

BALEAP PIM: EAP and Corpora
Coventry June 21 2014

Personal EAP Corpora:
What do Independent Users do?

Maggie Charles

Oxford University Language Centre, UK
maggie.charles@lang.ox.ac.uk

Context

Definitions

- **Personal corpora** are corpora built individually by students using research articles in their own field
- **Independent users** are students who use their personal corpus after the corpus course has finished

Outline

1. **Background:** the original corpus course
2. **Data for this study:** survey, participants, corpora
3. **Results and Conclusions**

The Corpus Course

Academic Writing Course for Graduates 2011/2012

- At Oxford University Language Centre
- Open-access, non-assessed course
- 6-7 parallel classes with 10 - 16 participants in each
- Multi-disciplinary, multi-national groups
- 1 weekly 2-hour session for 6 weeks in computer rooms
- Taught by 3 different tutors

Course Aims

- foster student autonomy
- provide a resource for future independent use (Charles 2012, 2014)

Corpus Competence of Students

Students

- built personal corpora from research articles in their own disciplines
- used their personal corpus
 - in class for discourse investigations
 - outside class for editing and revising written work
- used AntConc 3.2.4 (Anthony, 2011) for concordancing, clusters, collocations, plot, context searches, word list
- discussed and interpreted corpus data
- **Were 'corpus literate' becoming 'corpus proficient' (Charles 2011)**

Research Questions

- **To what extent do learners use their personal corpus after the course?**
- **What do independent users do?**
 - How frequently do they use the corpus?
 - Which tools and procedures do they use?
 - What search types do they perform?
- **Are there differences between frequent and infrequent users?**
- **What can we do in courses to encourage future independent use?**

Data for this Study

On-line survey

- 24 questions (e.g. corpus use, corpus and tools, search types, advantages/disadvantages)
- Sent 12 months after completion of course
- Sent to 127 students
 - 69 from 2011
 - 58 from 2012
- 72 replies received (57%) (2 incomplete)
 - 42 from 2011
 - 30 from 2012

