International Perspectives on Corpora for Language Learning
University of Queensland, 7 Oct 2022

UNIVERSITY OF OXFORD

Student Autonomy and Data-driven Learning in English for Academic Purposes

Maggie Charles
University of Oxford
maggiecharles_oxford@yahoo.com

Outline

1. Student Autonomy and Data-driven Learning (DDL)
2. Student Autonomy in the DDL Classroom
3. Student Autonomy outside the DDL Classroom: A Study
4. Future directions


Student
Autonomy
and DDL

Autonomy and Foreign Language Learning (Holec, 1979)

"In the learning of languages, autonomy is… ***the ability to take charge of one's own learning."***

(Holec, 1979: 3)

Language Learner Autonomy

Little (2022: 64)

"*For me, 'language learner autonomy' denotes a teaching/learning dynamic in which* ***learners plan, implement, monitor and evaluate their own learning***."

Benson (2013: 840)

"*Autonomy is manifested in the form of autonomous language learning, which here refers to* ***learning practices involving learners' control over aspects of their learning*** *or, more broadly,* ***learning that takes place outside the context of formal instruction***."

Data-driven Learning (DDL)

Johns (1991b: 2)

*"…the **language learner** is also essentially a **research worker** whose learning needs to be driven by access to linguistic data – hence the name **'data-driven learning'"***

- *"…we need to provide adequate opportunities for **students to raise problems and queries"***
- ***"inductive strategies** developed **in** the classroom should be equally applicable **outside** the classroom."*
- *"Our aim is to help our students to survive and **learn by themselves…"*** Johns (1991a: 29)

Autonomy and DDL
Aston: An Early Proponent

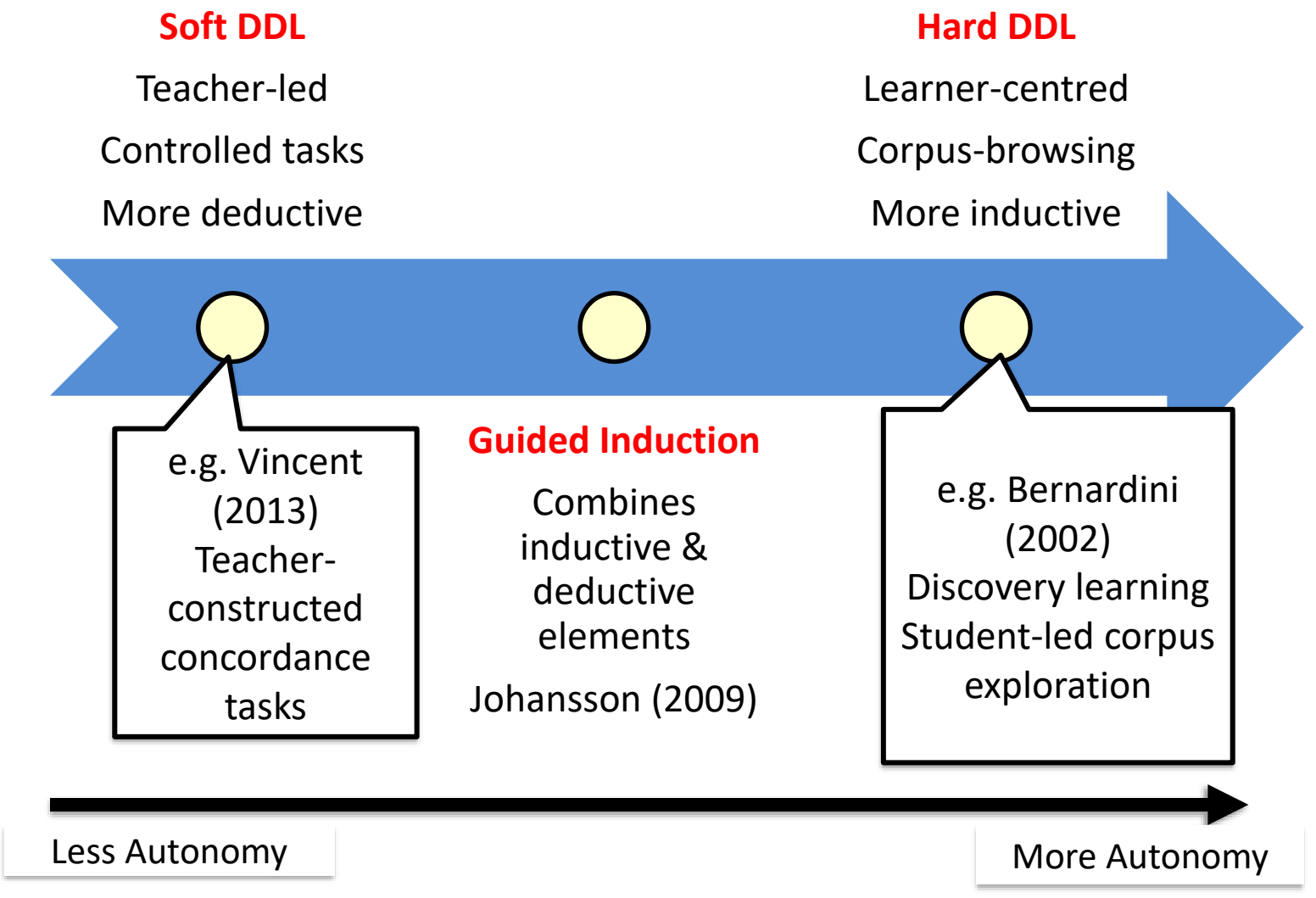
Aston (1997)
Involving learners in developing learning methods

Aston (2002)
The learner as corpus designer

Aston (2001: 41)
"…the greatest attraction of corpora for language pedagogy is ***their potential for autonomous learning"***

DDL on a Cline of Autonomy

(Gabrielatos, 2005; Mukherjee, 2006)



DDL in Autonomy Research

Reinders, 2022 Autonomy Bibliography
https://innovationinteaching.org/free-tools/autonomy-bibliography

Over 2,700 references

Search on "data-driven learning"
1 hit: Geluso & Yamaguchi (2014) EFL spoken fluency, **not EAP**

Search on "corpus" and "corpora"
5 hits
Kennedy & Miceli (2001, 2010) target language Italian, **not EAP**
St. John (2001) beginner level EFL, **not EAP**

Aston (1997) **EAP, interpreting & translation students**

Fan & Xu (2002) **EAP, legal English**

- Out of 2,700 references only 6 are related to DDL
- Of these only 2 are related to EAP
- Most references are to early work
- Only 1 reference in the last 10 years

What does this mean?

