Google Translate and DeepL: Breaking Taboos in Translator Training

Dr Ruth Burbat (University of Granada)
Dr María-José Varela Salinas (University of Malaga)

Google Translate and DeepL:
Breaking Taboos in Translator Training

1. Motivation

2. Theoretical Framework of Machine Translation (MT)

3. Goals and Methodology

4. Discussion of Results

5. Conclusions and Future Prospects

Motivation

- To use or not to use MT in translator training?

- Role of machine translation (MT): What has changed?

- Research focus:

  o Potential and limits of MT and students` handling

  o Didactic use

Theoretical Framework of MT

RBMT: Rule Based Machine Translation

(PB)SMT: (Phrase Based) Statistical Machine Translation

EBMT: Example Based Machine Translation

NMT: Neural Machine Translation

Goals and Methodology

Potential and limits of MT and students` handling for post-editing:

- Identify errors

- Classify error tendencies

Didactic use:

- Knowledge about grammatical features to be emphasized

- Improve metacognitive competence

Goals and Methodology

Post-editing:
Defined by ISO 18587:2017 as to "edit and correct machine translation output".

Here:
"**Light post-editing**": to correct only as far as "to obtain a merely comprehensible text without any attempt to produce a product comparable to a product obtained by human translation".

Goals and Methodology

Error categories:

- **Pragmatics**

- **Grammar**

- **Lexicon/ terminology**

- **Equivalence**

- **Spelling**

- **Style**

- **Culture**

Goals and Methodology

Observational study:

- with 32 – 35 students from the 4th year of Translation Studies (German as a second language after approximately 225 hours)

- MT translation of five texts (126 - 313 words) from the financial section of a non-specialized publication from Spanish into German + light post-editing

- MT tool: *Google Translate* and *DeepL*

Goals and Methodology

- 45 – 90 minutes for MT + light post-editing

- access to online and paperback resources

- use of own translations, previously corrected in class

Goals and Methodology

Google T.	Error Category (Number and Percentage)								
	DESEMPLEADOS 52+ (Google Translate)								
	Prag	Gr	Equ	Spell	Lex/ Term	Sty	Cult	Error rate	Total
1	7	17	8	1	19	3	1		
	12,5 %	30,36 %	14,29 %	1,79 %	33,93 %	5,36 %	1,79 %		

DeepL	Error Category (Number and Percentage)								
	DESEMPLEADOS + 52 AÑOS (DeepL)								
	Prag	Gr	Equ	Spell	Lex/ Term	Sty	Cult	Error rate	Total
	3	8	4		7	1			
	13,04 %	34,78 %	17,39 %		30,43 %	4,35 %			

Discussion of Results

- Google Translate performance is below that of DeepL

Discussion of Results

'Desempleados 52+': Error Categories
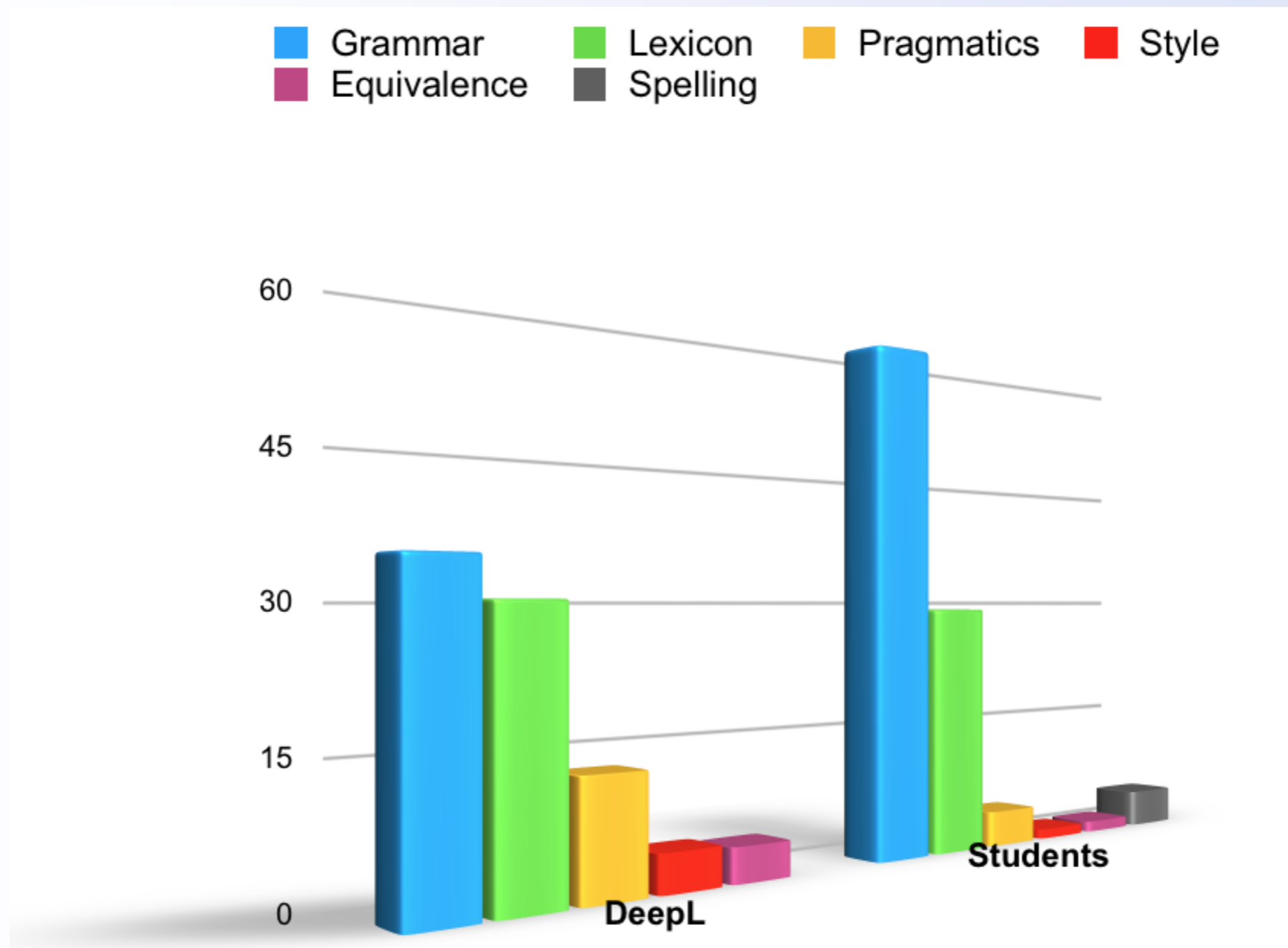
Discussion of Results

