

Second Language Learning and Teaching Research Group, Lancaster University
2 May 2012

CORPUS-BASED ANALYSIS OF THE USE OF THE PRESENT PROGRESSIVE CONSTRUCTION IN ORAL LANGUAGE PROFICIENCY EXAMS



Costas Gabrielatos and Judit Kormos

The rationale for the research

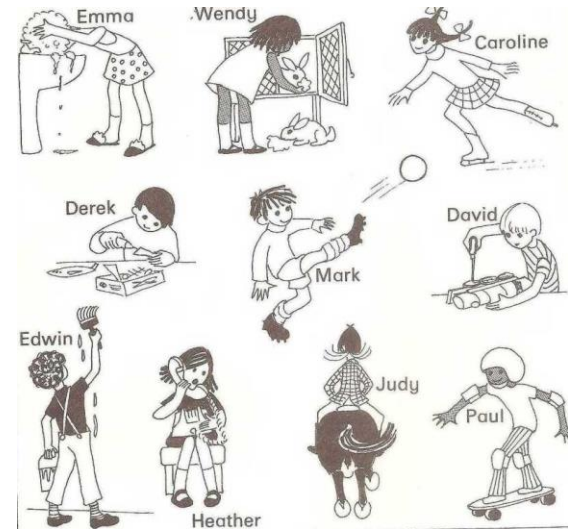
- There is little research on the productive use of grammatical constructions in the field of second language acquisition, especially with regard to young foreign language learners.
- We have little information on how oral communicative tasks can elicit targeted grammatical constructions.
- Little is known at what levels of proficiency certain grammatical structures with particular communicative functions emerge and can be meaningfully assessed.

Aims of the study

- To develop corpus-based queries for identifying the Present Progressive construction in second language data
- To analyze the frequency, semantic functions and accurate use of the Present Progressive in oral exam data across various proficiency levels
- To compare the frequency and communicative function of the Present Progressive construction in candidate and examiner's talk and naturally occurring native speaker conversation (spoken BNC)

Why the Present Progressive?

- Relatively easy to retrieve through corpus queries
- Relatively frequent and is targeted at different levels in the exam we worked with
- Has several communicative functions



The nature of the data

- Face to face oral interview conducted with a native speaker examiner
- 12 Grades, grouped into 4 Stages:

Stage	CEFR
Initial (Grades 1–3)	Basic User (A1-A2)
Elementary (Grades 4–6)	Basic to Independent User (A2-B1)
Intermediate (Grades 7–9)	Independent User (B1-B2)
Advanced (Grades 10-12)	Proficient User (C1-C2)

Test tasks

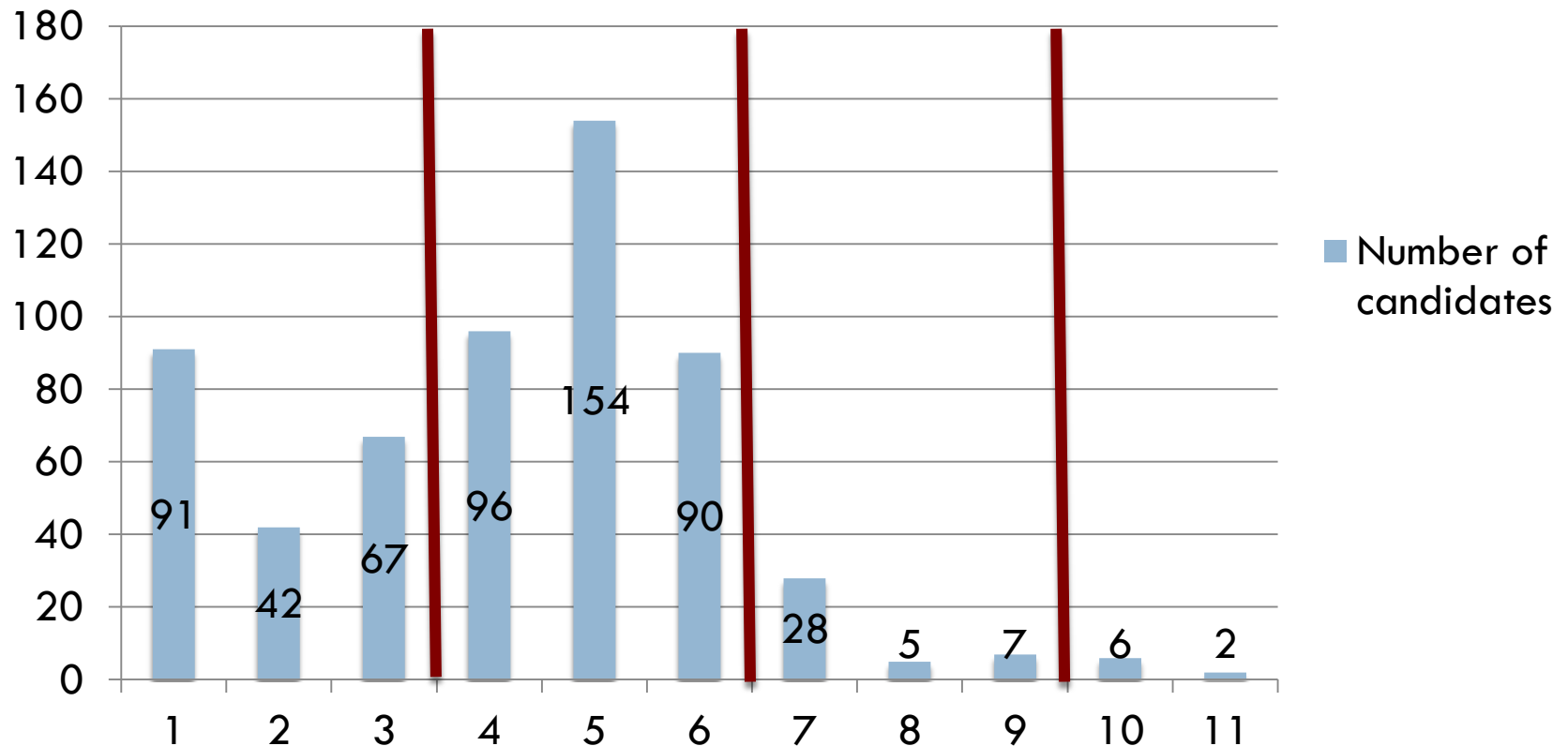
Initial	Elementary	Intermediate	Advanced
Grades 1-3 (5-7 minutes)	Grades 4-6 (10 minutes)	Grades 7-9 (15 minutes)	Grades 10-12 (25 minutes)
			Topic presentation Topic discussion
		Candidate-led discussion of topic	Interactive task
	Topic discussion	Interactive task	Listening task
Conversation	Conversation	Conversation	Conversation

Assessment focus

- Understanding of Present Progressive at Grade 2
- Productive use of the in progress meaning of Present Progressive at Grade 3
- Productive use of the future plans meaning of Present Progressive at Grade 6

Demographic characteristics of the participants

The distribution of candidates across the grades of the exam



Demographic characteristics cont.