Participants

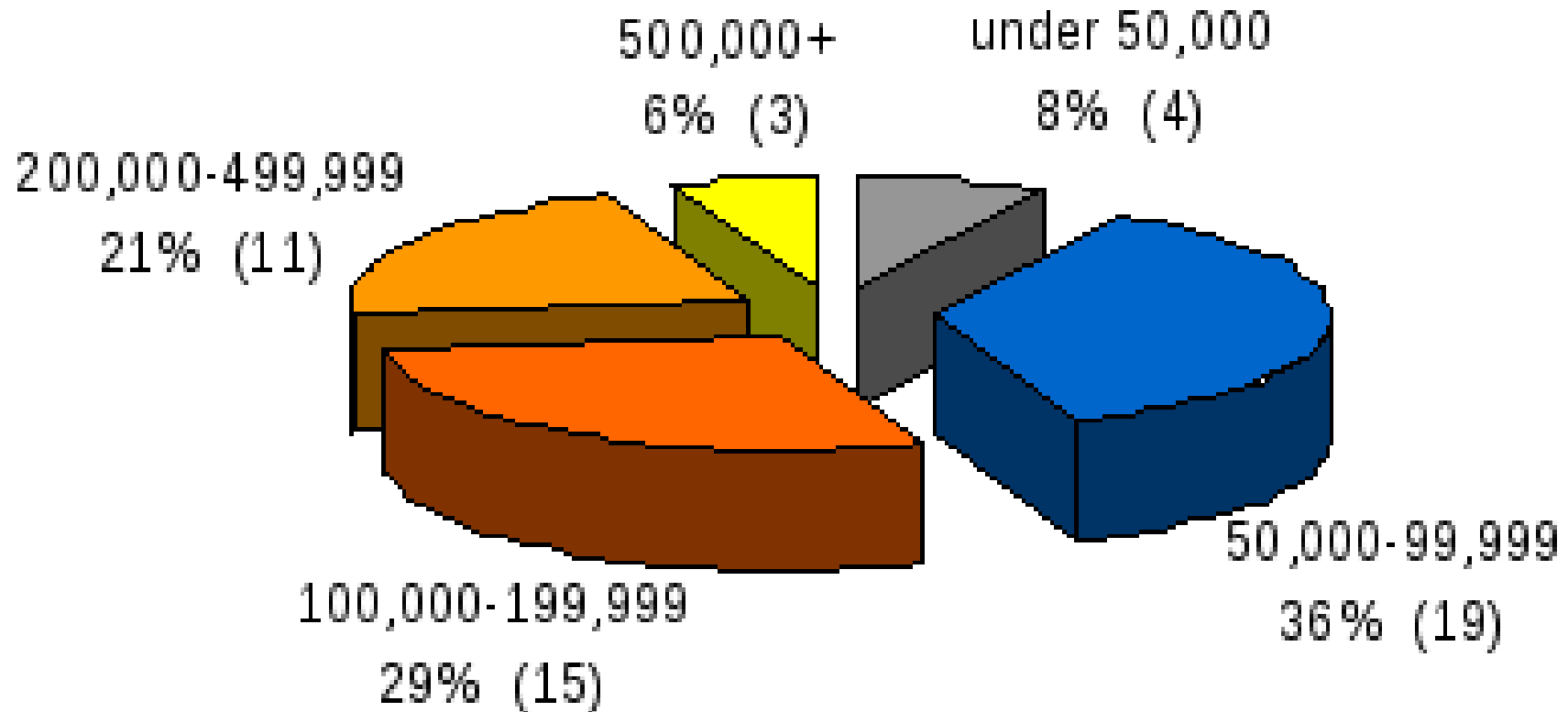
Degree level

- 47 Doctoral (65%)
- 16 Master's (22%)
- 7 Postdoc (10%)
- 2 Other graduates (3%)

Discipline

- 36 Different research fields
- 22 Social sciences (31%); 34 Natural sciences (47%); 16 Arts/humanities (22%)

Number of Words in Personal Corpora ($n = 52$)



Students' Personal Corpora

Number of Files (Research Articles) ($n = 65$)

Range: 5 to 200

Mean number of files per corpus: 23

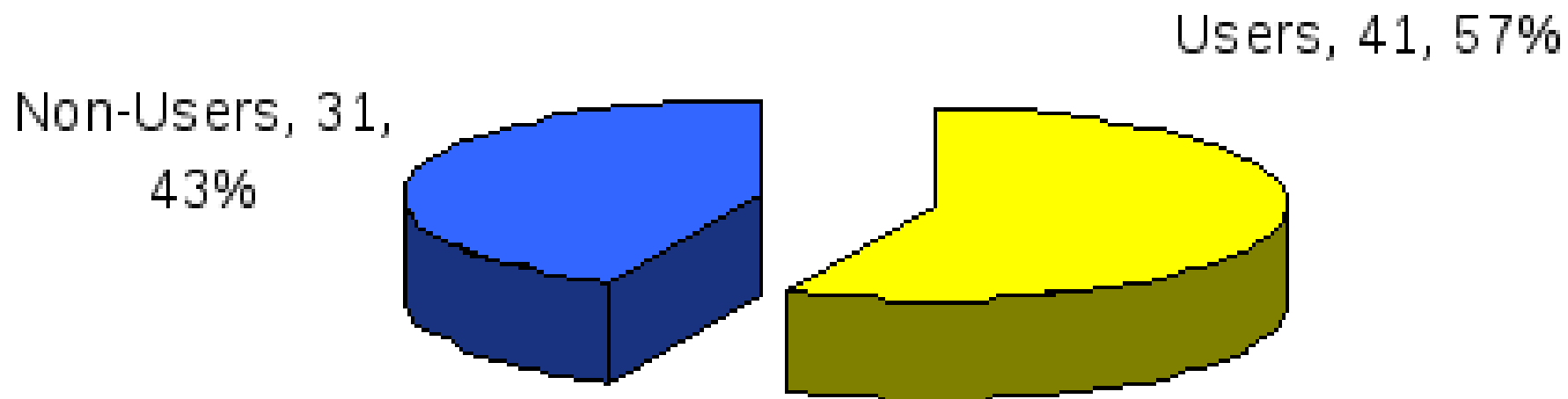
Number of Words ($n = 52$)