DeepL	Error Category (Number and Percentage) DESEMPLEADOS + 52 AÑOS (DeepL)								
	Prag	Gr	Equ	Spell	Lex/ Term	Sty	Cult	Error rate	Total
	3	8	4		7	1		⬛	⬛
	13,04 %	34,78 %	17,39 %		30,43 %	4,35 %			

Number of Students	Error Category (Number and Percentage) DESEMPLEADOS + 52 AÑOS (Estudiantes)									Media of Errors
	Prag	Gr	Equ	Spell	Lex/ Term	Sty	Cult	Error rate	Total	
34	33	475	10	35	232	9		⬛	794	⬛
	4,16 %	59,8 %	1,26 %	4,4 %	29,21 %	1,13 %				

Discussion of Results

'Desempleados 52
Students' error dis
within the group
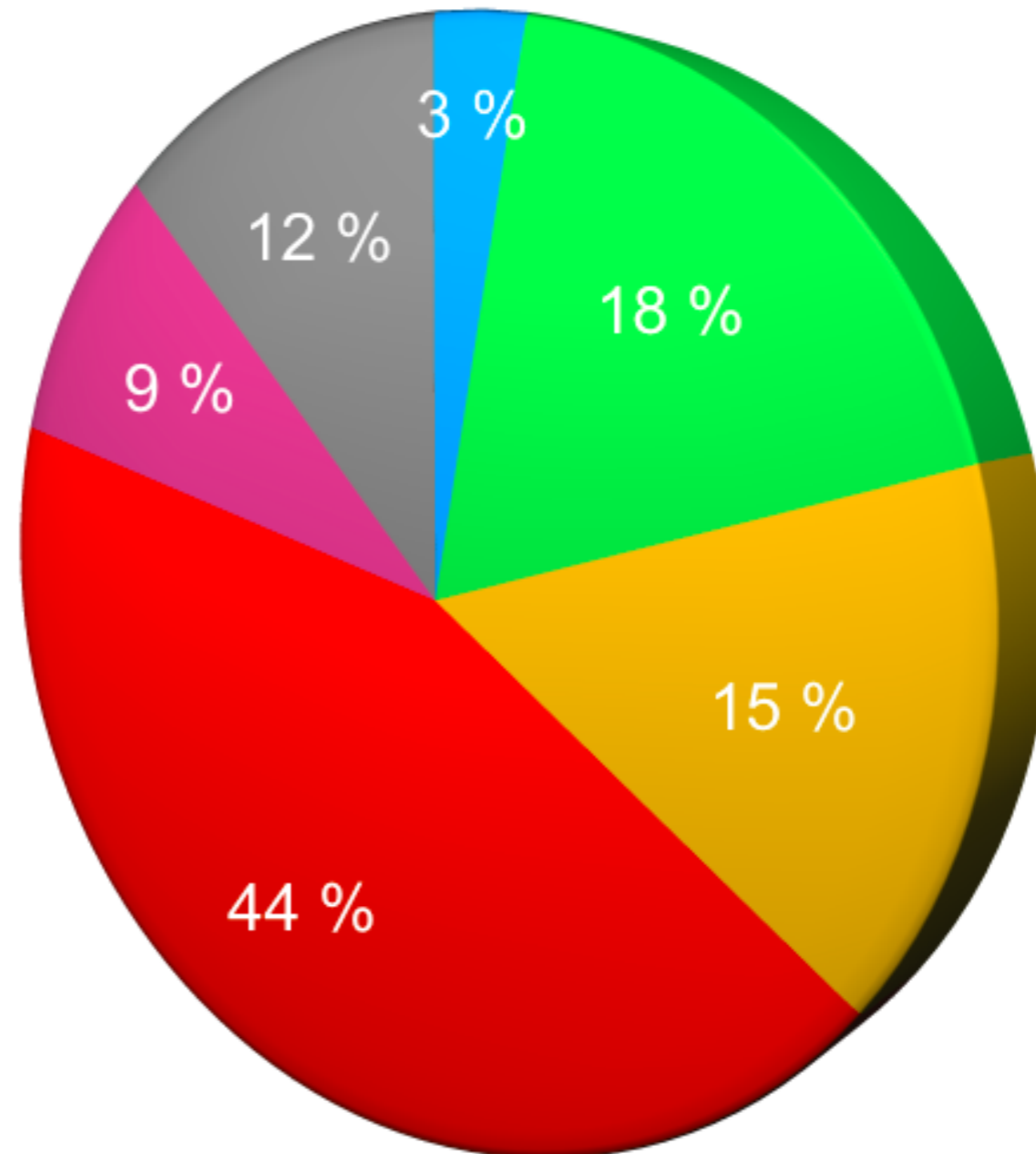
Discussion of Results