- The mean age: 12 years and 6 months, Range: 8-56, majority of the candidates aged between 9 and 12
- The gender of the participants was unknown in 44% of the cases, but the available gender data suggests that 48% of the candidates were female and 52% were male.
- Except for two candidates, all the participants passed the test, with 21% receiving a distinction, 27.5% a merit and 51% a pass mark.
- 38% Russian speaking and 62% Spanish as L1
- 3 examiners (2 male, one female)

The corpus query

```
[hw="be" & pos="VB[MRZ]"]
```

```
[pos="R.*" | pos="XX" | pos="AT.*" |  
pos="DB.*" | pos="DD.*" | pos="NN.*" |  
pos="NP.*" | pos="PN" | pos="PN1" |  
pos="PPGE" | pos="PPH1" | pos="PPHS1" |  
pos="PPHS2" | pos="PPIS1" | pos="PPIS2" |  
pos="PPY"] {0,3}
```

```
([pos=".*V.G.*" & word!="(gon|going)"%c]  
[] | [pos=".*V.G.*" & word=  
"(gon|going)"%c] [pos!="TO"])
```

- *am, are, is*

followed by (up to three + optional) ...

- **adverbs**
- **negation markers**
- **articles**
- **determiners (e.g. *this, any*)**
- ***there***
- **(proper) nouns**
- **personal/possessive pronouns**

followed by ...

- **the *-ing* form of a verb ...**
- **... but not the forms *going* or *gonna* when followed by an infinitive**

The accuracy of the corpus query

- **Precision:** The proportion of true positives out of the instances returned by the query.
- **Recall:** The proportion of true positives returned by the query out of the total of true positives in the corpus (sample).

$$F \text{ (accuracy)} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

Max. = 1

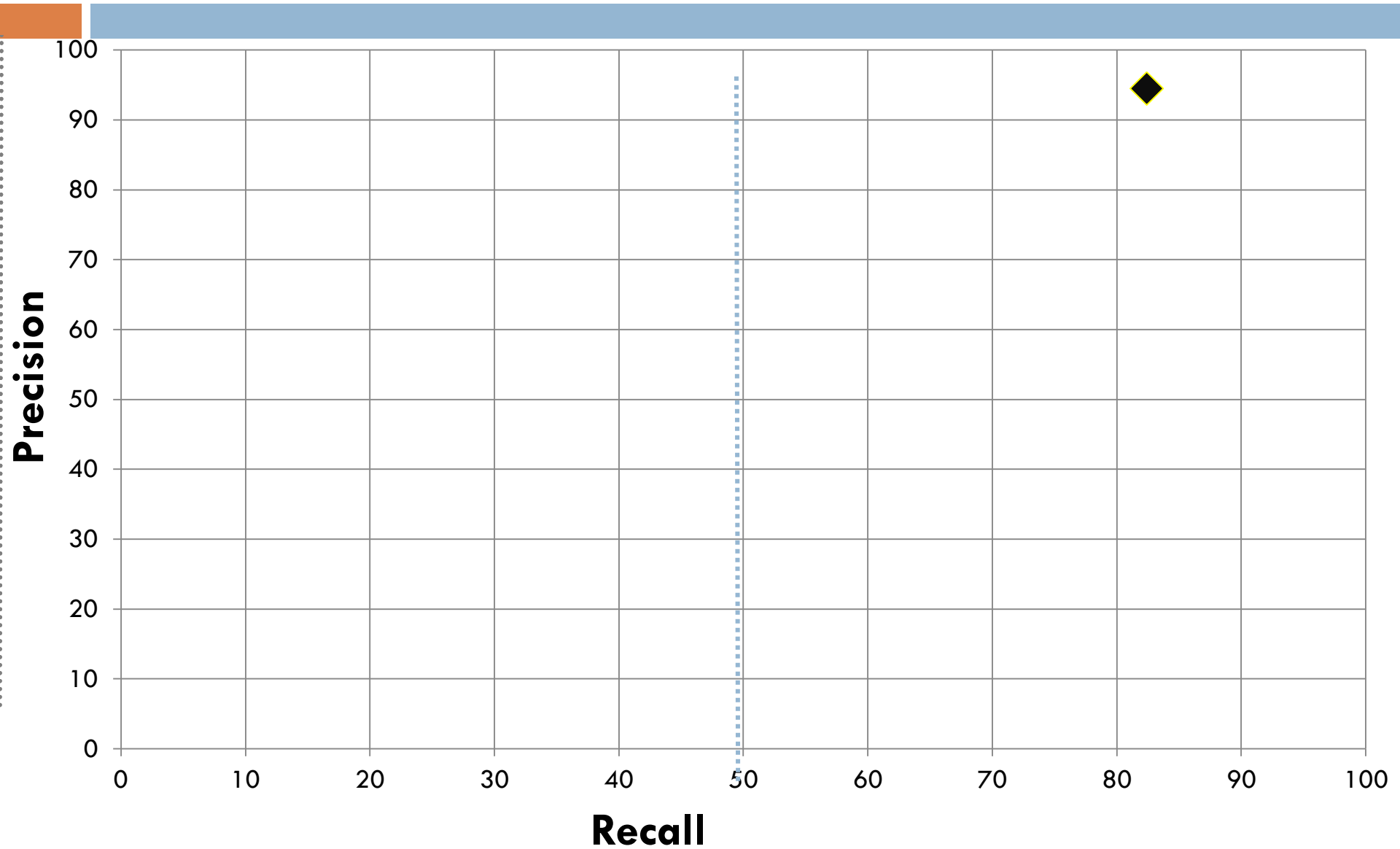
Min. = 0

Precision = 0.945 (631/668)

