When are grammatical structures best taught? Using learner corpora to discover the critical proficiency levels for each grammatical structure

Mick O'Donnell, Susana Murcia
Universidad Autónoma de Madrid

Contents

1. The TREACLE Project

- Project: TREACLE

  Teaching
  Resource
  Extraction from an
  Annotated
  Corpus of
  Learner
  English

- A cooperation between Universidad Autonoma de Madrid and University Politecnica de Valencia (Penny McDonald, Keith Stuart, Maria Boquera)
- Funded by Ministerio de Ciencia e Innovación 2010-2012 (FFI2009-14436/FILO)

2. Goals of Project

- To produce a syntactically analyzed learner corpora of English, with error annotations.
- Use this corpus to produce profiles of each proficiency level (A1, A2, B1, etc.)
- Use these profiles to redesign the teaching curriculum: determining which grammatical features need to be taught, in what order, and with what degree of emphasis.
- Extract teaching examples and exercises from the corpus.
- Provide a web-based language learning system which dynamically adapts exercises presented to the student by reference to the students current performance and the proficiency profiles derived above.

2. Goals of This Talk

- To produce a syntactically analyzed learner corpora of English, with error annotations.
- Use this corpus to produce profiles of each proficiency level (A1, A2, B1, etc.)
- Use these profiles to redesign the teaching curriculum: determining which grammatical features need to be taught, in what order, and with what degree of emphasis.
- Extract teaching examples and exercises from the corpus.
- Provide a web-based language learning system which dynamically adapts the materials and exercises presented to the student by reference to the students current performance and the proficiency profiles derived above.

2. Project Goals

- Error analysis is one way to explore the grammatical competence of students at each level (e.g. Dagneaux et al 1998).
- However, some students make few errors, because they avoid structures they are not sure about
- More adventurous students take risks and thus make more errors.
- We thus take a two-pronged approach:
  - Automatic syntactic tagging of corpus to see what structures students are attempting;
  - Manual error analysis to see what they do wrong.
- Only both together give the full picture.

3. The Corpus

- The project involves two corpora:

  The **WriCLE** corpus (UAM) - Written Corpus of Learner English. 700 essays of ~1000 words each, written by Spanish learners of English at University level. (Rollinson and Mendikoetxea 2008)

  The **UPV Learner Corpus** (UPV) containing 150,000 words of shorter texts by ESP students.

- Only the WriCLE corpus is involved in the study reported here.
- A 500,00 word subcorpus was used.

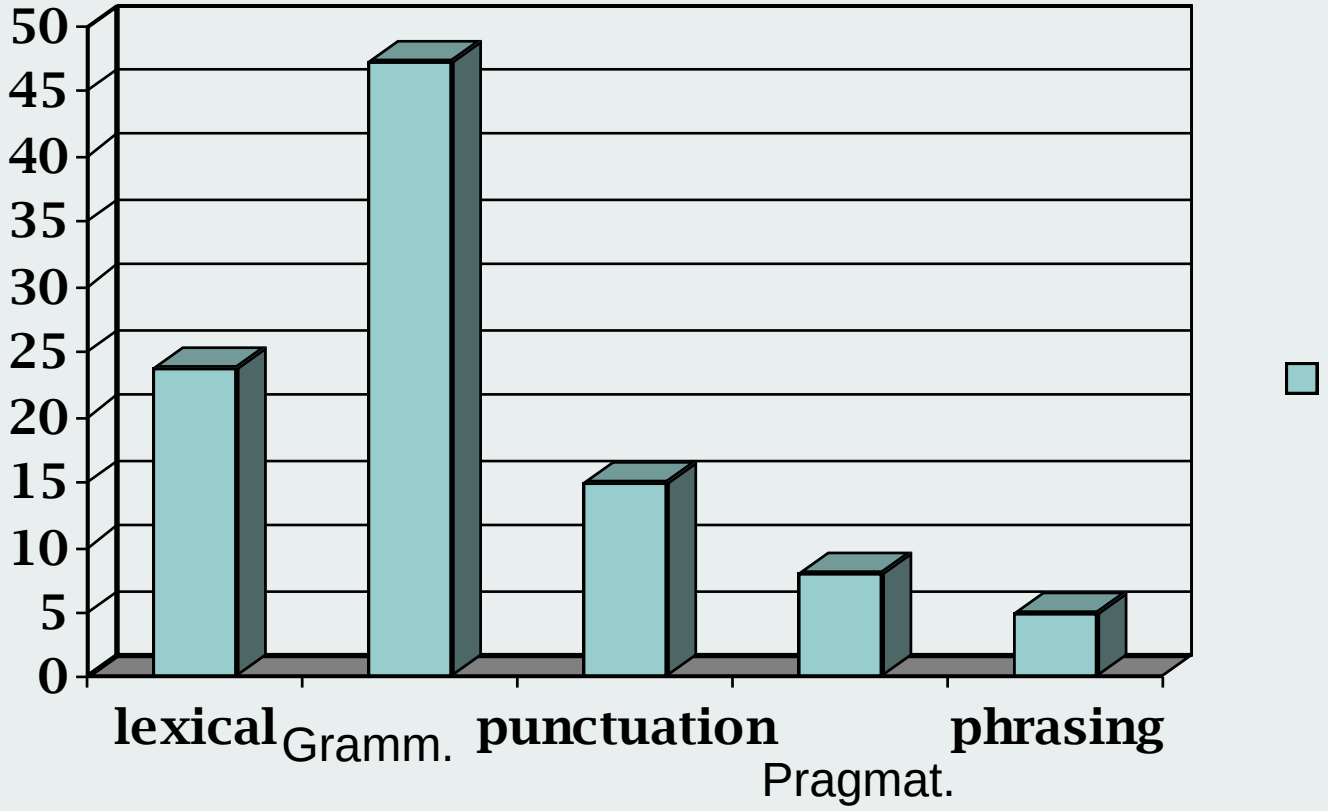
- Oxford Quick Placement test given at same time to measure proficiency

4. Error Anaysis

- Our error analysis still in an early phase (only 1800 errors coded, 28 texts)
- Currently we are in a series of inter-coder reliability studies to refine the error scheme and coding criteria document.
- However, current results give us some indications…