Distribution of grammatical errors in terms of error frequency:

DeepL

- Prepositions 25 %
- Gender 13 %
- (Un)determ. art. 13 %
- Syntax 13 %
- Inflection 13 %
- Time Use 25 %

Legend:
- Syntax
- Conjugation
- Passive Voice
- Subject Omission
- Inflection
- Gender
- (Un)determ. articles
- Prepositions
- Pronouns
- Time Use
- Punctuation
- Conjunctions
- Subjunctive

Students

- 0,4 %
- 1 %
- 1 %
- 1 %
- 2 %
- 2 %
- 3 %
- 7 %
- 10 %
- 10 %
- Syntax 26 %
- Inflection 22 %
- Prepositions 15 %
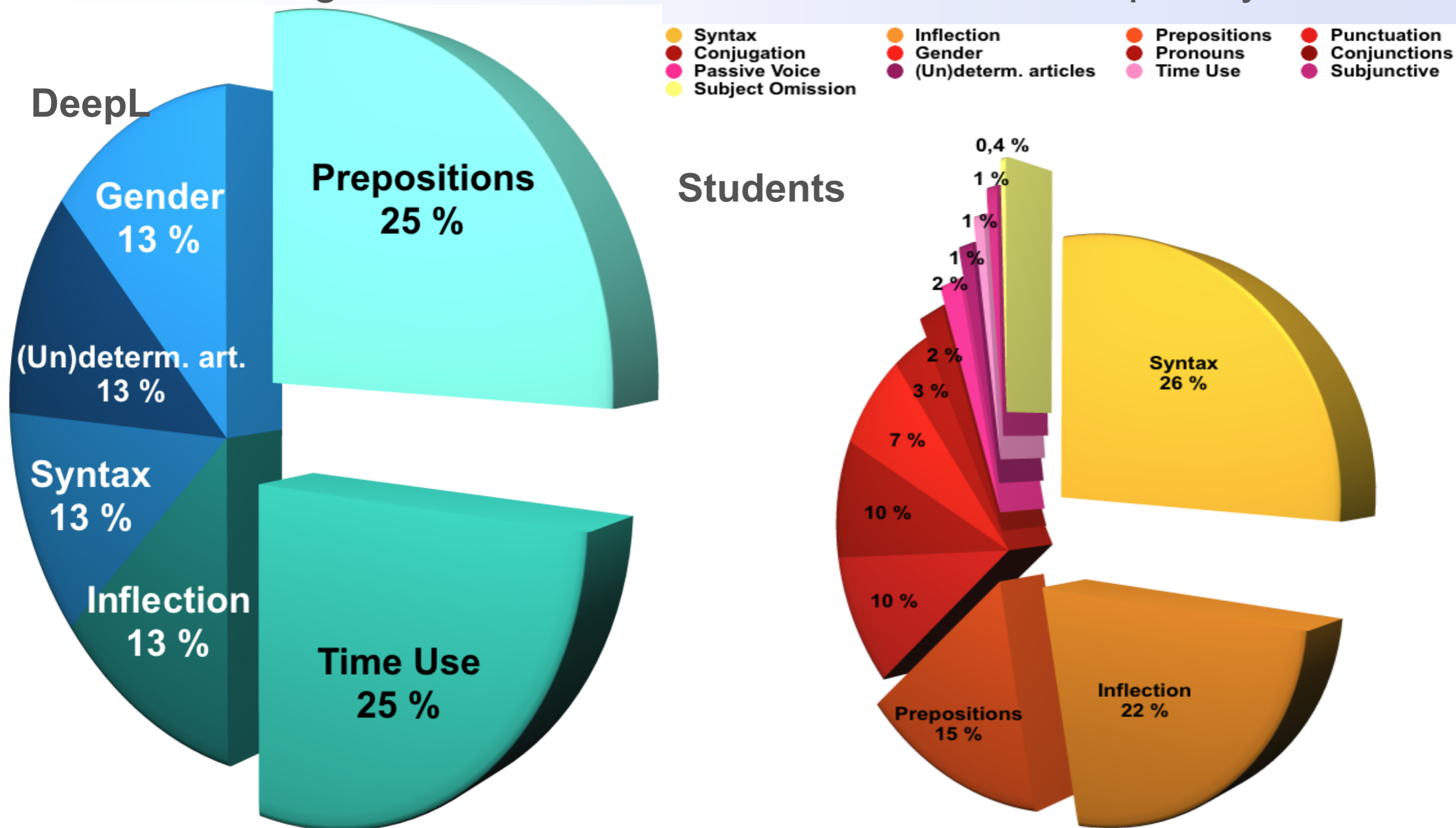
Discussion of Results

