A corpus-based grammatical analysis of modality in the writing of Spanish university students of English

Rebeca García González

UAM UNIVERSIDAD AUTONOMA DE MADRID

Financiado por el Ministerio de Ciencia e Innovación
(FFI2009-14436/FILO)

Introduction

Aim of paper:

- To explore how Spanish university students grammatically express modality in their writing across proficiency levels

Outline

- Research hypotheses
- Theoretical framework
- Corpus and procedures
- Findings
- Conclusions

Research Hypotheses

As students' level of proficiency rises, I expect:

1. A higher quantity of modal markers

2. A wider variety of modal markers

3. A movement away from *verbal* to *non-verbal* modal means

Theoretical framework

A typological account of modality:

Mood	vs.	**Modal systems**
Indicative – Modal		*A set of modal forms*
vs.		"Prototypical" means:
Subjunctive – Non-modal		Modal auxiliary verbs
(i.e. Romance languages)		(i.e. English)

(Palmer, F., 2001; 2003)

Halliday's views on modality:

COMMODITY EXCHANGED	SPEECH FUNCTION		TYPE OF INTERMEDIACY		TYPICAL REALIZATION	EXAMPLE
Information	Propo-sition	Statement Question	**Moda-lization**	**Proba-bility**	**-Finite Modal Operator** **-Modal Adjunct** **-Both the above combined**	*-They must have known* *-They certainly knew* *-They certainly must have known*
				Frequency	**-Finite Modal Operator** **-Modal Adjunct** **-Both the above combined**	*-It must happen* *-It always happens* *-It must always happen*
Goods and Services	Propo-sal	Com-mand	**Modu-lation**	**Obliga-tion**	**-Finite Modal Operator** **-Passive Verb Predicator**	*-You must be patient!* *-You´re required to be patient!*
		Offer		**Inclina-tion**	**-Finite Modal Operator** **-Adjective Predicator**	*-I must win!* *-I´m determined to win!*

(from Halliday and Matthiessen, 2004)

Nuyts' model of modality:

- Cross-linguistic analysis of modality: German, Dutch and English
- Grammatical elements expressing modality:

Modal adverbs

Modal adjectives

Modal auxiliaries

Mental state predicates:

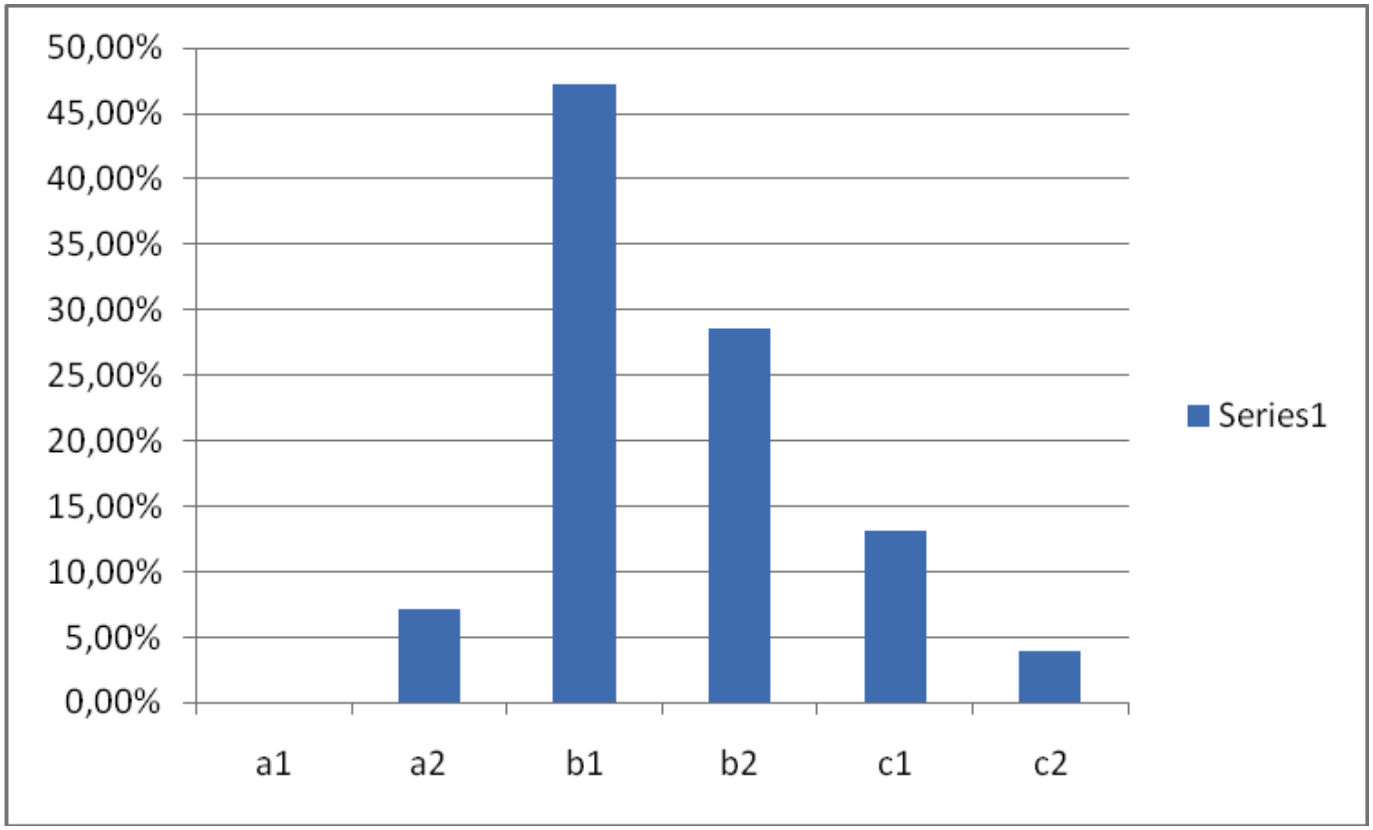
- Descriptive vs Performative
- Qualificational vs. Non-qualificational

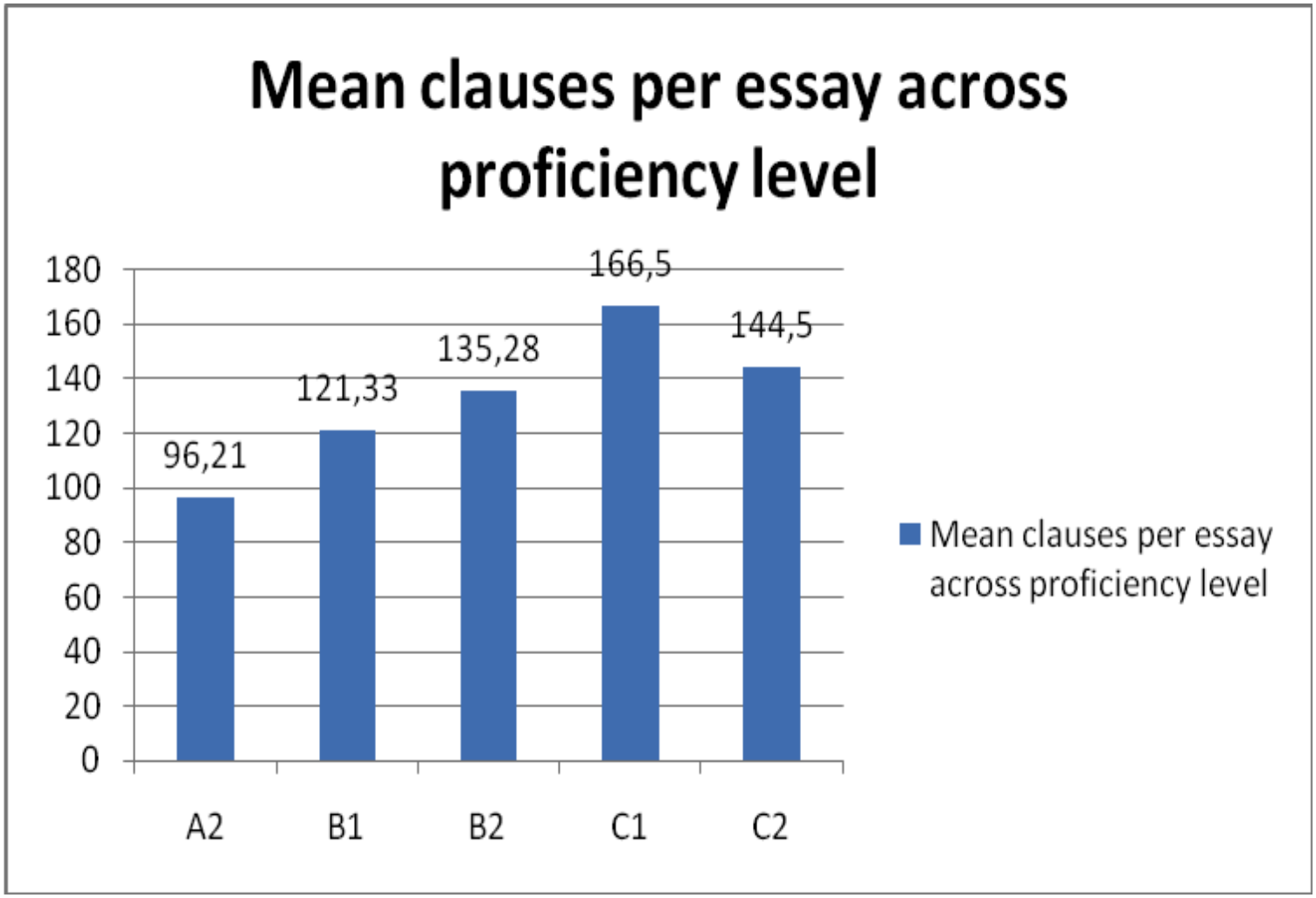
(from Nuyts, 2000)

Corpus

- The **WriCLE** corpus: a **learner corpus** of Spanish university students (Rollinson and Mendikoetxea, 2010)
- A section of **458 essays** analyzed: 445,776 words (263 first-year-student essays; 195 third-year-student essays)
- Metadata included: **Learner Profile**, with the students' level of proficiency –Oxford Quick Placement Test, UCLES(2001)