- That DDL/EAP research is not relevant to autonomy research and vice versa?
- That DDL/EAP research does not engage with autonomy research and vice versa?
- Or?

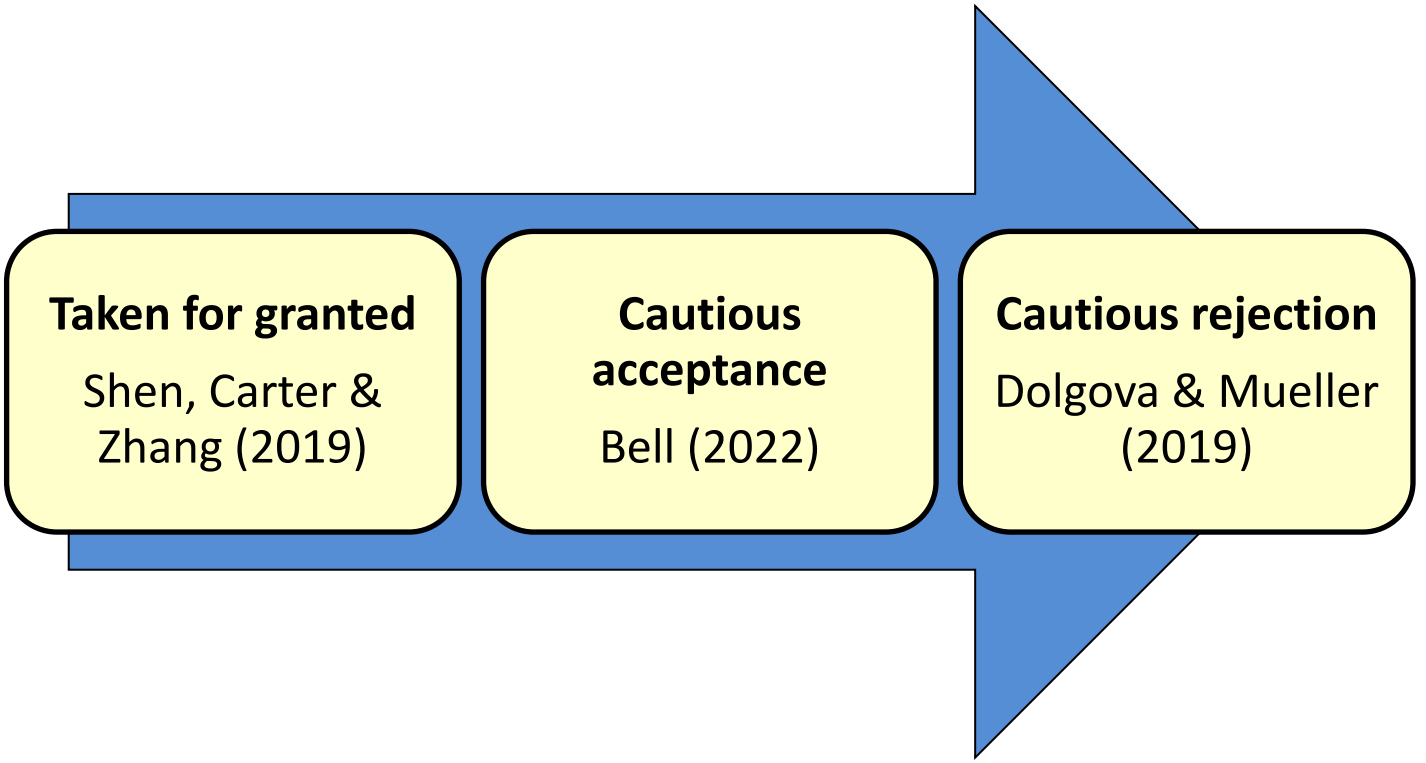
Autonomy, EAP and DDL

Shen, Carter & Zhang (2019: 84)
…a corpus based approach profits teaching academic writing with respect to enhancing learning input and **fostering students' autonomy** and noticing strategies.

Bell (2022: 6)
Encouraging students to become adept at **working things out for themselves** remains a **dominant methodological hallmark of the EAP classroom.** One particular manifestation of this can ***arguably*** be found in the practical applications of Data Driven Learning

Dolgova & Mueller (2019: 106)
…increasing one's reliance on **teacher-created corpus-informed materials may provide greater benefits** to learners than pedagogical materials or **approaches that rely solely on autonomous data-driven learning.**

Attitudes to Student Autonomy and DDL in EAP



Autonomy, EAP and DDL: Where are we now?

Bell (2022: 7)

"…most EAP teachers would probably recognize ***the promotion of learner autonomy as an inherent pedagogical principle of EAP."***

Chen and Flowerdew (2018: 357)

"Researchers and practitioners need to come up with ***ways of helping students to become more autonomous in their use of corpora."***

Boulton & Vyatkina (2021)

Such skills as critical thinking, independent learning, and **learner autonomy** feature prominently in article Conclusions as **alleged benefits from DDL**, but we found **virtually no direct exploration of these concepts** as research objectives.

While operationally defining these abstract constructs certainly is not straightforward, **we hope** that future innovative research will pursue this challenging yet most promising direction


Student
Autonomy in
the DDL
Classroom

Holec's 2 Conditions for Autonomy

1. *Learners must know how to make decisions*

Autonomy is an ability which can be developed

Consequences for Teachers

- Teachers can offer students opportunities for autonomy
- Teachers have to give up some control

2. *Learners must have the possibility of exercising autonomy*

Consequences for Institutions

- Institutions have to offer teachers opportunities for autonomy
- Institutions have to give up some control

The 3 P's and the 3(4) I's

Traditional Approach

The 3 P's

Presentation: teacher presents target language

Practice: students do controlled tasks

Production: students use target language in freer tasks

DDL Approach

The 3 I's

Illustration: students look at real data

Interaction: students discuss, share opinions and observations

Induction: students make their own rule for a particular feature

Carter & McCarthy (1995: 155)

Flowerdew (2009) adds **Intervention**: teacher provides clues and prompts where necessary

Stages in the development of learner autonomy ***in*** the classroom: Stages 1-3

Learning Stage (Reinders, 2010: 46)	DDL Research (Charles, 2022a)
1. Identifying needs	• Stages 1-3 tend to be teacher-directed at the class level • Scope for student autonomy at the individual level • Responsibility could be shared between teachers and learners
2. Setting goals	
3. Planning learning	