4. Error Anaysis

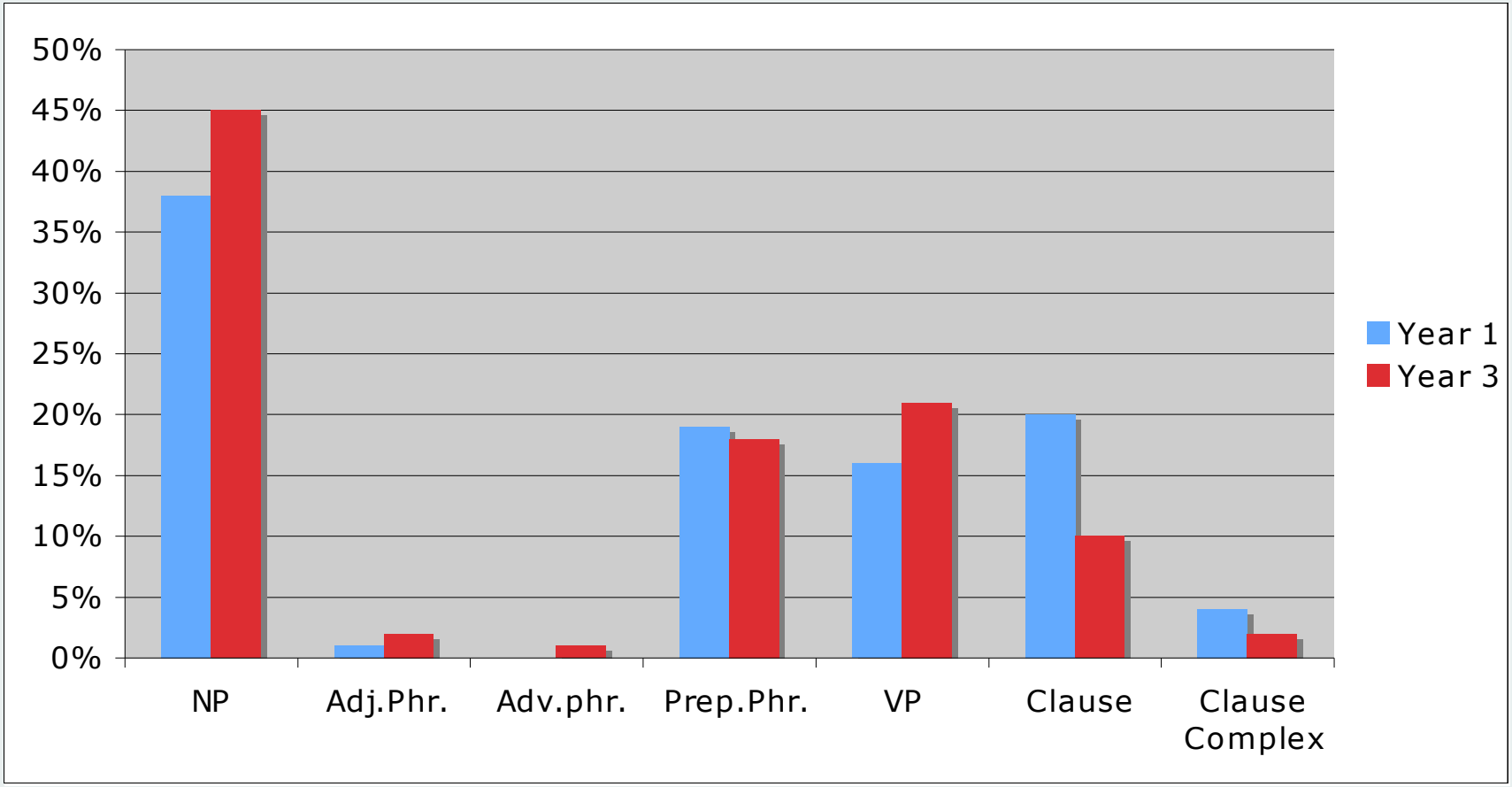
- By examining the types of errors made at each proficiency level, we can determine how much teaching time to spend on each area.



4. Error Anaysis

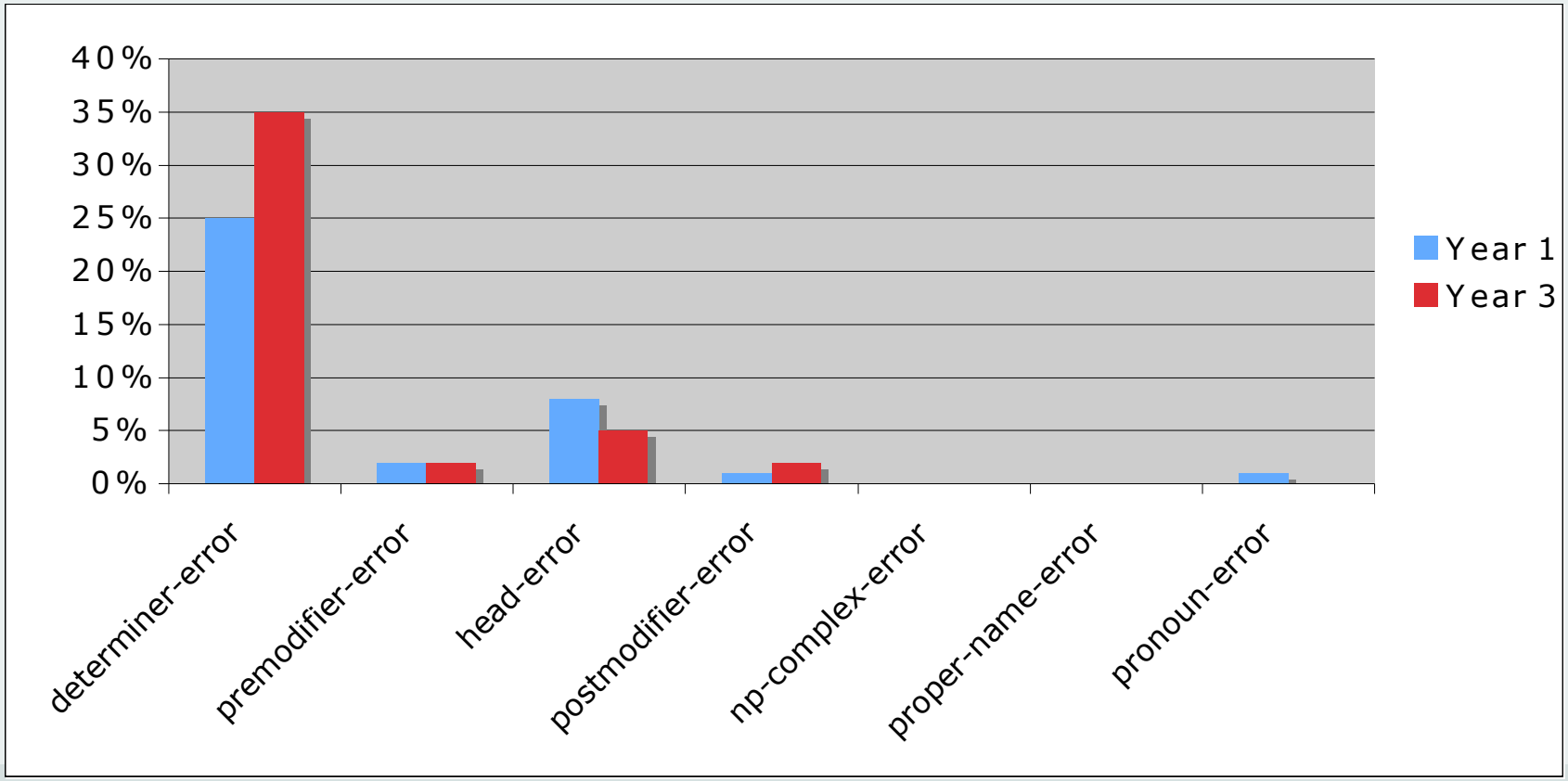
- Within grammar:

Contrasting 1st and 3rd year students



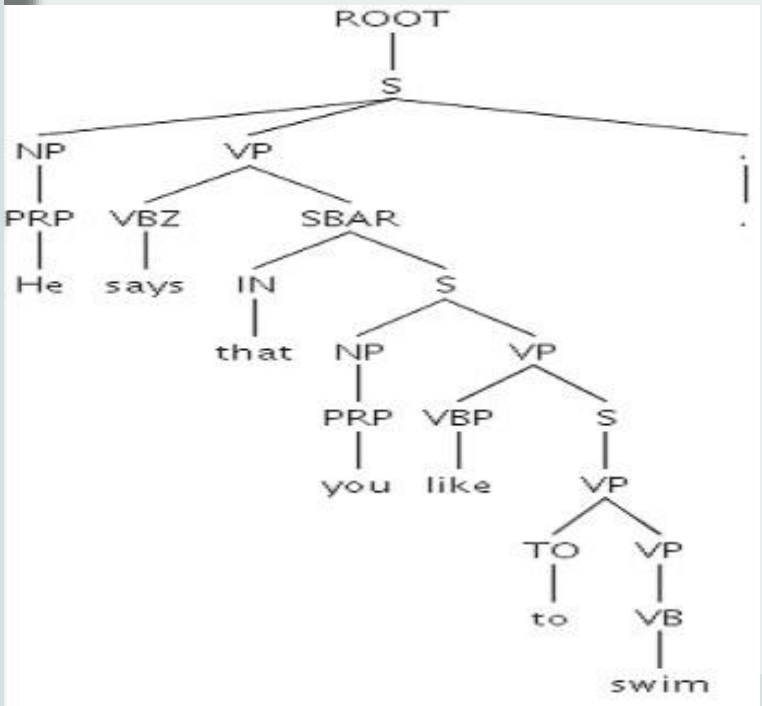
4. Error Analysis

- Within NP Errors
  - Determiner errors most common (inserted-not-required, absent-but-required)
  - All vocab choice errors which are not syntactically incorrect coded as lexical-selection-error



5. Syntactic Analysis

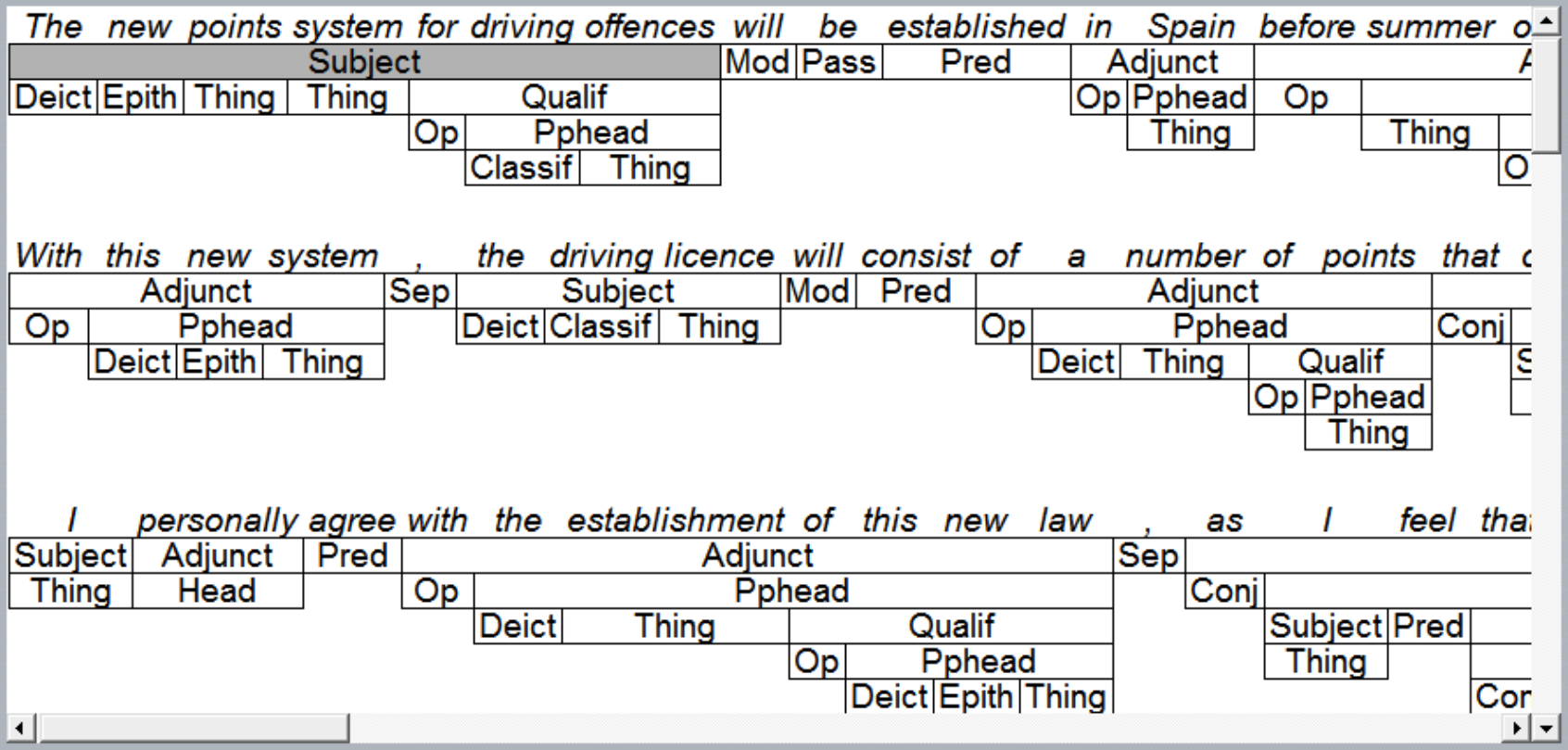
- The Stanford parser produces phrase structure trees (Klein and Manning 2003)
- For ESL research traditional grammar categories are more appropriate (Subj/Pred/Obj, active/passive, relative-clause, etc.)
- UAM CorpusTool thus transforms PSG trees into traditional grammar