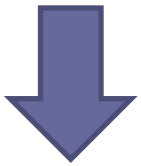
- Writers' levels of proficiency: from **A2 to C2**, following the CEFRL (Council of Europe, 2001)



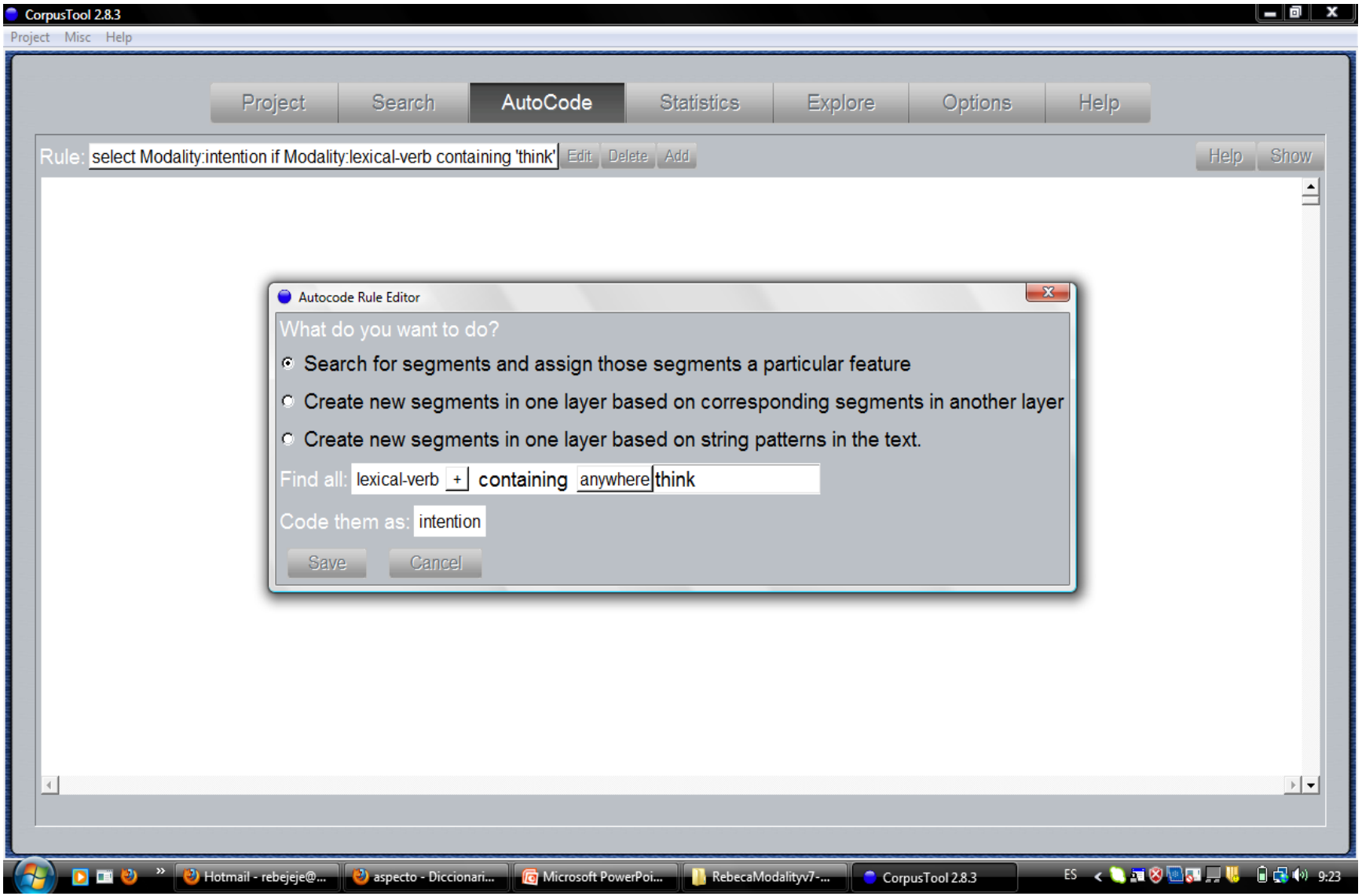

Mean clauses per essay across proficiency level
180
160
140
120
100
80
60
40
20
0
96,21
121,33
135,28
166,5
144,5
A2
B1
B2
C1
C2
Mean clauses per essay across proficiency level

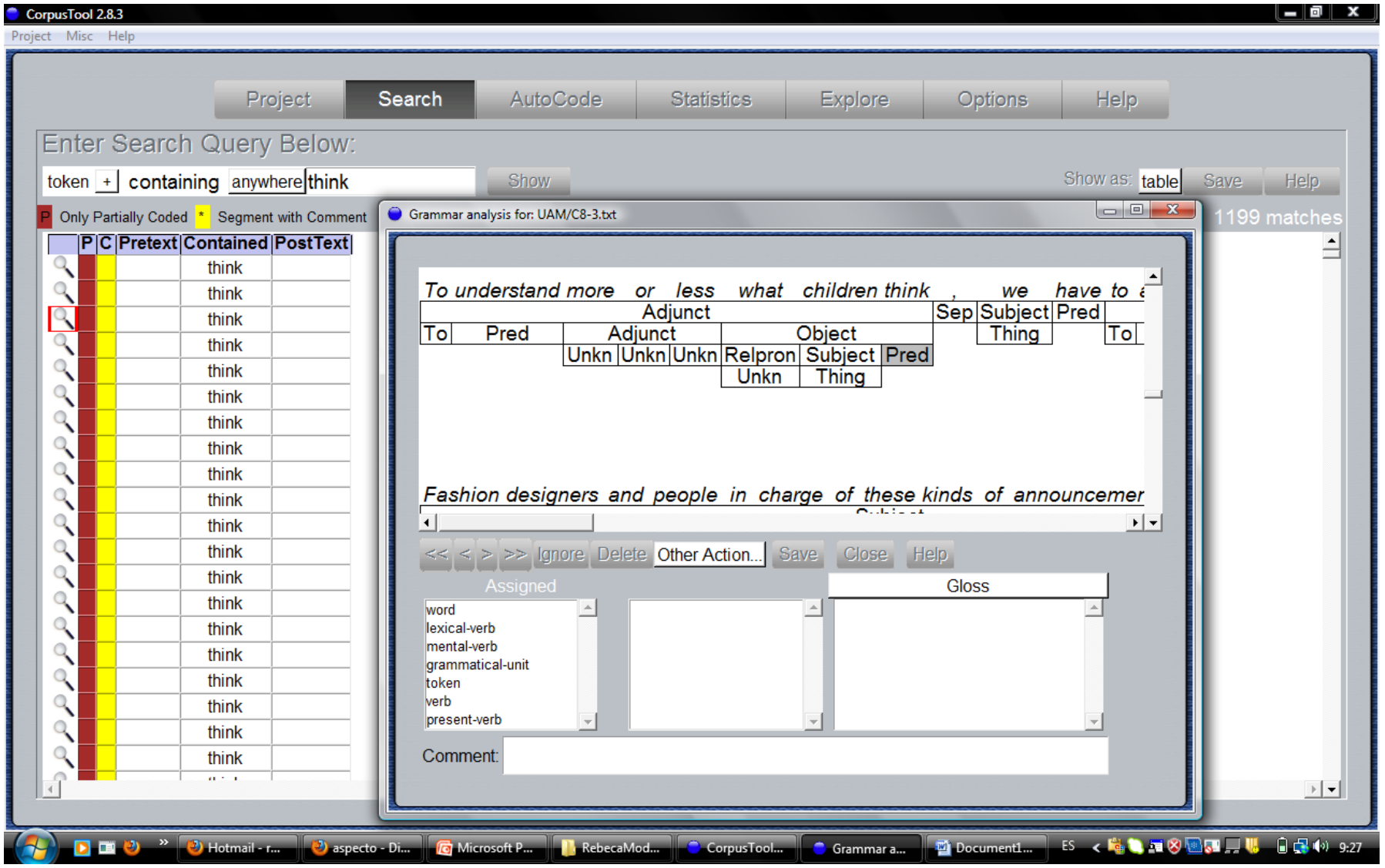
Procedures:

- **UAM CorpusTool** allows manual and automatic annotation of collections of text: pattern matching (O'Donnell, 2008)

