Evaluating the effect of data-driven learning (DDL) on the acquisition of academic collocations by advanced Chinese learners of English

Tanjun Liu (t.liu2@lancaster.ac.uk)

Lancaster University, UK

Lancaster University

CASS Corpus Approaches to Social Science

LITERATURE REVIEW

- Collocations are considered to be a crucial component in the language use and a challenge to L2 learners of all proficiency levels (e.g. Granger & Bestgen, 2014; Schmitt, 2010).
- Data-driven learning (DDL) approach: effective in language learning; but has so far not become one of mainstream teaching practice (e.g. Boulton 2010; Tribble, 2015).
- Large-scale, quantitative studies evaluating the effectiveness and assessed the benefits of DDL in the acquisition of academic collocations are still limited in number when compared to a different method of teaching of collocations; but DDL-based results overall promising (Boulton & Cobb, 2017).

RESEARCH QUESTIONS AND RESEARCH DESIGN

Research aim: This study seeks to evaluate the effect of corpus-based tools (#LancsBox and OCD) on the acquisition of academic collocations by advanced Chinese learners of English.

Research questions:

1. What is the effect of using #LancsBox (inductive DDL) and online Oxford Collocations Dictionary (OCD) (online dictionary) on the acquisition of academic collocations?
2. Do corpus-based tools (#LancsBox and online OCD) have an effect on learners' understanding of academic collocations?
3. What are learners' attitudes towards using #LancsBox (inductive DDL) and online OCD (online dictionary)?

Research design:

➢ **Participants:** 27 third-year English-major undergraduates from a mid-east Chinese university (14 learners in DDL group; 13 in OCD group).
➢ **Tools:** #LancsBox (GraphColl module) Version 2.0 (Brezina, McEnery & Wattam, 2015) and online Oxford Collocations Dictionary (OCD)
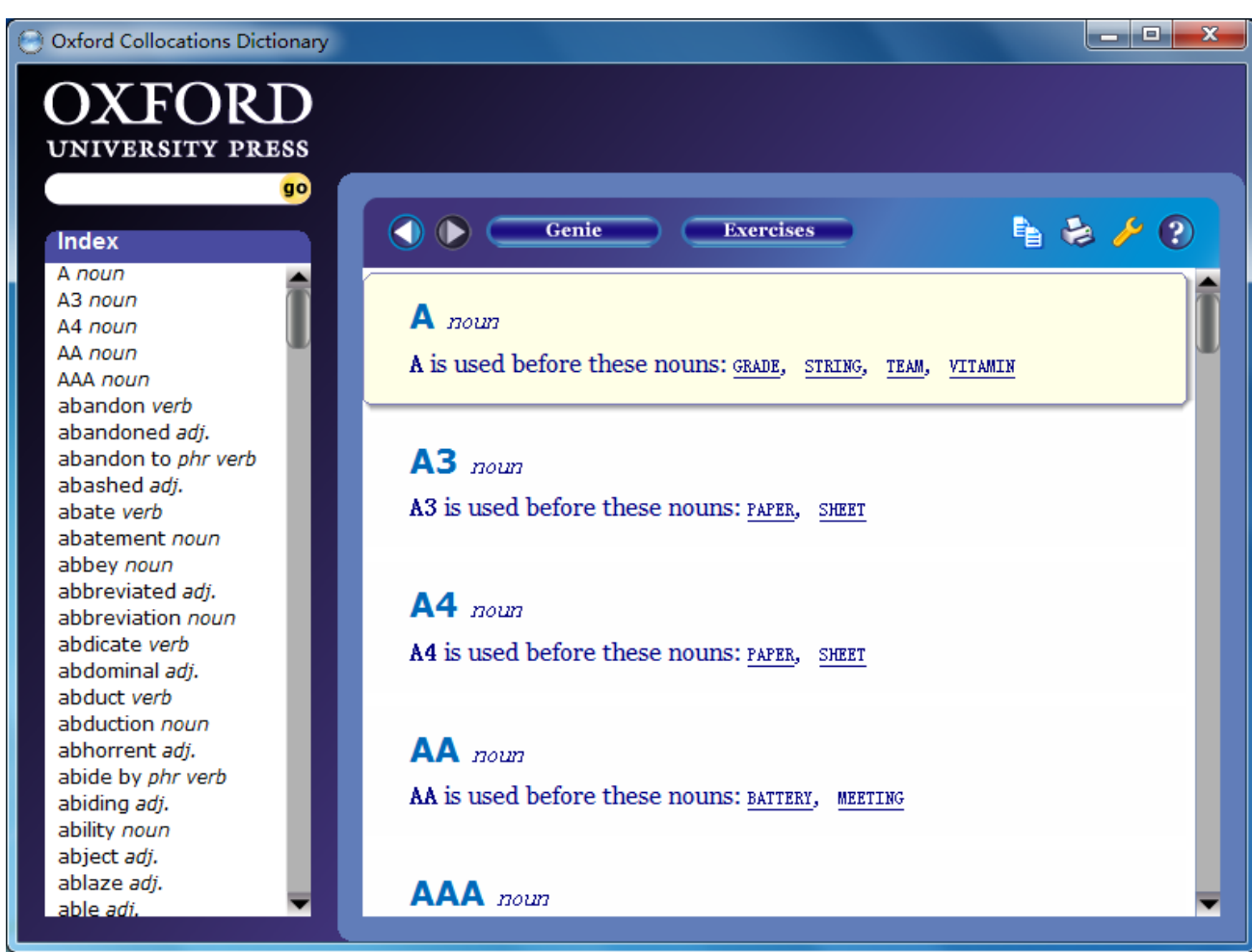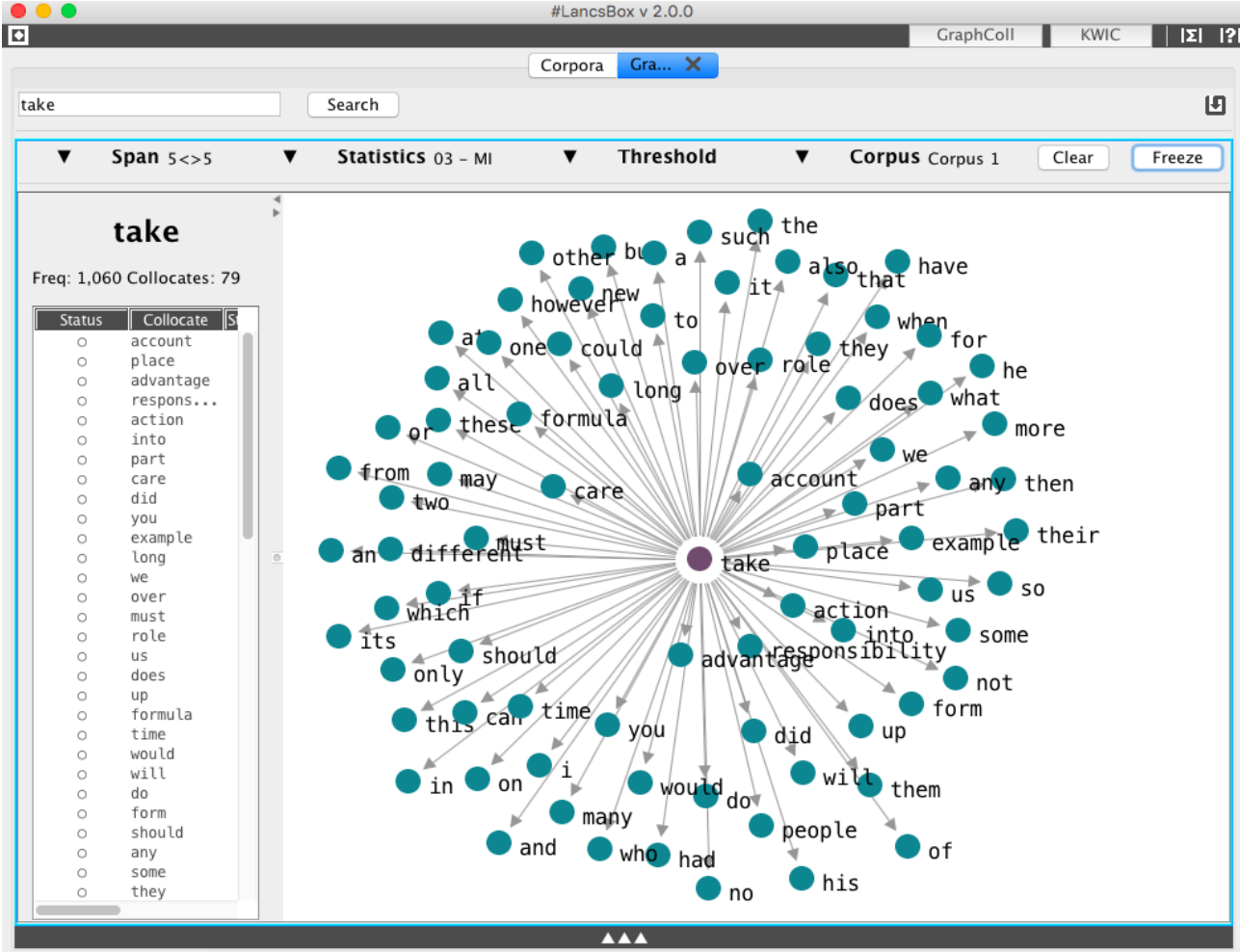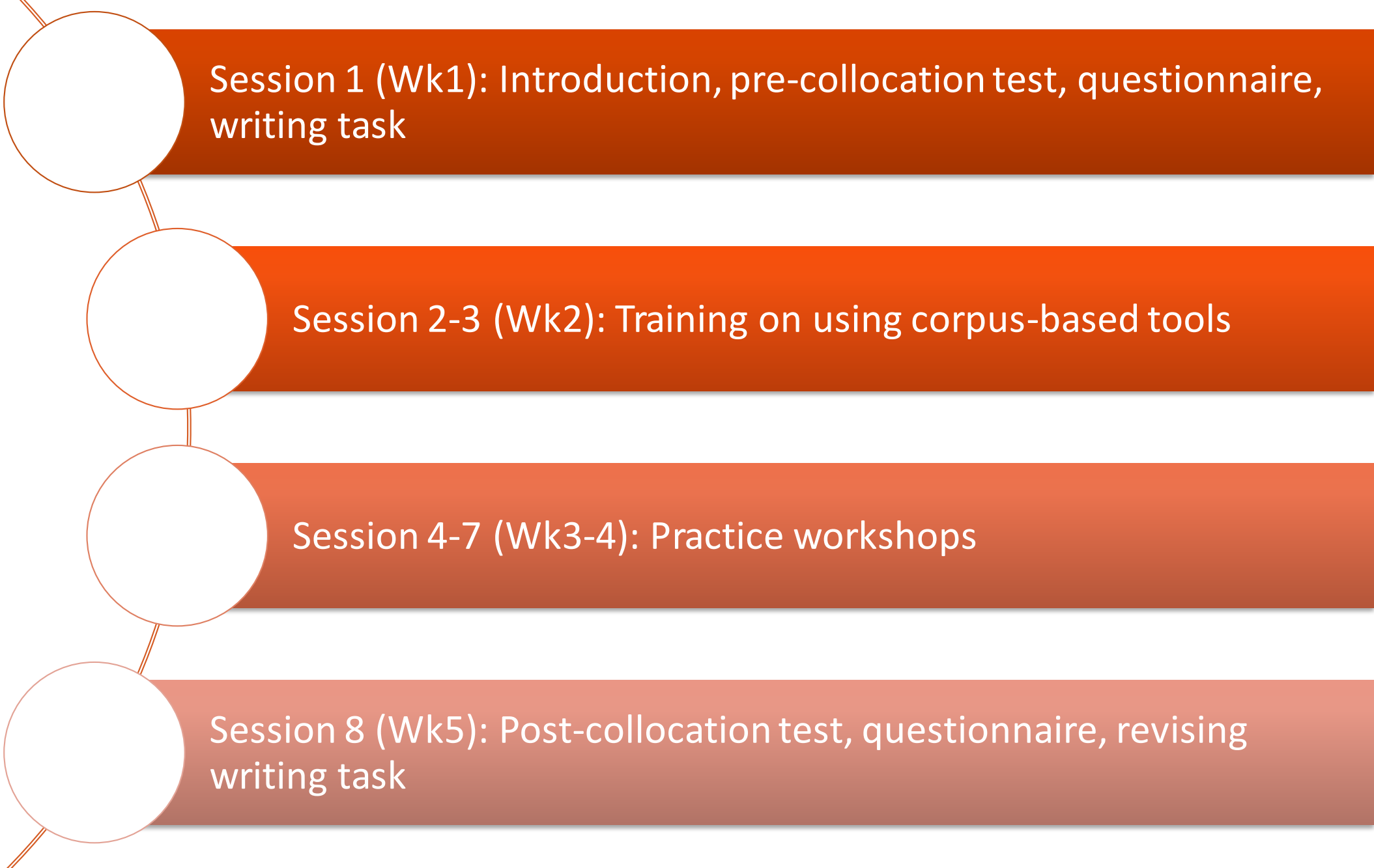

