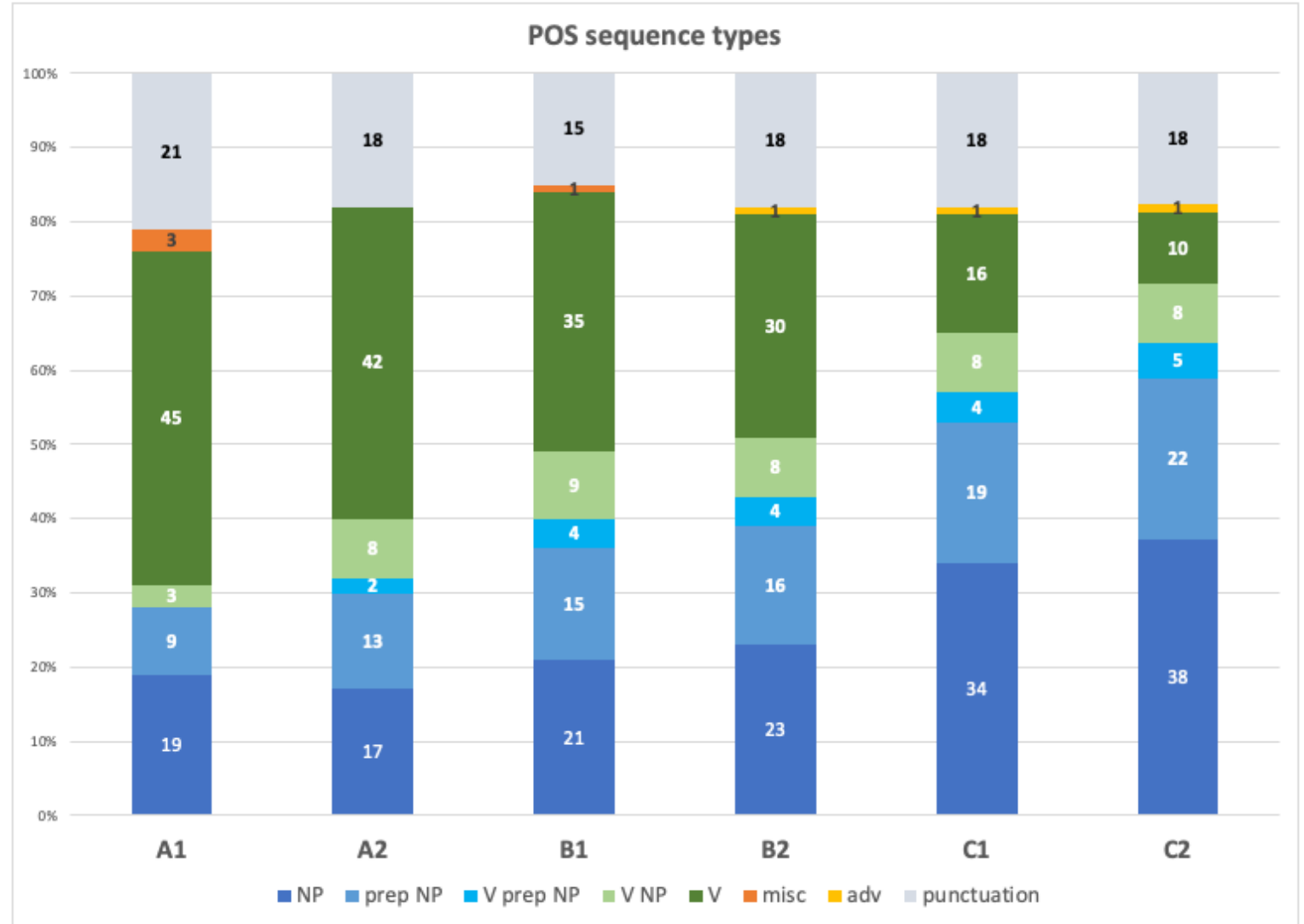| C2 rank		C1
26	IN DT NN NN	6
36	VV IN DT NN	6
44	IN DT NN PP	7