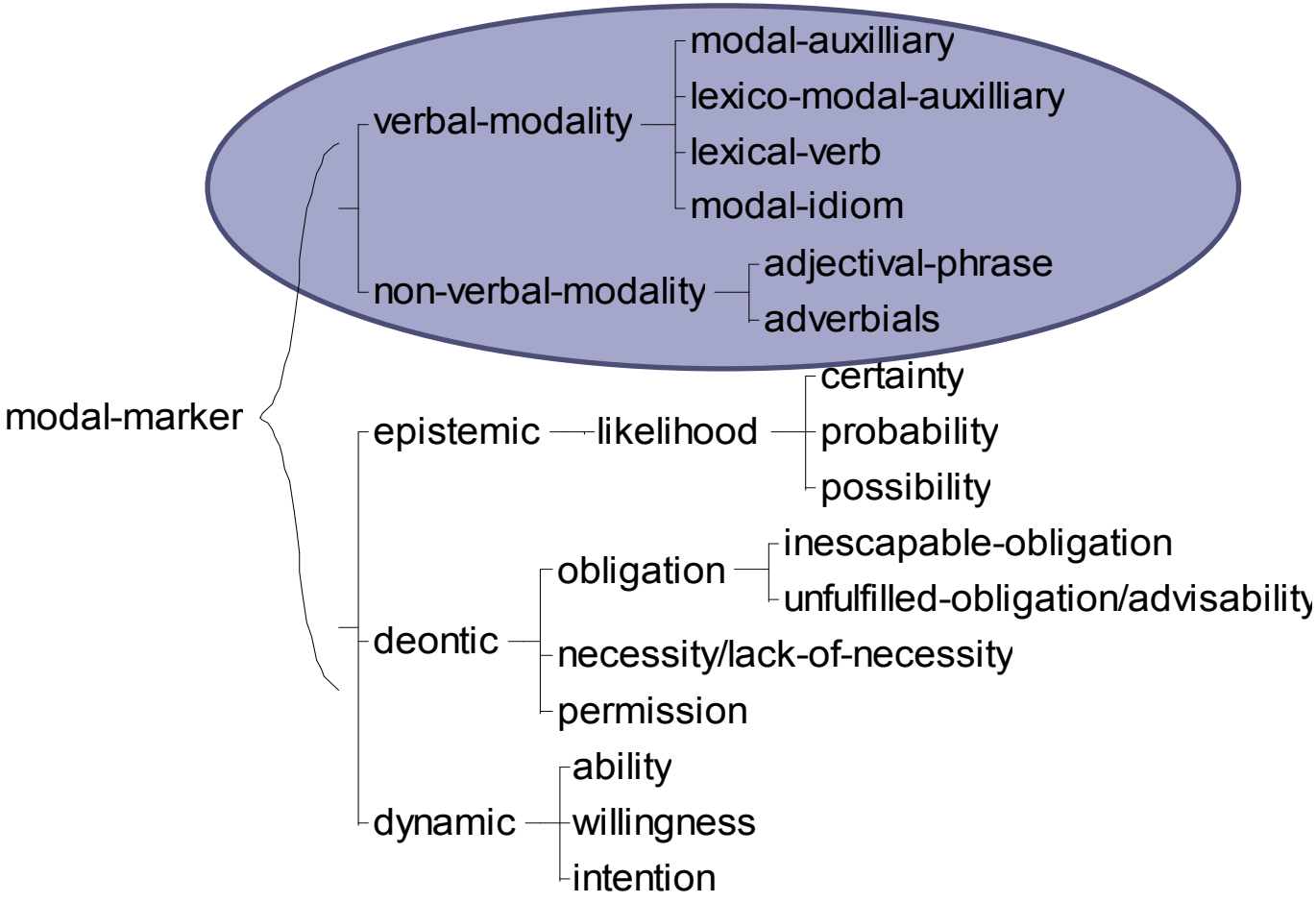


the software automatically assigns grammatical categories to modal markers following my taxonomy

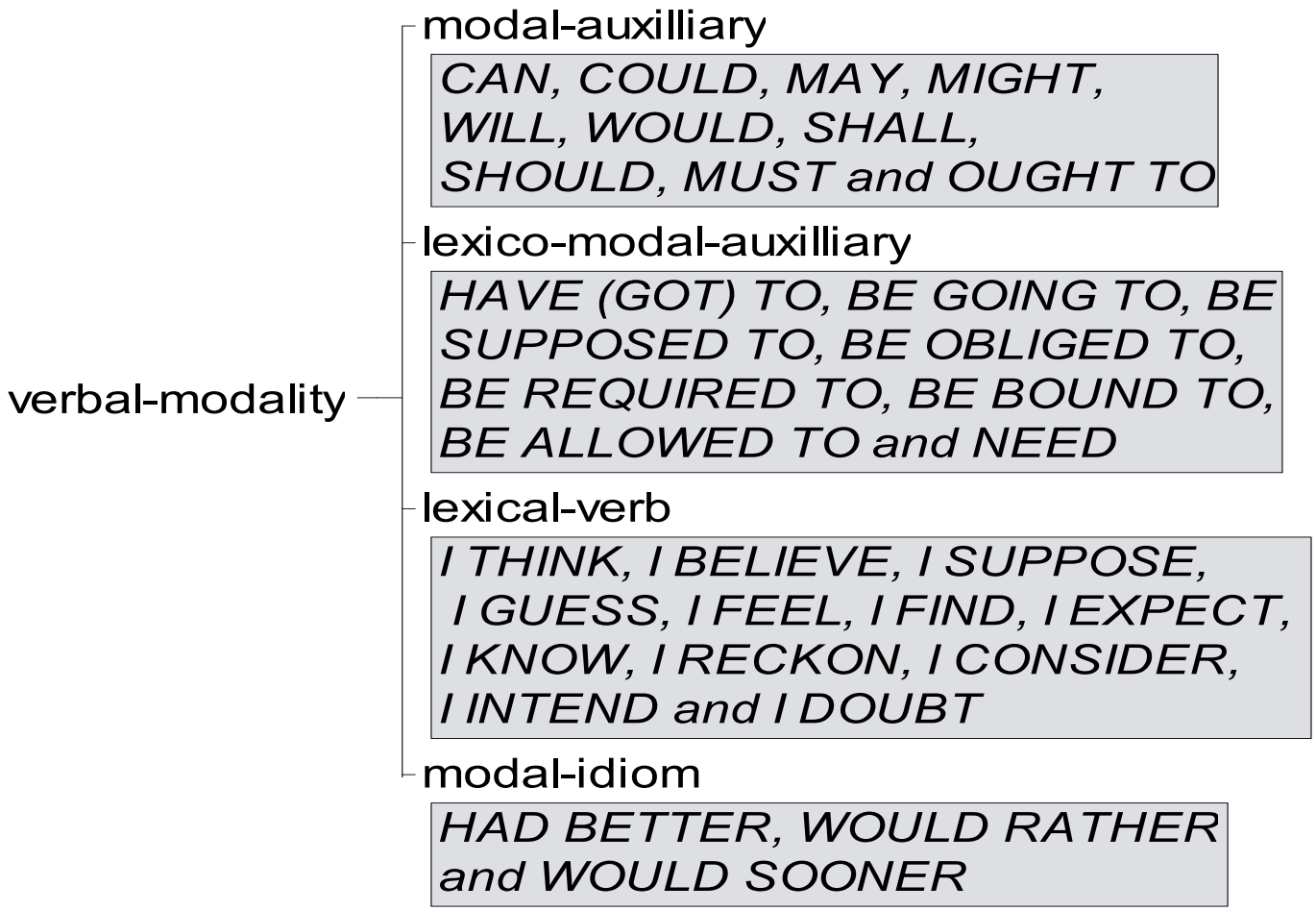

CorpusTool 2.8.3
Project Misc Help
Project Search AutoCode Statistics Explore Options Help
Rule: select Modality:intention if Modality:lexical-verb containing 'think'
Edit Delete Add
Help Show
Autocode Rule Editor
What do you want to do?
Search for segments and assign those segments a particular feature
Create new segments in one layer based on corresponding segments in another layer
Create new segments in one layer based on string patterns in the text.
Find all: lexical-verb + containing anywhere think
Code them as: intention
Save Cancel
Hotmail - rebejeje@...
aspecto - Diccionari...
Microsoft PowerPoi...
RebecaModalityv7-...
CorpusTool 2.8.3
ES
9:23


CorpusTool 2.8.3
Project Misc Help
Project Search AutoCode Statistics Explore Options Help
Enter Search Query Below:
token + containing anywhere think
Show
Show as: table Save Help
P Only Partially Coded * Segment with Comment
1199 matches
P C Pretext Contained PostText
think
Grammar analysis for: UAM/C8-3.txt
To understand more or less what children think , we have to
Adjunct Sep Subject Pred
To Pred Adjunct Object Thing To
Unkn Unkn Unkn Relpron Subject Pred
Unkn Thing
Fashion designers and people in charge of these kinds of announcemer
<< < > >> Ignore Delete Other Action... Save Close Help
Assigned
word
lexical-verb
mental-verb
grammatical-unit
token
verb
present-verb
Gloss
Comment:
Hotmail - r...
aspecto - Di...
Microsoft P...
RebecaMod...
CorpusTool...
Grammar a...
Document1...
ES
9:27

Taxonomy of modal markers used:



Verbal modal elements:



(from Greenbaum et al., 1985)

Non-verbal modal elements:

non-verbal-modality

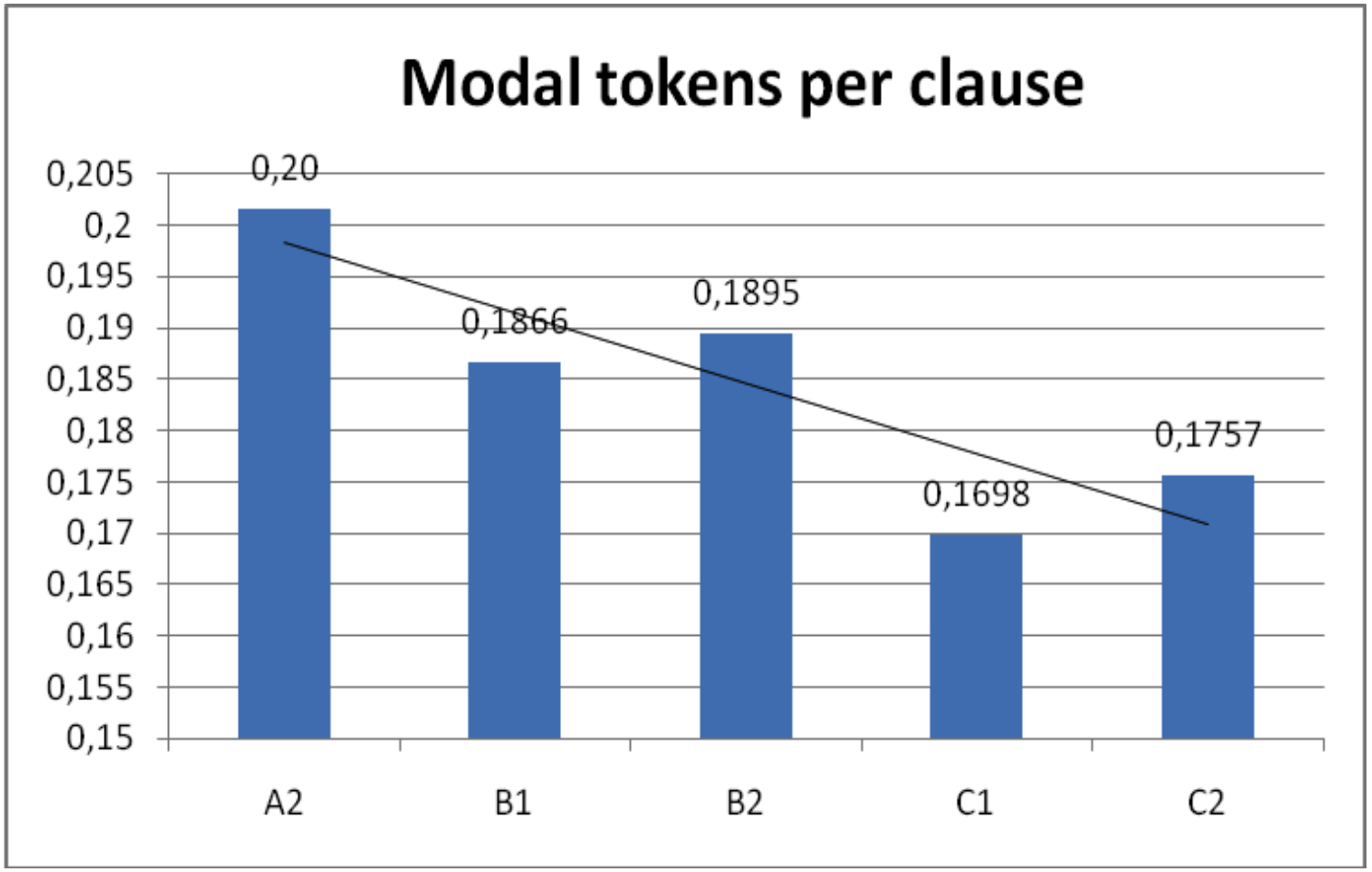
adjectival-phrase

ABLE, UNABLE, TRUE, FALSE, NECESSARY, UNNECESSARY, POSSIBLE, IMPOSSIBLE, PROBABLE, IMPROBABLE, CLEAR, LIKELY, UNLIKELY, SURE, UNSURE, CERTAIN, UNCERTAIN, PERMISSIBLE, WILLING, UNWILLING, DETERMINED, IMPERATIVE, BOUND