(Screenshots of Graphcoll in #LancsBox and online OCD)

	DDL (using #LancsBox and BAWE)	Oxford Collocations Dictionary
Similarities	• corpus-based • computer-based • showing collocations clearly • searching collocations by inserting node word	
Differences	• including additional functionalities (e.g. KWIC) • requiring importing the corpus before generating the collocations network • generating two-word combinations by association measures, requiring learners to find acceptable collocations by themselves • focusing on academic collocations	• having a paper version • providing ready-to-use collocations • including both informal and formal collocations • having been checked by native speakers

➢ **Instrument**

Research questions (RQs)	Instruments
RQ1	Pre- and post-collocation tests; the writing tasks
RQ2	Pre-treatment questionnaire; post-treatment questionnaire
RQ3	Post-treatment questionnaire; semi-structured interview

➢ **Procedure:** five weeks, 8 sessions

- Session 1 (Wk1): Introduction, pre-collocation test, questionnaire, writing task
- Session 2-3 (Wk2): Training on using corpus-based tools
- Session 4-7 (Wk3-4): Practice workshops
- Session 8 (Wk5): Post-collocation test, questionnaire, revising writing task

RESULTS

1. Comparison of the corpus-based tools from the results of the collocation tests: no significantly difference between using #LancsBox and online OCD in collocation learning $t(25)=.852$, $p=.403$.

Group	N	Pre-collocation test		Post-collocation test		df	t	95% Confidence interval		p
		M	SD	M	SD			Lower	Upper	
DDL	14	59.8	8.54	61.6	7.44	13	-.848	-6.33	2.76	.412
OCD	13	66.4	6.64	63.9	6.86	12	1.448	-1.28	6.36	.173

2. The learners' attitudes towards using #LancsBox and online OCD in collocation learning were mostly positive. While DDL group had some reservation about learning collocations during writing, the majority of learners in OCD group showed the positive attitudes.

Category	DDL group (#LancsBox)				OCD group (online Oxford Collocations Dictionary)			
	Agree or strongly agree (%)	Disagree or strongly disagree (%)	M*	SD	Agree or strongly agree (%)	Disagree or strongly disagree (%)	M	SD
I think the corpus-based tools are helpful in learning collocations	85.7	0	4.07	0.62	69.3	7.7	3.85	0.90
I think the corpus-based tools are helpful in learning collocations during my writing	57.2	14.3	3.71	1.07	84.5	7.7	3.92	0.76
I will keep using the corpus-based tools in my future language learning	71.4	7.1	3.79	1.05	84.7	7.7	4.15	0.90
I still have problems in using the corpus-based tools	28.6	42.9	2.86	0.86	30.8	53.8	1.03	2.69