Stages in the development of learner autonomy ***in*** the classroom: Stages 4-5

Learning Stage (Reinders, 2010: 46)	DDL Research (Charles, 2022a)
4. Selecting resources	• Stages 4-5 offer increased scope for learner autonomy • e.g. Choice of resource (Bridle, 2019; Gilmore, 2008) • Compiling Do-It-Yourself (DIY) corpora (Charles, 2012; Lee & Swales, 2006)
5. Selecting learning strategies	

Stages in the development of learner autonomy *in* the classroom: Stage 6

Learning Stage (Reinders, 2010: 46)	**DDL Research** (Charles, 2022a)
6. Practice	• Many DDL studies • Experimental DDL e.g. Boulton (2012): Little scope for learner autonomy • Error correction e.g. Crosthwaite (2017): Scope for autonomy in identifying errors and formulating queries

Stages in the development of learner autonomy ***in*** the classroom: Stages 7-8

Learning Stage (Reinders, 2010: 46)	**DDL Research** (Charles, 2022a)
7. Monitoring progress	• Learner evaluation of DDL in many DDL studies e.g. (Ackerley, 2021; Yoon & Hirvela, 2004). Mizumoto *et al (*2015): a scale to measure perceived benefits & preferences of DDL • Little quantification of extent of autonomy. • Little evidence of learner input into revised courses
8. Assessment and revision	

So, how much student autonomy is there in the DDL Classroom?

Stages 1-3

- Little evidence. **Scope for increased student autonomy**

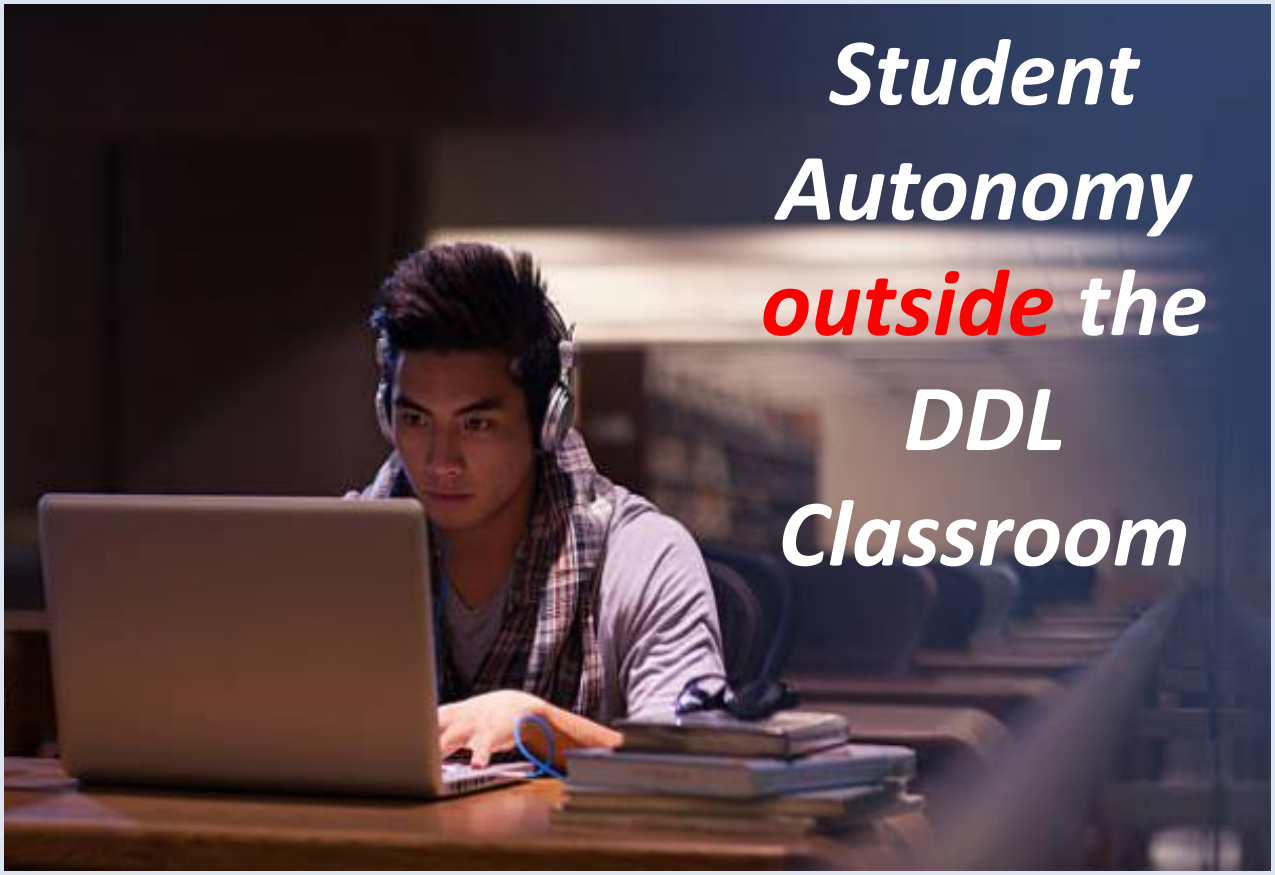
Stages 4-6

- Some evidence. Offering students a degree of choice in selecting resources and strategies & in the practice stage
  Increase possible

Stages 7-8

- Little evidence. **Scope for increased student autonomy**

- Emphasis on establishing the effectiveness of DDL has at the expense of attention to student autonomy.


Student
Autonomy
outside the
DDL
Classroom

Students' Intentions to use Corpora Autonomously

Yoon & Hirvela (2004)
88% will use corpus in **future** writing **(n=22)**

Charles (2012)
94% of students **intended to use their corpus** for help with their English in the **future (n=50)**

Karpenko-Seccombe (2018)
100% of respondents **declared their determination to use the software** in **future (n=30)**

Jablonkai & Čebron (2021)
78% will use a corpus for my English writing in the **future (n=9)**

***BUT* most numbers are small**

Can we rely on students' intentions?