adverbials

MAYBE, PERHAPS, POSSIBLY, DEFINITELY, CERTAINLY, SUPPOSEDLY, SURELY, UNDOUBTEDLY, DOUBTLESS, DOUBTLESSLY, LIKELY, PRESUMABLY, ARGUABLY, CONCEIVABLY, INDEED, FOR SURE, FOR CERTAIN, OF COURSE, WITHOUT DOUBT, PROBABLY and NECESSARILY.

(from Greenbaum et al., 1985)

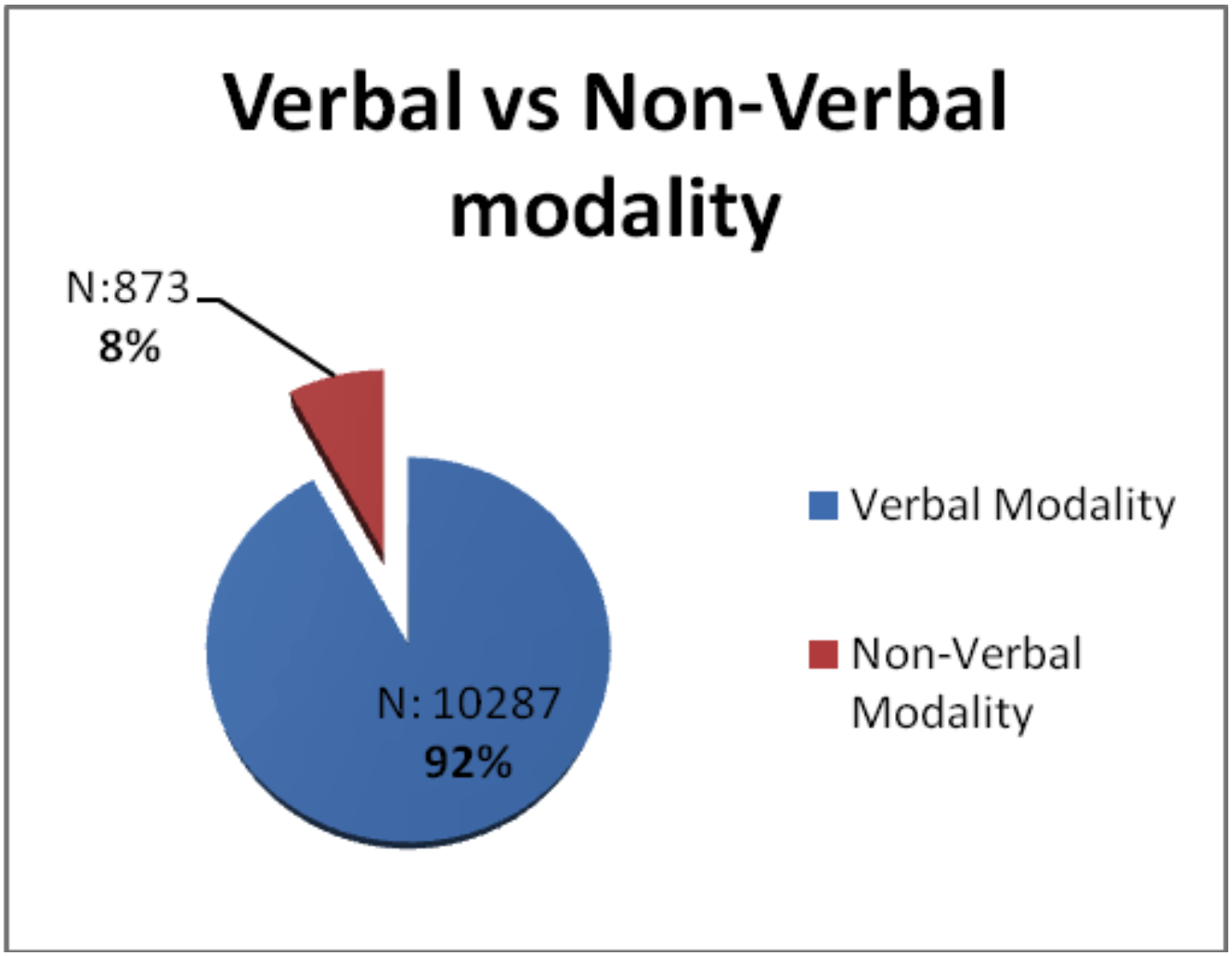

Modal tokens per clause
0,205
0,2
0,195
0,19
0,185
0,18
0,175
0,17
0,165
0,16
0,155
0,15
0,20
0,1866
0,1895
0,1698
0,1757
A2
B1
B2
C1
C2

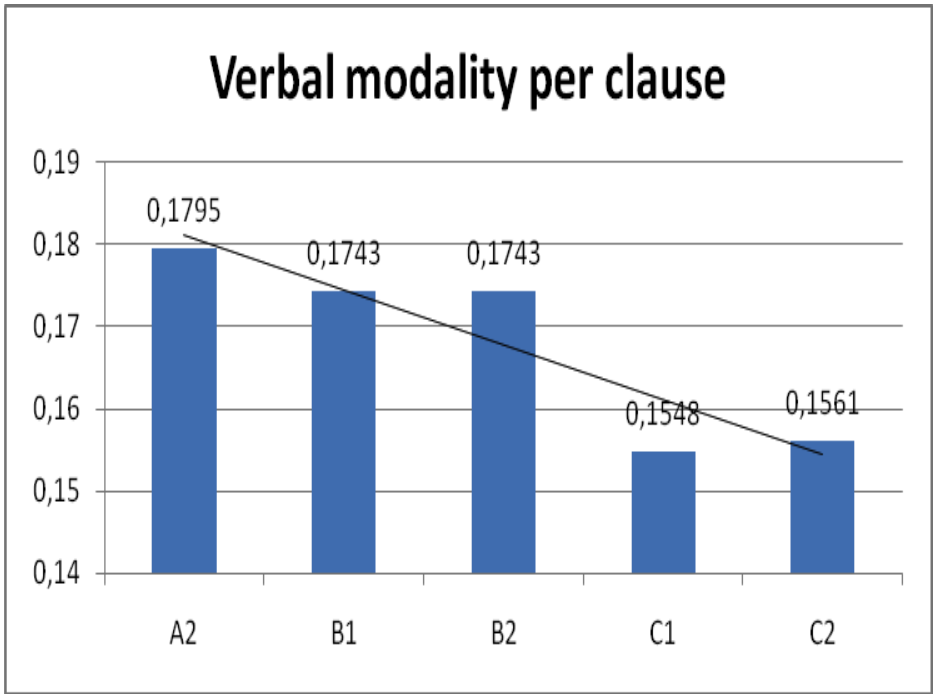
Hypothesis 1

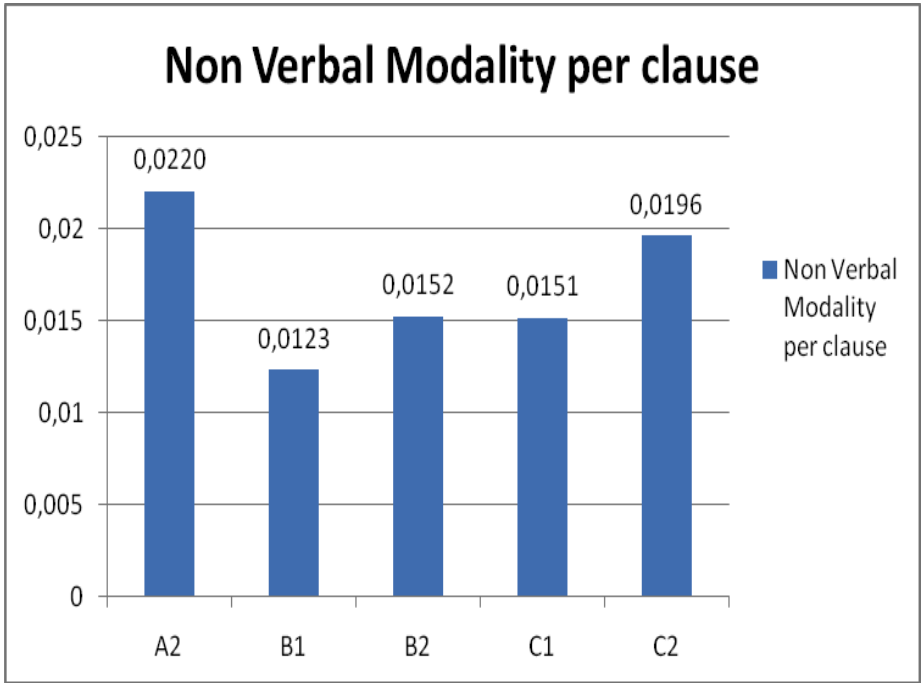
- As students' level of proficiency rises, they will use a higher quantity of modal markers – PROVEN WRONG