Distribution of grammatical errors in terms of error category:
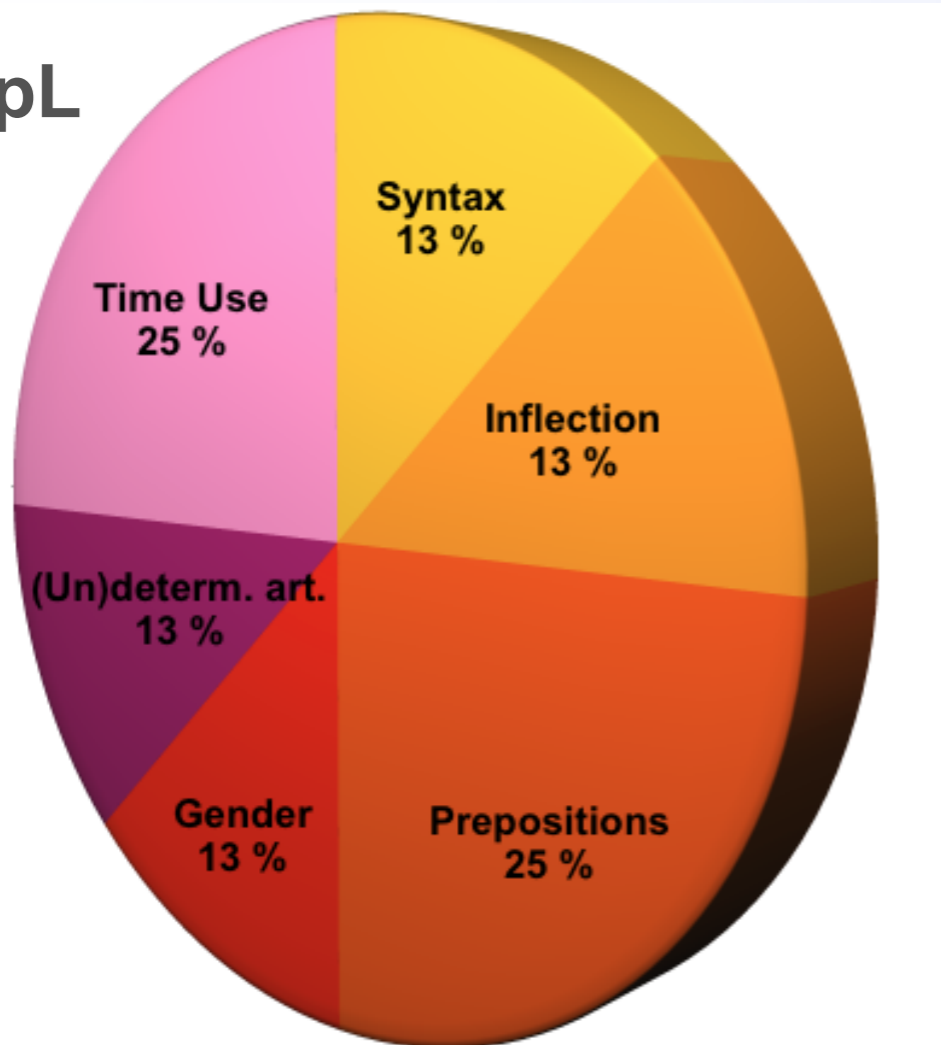
Discussion of Results