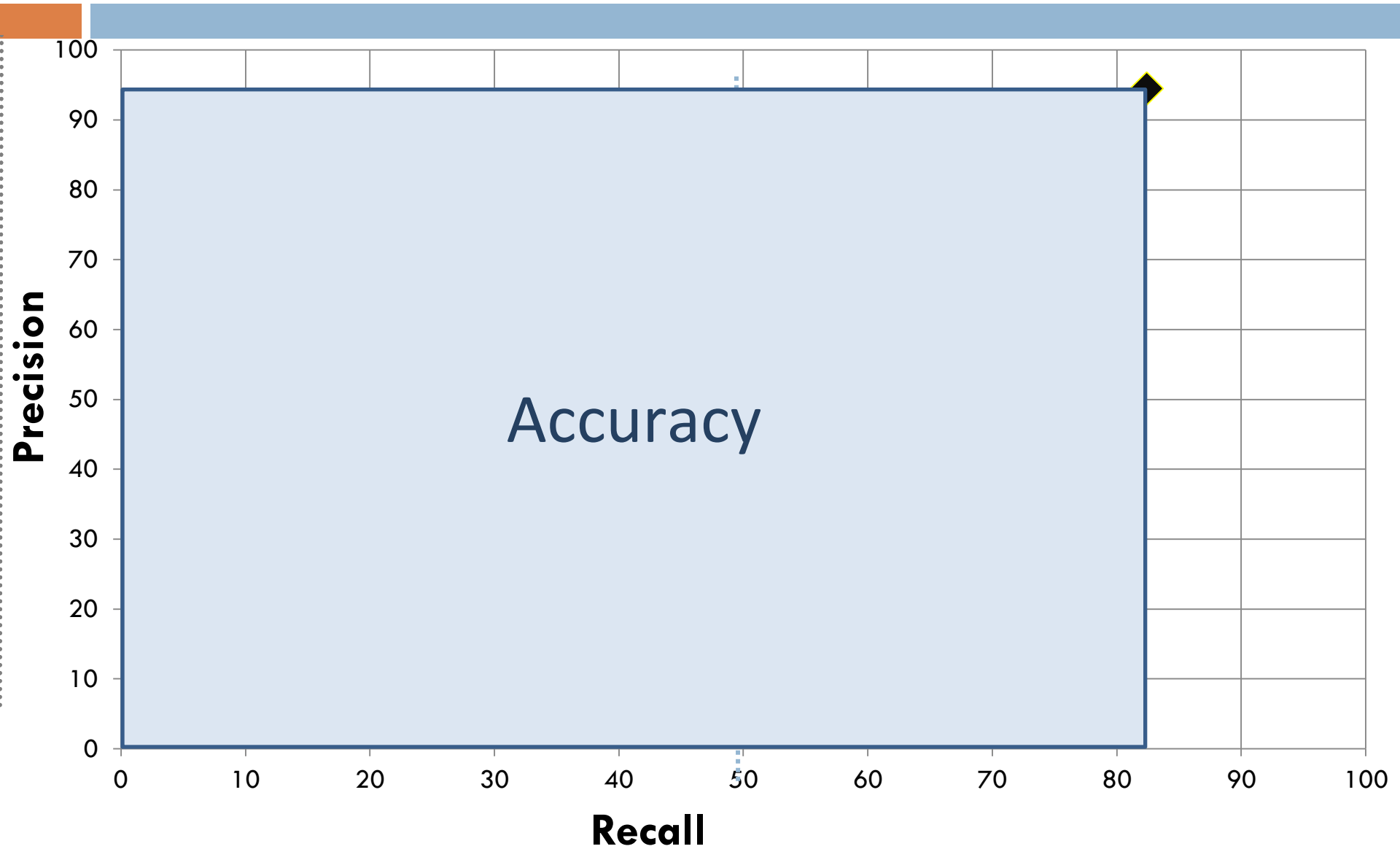
Recall = 0.824 (631/766)

F = 0.880

A visual representation of query accuracy



A visual representation of query accuracy



Methods of analysis

- Manual analysis of the instances of the Present Progressive used by the learners (766) and examiners (933)
- A random sample of 638 instances of the Present Progressive from the spoken BNC, derived after the manual cleaning of a sample of 1000 random instances.
- Establishing the communicative function of each instance of Present Progressive
- Double-coding and checking the reliability of the analysis
- Analysis of errors (errors of meaning and form)

The communicative functions of the Present

Progressive 1. (after Quirk et al., 1985: 1290-1292)

□ In_progress

An event/state/action is presented as being ongoing in the immediate present (i.e. around 'now', as we speak).

□ In_progress=narrative_past

An event/state/action that was ongoing around a past time point, is presented as if it were happening 'now' – as a narrative technique.

□ In_progress=narrative_present

An instance of a habitual event/action is presented as ongoing for effect; that is, in order to make the narrative more vivid. Usually used with adverbs of time denoting recurrence (e.g. *when, whenever, every time*).

The communicative functions of the Present Progressive 2.

□ Present=Temporary

A present event/state/action is presented as temporary, or in progress, but with interruptions, over a long period of time.