(*Mean score based on 5-point scale: 1 = strongly disagree, 2 = disagree, 3 = neutral, 4 = agree, 5 = strongly agree)
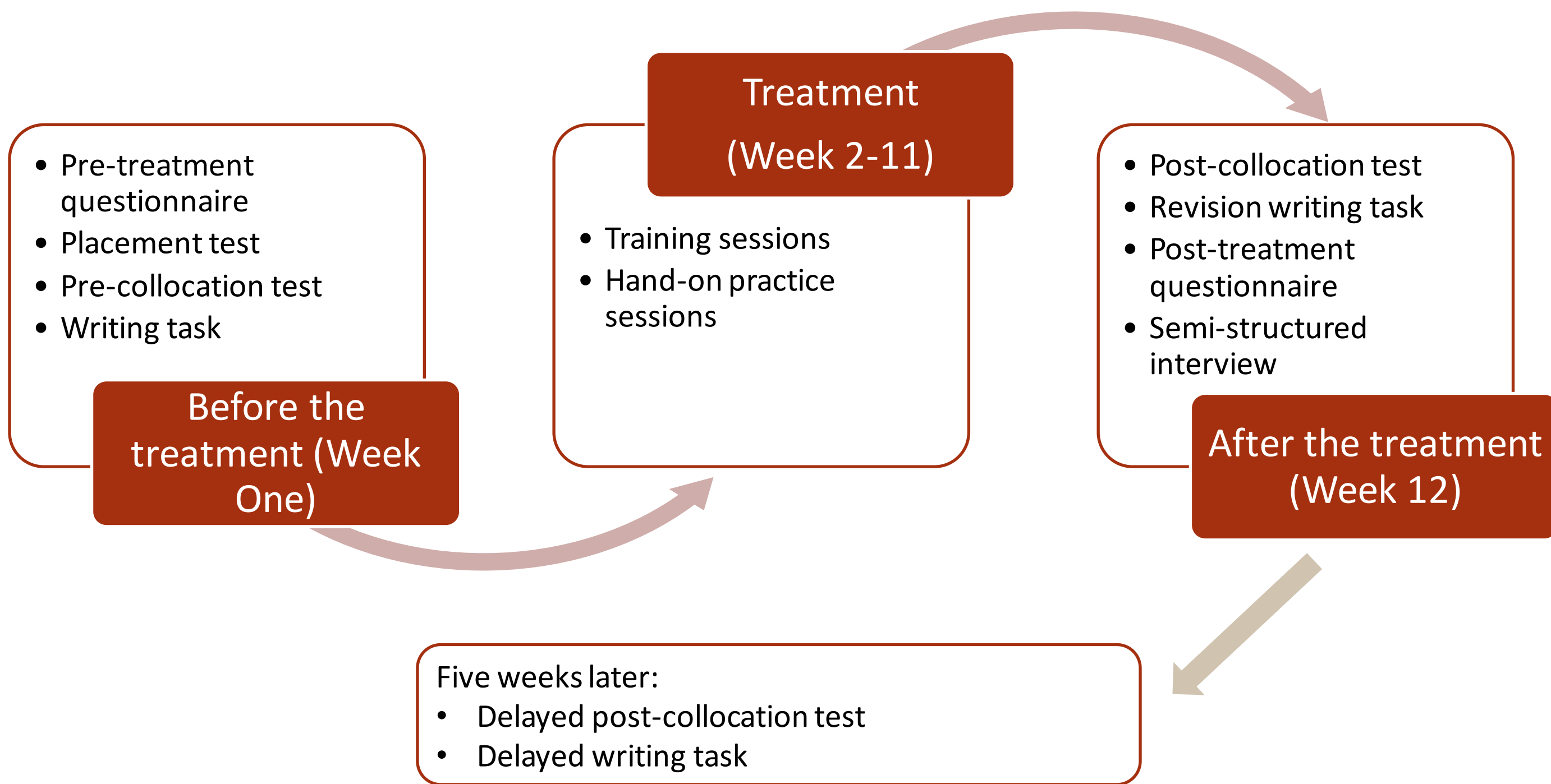
WHAT NEXT – MAIN STUDY

The present small-scale study served as a pilot for the large-scale (main) study to be conducted soon (from September, 2017)

➢ **Research aim:** Evaluating effectiveness of using DDL (#LancsBox) & the collocation dictionary (OCD) on the acquisition of academic collocations, compared with a control group (traditional writing class)
➢ **Participants:** 120 third-year English-major undergraduates from a middle-east Chinese university.

DDL group	OCD group	Control group
40 EFL learners	40 EFL learners	40 EFL learners

➢ **Procedure:**

- Pre-treatment questionnaire
- Placement test
- Pre-collocation test
- Writing task

Before the treatment (Week One)

Treatment (Week 2-11)
- Training sessions
- Hand-on practice sessions

- Post-collocation test
- Revision writing task
- Post-treatment questionnaire
- Semi-structured interview

After the treatment (Week 12)

Five weeks later:
- Delayed post-collocation test
- Delayed writing task

REFERENCES

Boulton, A. (2010). Data-driven learning: Taking the computer out of the equation. *Language Learning, 60*, 534-572. DOI: 10.1111/j.1467-9922.2010.00566.x

Boulton, A., & Cobb, T. (2017). Corpus use in language learning: A meta-analysis. *Language Learning, 67*(2), 348-393. DOI: 10.1111/lang.12224

Brezina, V., McEnery, T., & Wattam, S. (2015). Collocations in context: A new perspective on collocation networks. *International Journal of Corpus Linguistics,* 20(2), 139-173. DOI: 10.1075/ijcl.20.2.01bre

Granger, S., & Bestgen, Y. (2014). The use of collocations by intermediate vs. advanced non-native writers: A bigram-based study. *International Review of Applied Linguistics in Language Teaching, 52*(3), 229-252. DOI: https://doi.org/10.1515/iral-2014-0011

Schmitt, N. (2010). *Researching vocabulary: A vocabulary research manual*. New York: Palgrave Macmillan.

Tribble, C. (2015). Teaching and language corpora: Perspectives from a personal journey. In A. Leńko-Szymańska, & A. Boulton (Eds.), *Multiple affordances of language corpora for data-driven learning* (pp. 37-62). Amsterdam, the Netherlands: John Benjamins.