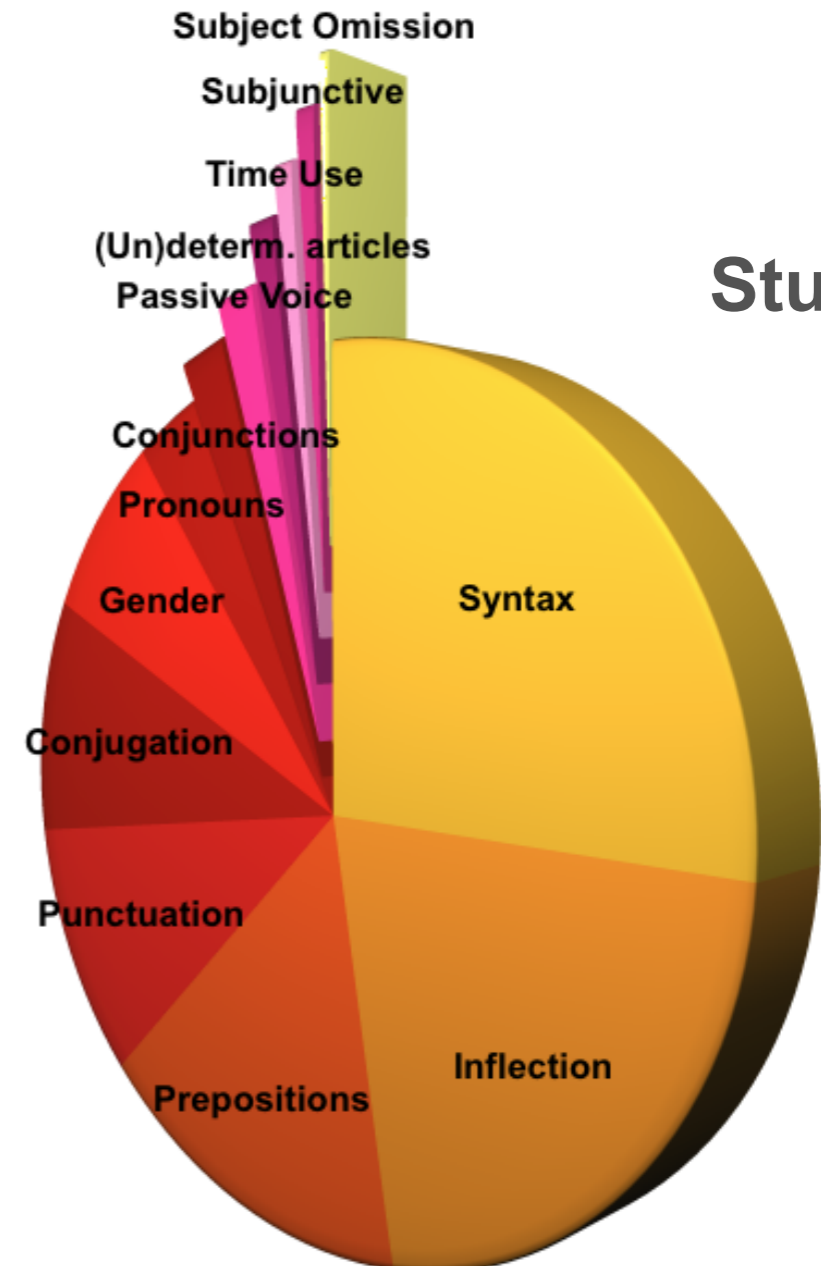
'Inflación subyacente': Error Categories
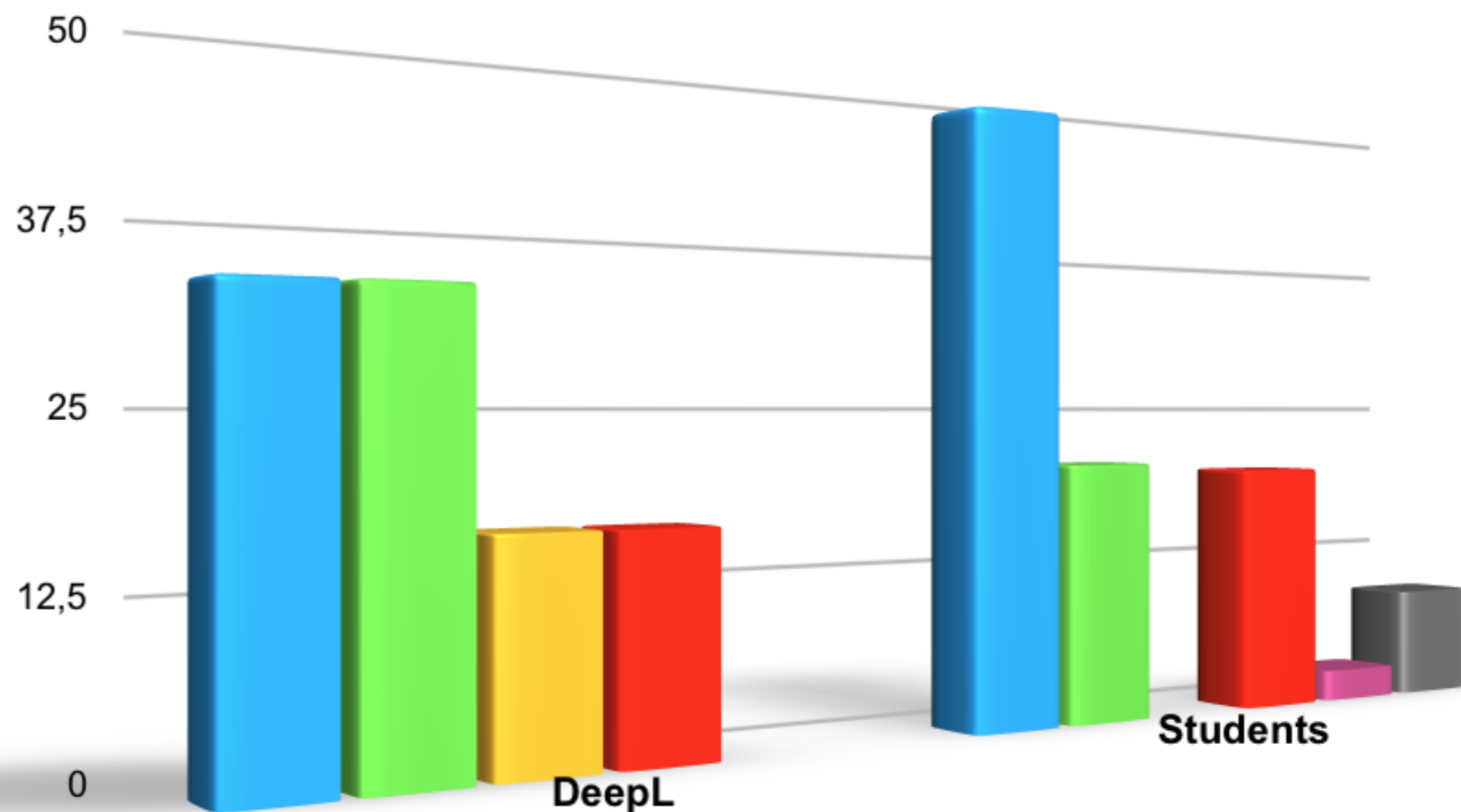
Discussion of Results