- Decreasing sequences decrease as proficiency increases
- Verb phrases continue to decrease
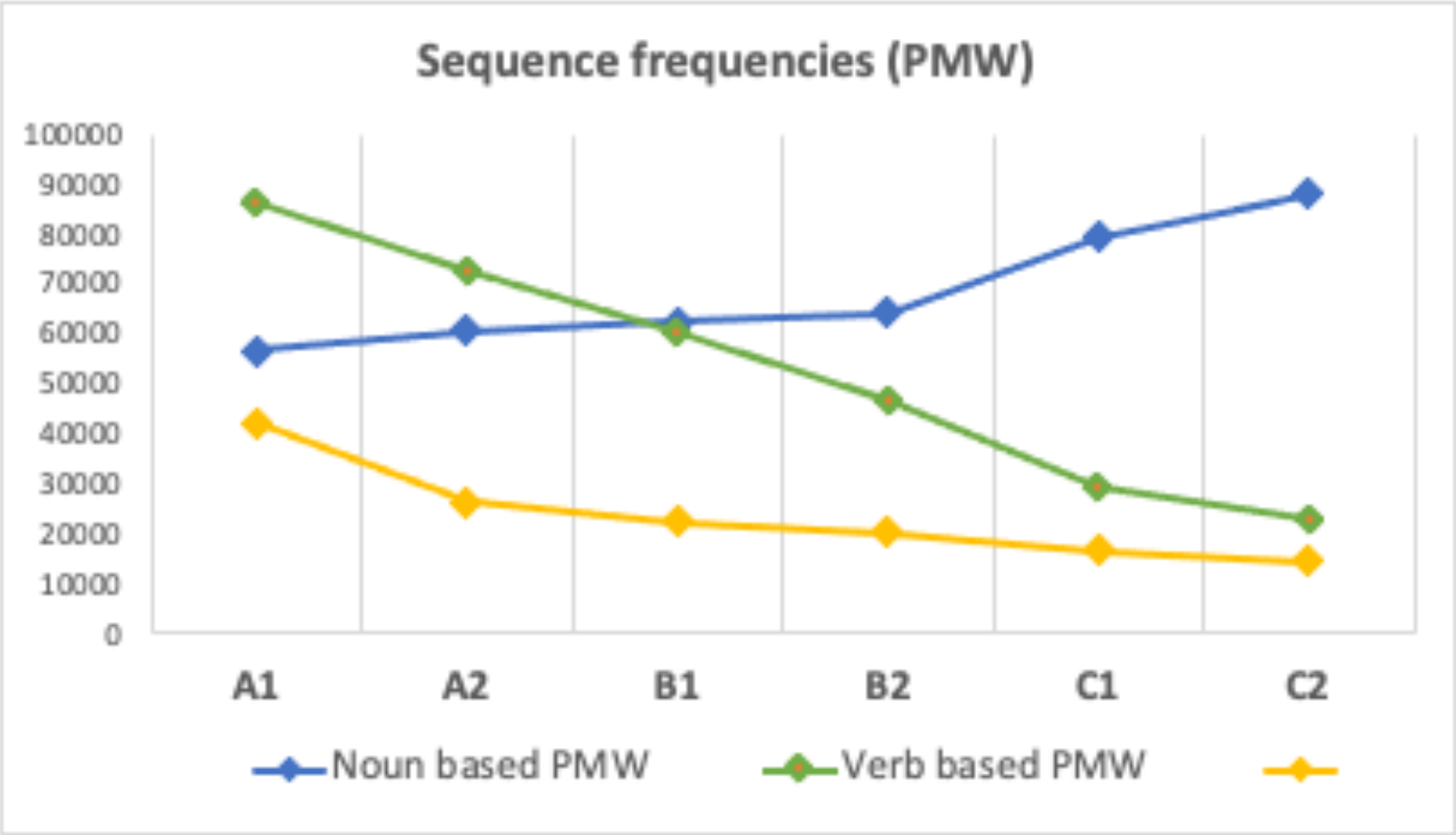- Sequences start to stabilise between A1 and A2

Noun sequences increase as proficiency increases

Top 100 sequences

Categorised based on Gray & Biber (2013)

POS sequence types

B1 level is a pivotal point in sequence usage

Sequence frequencies (PMW)

