

Corpora in teacher training through Pedagogical Grammar – A case study of two workshops

Balázs Vida (Eötvös Loránd University, Hungary)

Background

As several surveys conducted with practising teachers of English revealed (see Mukherjee, 2004; Römer, 2009; Tribble, 2015), even if corpora are present in language education, direct uses of corpora in secondary, let alone primary, school settings are still uncommon (McEnery & Xiao, 2010).

One possible way to accomplish the mission of spreading the word about corpora is to train future language teachers in their pre-service training, as expressed by many scholars of the field, cf. Meunier (2002); O’Keeffe & Farr (2003); Mauranen (2004); Römer (2006); Granath (2009); Boulton & Tyne (2014). Student (or trainee) teachers, as the very name suggests, are in the peculiar position of still being students while progressing to becoming teachers. Thus, not only do they have first-hand language learning experience but they can already see the pedagogical value of using corpora. This fact has been recognised by a growing number of teacher training programmes across the globe. They incorporated corpora in their curriculum either as part of pre-existing *Pedagogical Grammar* courses (Helt & Reppen, 2008; Heather & Helt, 2012; Samburskiy, 2014; Zareva, 2016), or in the form of introductory training on corpus applications in language education (Farr, 2008; Breyer, 2009; Leńko-Szymańska, 2014, 2015; Callies, 2016; Naismith, 2016).

At the same time, while some professors mentioned through personal communication that they sometimes use corpora to enhance their academic courses, a systematic and practical introduction to corpora is still to be done in the Hungarian teacher training context.

For this reason – drawing on the results of the aforementioned studies –, with my thesis supervisor’s consent, I decided to bring a change to the status quo in Hungary. Two practical workshop sessions were organised on the uses of corpora in language teaching through pedagogical grammar-related tasks at a large public university in Budapest. The workshops were held within the framework of *Pedagogical Grammar* courses in the MA in English Language Teaching programme with 33 Hungarian trainee teachers of English.

Purpose of the study

The purpose of this pioneering and exploratory case study was to assess trainee teachers’ prior knowledge about corpora and corpus linguistics and, after the two workshop sessions, to evaluate the effectiveness of these workshops regarding the future, based on questionnaire data from trainees and their home assignments. These aims motivated the research questions of the study:

- How can trainee teachers’ prior knowledge about corpora and corpus linguistics be described?
- How effective were the workshops as perceived by trainee teachers?

The case

The two workshops represented the case of the study. They were offered in two separate, 90-minute-long sessions to two groups of trainee teachers, and taught by the author. The

two sessions were identical in content, namely both of them revolved around the following three major topics: corpus linguistics, language teaching and pedagogical grammar.

The design of the workshops followed Mukherjee's (2004) proposed framework for a one-day workshop, and they consisted of three main, sequential thematic modules:

- (1) teaching about corpora – a brief presentation of the basic concepts;
- (2) exploiting corpora to teach language – various hands-on explorations in the BYU corpora¹ with the help of five topics from prescriptive grammar (e.g., *the majority is or are, according to me...*) and five from descriptive grammar (e.g., *help sb (to) verb, double negatives*)²;
- (3) teaching to exploit corpora – discussion about classroom applications of corpora, supported by some relevant books, such as O'Keeffe, McCarthy & Carter (2007), Tribble & Jones (1997) and Willis & Willis (1988).

Methods of data collection and analysis

As the research is a case study, one of the top priorities was to use multiple data sources. The data analysed in the study came from the following three major sources:

- (1) a questionnaire before the workshop – to explore trainees' level of knowledge about corpora and corpus linguistics prior to the instruction;
- (2) participants' home assignments – to assess trainees' ability to make targeted searches in the corpus, interpret corpus data and draw pedagogically relevant conclusions;
- (3) a questionnaire after the workshop – to get feedback from trainees on their perception of their own development after the workshop.

Both questionnaires included 5-point Likert scale items and open-ended questions, and their structure was based on those administered by Leńko-Szymańska (2015). In the case of each Likert scale item, after assigning each response option a numerical value (i.e., 5 to strongly agree, 4 to agree, ..., 1 to strongly disagree), the mean and the standard deviation were calculated to specify the mean trend. With open-ended questions, the ocular scan method was used (Russell Bernard, 2000).

Home assignments as principal qualitative documents were analysed in an inductive way. All home assignments were eyeballed and labelled. At this point, several common themes started to emerge, which were used to identify the main trends.³

During the process of data analysis, sources of data were triangulated in order to determine whether data from the questionnaires and those from the home assignments pointed to the same conclusions.

¹ <http://corpus.byu.edu>

² These task ideas were inspired by Professor Mark Davies' teaching materials at Brigham Young University, publicly accessible via <http://davies-linguistics.byu.edu/ling485/>.

³ Here, as in Zareva's (2016) study, a trend was defined as at least three responses with identical content.