INFLACIÓN SUBYACENTE (DEEPL)

	Prag	Gr	Equ	Spell	Lex/ Term	Sty	Cult	Error rate	Total
1	2	4			4	2		🟧	🟧
	16,7 %	33,4 %			33,4 %	16,7 %			

DeepL.	Error Category (Number and Percentage) INFLACIÓN SUBYACENTE (STUDENTS)									Media of Errors
	Prag	Gr	Equ	Spell	Lex/ Term	Sty	Cult	Error rate	Total	
31	0	136	7	25	58	56	0	🟧	282	🟧
		48,2 %	2,5 %	8,9 %	20,6 %	19,9 %				

Discussion of Results

Distribution of grammatical errors in terms of error frequency:

DeepL

Students

● Syntax	● Inflection	● Prepositions	● Conjunctions
● Conjugation	● Pronouns	● Punctuation	● (Un)determ. articles
● Gender	● Subject Omission	● Passive Voice	● Time Use
● Subjunctive			
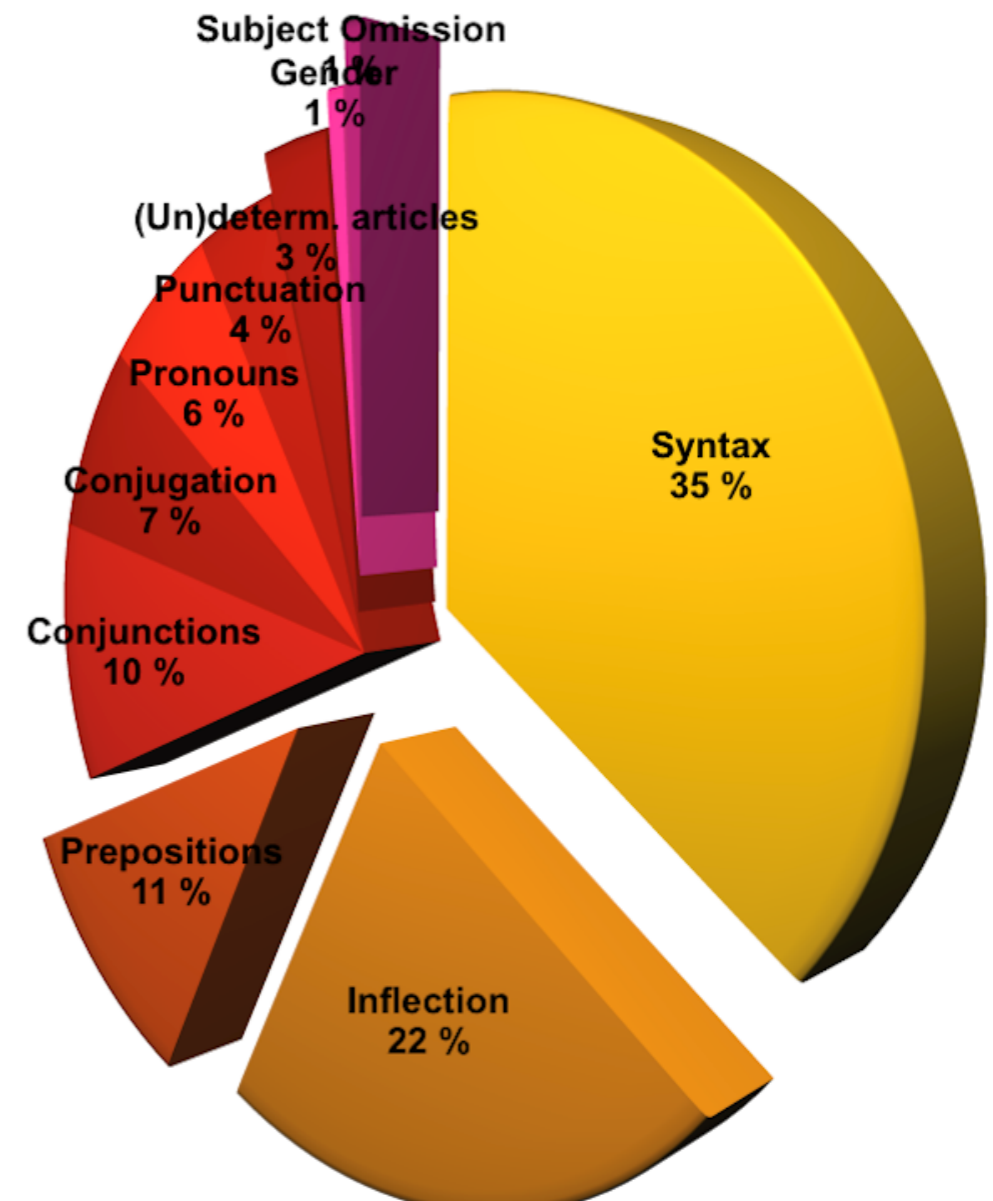
Discussion of Results