□ Present=Repeated action/event

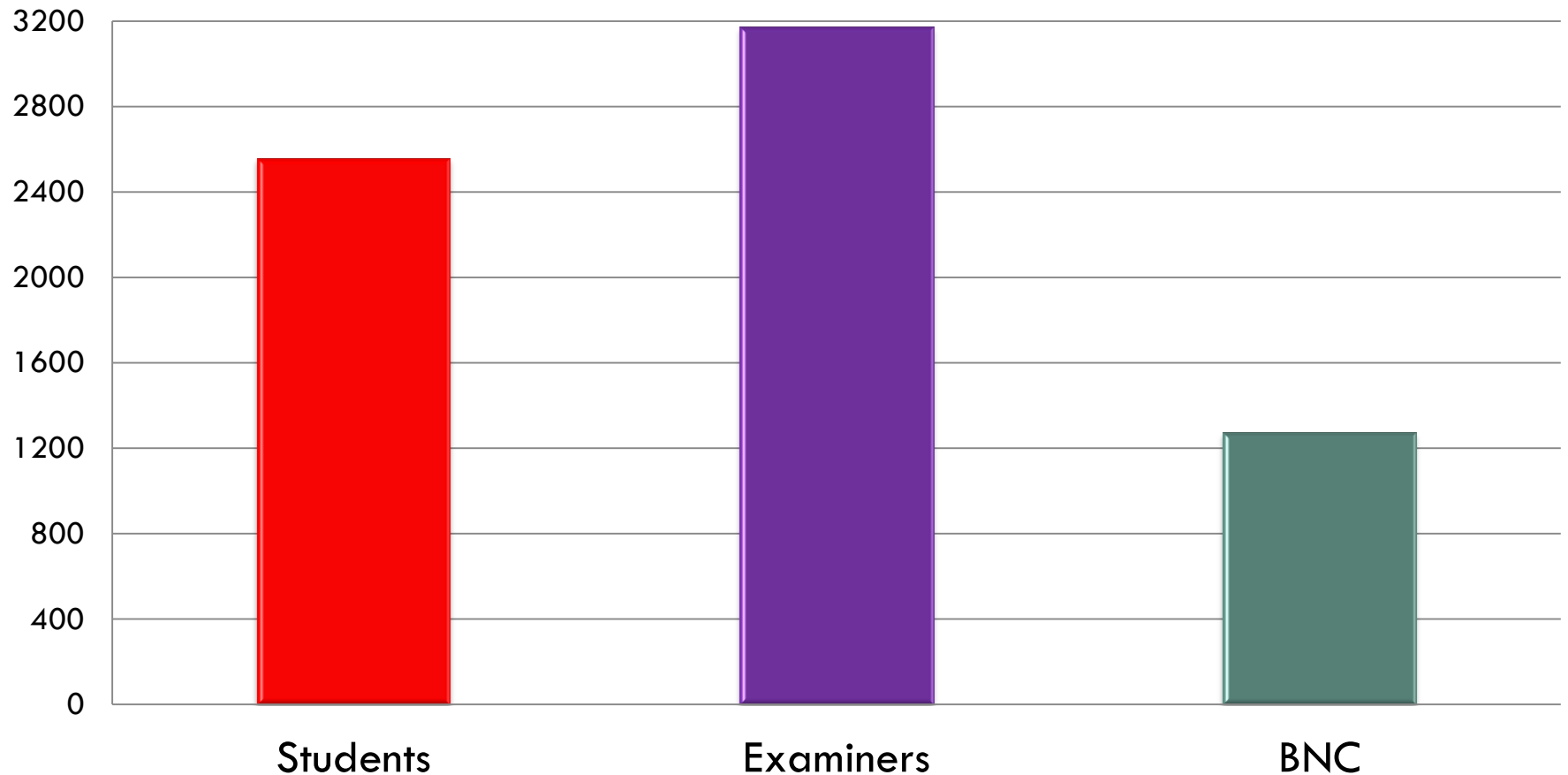
The repetition of a habitual action/event is stressed – normally accompanied by *always*, to imply that the speaker has strong (usually, but not necessarily, negative) feelings towards the habit.

□ Planned future:

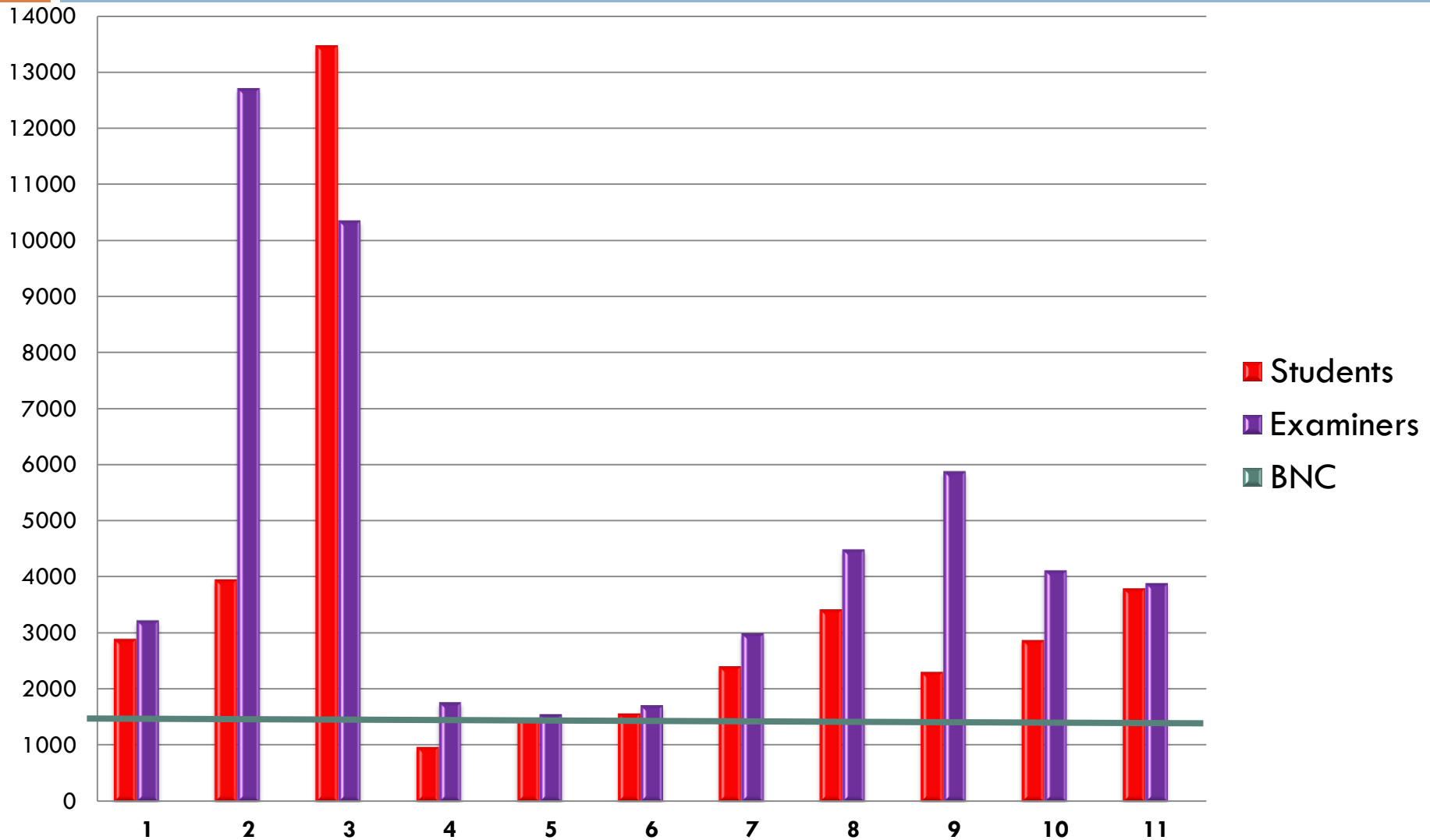
future events for which arrangements have been made in the present.