Range: 39,859 – 1,631,564

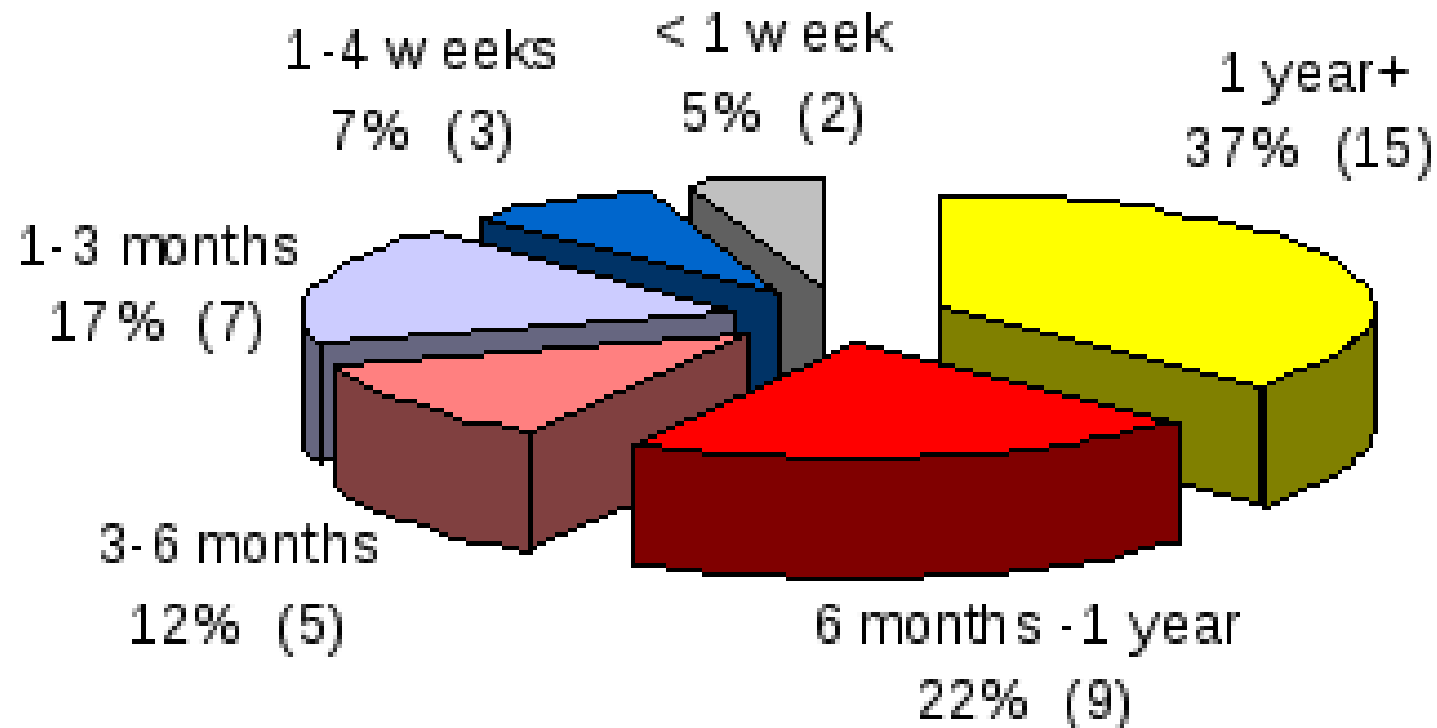
Mean number of words per corpus: 192,469

- **Small specialised corpora**
- **For editing and revising purposes**

Have you used your own corpus at any time since the course ended?