Distribution of grammatical errors in terms of error category:
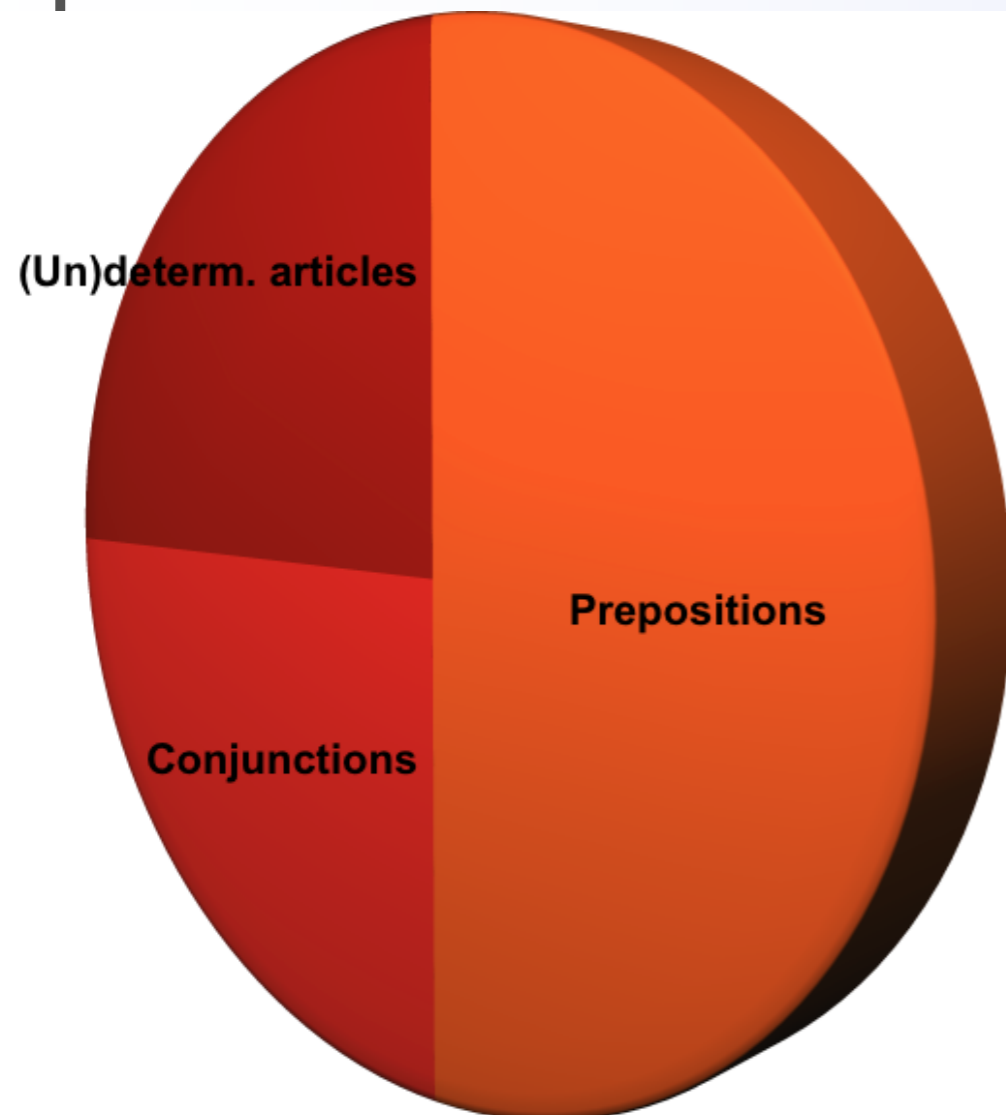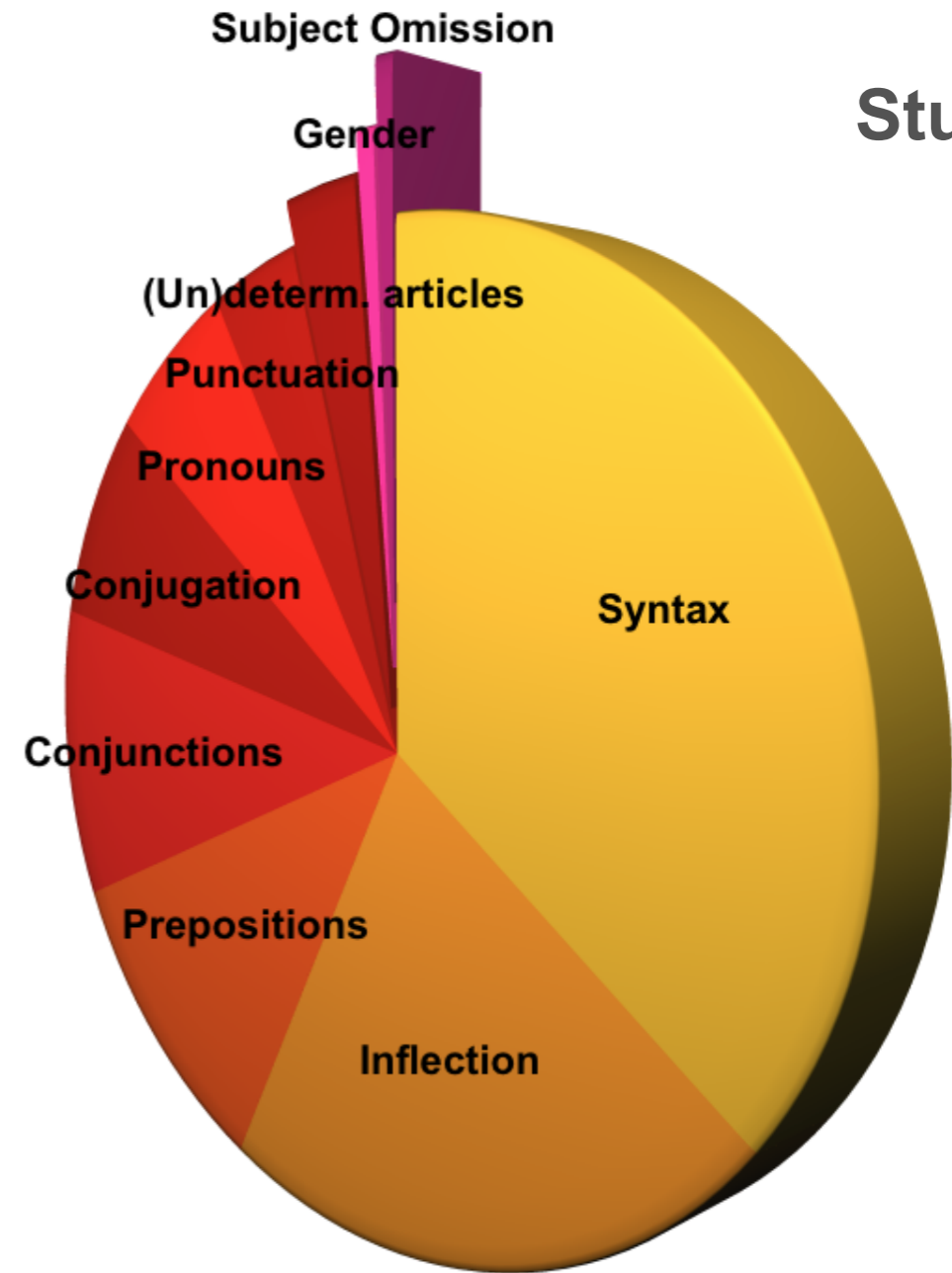
Conclusions and Future Prospects