The frequency of the Present Progressive construction

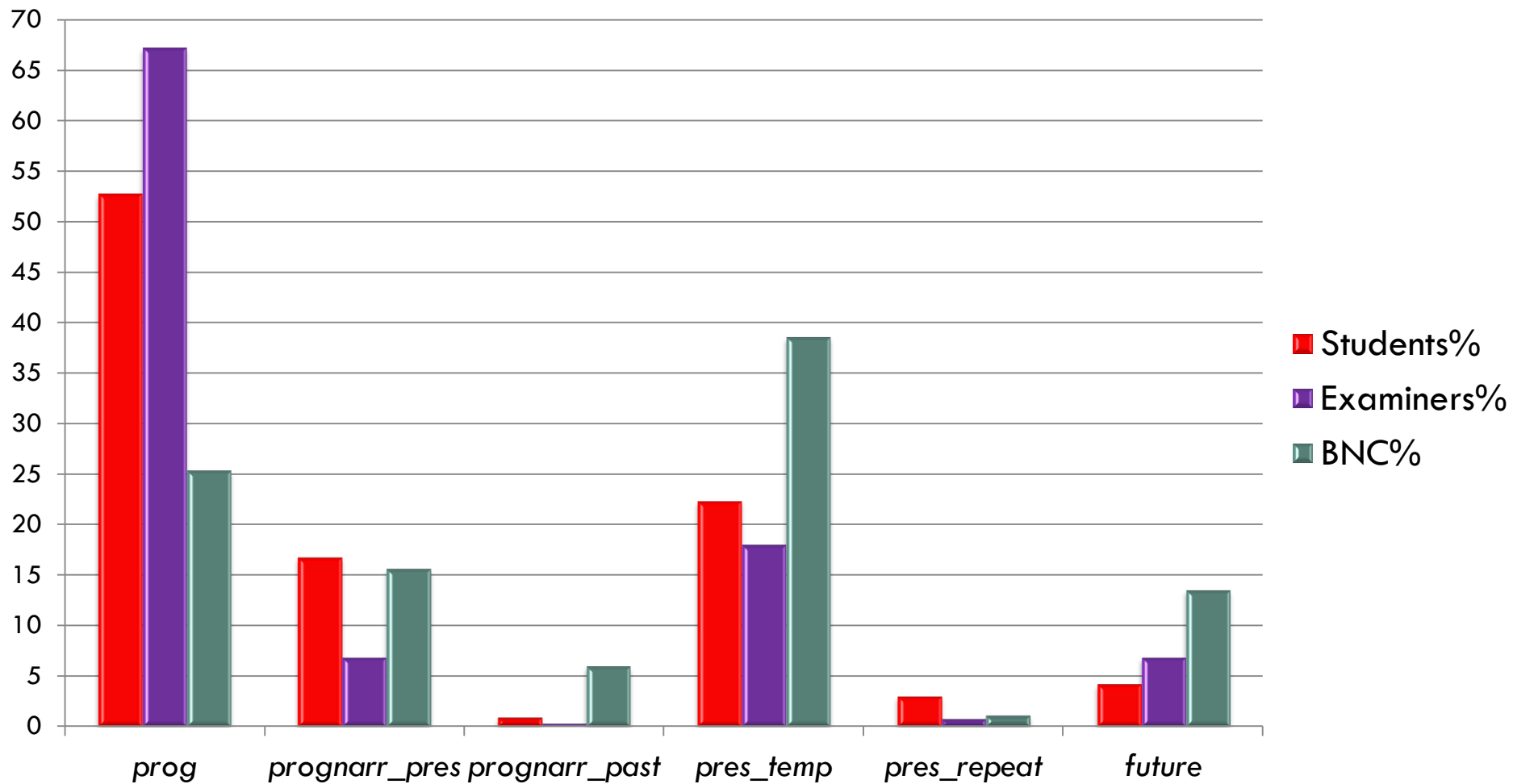
Normalised frequencies (per million words)



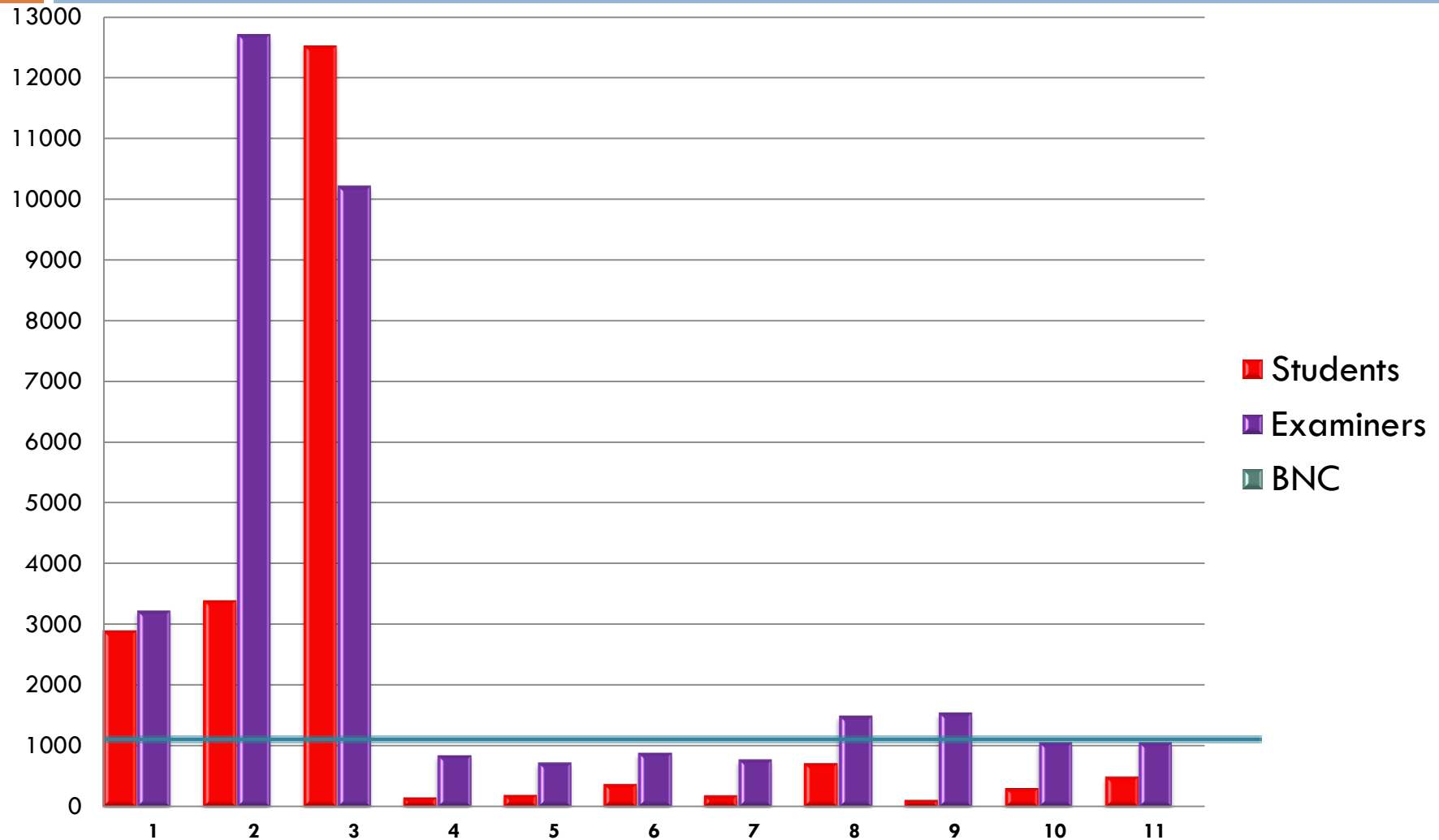
Frequency of use across the grades of the exam