Study Context (Charles, 2022b)

Course: 'Writing in your field with corpora'

- Part of academic writing programme
- 4-7 classes per year
- One 2-hour session/week
- 6 weeks

Picture by Ian Wallman © University of Oxford

Participants

- 182 students (2009-2017)
- 69% doctoral; 21% master's
- 46% STEM; 30% social sciences; 25% humanities
- 32 different L1s (30% Chinese, 10% Spanish, 7% Italian)
- Equal numbers of male & female students

Course Details

Approach

- Students built **Do-It-Yourself (DIY) corpora of research articles** in their fields (Lee & Swales, 2006)

Aims

- to explore **disciplinary discourse** functions in class
- to provide a **tailor-made corpus resource for autonomous use**

Software

- AntConc (Anthony, 2020)
- AntFileConverter (Anthony, 2017)

Data 1: Immediate post-course questionnaire

1. How often do you use your corpus outside class?

- ***several times a day, about once a day, about five times a week, about once a week*** — **Regular use**
- ***about once a month, seldom*** — **Occasional use**
- ***never*** — **Non-use**
- ***other*** — **Variable use**

2. I intend to use my corpus for help with my English in the future

strongly disagree – somewhat disagree – *neither agree nor disagree* – ***somewhat agree - strongly agree***

Data 2: Delayed post-course questionnaire

Completed one year after the end of the corpus course

1. **Have you used your own corpus at any time since the academic writing course ended?**

- *Yes/no*

- If 'yes', how often do/did you use your own corpus?
- Same options as immediate post-course questionnaire

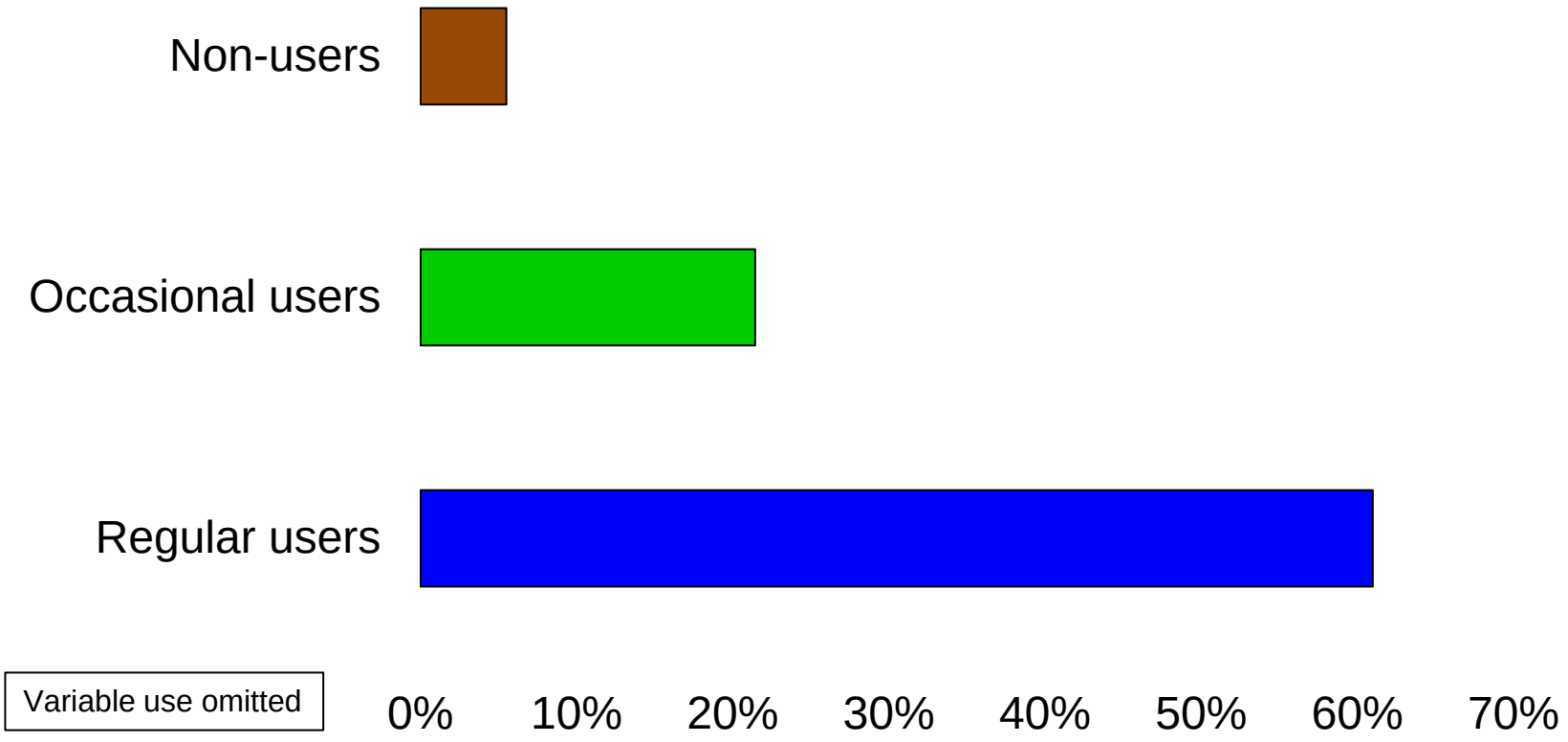
3 Research Questions

RQ1 What is the **extent** of autonomous corpus use immediately post-course?

RQ2 What **intentions** as to future corpus use are expressed by respondents immediately post-course?

RQ3 What is the **extent** of autonomous corpus use after 1 year (delayed post-course)

RQ1 The **extent** of autonomous corpus use immediately post-course (n = 182)