- Post-editing machine translated texts can be useful if

  - target language is mastered

  - the quality of NMT-System is acceptable

Conclusions and Future Prospects

- MT allows to detect which grammatical aspect needs to be deepened => didactic use

- The error tendencies we found in our research in 2006 were confirmed.

- Students must learn about possibilities and limits of MT, and how they work.

- Context is crucial for correct translation, and this is currently not being considered in MT.

Conclusions and Future Prospects

- Professional translator profile is changing: Syllabus in translator training at universities should include more MT and PE

Conclusions and Future Prospects

Future Prospects

- Studies on comparing translation quality of same target texts with and without MT

- MT into the mother language

- Examples of how MT can be used to promote metacognition among translation trainees and to strengthen their grammar skills

Google Translate and DeepL: Breaking Taboos in Translator Training

Dr María-José Varela Salinas

mjvs@uma.es
https://uma.academia.edu/DrMaríaJoséVarela
https://www.researchgate.net/profile/Maria_Jose_Varela_Salinas

Dr Ruth Burbat

rburbat@ugr.es
http://ugr.academia.edu/RuthBurbat
https://www.researchgate.net/profile/Ruth_Burbat

Google Translate and DeepL: Breaking Taboos in Translator Training

Some of the research presented was conducted in cooperation with:

First International Conference on Research in Multilingualism:
Innovation and New Challenges
Universidad de Oviedo, 14-16/03/2018

Google Translate and DeepL: Breaking Taboos in Translator Training

Dr María-José Varela Salinas (University of Malaga)
Dr Ruth Burbat (University of Granada)