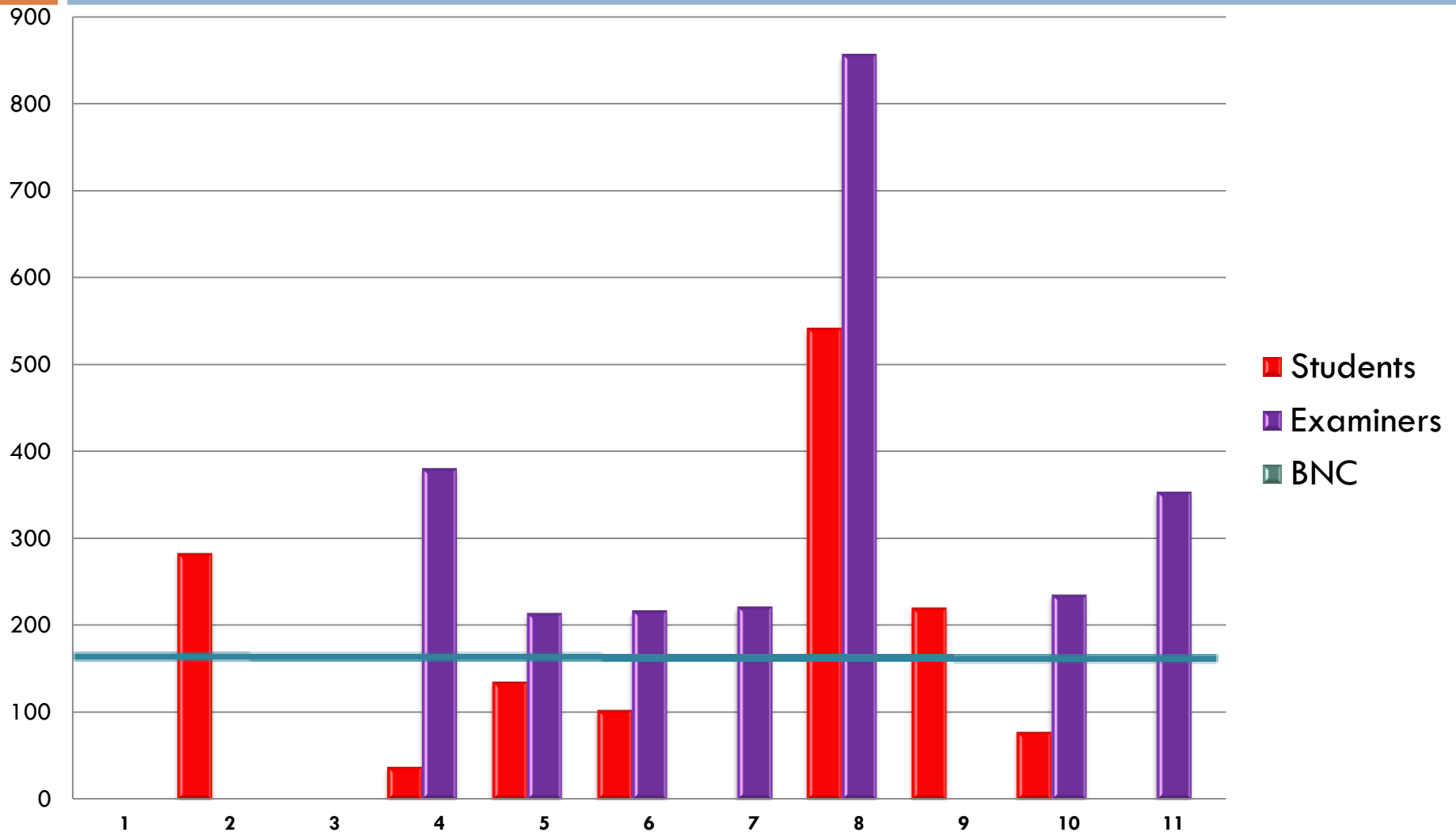
Proportion of communicative functions



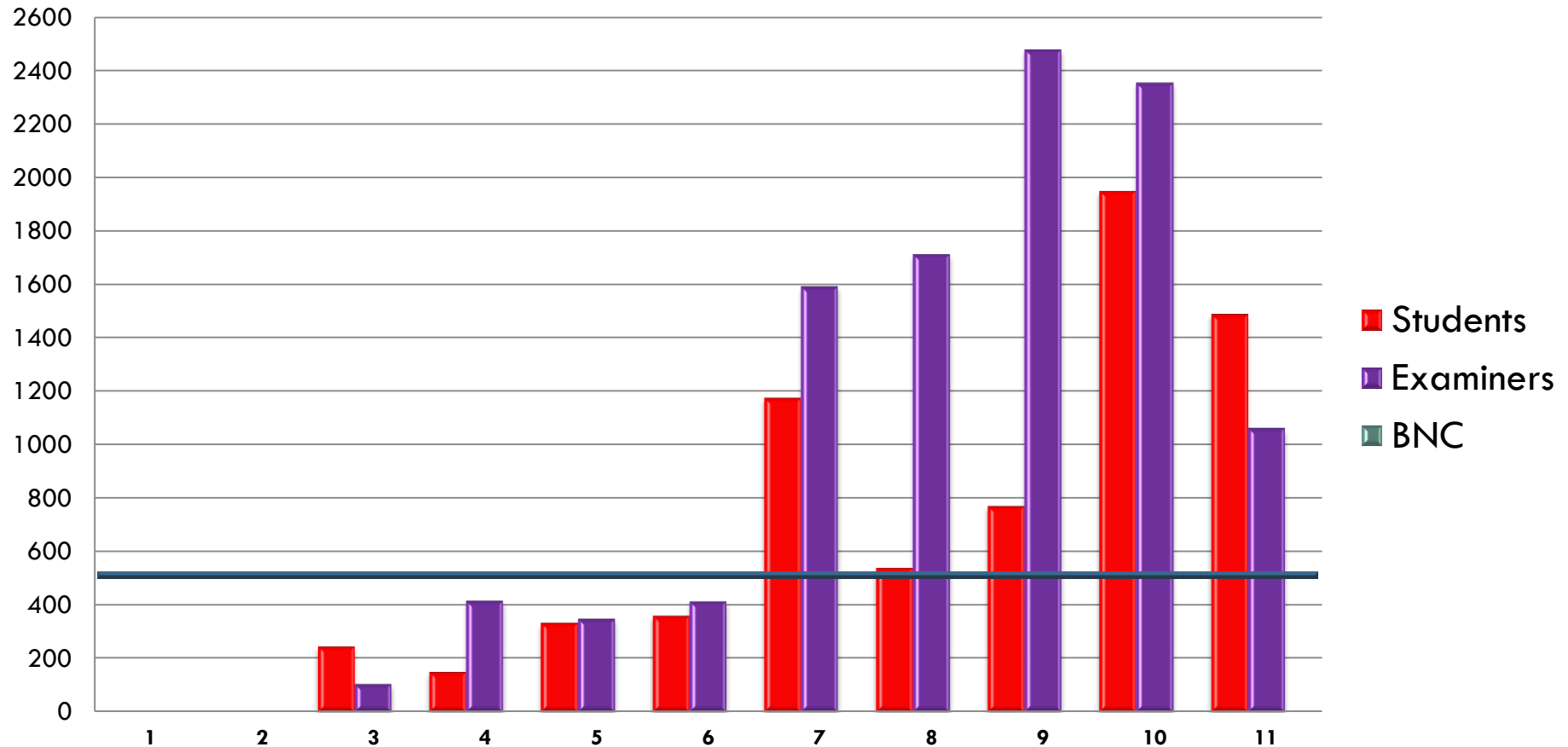
Frequency of the “in-progress now” meaning across grades