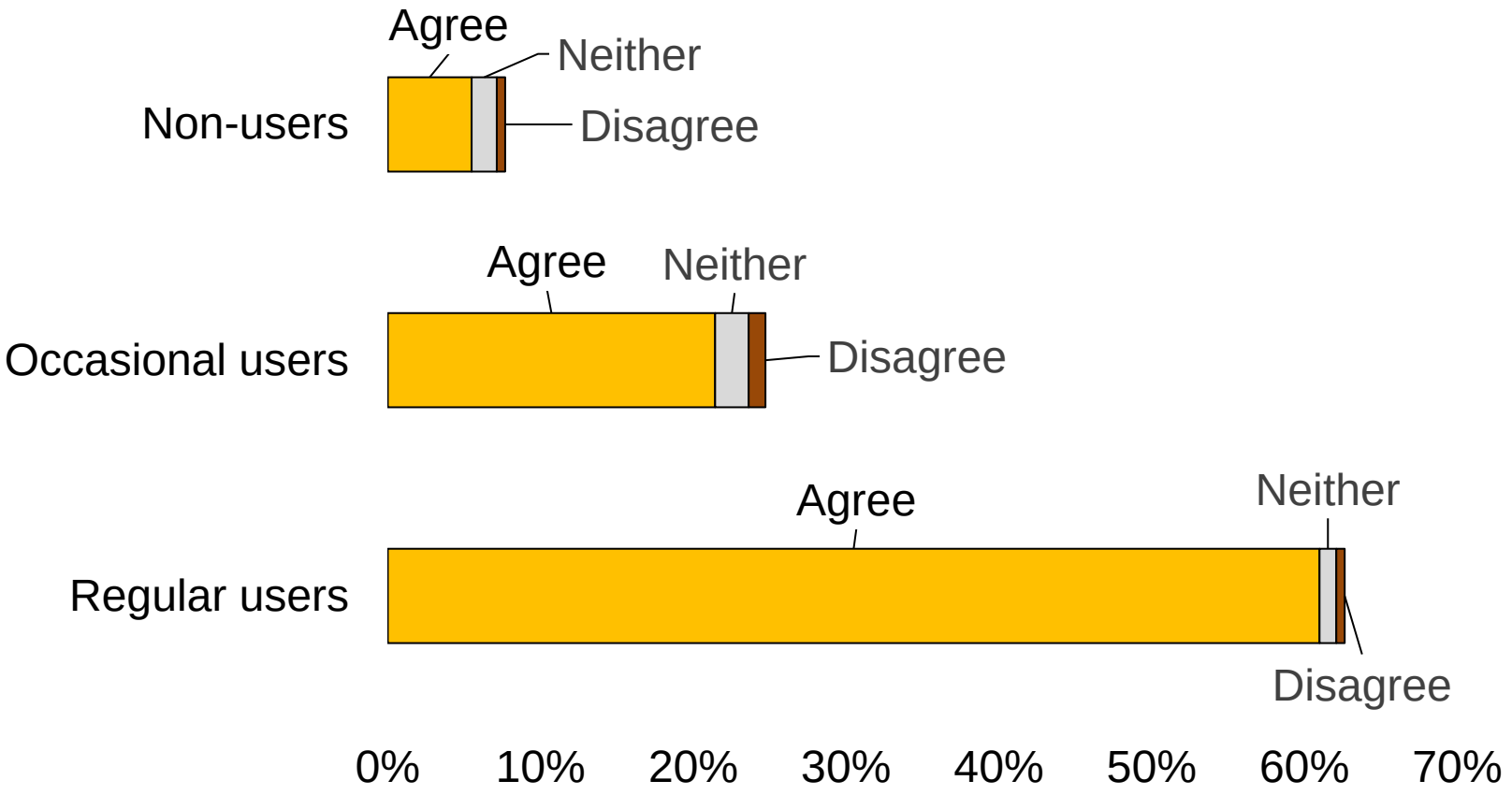


RQ1 The **extent** of autonomous corpus use immediately post-course

User Category	Percentage
Regular	63%
Occasional	25%
Non-use	8%

- **High regular** autonomous use of DDL
- DDL course seems to have been very **effective**

RQ2 **I intend** to use my corpus in future (Immediate post-course stage)