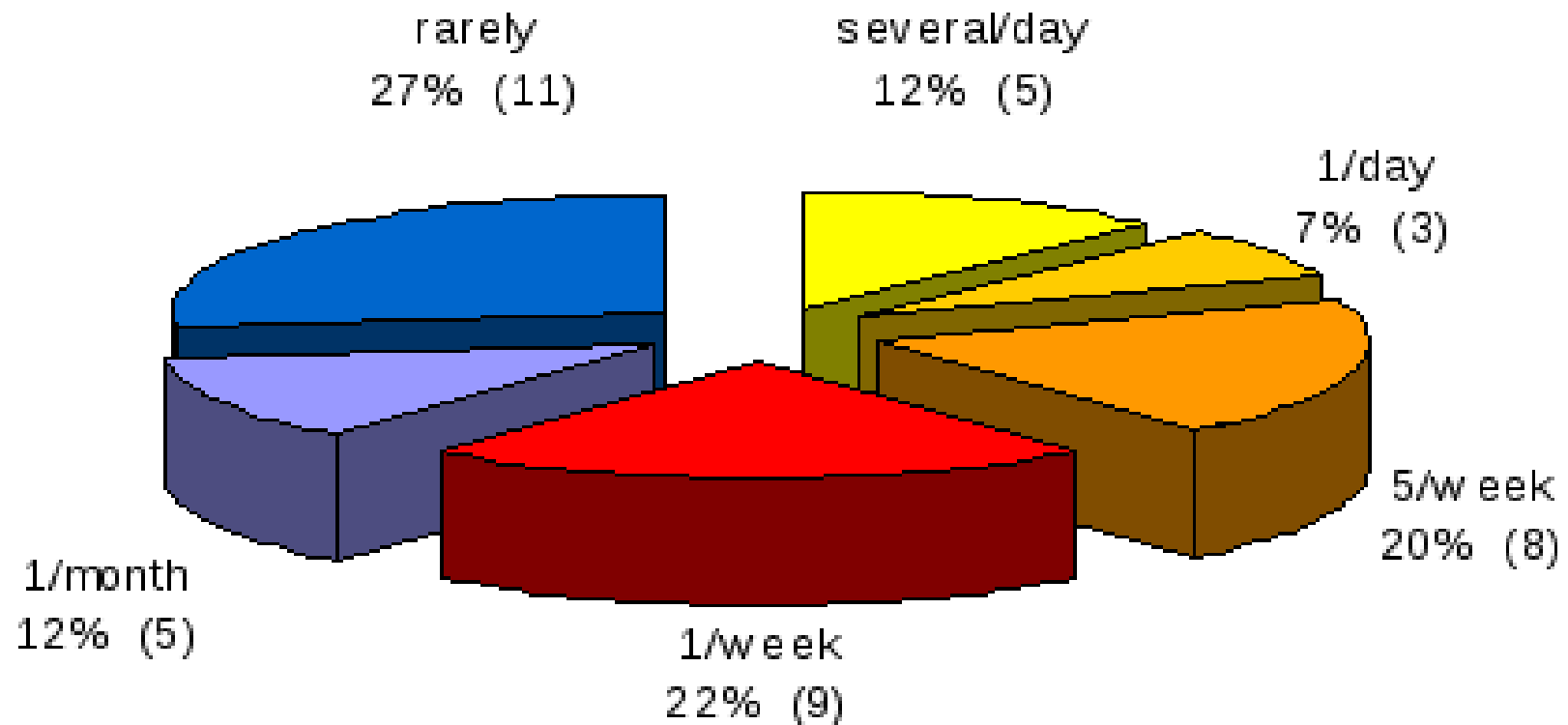


Duration of Personal Corpus Use ($n = 41$)