Hypothesis 3:

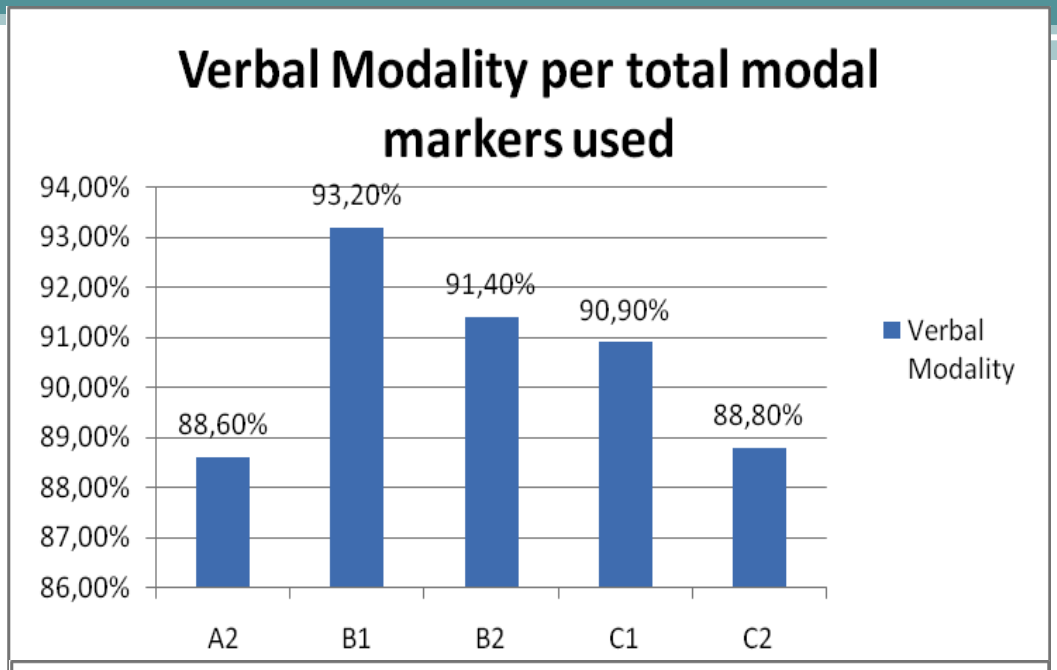
- As students' level of proficiency rises, I expect a movement away from *verbal* to *non-verbal* modal means

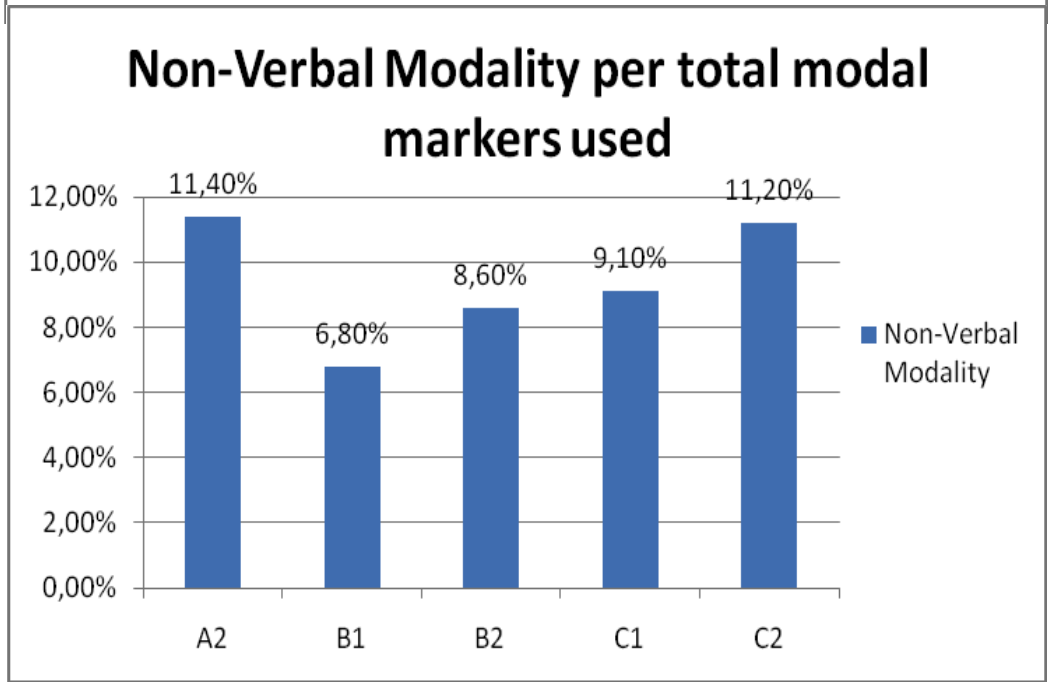

Verbal vs Non-Verbal modality
N:873
8%
N: 10287
92%
Verbal Modality
Non-Verbal Modality





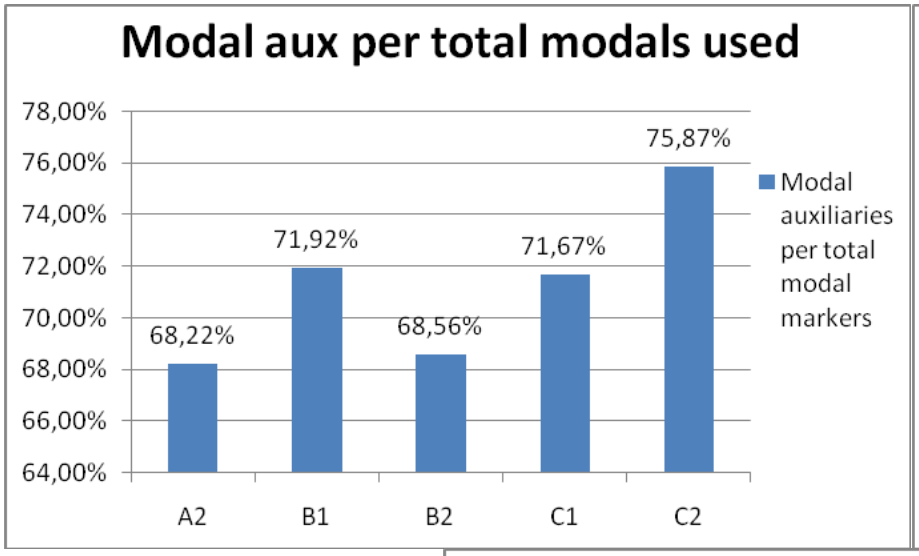
These results show that hypothesis 3 is RIGHT

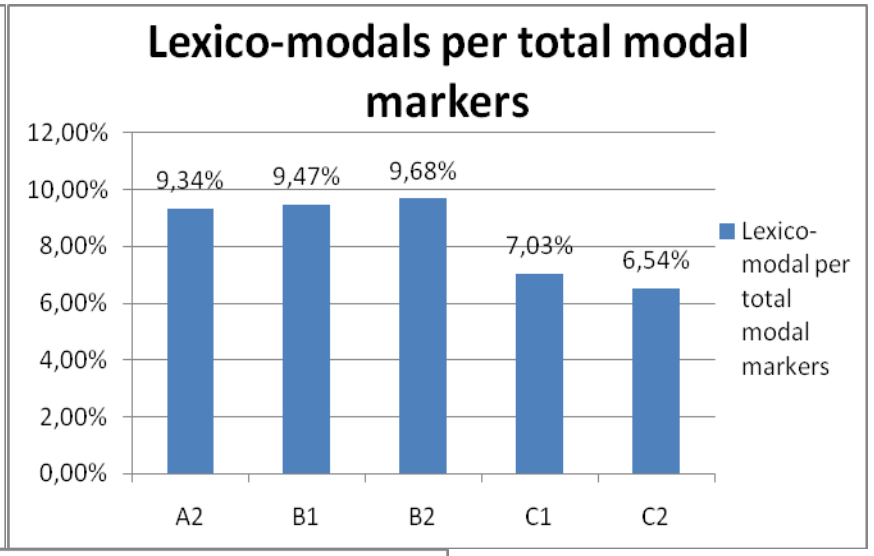

Verbal Modality per total modal markers used
94,00%
93,00%
92,00%
91,00%
90,00%
89,00%
88,00%
87,00%
86,00%
88,60%
93,20%
91,40%
90,90%
88,80%
A2
B1
B2
C1
C2
Verbal Modality

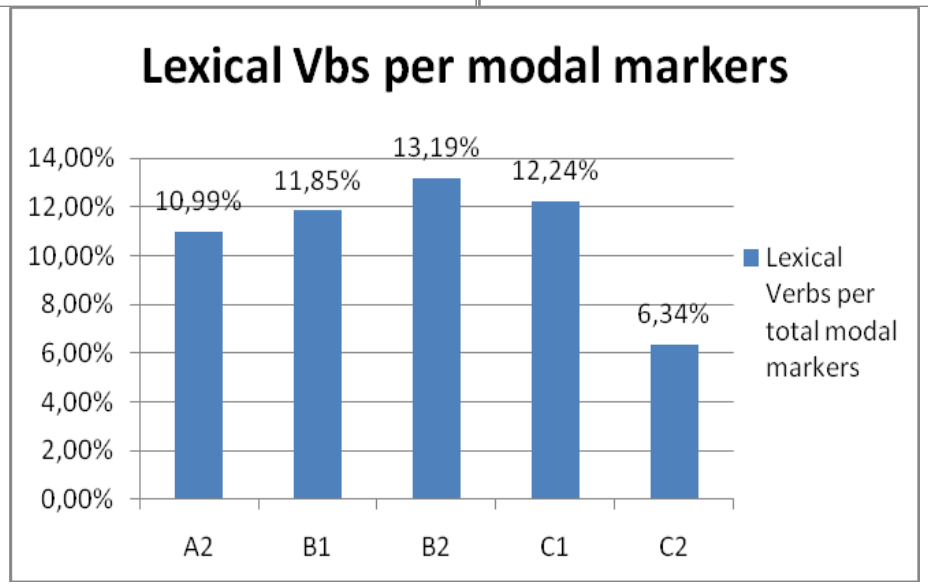

Non-Verbal Modality per total modal markers used
12,00%
10,00%
8,00%
6,00%
4,00%
2,00%
0,00%
11,40%
6,80%
8,60%
9,10%
11,20%
A2
B1
B2
C1
C2
Non-Verbal Modality