Results

The findings of the study can teach us several invaluable lessons as regards future integration of corpora and corpus linguistics into teacher training programmes. In the following, results of the three data sources will be presented and discussed in turn.

1 Pre-workshop questionnaire

The pre-workshop questionnaire consisted of three statements (Q1 *I have already heard about corpora and corpus linguistics*, Q2 *I have a fair understanding of corpora and corpus linguistics*, Q3 *I have already had a chance to use corpora*). The means of responses to these statements showed a downward trend, which can be visualised as in Figure 1.

What does this graph tell us? Over 77% of the respondents had some prior knowledge about corpora, but only around half of them indicated that they knew corpora reasonably well. On the other hand, only 3 trainees, out of 33, signalled that they had had the chance to try corpora out, resulting in a mean of 2.23. Furthermore, none of them had exploited corpora previously for searching for any kind of grammatical structures.

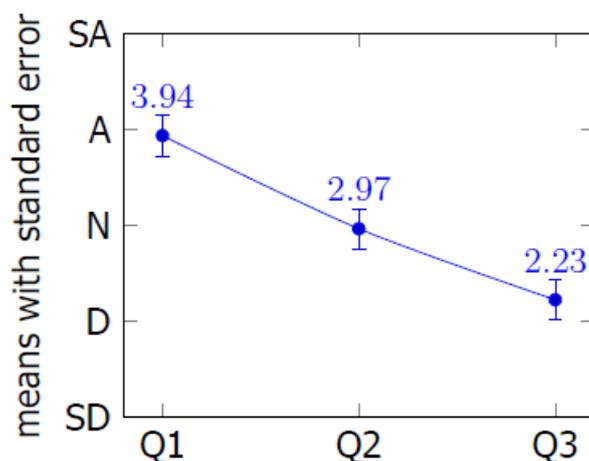


Figure 1. The downward trend of means from statements of the pre-workshop questionnaire.

2 Home assignments

Trainees' home assignments showed a more mixed picture. Using the eyeballing and labelling technique, four general themes had been established, which are shown in Table 1.

Trends	No of assignments
incomplete assignments	7
diverse approaches (e.g., in the <i>double negation</i> task)	4
critical thinking and sophisticated solutions	3
decent job	3

Table 1. Trends emerging from the home assignments.

A group of trainees submitted incomplete home assignments. Even if they could obtain raw data from the corpus, they did not attempt to draw conclusions from them. What is to be learnt from this group of assignments? To help trainees overcome these initial struggles, instructors of future offerings should emphasise the goal-orientedness of corpus searches and offer trainees even more guidance throughout the process.

Nevertheless, there were a handful of trainees who solved their tasks in a scientific and professional way, providing a balanced opinion about grammatical topics. They exhibited not only deep knowledge about the particular topic but also advanced research skills, which corresponds to what was found by Farr (2008) and Callies (2016). As one trainee noted in connection with the *according to me... vs in my opinion* problem, "In my opinion *is much more frequent than according to me*. In COCA, there are very few instances [when] according to me is used. I would not enforce teaching according to me, as in everyday life the students will come across in my opinion."

3 Post-workshop questionnaire

Thanks to the workshops, trainees expressed that they had become more aware of available corpus resources and tools (M = 3.91) and that they had come to know how to use corpus resources fairly well (M = 3.41). However, some of the trainees are still not very confident to interpret corpus data on grammar (M = 3.05), which is congruent with the most common trend from the home assignments.

Furthermore, over three thirds of the trainees were keen on learning more about using corpus tools in education (M = 3.68), but some turned out to be still hesitant to exploit corpora in their own teaching (M = 2.36). This was due to the limitation of the study, as two one-off workshop sessions could not possibly show all aspects of using corpora in classrooms. Nevertheless, future sessions will be organised to accommodate this need as well.

Conclusions

The findings of the study have practical implications for any teacher training programme aiming at having corpus linguistics as its component. There are at least three possible paths that can be taken from here. One option is to fully integrate corpora into *Pedagogical Grammar* courses, as was done by Helt & Reppen (2008), Heather & Helt (2012), Samburskiy (2014) and Zareva (2016). An alternative to this is to introduce a separate *Introduction to Corpora and Corpus Linguistics* course to the teacher training programme (either obligatory or elective), which would be akin to the courses described in the studies by Farr (2008), Breyer (2009), Leńko-Szymańska (2014, 2015), Callies (2016) and Naismith (2016). A third route would include a series of workshop sessions organised for trainees or even for practising teachers, an example of which was described in this study.

My vision, in line with that of the authors of previous research studies, is about corpus literate language teachers. Teachers who are able to make better decisions in their teaching as a result of their ability to manipulate corpora skilfully and provide more appropriate answers to their students based on authentic language data. To achieve this, we need to make the first step forward on either of the above paths.