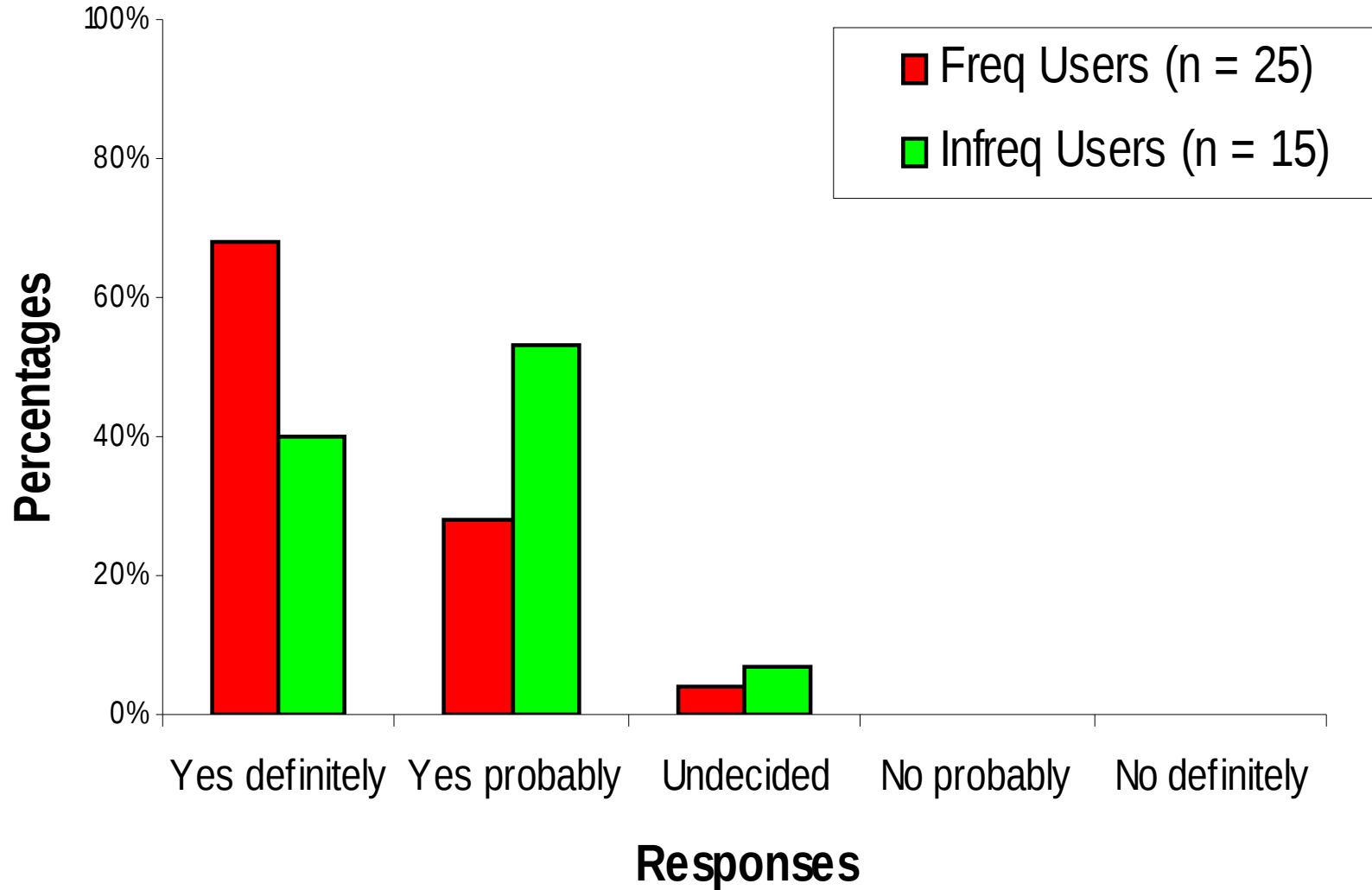


Frequency of Personal Corpus Use

(*n* = 41)



Improvement in Writing ($n = 40$)



Improvement in Writing

- **Frequent users are more confident of their improvement** (*Yes definitely* responses)

Frequent Users: 68% (17)