Lexical and functional usage changes as proficiency increases
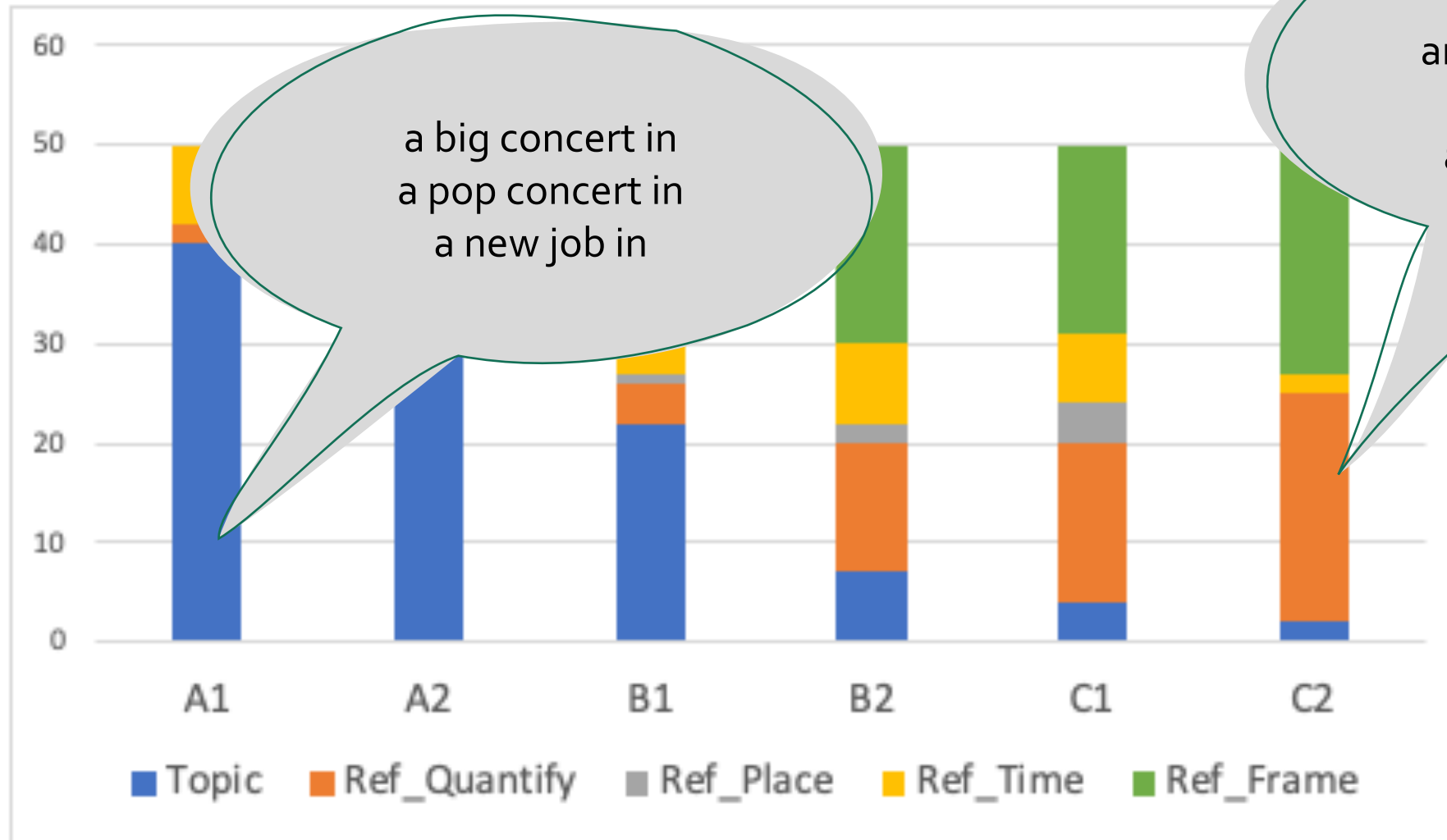
DT JJ NN IN

det + adj + noun + prep

a new pair of, a new job in
a wide range of, a great deal of

- Referential: quantity, framing, place, time
- Stance
- Discourse

Biber et al. 2014

From Topics to References & Frames

rank	A1	A2	B1	B2	C1	C2
1	a new pair of	a large school in	a large school in	a wide range of	a wide range of	a wide range of
2	this mobile phone because	a new shop in	a small school in	a new shop in	a great deal of	a great deal of
3	a new job in	a small school in	the large school in	a new collection of	a wide variety of	a great number of
4	a big concert in	the large school in	a long time since	a special day in	a great number of	a large number of
5	a good time with	a new shop near	a special day in	a free copy of	a large number of	a wide variety of
6	a pop concert in	the small school in	a great time with	a great deal of	an important role in	an important role in
7	a pop concert on	the new shop in	a good time with	the other side of	the other side of	the other side of
8	a new house in	a long time since	the new class because	a long time since	a great variety of	a great amount of
9	a pop concert with	a new bed for	the small school in	an essential part of	a great opportunity for	the vast majority of
10	a mobile phone for	a good time with	the other side of	a great number of	the back row in	an important part of