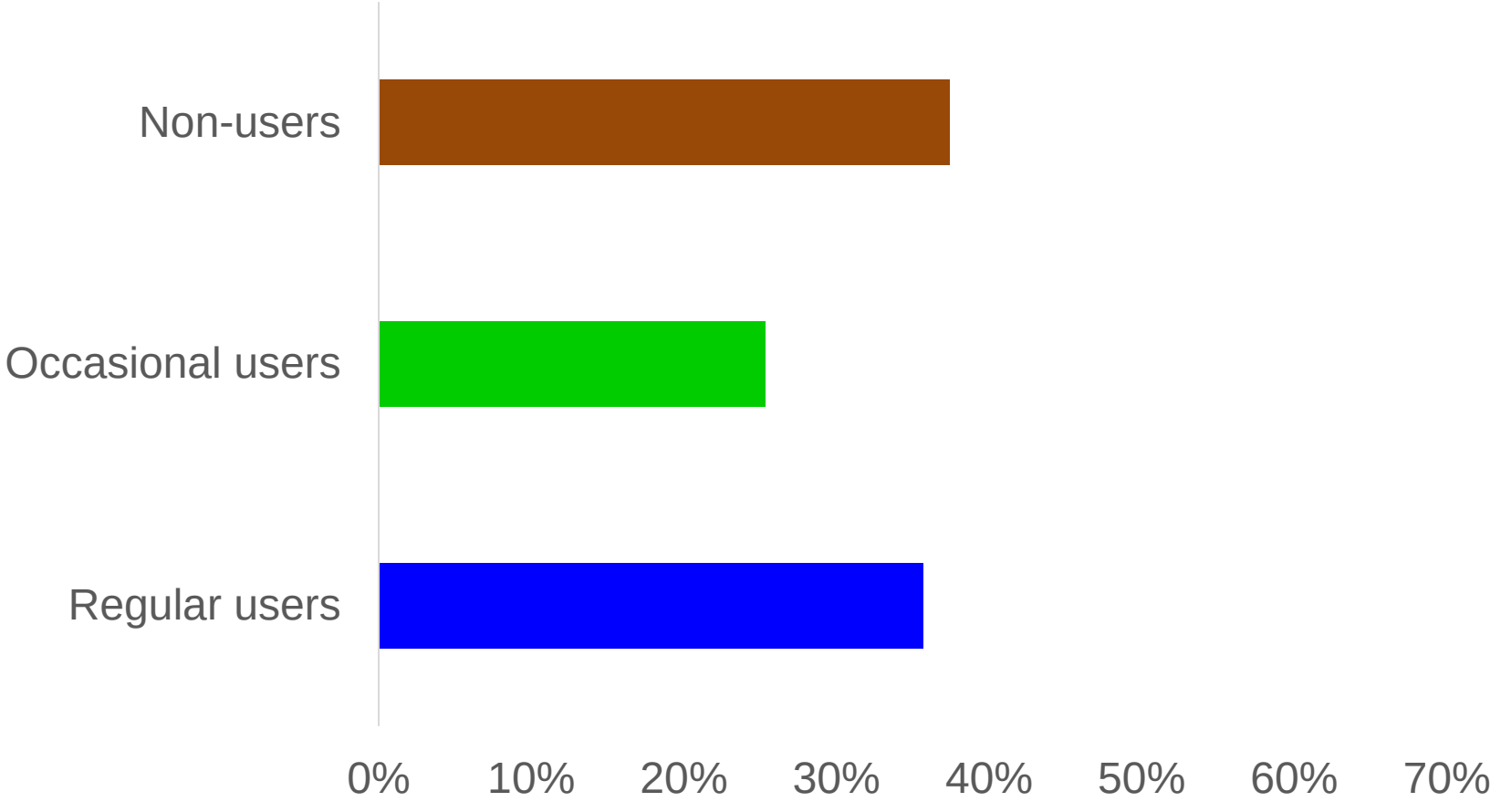


RQ2 Intention to use corpus immediately post-course

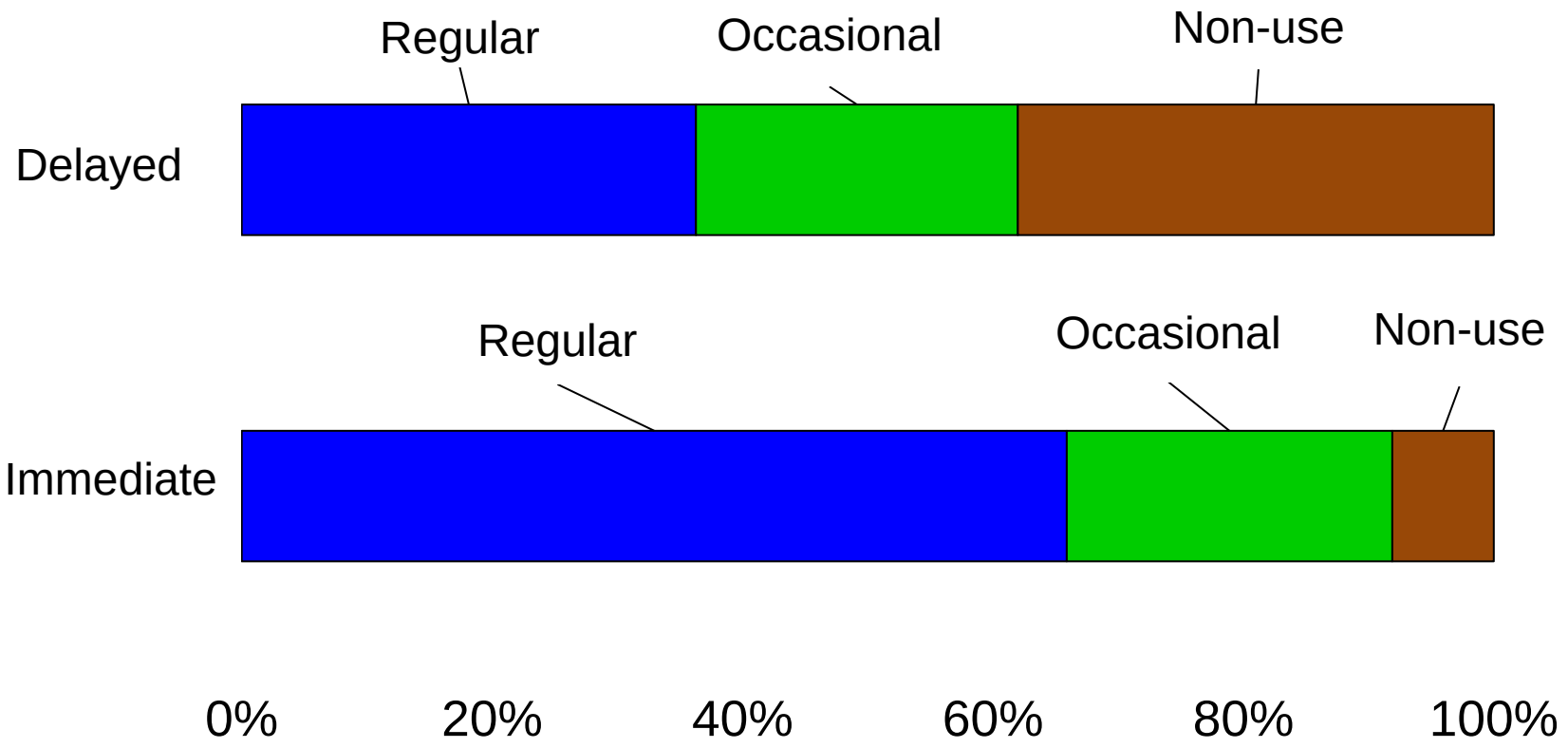
User Category	Percentage of 'Agree'
Regular	61%
Occasional	21%
Non-use	5%

- In each category of user, a **large majority** of students **intend to use** their corpus autonomously in the future
- In total, **87%** of all students **intend to use** their corpus autonomously in the future

RQ3 The **extent** of autonomous corpus use at delayed post-course



Comparison of autonomous use at immediate and delayed post-course stages



RQ3 The **extent** of autonomous corpus use at delayed post-course

User Category	Immediate Use	Intention to Use (Agree)	Delayed Use
Regular	63%	61%	36%
Occasional	25%	21%	25%
Non-use	8%	5%	37%

- **Large decrease in regular autonomous use**
- **Intentions do not match the reality of long term autonomous use**