⇒

He	*says*	*that you like to swim*
Subj	Pred	Obj

Grammar analysis for: Files/A101-2.txt



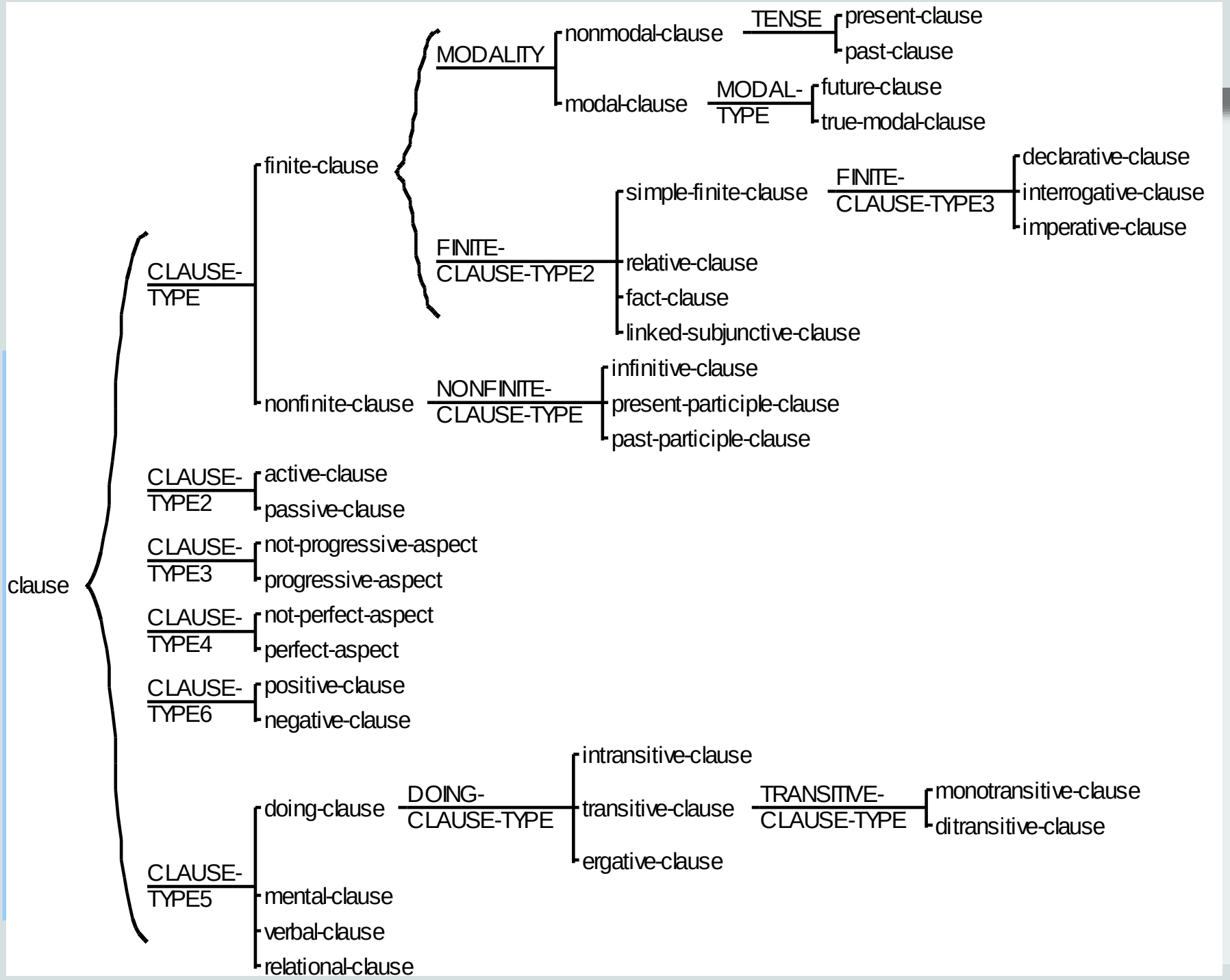
<< < > >> Ignore Delete Other Action... Save Close Help

Assigned

grammatical-unit
group
np
common-phrase
singular-phrase
nonwh-noun-phrase

Gloss

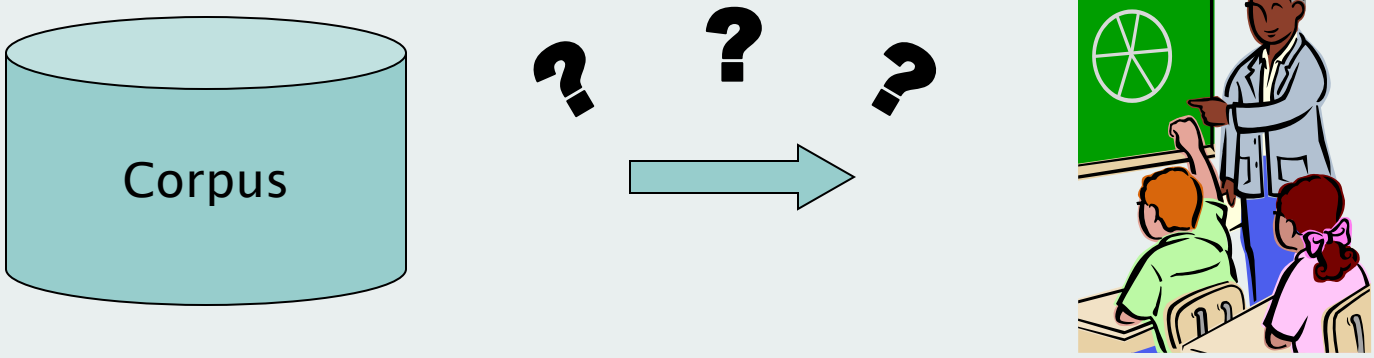
Comment:


clause
CLAUSE-TYPE
finite-clause
MODALITY
nonmodal-clause
TENSE
present-clause
past-clause
modal-clause
MODAL-TYPE
future-clause
true-modal-clause
FINITE-CLAUSE-TYPE2
simple-finite-clause
FINITE-CLAUSE-TYPE3
declarative-clause
interrogative-clause
imperative-clause
relative-clause
fact-clause
linked-subjunctive-clause
nonfinite-clause
NONFINITE-CLAUSE-TYPE
infinitive-clause
present-participle-clause
past-participle-clause
CLAUSE-TYPE2
active-clause
passive-clause
CLAUSE-TYPE3
not-progressive-aspect
progressive-aspect
CLAUSE-TYPE4
not-perfect-aspect
perfect-aspect
CLAUSE-TYPE6
positive-clause
negative-clause
CLAUSE-TYPE5
doing-clause
DOING-CLAUSE-TYPE
intransitive-clause
transitive-clause
TRANSITIVE-CLAUSE-TYPE
monotransitive-clause
ditransitive-clause
ergative-clause
mental-clause
verbal-clause
relational-clause

6. Extracting Profiles from the Corpus

- After the parsing process, we have a corpus of 500 texts, 500,000 words, 66,000 clauses, 120,000 NPs.
- Each clause provided with syntactic function and a range of syntactic features.
- So, what do we do with it?

How do we use the corpus to inform us about **what** students need to learn and **when**?

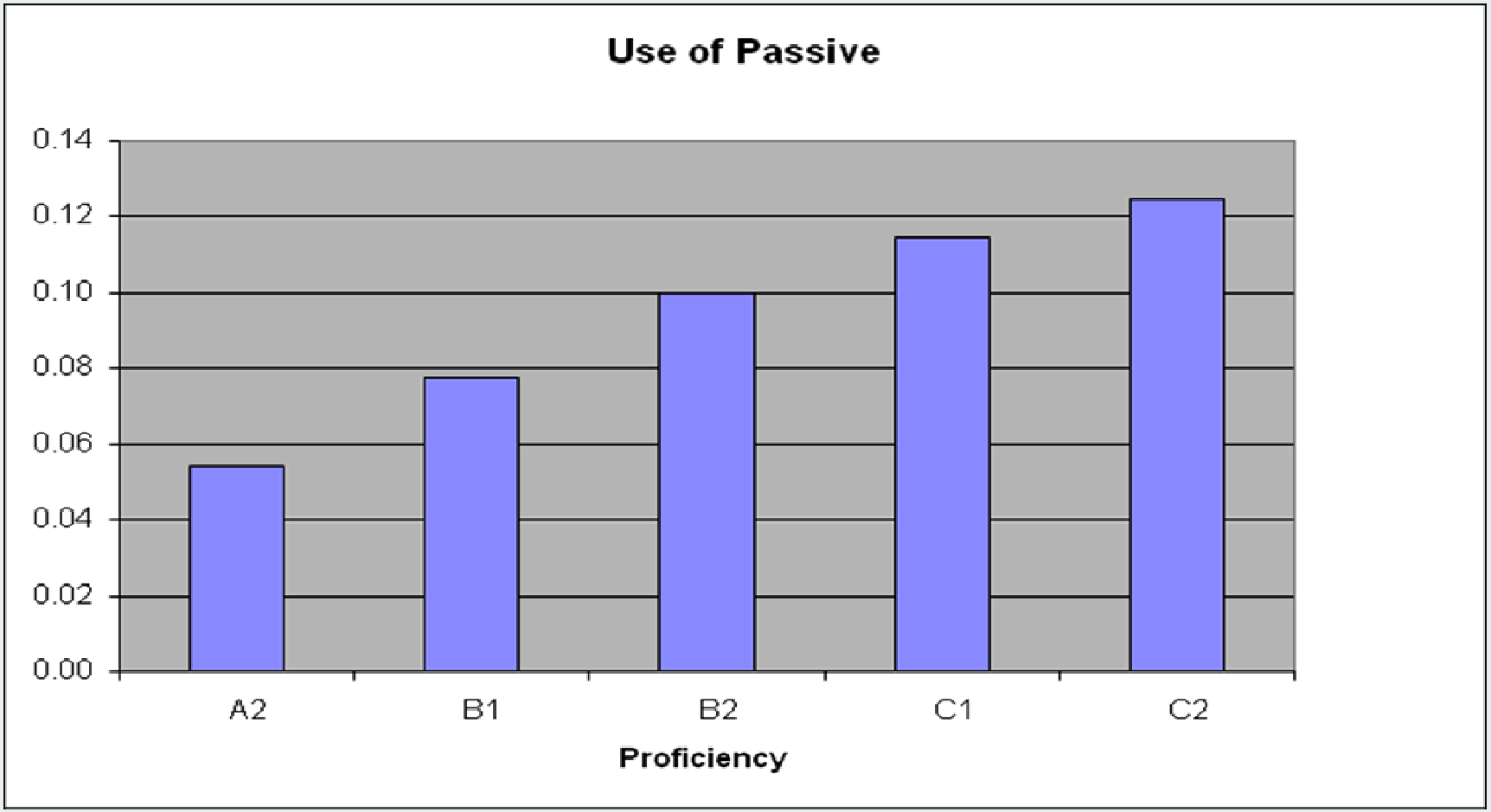


6. Extracting profiles (i): simple frequencies

- Some researchers contrast the learner's degree of usage of a syntactic feature with the degree of usage of natives
- Where students under-use the feature, more emphasis is needed in teaching.
- Over-usage also needs to be corrected (perhaps by teaching alternative lexico-grammatical strategies, or teaching appropriate contexts of use).