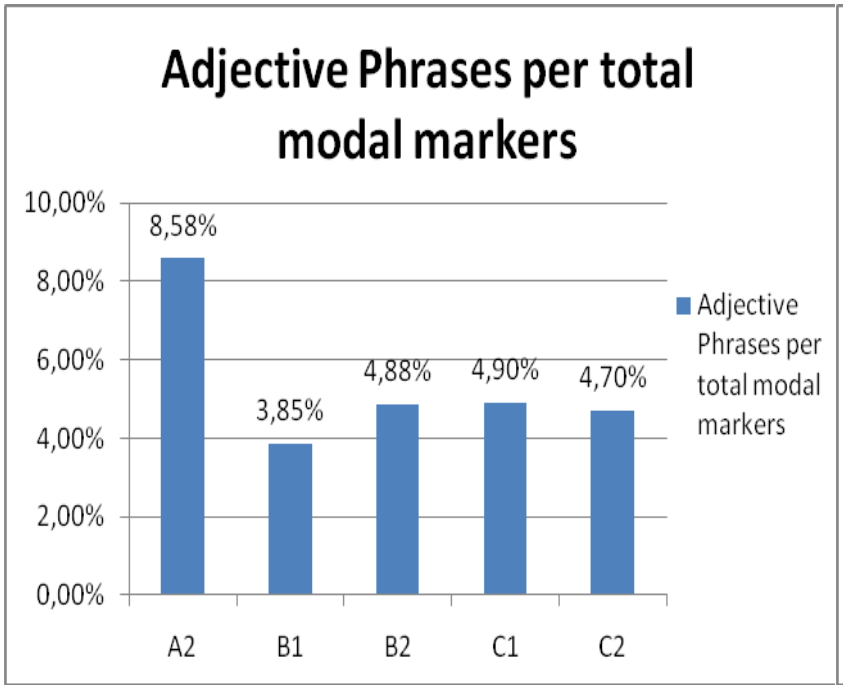
Hypothesis 2:

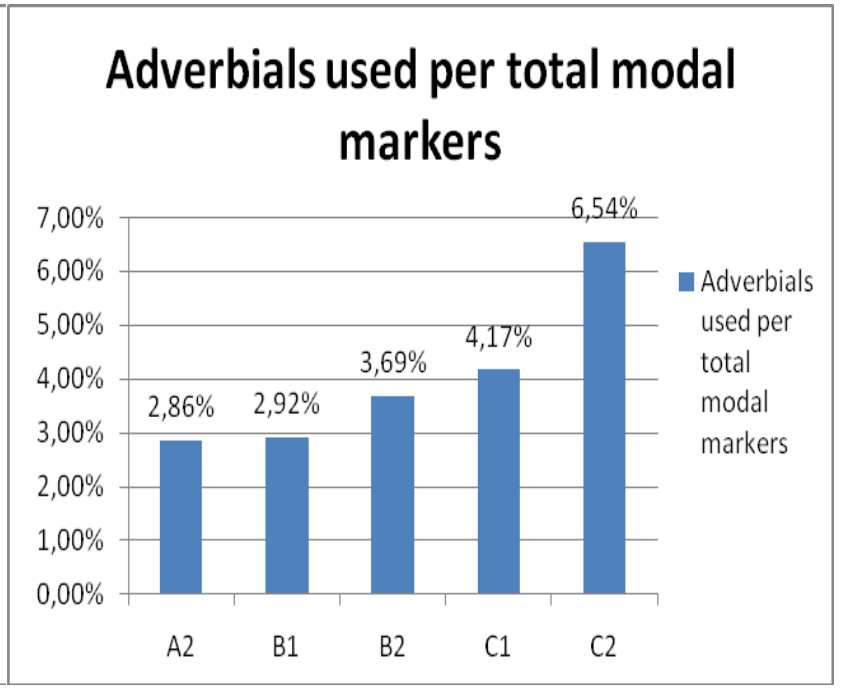
- As students' level of proficiency rises, I expect them to use a wider variety of modal markers


Modal aux per total modals used
78,00%
76,00%
74,00%
72,00%
70,00%
68,00%
66,00%
64,00%
68,22%
71,92%
68,56%
71,67%
75,87%
A2
B1
B2
C1
C2
Modal auxiliaries per total modal markers


Lexico-modals per total modal markers
12,00%
10,00%
8,00%
6,00%
4,00%
2,00%
0,00%
9,34%
9,47%
9,68%
7,03%
6,54%
A2
B1
B2
C1
C2
Lexico-modal per total modal markers

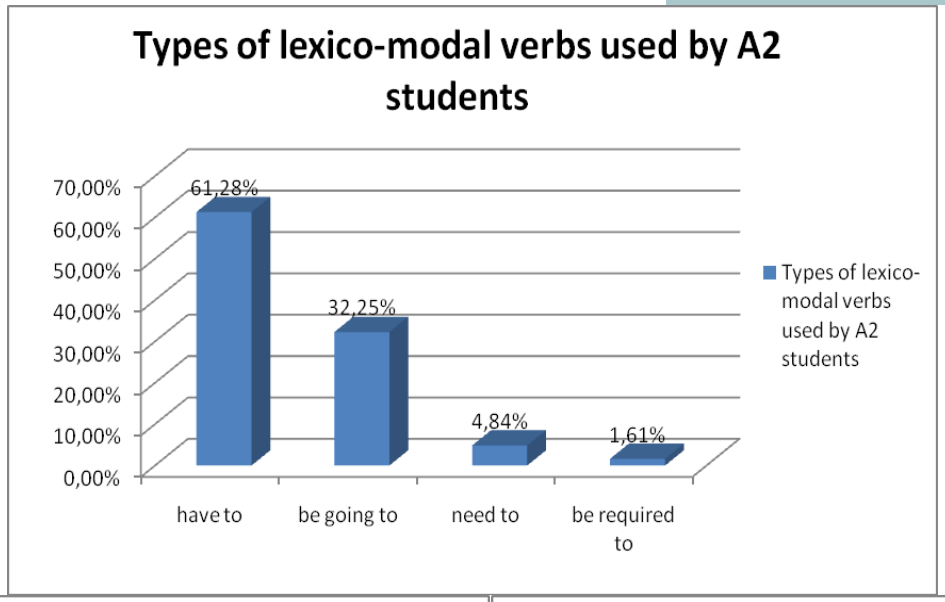

Lexical Vbs per modal markers
14,00%
12,00%
10,00%
8,00%
6,00%
4,00%
2,00%
0,00%
10,99%
11,85%
13,19%
12,24%
6,34%
A2
B1
B2
C1
C2
Lexical Verbs per total modal markers

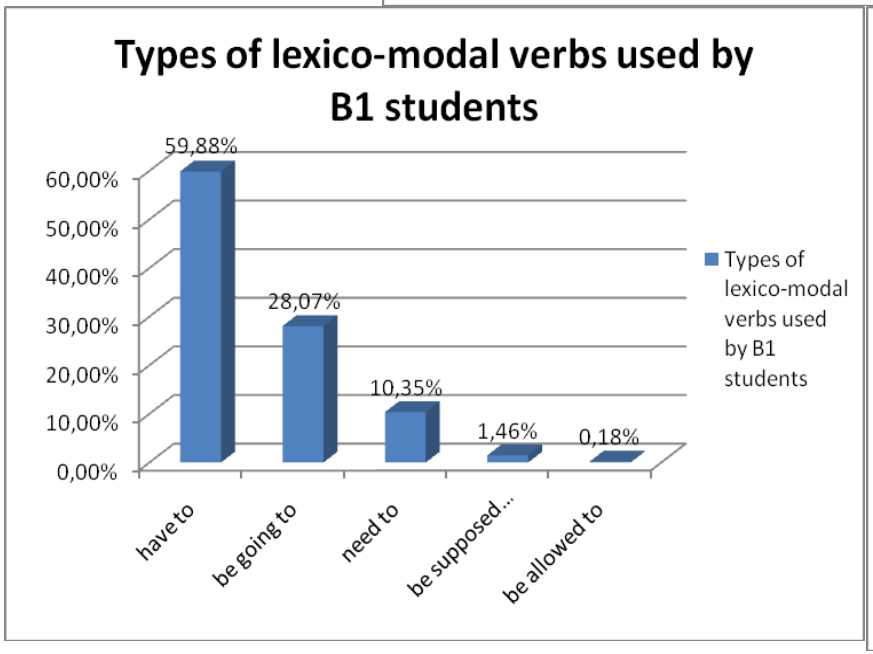

Adjective Phrases per total modal markers
10,00%
8,00%
6,00%
4,00%
2,00%
0,00%
8,58%
3,85%
4,88%
4,90%
4,70%
A2
B1
B2
C1
C2
Adjective Phrases per total modal markers

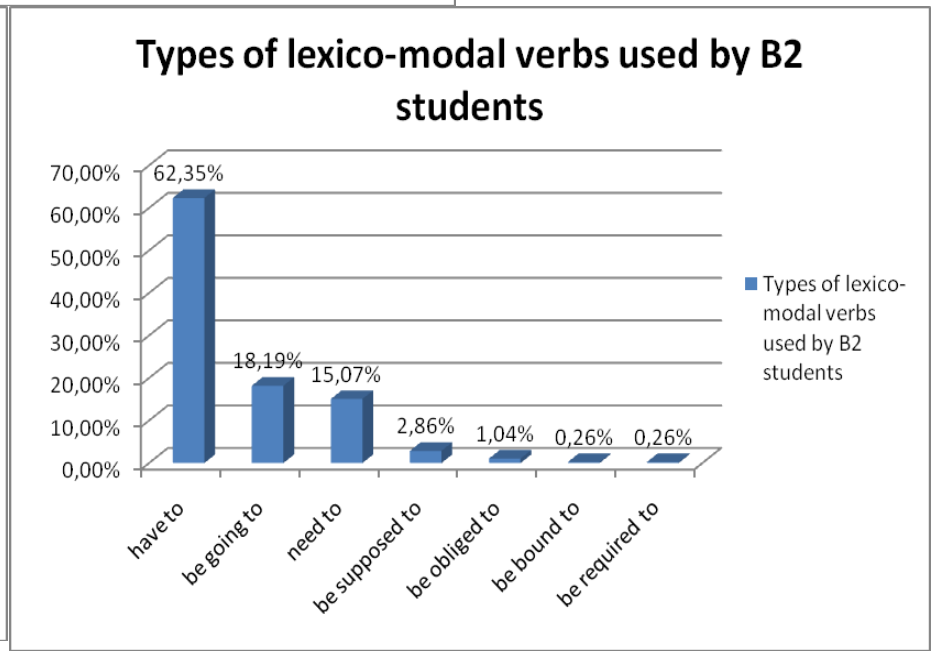

Adverbials used per total modal markers
7,00%
6,00%
5,00%
4,00%
3,00%
2,00%
1,00%
0,00%
2,86%
2,92%
3,69%
4,17%
6,54%
A2
B1
B2
C1
C2
Adverbials used per total modal markers