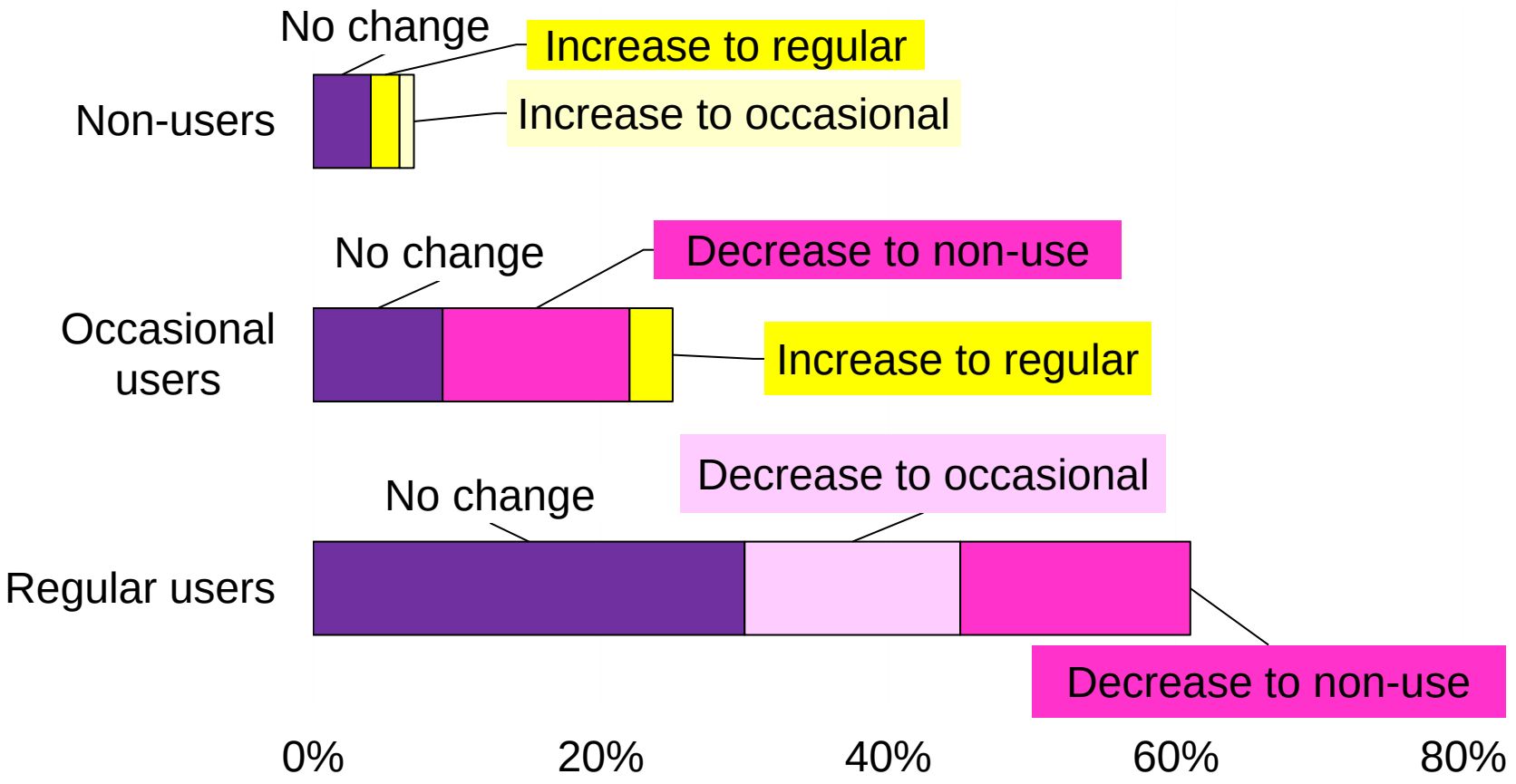
Regular Delayed Use

Another Question

User Category	Delayed Use
Regular	36%
Occasional	25%
Non-use	37%

- These are **aggregate** figures.
- How do they break down by each **category of user**?
- Once a regular autonomous user, always a regular autonomous user?

Changes in use for each category of autonomous user



Changes in Autonomous Use

50% of regular users remain **regular users** ☺

50% of non-users remain **non-users** ☹

25% of regular users become **non-users** ☹

Regular autonomous users do *NOT* always remain regular autonomous users

50% of occasional users become **non-users** ☹

Occasional users are likely to become non-users

What can be done?

Future Directions

Pedagogical Suggestions (1)

- **Help learners to turn intentions for autonomous use into reality**
  - By discussing **realistic goals** for autonomous DDL use with them
  - **Planning** autonomous DDL use together
  - Promoting **reflection** on DDL use
  - Encouraging them to **monitor** and **assess** their DDL use

Picture by Ian Wallman © University of Oxford

Pedagogical Suggestions (2)

- **Help occasional autonomous users to become regular autonomous users**

- **Increasing** amount of **student corpus practice** in class
- Relating **in-class** practice closely to **autonomous use**
- Increasing **engagement with teachers** (pre-service and in-service)

Picture by Ian Wallman © University of Oxford

Possible Research Directions

More research needed on

- **Developing student autonomy through DDL**
  - How to best help students

    - identify **needs**, set **goals**, **plan** DDL learning?

    - monitor **progress?**
  - What role can **local learner corpora** play?

Possible Research Directions

More research needed on

- **Assessing the extent and effects of autonomy**
  - How to **quantify** in-class autonomy?
  - How to **assess** the effects of autonomy?
  - Does **greater autonomy** lead to **proficiency gains?**

Thank you!

Any questions?

References

Ackerley, K. (2021). Exploiting a genre-specific corpus in ESP writing: Students' preferences and strategies. In M. Charles & A. Frankenberg-Garcia (Eds.), *Corpora in ESP/EAP writing instruction: Preparation, exploitation, analysis* (pp. 78–99). Routledge.

Anthony, L., (2017). AntFileConverter (1.2.1). [computer program] Tokyo, Japan: Waseda University. Available at: www.laurenceanthony.net/

Anthony, L. (2020). AntConc (Version 3.5.9) [Computer Software]. Tokyo, Japan: Waseda University. Available at www.laurenceanthony.net/

Aston, G. (1997). Involving learners in developing learning methods: Exploiting text corpora in self-access. In P. Benson & P. Voller (Eds.), *Autonomy and independence in language learning* (pp. 204–214). Pearson Education.

Aston, G. (2001). Learning with corpora: An overview. In G. Aston (Ed.), *Learning with corpora* (pp. 7–45). CLUEB.

Aston, G. (2002). The learner as corpus designer. In B. Kettemann & G. Marko (Eds.), *Teaching and learning by doing corpus analysis* (pp. 9–25). Rodopi.

Bell, D. E. (2022). Methodology in EAP: Why is it largely still an overlooked issue? *Journal of English for Academic Purposes*, *55*, 101073. doi:10.1016/j.jeap.2021.101073

Benson, P. (2013). Learner Autonomy. *TESOL Quarterly*, *47*(4), 839–843. doi:10.1002/tesq.134

Bernardini, S. (2002). Exploring new directions for discovery learning. In B. Kettemann & G. Marko (Eds.), *Teaching and learning by doing corpus analysis* (pp. 165–182). Rodopi. doi:10.1163/9789004334236_015

Boulton, A. (2012). Hands-on/hands-off: Alternative approaches to data-driven learning. In J. Thomas & A. Boulton (Eds.), *Input, process and product: Developments in teaching and language corpora* (pp. 152–168). Masaryk University Press.