Frequency of the “future” meaning across grades



Frequency of the “temporary present” meaning across grades

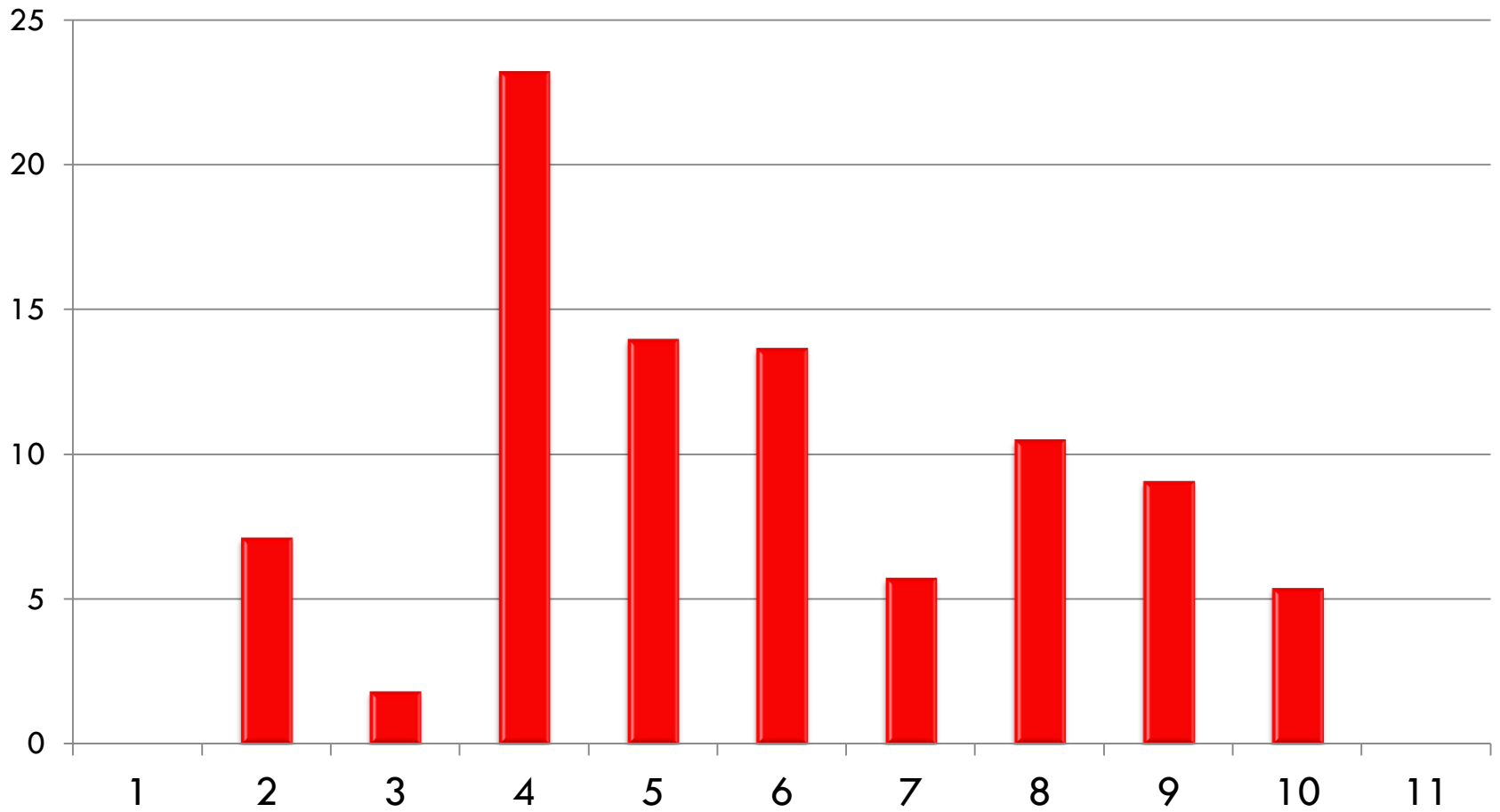


Learner errors – Errors of meaning

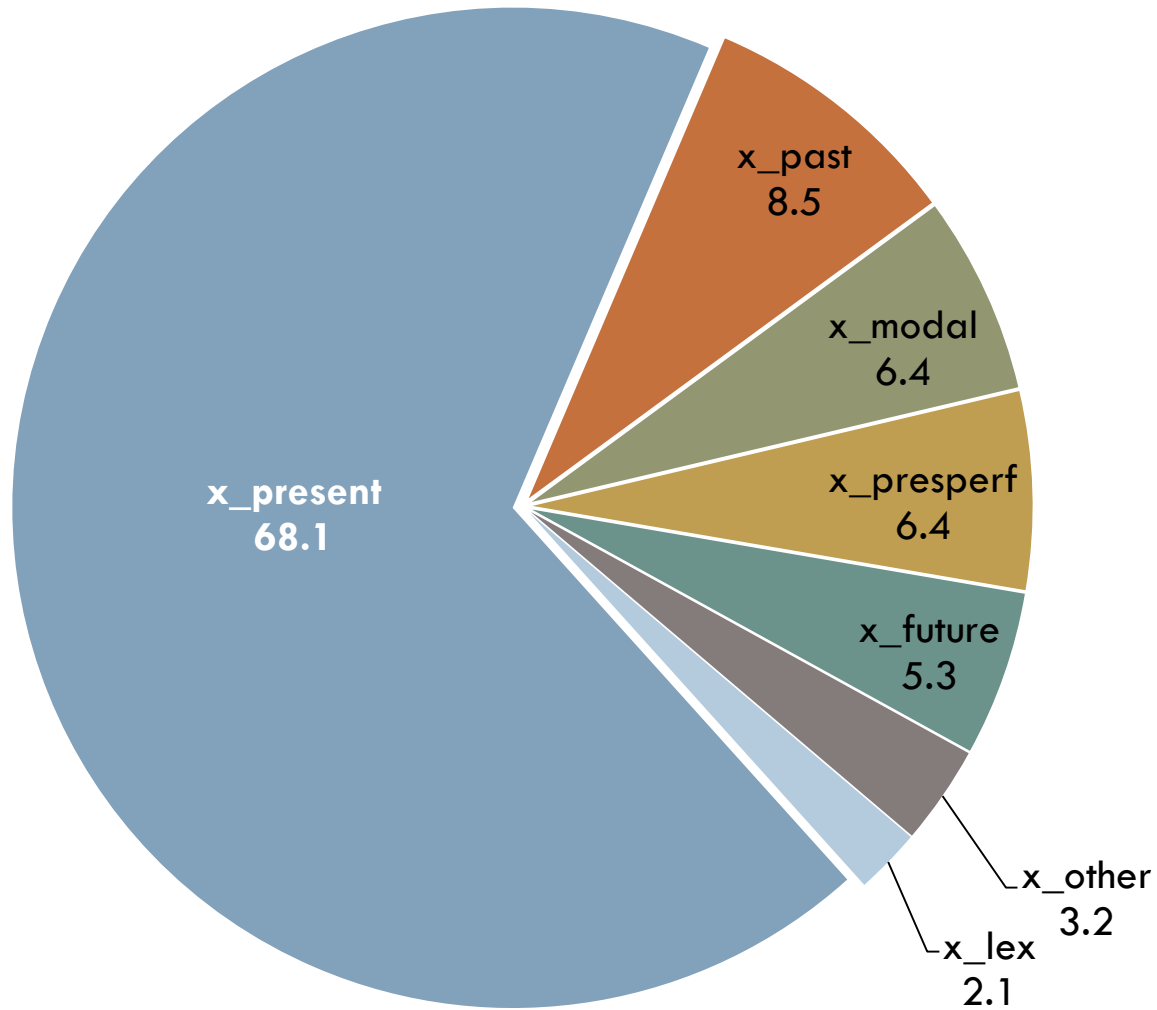
- 12.4%: 94 functional errors out of 756 instances (in which function could be clearly established).
- Most frequent error - 64 instances (68.1% of functional errors)
- A habitual event/action/(mental) state is presented as temporary — accurately expressed through the **Present Simple**.

I think er my first question will be er **are you eating** hamburgers yes you should stop [9265526_1 disc]

Distribution of $x_present$ across grades



Other errors (proportions)



Implications for corpus-research

- Corpus research can identify, and quantify, patterns of use that can inform ...
 - ▣ Exam syllabi (what is examined at which level)
 - ▣ Exam tasks (eliciting the target features)
 - ▣ Examiner training (prompting techniques)
- Corpus queries can be expected to have lower accuracy on learner data – particularly in spoken production.
 - ▣ Errors, false starts, repetition, interrupted turns
- Target of corpus queries in learner corpora:
 - ▣ Correct uses or errors?
- For fine distinctions (e.g. cross-tabulations) the corpus needs to be fairly large.
 - ▣ About 5 million for current data and cross-tabulation detail.

Implications of the findings for SLA research

- The ‘in-progress’ meaning of Present Progressive emerges first in L2 learning, but there is a back-slide in the accuracy of the construction at A2 level.
- The ‘future’ meaning of the Present Progressive emerges early, but it is rarely used.
- The most typical meaning of Present Progressive in BNC data is the ‘temporary present’ function.
- The ‘temporary present’ function is also used frequently by students and examiners and it emerges at B1 level

Implications for assessment

- Grammatical constructions where there is a one-to-one correspondence between the communicative meaning and the construction used to express it can be easily elicited in oral proficiency interviews and in free production tasks.
- Grammatical constructions where there is no such one-to-one correspondence, i.e. when a communicative function can be expressed by a number of alternative constructions, cannot be easily elicited successfully.