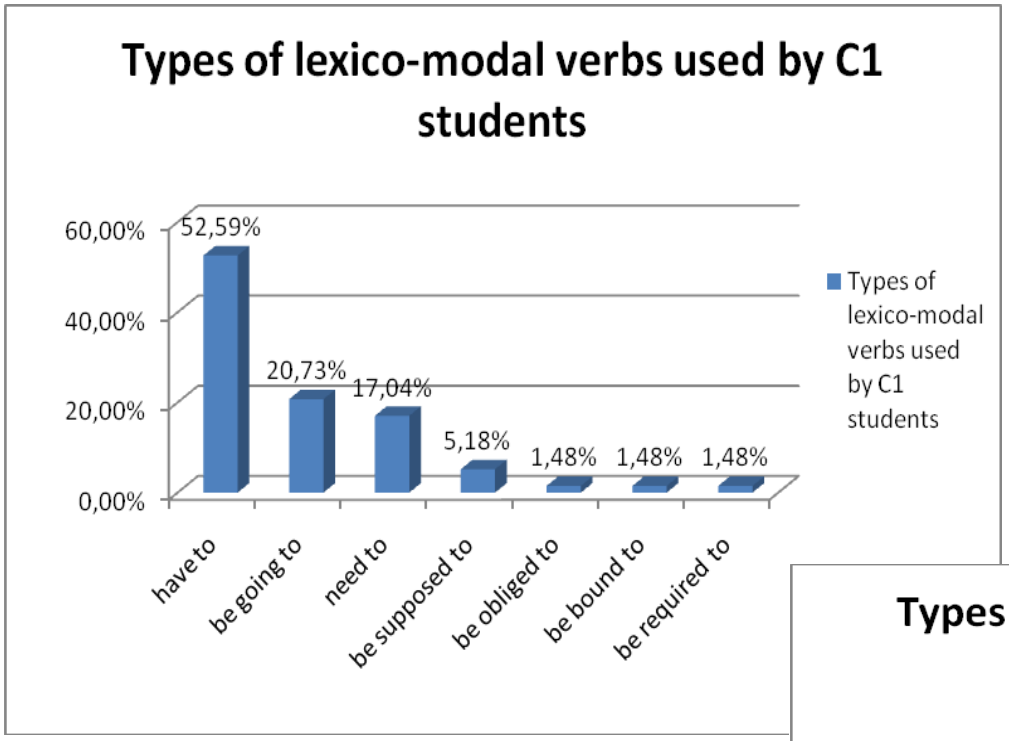
- There are no clear patterns, because all the grammatical categories established as modal are used to some degree at all levels.

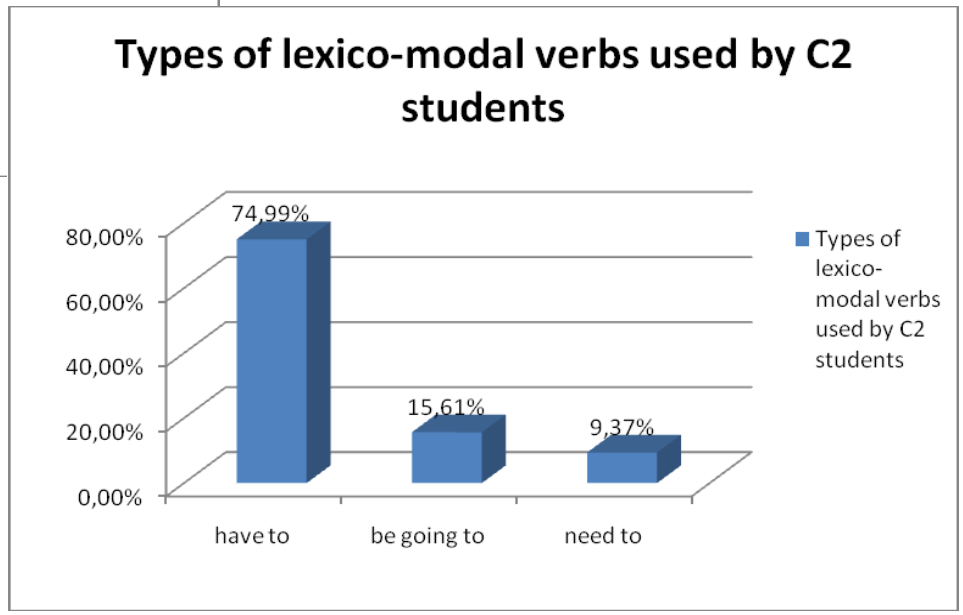
- What does change is that the more advanced levels, especially B2 and C1, use a wider range of tokens within each of these categories

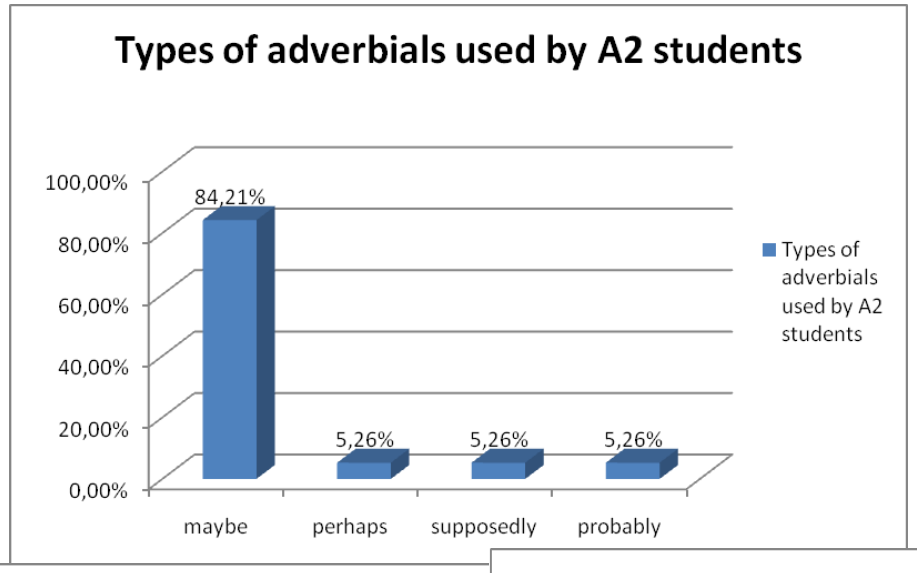

Types of lexico-modal verbs used by A2 students
70,00%
60,00%
50,00%
40,00%
30,00%
20,00%
10,00%
0,00%
61,28%
32,25%
4,84%
1,61%
have to
be going to
need to
be required to
Types of lexico-modal verbs used by A2 students

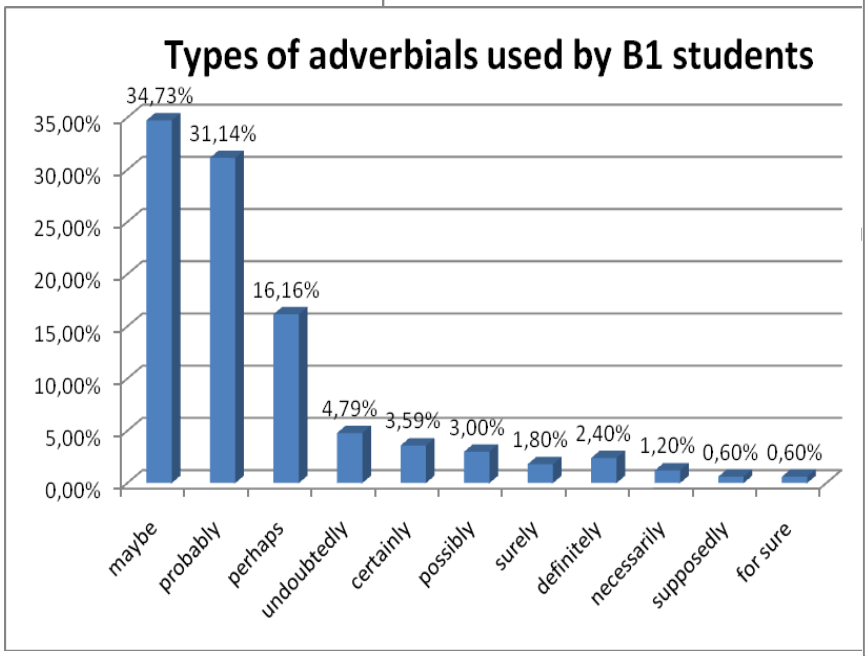

Types of lexico-modal verbs used by B1 students
60,00%
50,00%
40,00%
30,00%
20,00%
10,00%
0,00%
59,88%
28,07%
10,35%
1,46%
0,18%
have to
be going to
need to
be supposed...
be allowed to
Types of lexico-modal verbs used by B1 students