Boulton, A., & Vyatkina, N. (2021). Thirty years of data-driven learning: Taking stock and charting new directions over time. *Language Learning and Technology*, *25*(3), 66–89.

Bridle, M. (2019). Learner use of a corpus as a reference tool in error correction: Factors influencing consultation and success. *Journal of English for Academic Purposes*, *37*, 52–69. doi:10.1016/j.jeap.2018.11.003

Carter, R., & McCarthy, M. (1995). Grammar and the spoken language. *Applied Linguistics*, *16*(2), 141–158. doi:10.1093/applin/16.2.141

Charles, M. (2012). 'Proper vocabulary and juicy collocations': EAP students evaluate do-it-yourself corpus-building. *English for Specific Purposes*, *31*(2), 93–102. doi:10.1016/j.esp.2011.12.003

Charles, M. (2018). Corpus-assisted editing for doctoral students: More than just concordancing. *Journal of English for Academic Purposes*, *36*, 15–25. doi:10.1016/j.jeap.2018.08.003

Charles, M. (2022a). Corpora and autonomous language learning. In R. Jablonkai & E. Csomay (Eds.), *The Routledge handbook of corpora in English language teaching and learning*. Routledge.

Charles, M. (2022b). The gap between intentions and reality: Reasons for EAP writers' non-use of corpora. *Applied Corpus Linguistics,* Special Issue in honour of Dr Randi Reppen. 100032. doi:10.1016/j.acorp.2022.100032.

Charles, M., & Hadley, G. (2022). Autonomous corpus use by graduate students: A long-term trend study (2009–2017). *Journal of English for Academic Purposes*, *56*, 101095. doi:10.1016/j.jeap.2022.101095

Chen, M., & Flowerdew, J. (2018). A critical review of research and practice in data-driven learning (DDL) in the academic writing classroom. *International Journal of Corpus Linguistics*, *23*(3), 335–369. doi:10.1075/ijcl.16130.che

Crosthwaite, P. (2017). Retesting the limits of data-driven learning: Feedback and error correction. *Computer Assisted Language Learning*, *30*(6), 447–473. doi:10.1080/09588221.2017.1312462

Dolgova, N., & Mueller, C. (2019). How useful are corpus tools for error correction? Insights from learner data. *Journal of English for Academic Purposes*, *39*, 97–108.

Fan, M. & Xu, X. F. (2002). An evaluation of an online bilingual corpus for the self-learning of legal English. *System,* 30(1), 47-63.

Flowerdew, L. (2009). Applying corpus linguistics to pedagogy. *International Journal of Corpus Linguistics*, *14*(3), 393–417. doi:10.1075/ijcl.14.3.05flo

Gabrielatos, C. (2005). Corpora and language teaching: Just a fling or wedding bells? *TESL-EJ*, *8*(4), 1–35.

Geluso, J. & Yamaguchi, A. (2014). Discovering formulaic language through data-driven learning: Student attitudes and efficacy. *ReCALL*, *26*(2), 225–242. doi:10.1017/S0958344014000044

Gilmore, A. (2008). Using online corpora to develop students' writing skills. *ELT Journal*, *63*(4), 363–372. doi:10.1093/elt/ccn056

Holec, H. (1979). *Autonomy and foreign language learning*. Pergamon Press.

Jablonkai, R., & Čebron, N. (2021). Undergraduate students' responses to a corpus-based ESP course with DIY corpora. In M. Charles & A. Frankenberg-Garcia (Eds.), *Corpora in ESP/EAP writing instruction: Preparation, exploitation, analysis* (pp. 100–120). Routledge.

doi:10.4324/9781003001966-5-8

Johansson, S. (2009). Some thoughts on corpora and second-language acquisition. In K. Aijmer (Ed.), *Corpora and language teaching* (pp. 34–44). Benjamins. doi:10.1075/scl.33.05joh

Johns, T. (1991a). From printout to handout: Grammar and vocabulary teaching in the context of data-driven learning. In T. Johns & P. King (Eds.), *Classroom concordancing* (pp. 27–37). University of Birmingham. doi:10.1017/cbo9781139524605.014

Johns, T. (1991b). Should you be persuaded: Two samples of data-driven learning materials. In T. Johns & P. King (Eds.), *Classroom concordancing* (pp. 1–16). University of Birmingham.

Karpenko-Seccombe, T. (2018). Practical concordancing for upper-intermediate and advanced academic writing: Ready-to-use teaching and learning materials. *Journal of English for Academic Purposes*, *36*, 135–141. doi:10.1016/j.jeap.2018.10.001

Kennedy, C. & Miceli, T. (2001). An evaluation of intermediate students' approaches to corpus investigation. *Language Learning and Technology,* 5(3), 77-90.

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*(1), 56–75. doi:10.1016/j.esp.2005.02.010

Little, D. (2022). Language learner autonomy: Rethinking language teaching. *Language Teaching*, *55*(1), 64–73. doi: 10.1017/S0261444820000488

Mizumoto, A., Chujo, K., & Yokota, K. (2015). Development of a scale to measure learners' perceived preferences and benefits of data-driven learning. *ReCALL*, *28*(2), 227–246. doi:10.1017/S0958344015000208

Mukherjee, J. (2006). Corpus linguistics and language pedagogy: The state of the art – and beyond. In S. Braun, K. Kohn, & J. Mukherjee (Eds.), *Corpus technology and language pedagogy* (pp. 5–24). Peter Lang.

Reinders, H. (2010). Towards a classroom pedagogy for learner autonomy: A framework of independent language learning skills. *Australian Journal of Teacher Education*, *35*(5), 39–55. doi:10.14221/ajte.2010v35n5.4

Shen, L., Carter, S., & Zhang, L. J. (2019). EL1 and EL2 doctoral students' experience in writing the discussion section: A needs analysis. *Journal of English for Academic Purposes*, *40*, 74–86. doi:10.1016/j.jeap.2019.06.004

St. John, E. (2001). A case for using a parallel corpus and concordancer for beginners of a foreign language. *Language Learning and Technology,* 5(3), 185-203.

Vincent, B. (2013). Investigating academic phraseology through combinations of very frequent words: A methodological exploration. *Journal of English for Academic Purposes*, *12*(1), 44–56. doi:10.1016/j.jeap.2012.11.007

Yoon, H., & Hirvela, A. (2004). ESL student attitudes toward corpus use in L2 writing. *Journal of Second Language Writing*, *13*(4), 257–283. doi:10.1016/j.jslw.2004.06.002