6. Extracting profiles (i): simple frequencies

Increased use of passive with proficiency



5. Extracting profiles (i): simple frequencies

Problems with under/over-usage comparisons:

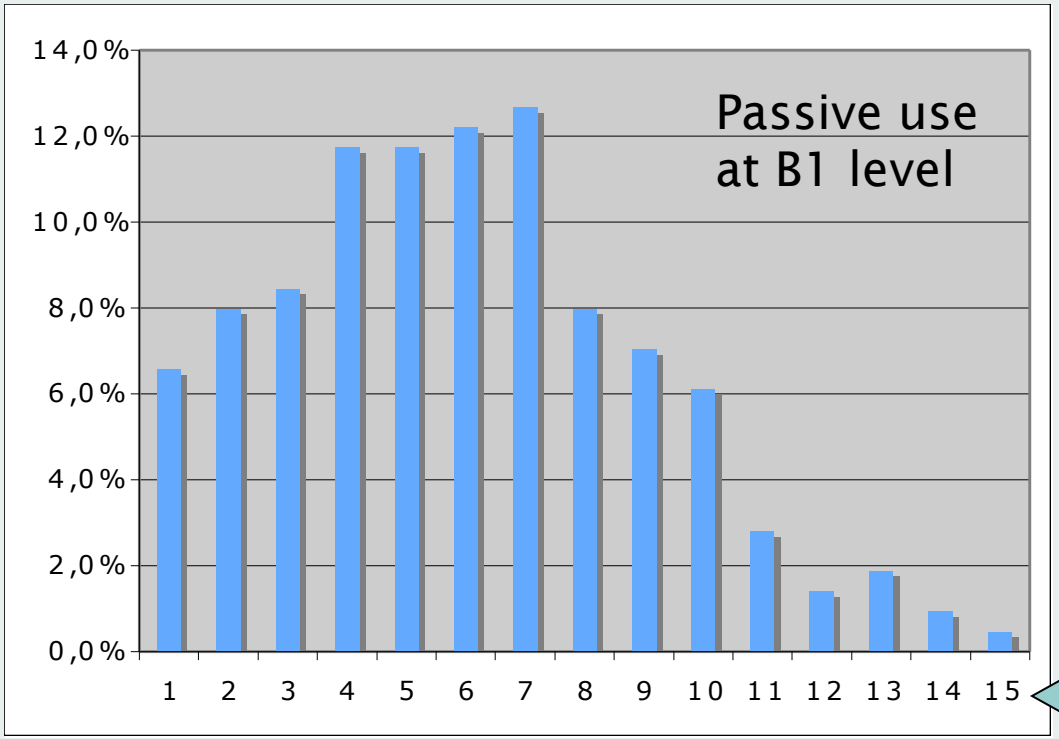
- **When dealing with individual students**: the degree of usage of many features is **register**-dependent, so we cannot really compare with native corpus unless we have a register-matched native corpus.
- **Treating all students in a proficiency band as homogenous**: if we say that average usage of passives at a particular level is 10%, that ignores the fact that some students will over-use passives, and others will not use them at all.

Any teacher will tell you that the students within a proficiency band can have different strengths and weaknesses.

Taking the average of non-homogenous students is like averaging apples and oranges!!

5. Extracting profiles (ii): Signatures

- Rather than averaging the students in a proficiency band, we could instead look at the distribution of students **within** the band.
- The distribution graph within each band shows us the levels of proficiencies **with this feature** at this proficiency level