PIONEER

- N + of
- Fixedness of patterning
- Specialised function: quantity

- Task effect

Formula to slot and frame to formulaic

VVD IN DT NN   A2
past simple + preposition + determiner + noun

went	~~into~~	~~the~~	~~restaurant~~
looked	in	the	shop

A2	B2	C2	BNC
went to the cinema	turned on the television	did in the past	came to an end
looked in the shop	played for an hour	came to the conclusion	came to a head
went to the shopping	walked into the house	went to the kitchen	said that the government
went to a shop	went to the theatre	went to the hospital	came to the conclusion
went to the shop	went to the beach	died at the age	drew to a close
went to the theatre	went to a shop	went for a walk	came to the rescue
went to the mall	went to the festival	happened in the past	said in a statement
went to the house	went to the kitchen	died in a car	said at the time
went to a restaurant	went to the cinema	came as a surprise	died at the age
went to a pub	included in the offer	included in the price	came in the form
went to the police	went for a walk	came as no surprise	came in the wake
went to the city	went to the restaurant	came as a shock	came to the fore
went to the disco	went to the hospital	arrived at the museum	came as no surprise
went to the beach	arrived at the airport	arrived at the airport	said that the company

VVD IN DT NN Top 60 exponents

	A2	B2	C2
	went to the/a	went to the/a	went to the
FUNCTION			
movement to thing	came to a ran to the entered in the	walked into came into the arrived at the	came into moved to a arrived at the
action on thing	looked in the window heard about the ... knocked o...	looked through the window ... the TV ... the door stayed in a hotel	looked at the knocked on the looked through the listened to the
state / action	lived in that house went for a walk	worked as a teacher played for a hour included in the offer/price	went for a walk did in the past died at the age died in a car lived in the participated in the
formulaic evaluative		came to the conclusion	came as no surprise came as a shock came to an end added to the fact that came to the conclusion

FORMULA

SLOT AND FRAME

FORMULAIC FULLY ABSTRACTED

VBD RB VVN IN
was/were + adverb + -ed form + preposition

A2
was very pleased with
was very excited because
was very surprised about
was very surprised because
was really surprised because
was very surprised with
was very pleased about
was very excited about
was really excited because
was completely lost in
was really worried about
was really surprised about
was so excited because
were really scared about
was very worried because
was very upset because
was very surprised for
was very surprised at
was very suprised because
was very impressed about

B2
was very satisfied with
was very disappointed with
was very impressed by
was not satisfied with
was very pleased with
was very disappointed because
was very excited because
was very disappointed about
was very surprised because
were very satisfied with
were very pleased with
were not satisfied with
were not ruined by
was very worried about
was really surprised by
was not ruined by
was not pleased with
was n''t entered for
were very impressed by
were very excited because

C2
was recently published in
was very impressed by
were very satisfied with
were not satisfied with
was very satisfied with
was very pleased with
was n''t used to
was not satisfied with
was very impressed with
was particularly impressed by
was not included in
was n''t satisfied with
were very excited about
were so impressed by
were not used to
were not included in
were generally satisfied with
were also satisfied with
was very surprised of
was very pleased by

VBD RB VVN IN
was/were + adverb + -ed form + preposition

A2
Thank you for your letter. I **was very pleased with** *it*.

B2
I think that is a good idea to inform you *that I* **was very pleased with** *the guide*, Rebecca Smith.

C2
… she **was very excited about** *the idea of telling her mother all the wonderful things that had happened to her*.

Complexity & dexterity

A2	was	very	pleased	about
	were	really	surprised	with
			excited	
			worried	
B2			disappointed	
			impressed	by
			satisfied	
			ruined	
C2		particularly		
		so		
		generally		

Adverb modifying VVN in relative clause:
backgrounding
evaluating

out being young, which **was recently published in** the magazine "Feel
realms afgnosis which **were completely lost to** humanity a couple
heir country and which **was openly displayed by** means of flags, T-sh
some animations which **were sometimes done at** the chalet, they wei
isit the buildings which **were formerly known as** "Educational Institu
the town centre which **was fully equipped with** many amenities. St
The Activity Day which **was recently held by** our Social Club was
ies into domains which **were formerly characterized by** a total lack of medic
oomy farmhouse which **was also occupied by** my parents an seve
y watched a film which **was first released on** January 1992 and it
veloping the land which **was kindly offered to** our town by an inte
ndless language which **was only known to** us. Thus my childho
ment on a letter which **was recently published in** your newspaper, ab
gramme" neither which **was expressively mentionned in** the advertisement.
e official number which **was also published by** our organisation ha
goods, many of which **were simply dumped on** the market, Russian
wo books, one of which **was recently published in** Germany. This woul
ur door, old ones which **was n't paid for** and whenever I con
rare opportunity which **was kindly offered to** us. </p><p> Finall
ntimental person which **was unfortunately hidden behind** his huge eye glasse
reinforcements which **were permanently used during** my education helpe
ent of a resource which **was not exploited at** all. </p><p> All in
to a little school which **was normally situated in** little villages or rura
opening speech which **was unfortunatelly canceled at** the last minute and
reat deal of time which **was not organised by** my parents but by r
nd and the walls which **were specially designed at** absorbing the noise
ortance of work which **was recently raised in** your radio program
s. ago. The Work which **was previously done by** peope is being grad
ed superb works which **were only appreciated after** his death. His storie