References

- Boulton, A., & Tyne, H. (2014). Corpus-based study of language and teacher education. In M. Bigelow & J. Enns-Kananen (Eds.), *The Routledge handbook of educational linguistics* (pp. 301–312). New York, NY: Taylor & Francis.
- Breyer, Y. (2009). Learning and teaching with corpora: Reflections by student teachers. *Computer Assisted Language Learning*, 22(2), 153–172. doi:10.1080/09588220902778328
- Callies, M. (2016). Towards corpus literacy in foreign language teacher education: Using corpora to examine the variability of reporting verbs in English. In R. Kreyer, S. Schaub & B. Güldenring (Eds.), *Angewandte Linguistik in Schule und Hochschule* [Applied linguistics in secondary school and at university] (pp. 391–415). Frankfurt a. M., DE: Peter Lang.
- Farr, F. (2008). Evaluating the use of corpus-based instruction in a language teacher education context: Perspectives from the users *Language Awareness*, 17(1), 25–43. doi:10.2167/la414.0
- Granath, S. (2009). Who benefits from learning how to use corpora? In K. Aijmer (Ed.), *Corpora and language teaching* (pp. 47–65). Amsterdam, NL: John Benjamins.
- Heather, J., & Helt, M. (2012). Evaluating corpus literacy training for pre-service language teachers: Six case studies. *Journal of Technology and Teacher Education*, 20(4), 415–440.
- Helt, M., & Reppen, R. (2008). Training teachers to use corpus resources. *TESOL Teacher Education Interest Section*, 23(2). Online publication. Available from [http://www.tesol.org/news-landing-page/2011/10/31/teis-news-volume-23-2-\(2008\)](http://www.tesol.org/news-landing-page/2011/10/31/teis-news-volume-23-2-(2008))
- Leńko-Szymańska, A. (2014). Is this enough? A qualitative evaluation of the effectiveness of a teacher-training course on the use of corpora in language education. *ReCALL*, 26(2), 260–278. doi:10.1017/S095834401400010X
- Leńko-Szymańska, A. (2015). A teacher-training course on the use of corpora in language education: Perspectives of the students. In A. Turula, B. Mikołajewska & D. Stanulewicz (Eds.), *Insights into Technology Enhanced Language Pedagogy* (pp. 129–143). Frankfurt a. M., DE: Peter Lang.
- Mauranen, A. (2004). Spoken corpus for an ordinary learner. In J. M. Sinclair (Ed.), *How to use corpora in language teaching* (pp. 89–105). Amsterdam, NL: John Benjamins.
- McEnery, T., & Xiao, R. (2010). What corpora can offer in language teaching and learning. In E. Hinkel (Ed.), *Handbook of Research in Second Language Teaching and Learning*. (Vol. 2, pp. 364–380). London, UK: Routledge.
- Meunier, F. (2002). The pedagogical value of native and learner corpora in EFL grammar teaching. In S. Granger, J. Hung & S. Petch-Tyson (Eds.), *Computer learner corpora, second language acquisition and foreign language teaching* (pp. 119–141). Amsterdam, NL: John Benjamins.
- Mukherjee, J. (2004). Bridging the gap between applied corpus linguistics and the reality of English language teaching in Germany. In U. Connor & T. A. Upton (Eds.), *Applied corpus linguistics: A multidimensional perspective* (pp. 239–250). Amsterdam, NL: Rodopi.
- Naismith, B. (2016). Integrating corpus tools on intensive CELTA courses. *ELT Journal*. doi:10.1093/elt/ccw076
- O’Keeffe, A., & Farr, F. (2003). Using language corpora in initial teacher education: Pedagogic issues and practical applications. *TESOL Quarterly*, 37 (3), 389–413. doi:10.2307/3588397

- O'Keeffe, A., McCarthy, M., & Carter, R. (2007). *From corpus to classroom: Language use and language teaching*. Cambridge, UK: Cambridge University Press.
- Römer, U. (2006). Pedagogical applications of corpora: Some reflections on the current scope and a wish list for future developments. *Zeitschrift für Anglistik und Amerikanistik*, 54(2), 121–134. doi:10.1515/zaa-2006-0204
- Römer, U. (2009). Corpus research and practice: What help do teachers need and what can we offer? In K. Aijmer (Ed.), *Corpora and language teaching* (pp. 83–98). Amsterdam, NL: John Benjamins.
- Russell Bernard, H. (2000). *Social research methods: Qualitative and quantitative approaches*. Thousand Oaks, CA: Sage Publications.
- Samburskiy, D. (2014). Corpus-informed Pedagogical Grammar of English: Pros and cons. *Procedia – Social and Behavioral Sciences*, 154, 263–267. doi:10.1016/j.sbspro.2014.10.148
- Tribble, C., & Jones, G. (1997). *Concordances in the classroom*. Houston, TX: Athelstan.
- 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, NL: John Benjamins.
- Willis, J., & Willis, D. (1988). *Collins COBUILD English Course (Part 3)*. London, UK: Collins.
- Zareva, A. (2016). Incorporating corpus literacy skills into TESOL teacher training. *ELT Journal*. doi:10.1093/elt/ccw045