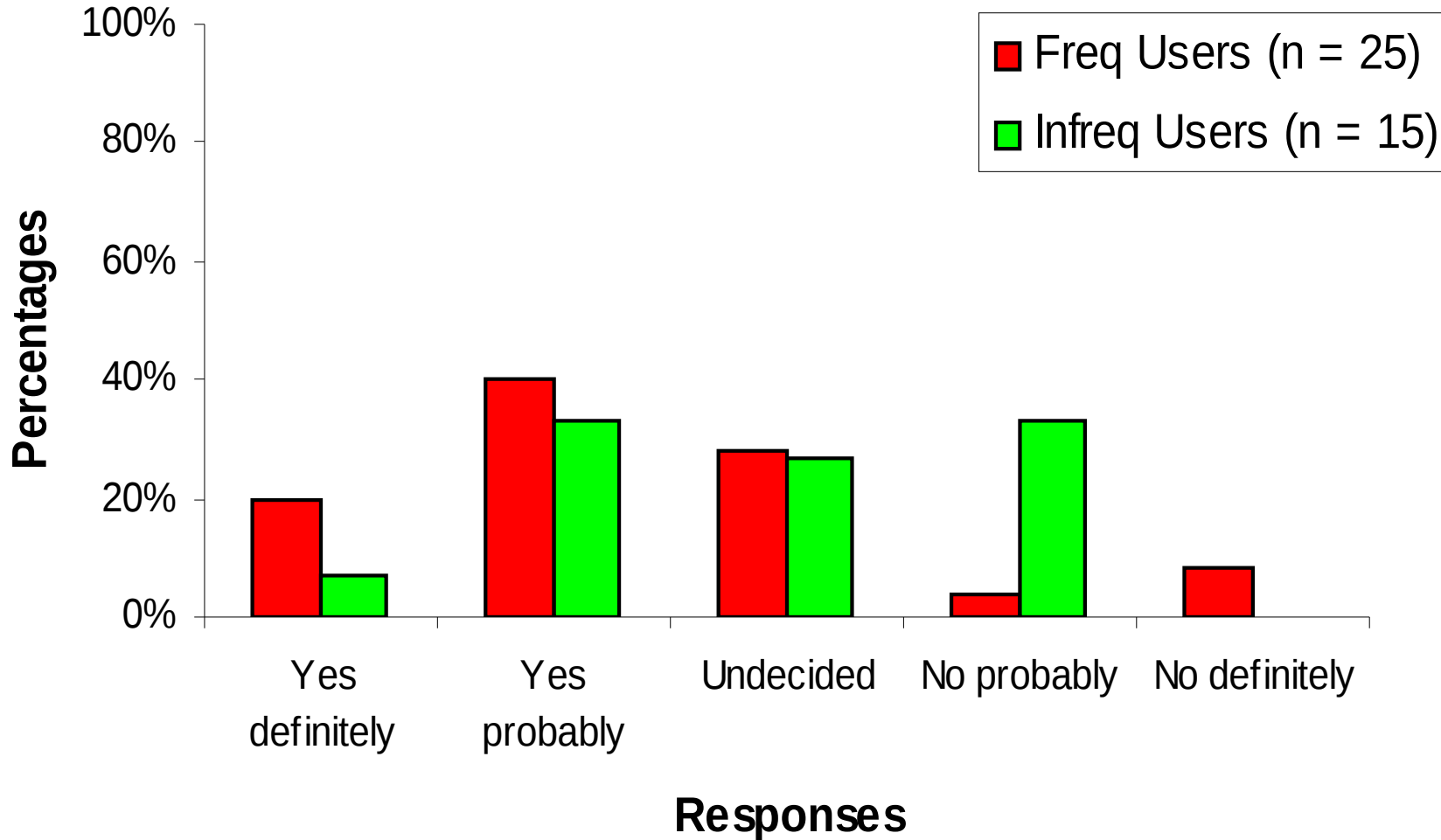
Infrequent Users: 40% (6)

- **Both sets of users equally likely to consider they have improved** (Total *Yes* responses)