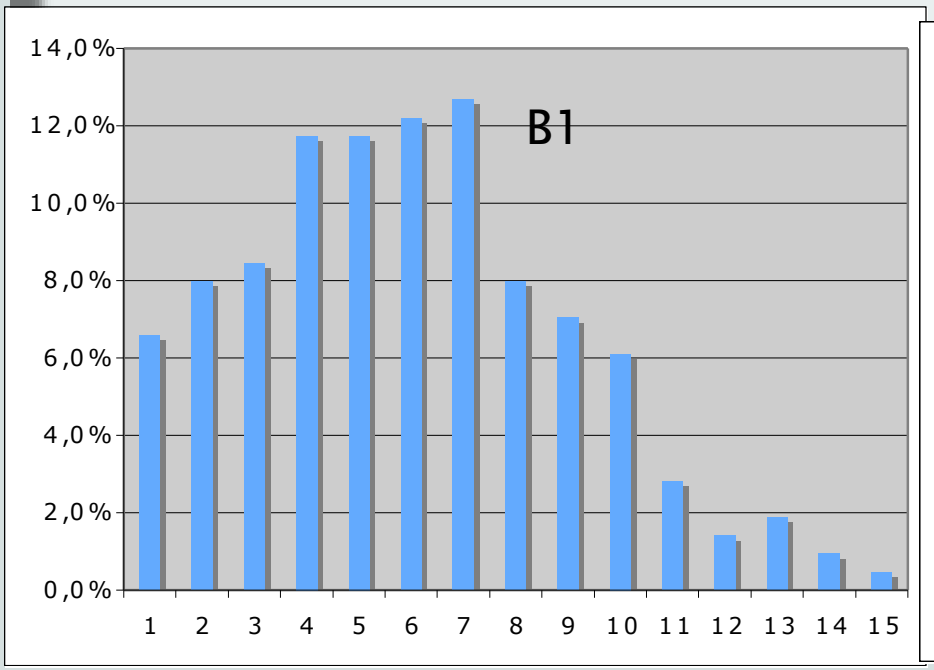


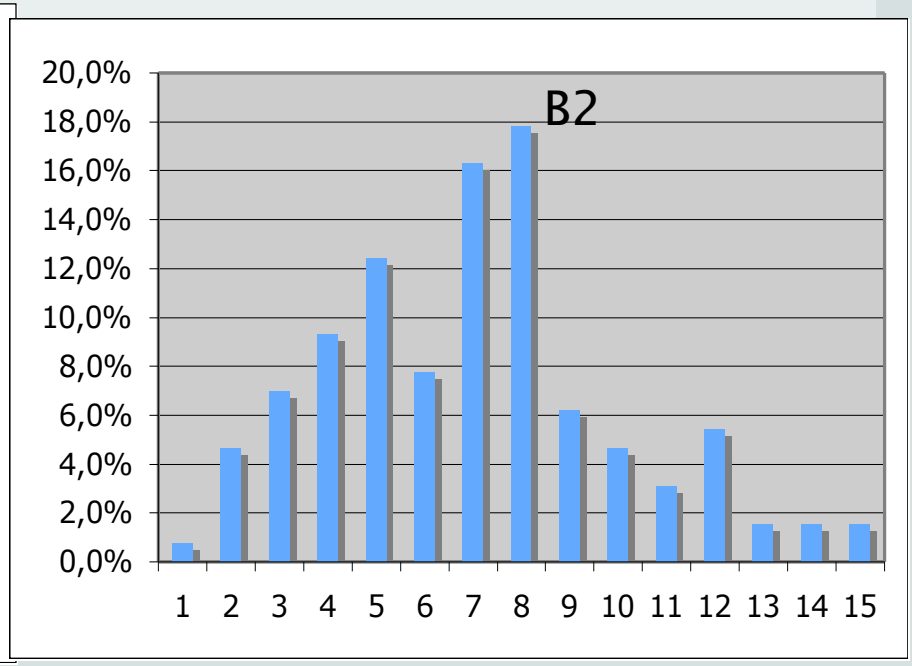
% of texts with this degree of usage

Degree of usage of passive (%)

5. Extracting profiles (ii): Signatures

- Main thing the graph reveals to us is that:
  - Students at a given proficiency level do not perform the same in regards to a particular structure.
  - Different proficiency bands have different profiles for this feature, but lots of overlap
  - E.g. Use of passive:

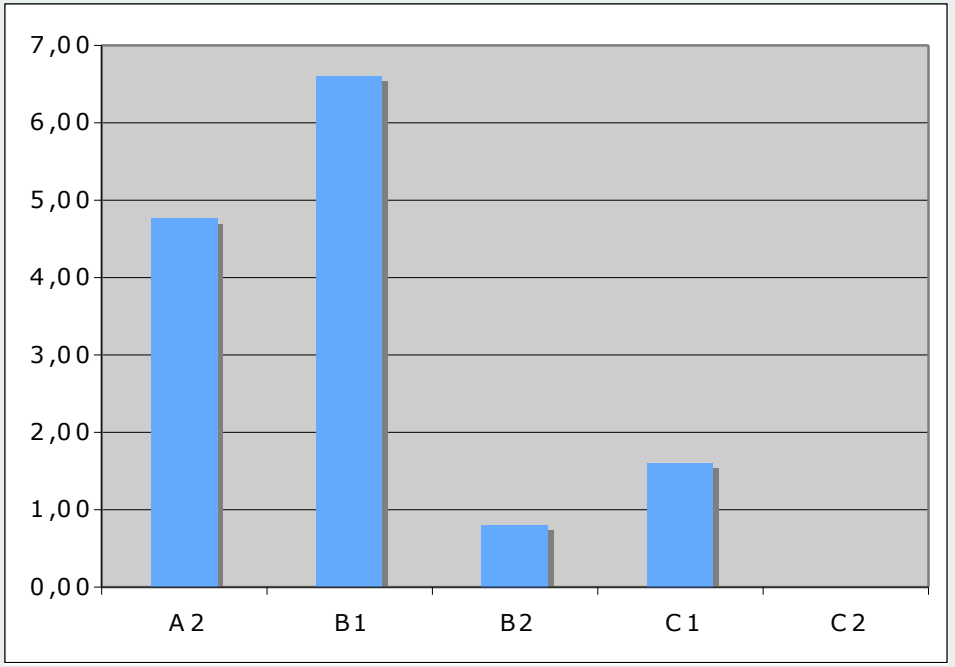




5. Extracting profiles (iii): Onset of Use

- Our belief is that a first concern should be with whether a leaner is capable of producing a structure at all.
- We thus look at each text individually, to see if the structure is present or not.
- We then measure the percentage of texts (~ no. of students) which use the feature **at all** (at each level)
- For this, a reasonably long text is needed (our texts are approx. 1000 words each).

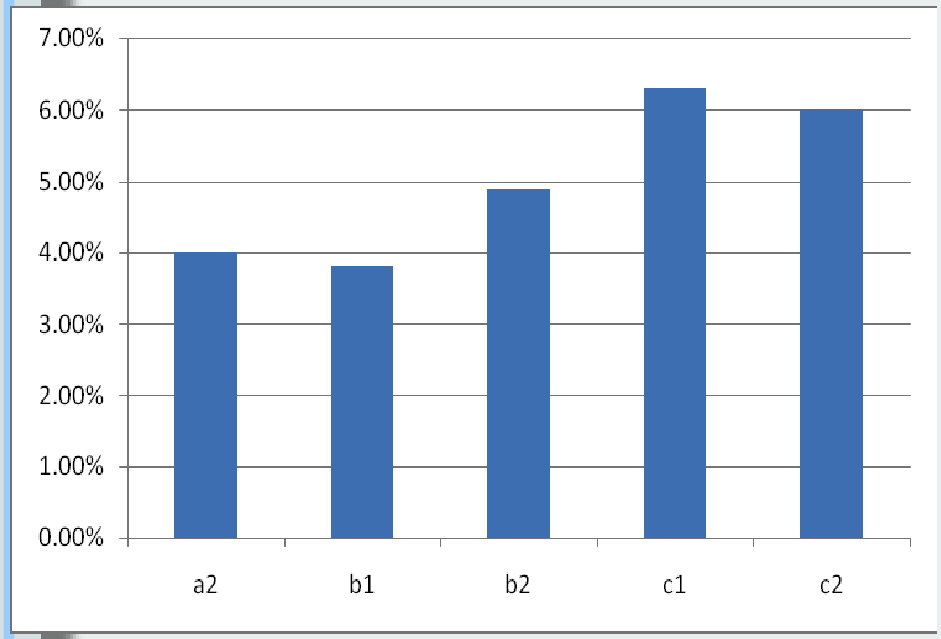
Texts which don't use passive (%)