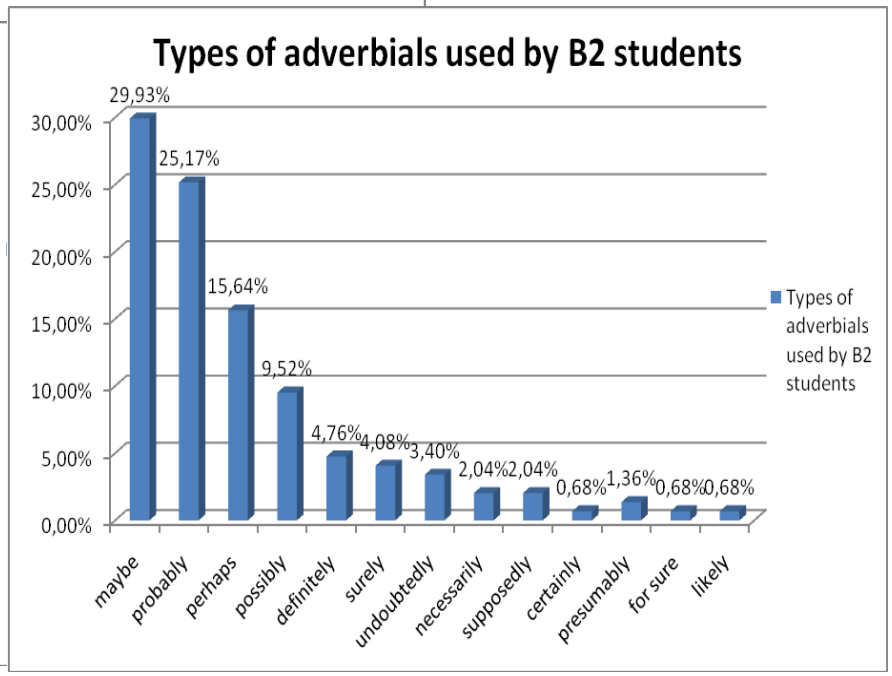

Types of lexico-modal verbs used by B2 students
70,00%
60,00%
50,00%
40,00%
30,00%
20,00%
10,00%
0,00%
62,35%
18,19%
15,07%
2,86%
1,04%
0,26%
0,26%
have to
be going to
need to
be supposed to
be obliged to
be bound to
be required to
Types of lexico-modal verbs used by B2 students

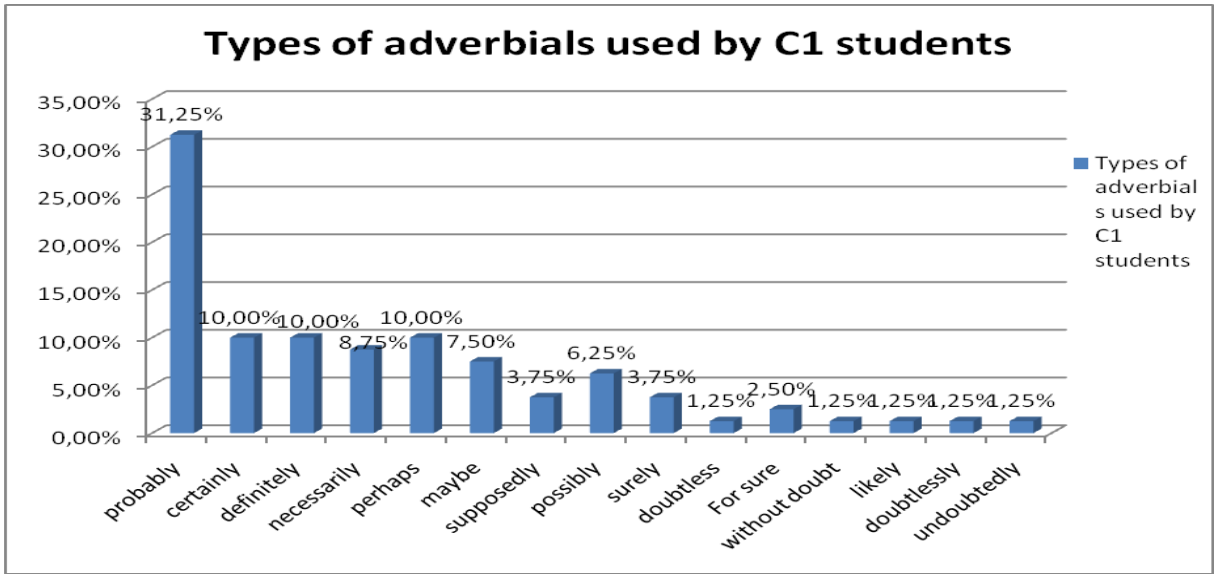

Types of lexico-modal verbs used by C1 students
60,00%
40,00%
20,00%
0,00%
52,59%
20,73%
17,04%
5,18%
1,48%
1,48%
1,48%
have to
be going to
need to
be supposed to
be obliged to
be bound to
be required to
Types of lexico-modal verbs used by C1 students

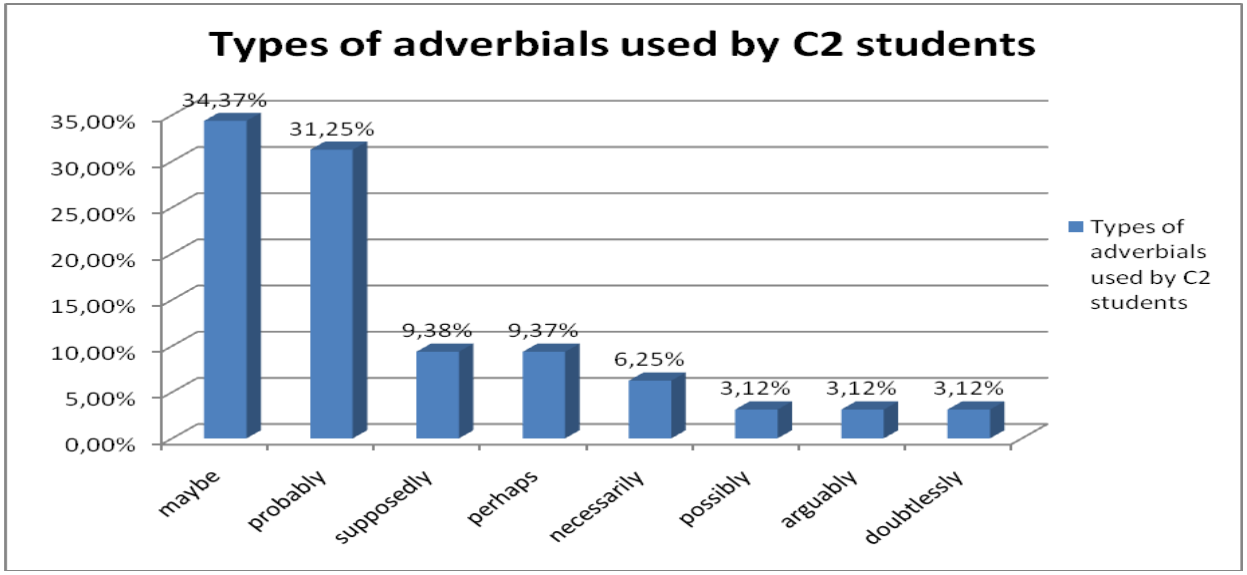

Types of lexico-modal verbs used by C2 students
80,00%
60,00%
40,00%
20,00%
0,00%
74,99%
15,61%
9,37%
have to
be going to
need to
Types of lexico-modal verbs used by C2 students


Types of adverbials used by A2 students
100,00%
80,00%
60,00%
40,00%
20,00%
0,00%
84,21%
5,26%
5,26%
5,26%
maybe
perhaps
supposedly
probably
Types of adverbials used by A2 students


Types of adverbials used by B1 students
35,00%
30,00%
25,00%
20,00%
15,00%
10,00%
5,00%
0,00%
34,73%
31,14%
16,16%
4,79%
3,59%
3,00%
1,80%
2,40%
1,20%
0,60%
0,60%
maybe
probably
perhaps
undoubtedly
certainly
possibly
surely
definitely
necessarily
supposedly
for sure


Types of adverbials used by B2 students
30,00%
25,00%
20,00%
15,00%
10,00%
5,00%
0,00%
29,93%
25,17%
15,64%
9,52%
4,76%
4,08%
3,40%
2,04%
2,04%
0,68%
1,36%
0,68%
0,68%
maybe
probably
perhaps
possibly
definitely
surely
undoubtedly
necessarily
supposedly
certainly
presumably
for sure
likely
Types of adverbials used by B2 students


Types of adverbials used by C1 students
35,00%
30,00%
25,00%
20,00%
15,00%
10,00%
5,00%
0,00%
31,25%
10,00%
10,00%
8,75%
10,00%
7,50%
3,75%
6,25%
3,75%
1,25%
2,50%
1,25%
1,25%
1,25%
1,25%
probably
certainly
definitely
necessarily
perhaps
maybe
supposedly
possibly
surely
doubtless
For sure
without doubt
likely
doubtlessly
undoubtedly
Types of adverbials used by C1 students


Types of adverbials used by C2 students
35,00%
30,00%
25,00%
20,00%
15,00%
10,00%
5,00%
0,00%
34,37%
31,25%
9,38%
9,37%
6,25%
3,12%
3,12%
3,12%
maybe
probably
supposedly
perhaps
necessarily
possibly
arguably
doubtlessly
Types of adverbials used by C2 students

Hypothesis 2:

- As students' level of proficiency rises, I expect them to use a wider variety of modal markers:

Increasing use of non-verbal modality and more variety of tokens within each category: it proves RIGHT.

Conclusions

As students's proficiency level rises,

- **A fall in the use of total modal markers:**

  - **Verbal modal** elements: **decreasing** presence

    ***BUT their number is the largest across proficiency levels

  - **Non-verbal** modal elements: **increasing** presence

- Proportion of modal markers per clause may be diluted in the advanced levels (they write more)
- A movement towards categories not included in my taxonomy? (i.e. evidentiality, frequency, reporting verbs, mental verbs different from the ones included)
- More quantity does not mean more quality, e.g. decreasing use of *be going to*

- The tokens used in each grammatical class show more grammatical and pragmatic accuracy in the advanced levels,

  e.g. low-proficient students overuse *can* and *will* at the expense of *would* and *should (modal aux.);* decreasing use of *maybe vs.* increasing use of *probably (adverbials)*

  → notions of ***register*** and ***genre***
- A wider range of tokens within each grammatical class as level of proficiency rises (except for C2 students)

Thank you!

Rebeca García González
rebeca.garciagonzalez@estudiante.uam.es