Frequent Users: 96% (24)

Infrequent Users: 93% (14)

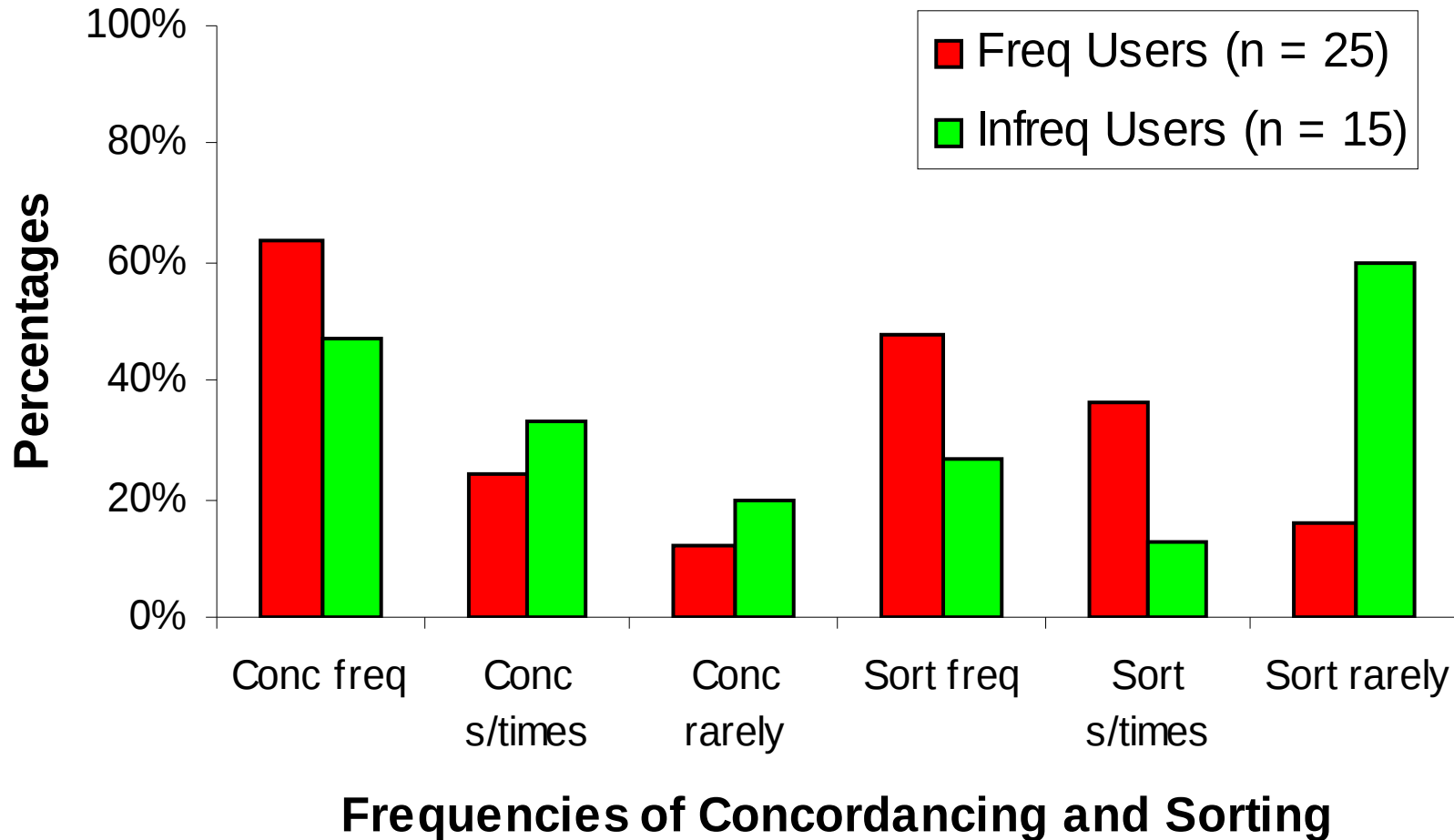
Improvement in Techniques ($n = 40$)