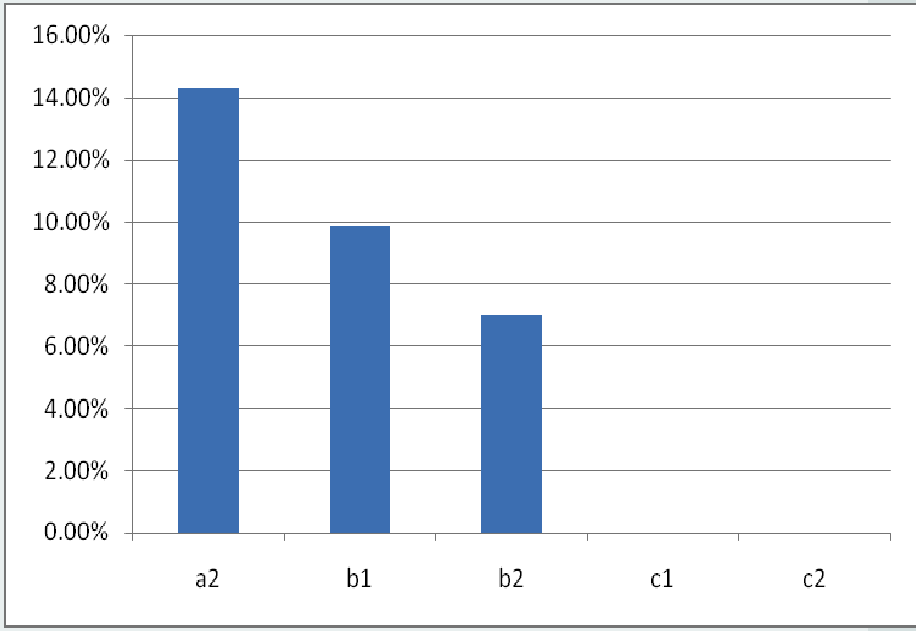
5. Extracting profiles (iii): Onset of Use

- Another Example: Use of Present-participle clauses:
  - *"He likes going to the zoo"*

Present participle clauses as % of all clauses

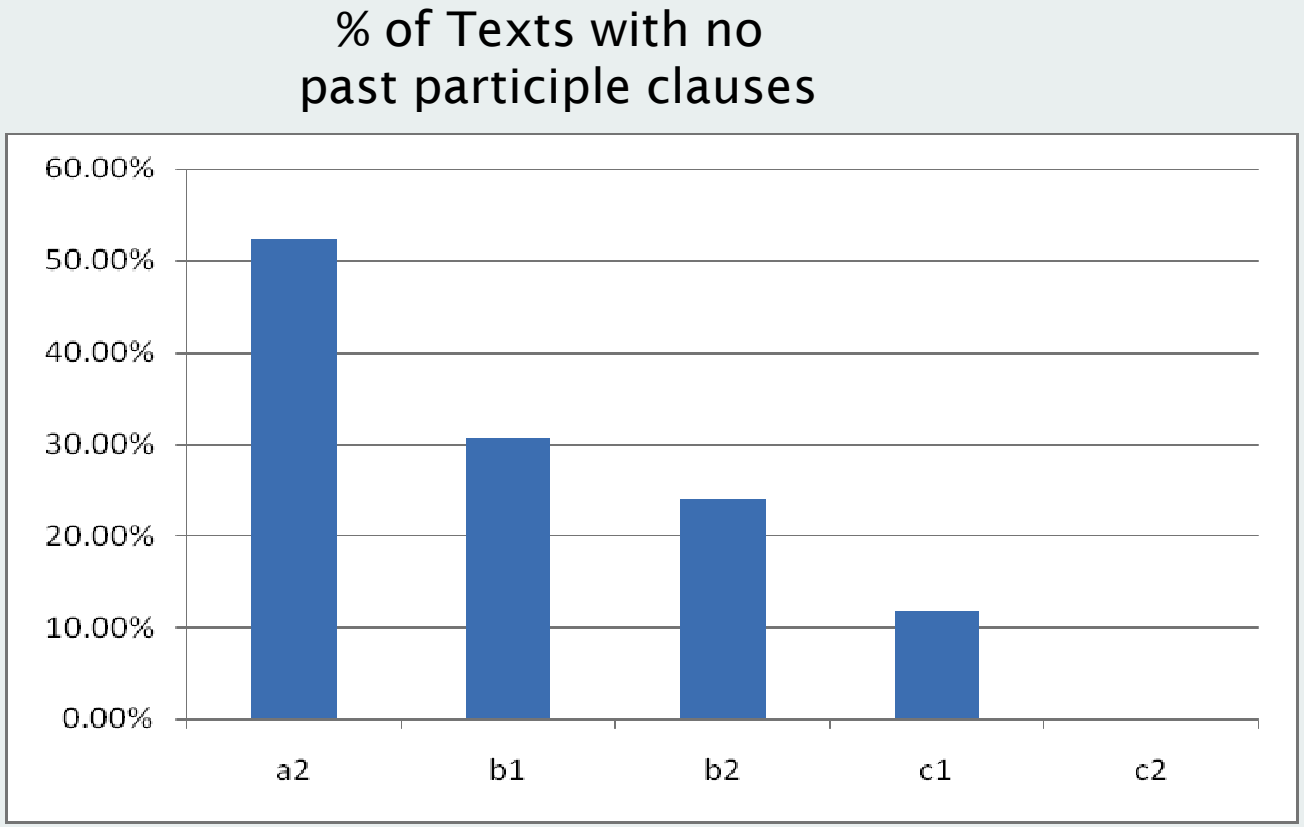


% of Texts with no present participle clauses



5. Extracting profiles (iii): Onset of Use

- Another Example: Use of Past-participle clauses:
  - *The man driven by hunger*
  - *Burnt by the sun, he marched on*



6. Conclusions for Curriculum design

- By analysing the degree of nonusage of each grammatical feature at each proficiency level, we can determine when the feature is most critical to the group as a whole
  - **When** the early adopters have started to use it
  - **Before** the cautious have started to use it
- Exactly where in this range a structure is best taught needs to be decided.
- Some flexibility good, to fit into a structured grammar teaching environment

6. Limitations

- Measuring “onset of use” of a feature requires a reasonable length of text per student.
- We have approx. 1000 words per essay.
- Fine for structures with native use in over 3% of clauses.
- For rarer structures (e.g., clefting), longer texts (or multiple texts by same student) needed to place critical proficiency level

6. Limitations

- We can determine at which proficiency level particular grammatical structures can most valuably be taught.
- But students in a class will be of mixed proficiency levels.
- Partial answers:
  - Curriculum designers can assume a particular target level for each class (e.g., assumed B1 level at university entrance)
  - Individual students provided with a sheet indicating their personal weaknesses, and where to find study resources on these issues.
  - Online teaching systems can target specific needs of each students based on their proficiency level.