Limitations and tagging

A2
was very pleased with
was very excited because
was very surprised about
was very surprised because
was really surprised because
was very surprised with
was very pleased about
was very excited about
was really excited because
was completely lost in
was really worried about
was really surprised about
was so excited because
were really scared about
was very worried because
was very upset because
was very surprised for
was very surprised at
was very suprised because
was very impressed about

B2
was very satisfied with
was very disappointed with
was very impressed by
was very satisfied with
was very surprised because
was disappointed about
was very surprised because
were very satisfied with
were very pleased with
were not satisfied with
were not ruined by
was very worried about
was really surprised by
was not ruined by
was not pleased with
was n''t entered for
were very impressed by
were very excited because

C2
was recently published in
was very impressed by
were very satisfied with
were not satisfied with
was very satisfied with
was very pleased with
was n''t used to
was not satisfied with
was very impressed with
was particularly impressed by
was not included in
was n''t satisfied with
were very excited about
were so impressed by
were not used to
were not included in
were generally satisfied with
were also satisfied with
was very surprised of
was very pleased by

The local newspaper has recently published an article ...

Your story must begin with this sentence: 'It was getting dark and I was completely lost.'

Summary: bottom-up approach

- Learners are sensitive to statistical frequencies in language usage.
- Sequence usage converges as proficiency increases.
- Learners demonstrate growing register awareness.
- Noun sequences increase as proficiency increases.
- B1 level is a pivotal point in sequence usage.
- Evidence of process from formula to slot and frame to formulaic, aligning with UB theory
- Lexical and functional usage changes as proficiency increases.
- There is a fixedness of patterning as proficiency increases.

Where SLA meets LCR

SLA
small scale studies
experimental
hypothesis-driven
top-down
morphological, syntactic
interlanguage
competence
longitudinal
qualitative

THIS STUDY
syntactic
interlanguage
competence
quasi-longitudinal
qualitative
quantitative
process through output
large data set
corpus-driven

LCR
larger data sets
corpus-based
phraseological or
    lexical focus
contrastive
cross-sectional
descriptions of output
    and performance
corpora built to
    address questions
quantitative

Thanks for listening

Any questions or suggestions?

References

Biber, D., & Gray, B. (2016). *Grammatical complexity in academic English: Linguistic change in writing*. Cambridge University Press.

Díez-Bedmar, M. B., & Pérez-Paredes, P. (2020). Noun phrase complexity in young Spanish EFL learners' writing: Complementing syntactic complexity indices with corpus-driven analyses. *International Journal of Corpus Linguistics*, *25*(1), 4-35.

Ellis, N. C., & Larsen-Freeman, D. (2006). Language emergence: Implications for applied linguistics—Introduction to the special issue. *Applied linguistics*, *27*(4), 558-589.

Ellis, N. C., Römer, U., & O'Donnell, M. B. (2016). Usage-based approaches to language acquisition and processing: Cognitive and corpus investigations of construction grammar.

Ellis, N. C. (2012). Formulaic language and second language acquisition: Zipf and the phrasal teddy bear. *Annual review of applied linguistics*, *32*, 17-44.

References

Ninio, A. (2006). *Language and the learning curve: A new theory of syntactic development*. Oxford University Press.

Pérez-Paredes, P., Mark, G. & O´Keeffe, A. (2020). The impact of usage-based approaches on second language learning and teaching. Cambridge Education Research Reports. Cambridge University Press.

Römer, U. (2019). A corpus perspective on the development of verb constructions in second language learners. International Journal of Corpus Linguistics, 24(3), 268-290.

Ryland Williams, Jake, Lessard, Paul R, Desu, Suma, Clark, Eric M, Bagrow, James P, Danforth, Christopher M, and Sheridan Dodds, Peter. "Zipf's Law Holds for Phrases, Not Words." Scientific Reports 5.1 (2015): 12209.

Wulff, S., & Ellis, N. C. (2018). *Usage-based approaches to second language acquisition* (Vol. 54, pp. 37-56). Amsterdam: John Benjamins

https://www.youtube.com/watch?v=fCn8zs912OE