Improvement in Techniques

- **Frequent users are more confident of their improvement** (*Yes, definitely* responses)
Frequent Users: 20% (5)
Infrequent Users: 7% (1)
- **Frequent users more likely to consider they have improved** (Total *Yes* responses)
Frequent Users: 60% (15)
Infrequent Users: 40% (6)
- **Both sets of users more negative about improvement in techniques than improvement in writing**

Consulting and Sorting Concordance Lines ($n = 40$)



Consulting and Sorting Concordance Lines

- **Frequent users consult concordance lines more often**

Frequent Users: 64% (16)

Infrequent Users: 47% (7)

- **Frequent users sort concordance lines much more often**

Frequent Users: 48% (12)

Infrequent Users: 27% (4)

Two Search Types

1. 'Pattern-defining' (Called here Checking)

'to edit text for lexico-grammatical accuracy'

e.g. Is '*capable to do...*' correct?

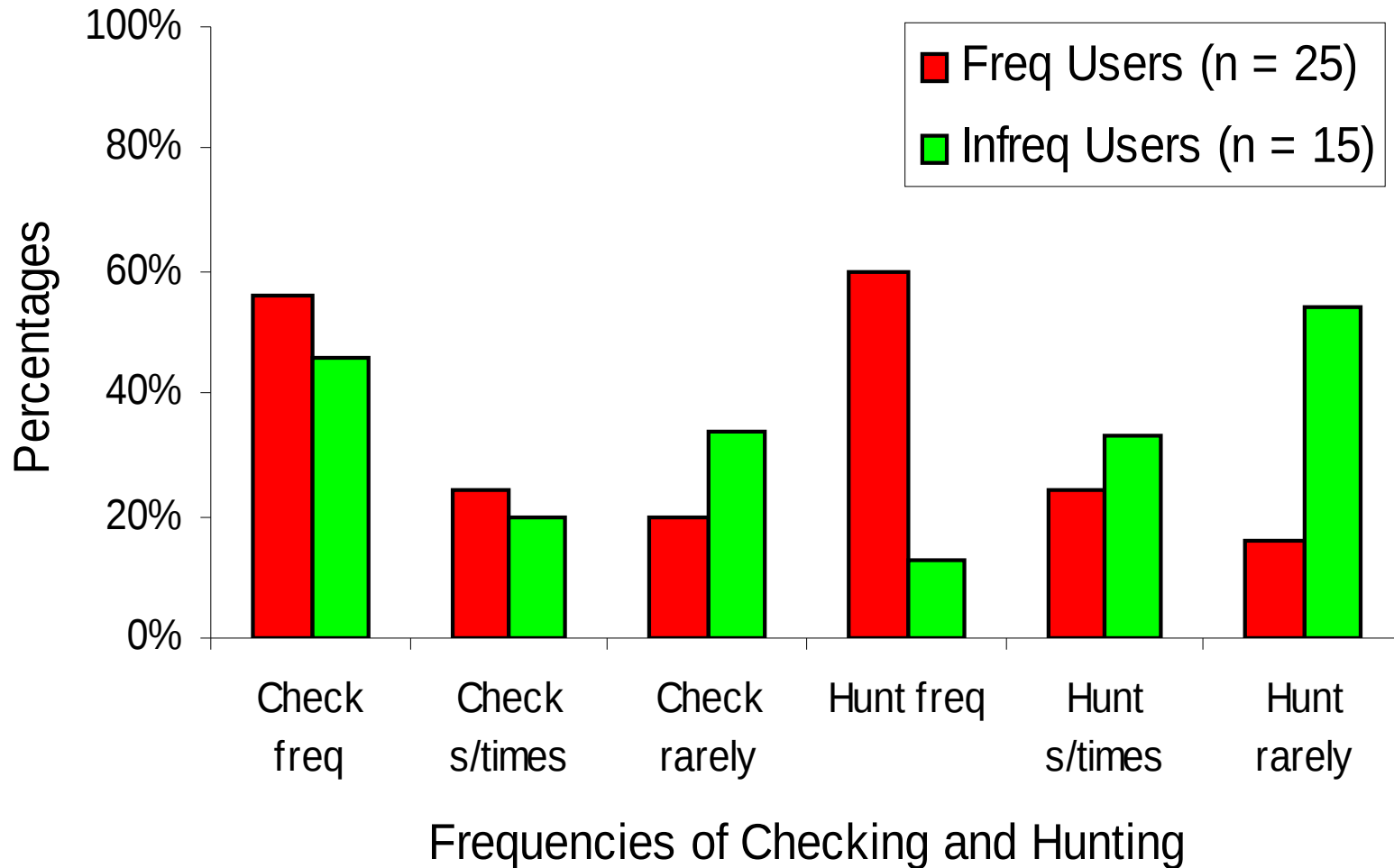
2. 'Pattern-Hunting' (Called here Hunting)

'to enrich the content and language of text'

e.g. How do writers in my field discuss future research?

Kennedy & Miceli (2010: 31)

Checking and Hunting: Comparative Frequencies ($n = 40$)



Checking and Hunting

- **Frequent users check for patterns slightly more often**

Frequent Users: 56% (14)

Infrequent Users: 46% (7)

- **Frequent users hunt for patterns much more often**

Frequent Users: 60% (15)

Infrequent Users: 20% (3)

Conclusions

Most independent learners

- use their personal corpus after the course
- consult their corpus frequently over an extended period

Frequent users

- are more confident about their improvement in writing and techniques
- perform more creative and complex corpus tasks

What can we do in courses?

- Devise (more) tasks for sorting and hunting
- Apply tasks to individual writing

References

- Anthony, L.** (2011), AntConc (Version 3.2.4) [Computer software]. Retrieved from: http://www.antlab.sci.waseda.ac.jp/antconc_index.html.
- Charles, M.** (2011), Using hands-on concordancing to teach rhetorical functions: Evaluation and implications for EAP writing classes. In *New Trends in Corpora and Language Learning*, A. Frankenberg-Garcia, L. Flowerdew, & G. Aston (eds.) 26-43. London: Continuum.
- Charles, M.** (2012), 'Proper vocabulary and juicy collocations': EAP students evaluate do-it-yourself corpus-building. *English for Specific Purposes* 31: 93-102.
- Charles, M.** (2014), Getting the corpus habit: EAP students' long-term use of personal corpora. *English for Specific Purposes* 35: 30-40.
- Kennedy, C., & Miceli, T.** (2010), Corpus-assisted creative writing: Introducing intermediate Italian learners to a corpus as a reference resource. *Language Learning and Technology* 14 (1): 28-44.
- Lee, D., & Swales, J.** (2006), A corpus-based EAP course for NNS doctoral students: moving from available specialized corpora to self-compiled corpora. *English for Specific Purposes* 25: 56-75.