

Corpus-driven Learning and Language Learning Theories

Lynne FLOWERDEW
Centre for Applied Language Studies
Hong Kong University of Science and Technology
Kowloon, Hong Kong
lclynn@ust.hk

Corpus-driven activities are underpinned, explicitly or implicitly, by particular language learning theories. However, it is only in a few accounts in the literature where these are discussed in depth. The aim of this paper is thus to bring to light those language learning theories underpinning corpus-based pedagogy, an aspect which is not always made explicit in reports of corpus-driven learning.

At the crux of corpus-based pedagogy is arguably the observation that corpus-based activities, by nature of their methodology, tend to belong to the inductive approach, i.e. students extrapolate rules/probabilistic tendencies from recurrent corpus examples through close scrutiny of specific linguistic features and their frequencies. The “noticing” hypothesis discussed in second language acquisition (SLA) studies thus underpins many corpus activities. The principle underlying this cognitive concept is that learners’ acquisition of linguistic input is more likely to increase if their attention is drawn to salient linguistic features, which can be either student-initiated or teacher-directed. Corpus-based materials highlighting recurrent phrases would thus seem to be an ideal means for enhancing learners’ input via ‘noticing’, leading to uptake.

In addition to corpus-driven learning being rooted in the “noticing” hypothesis, either through spontaneous noticing by learners or teacher-directed noticing activities, the approaches taken to some corpus-driven activities also draw on sociocultural and constructivist theories of learning. Vygotskian sociocultural theories (Vygotsky, 1986) hold that knowledge is co-constructed through collaborative dialogue and negotiation. With guidance and support through teacher “scaffolding”, students are gradually acculturated into disciplinary writing. “Learner agency” is another concept associated with sociocultural theory. O’Keeffe *et al.* (2009: 55) discuss the advantages of using corpora for enhancing learner agency, stating that this notion involves learners being trained to operate independently to develop a set of skills and strategies for processing and using new vocabulary. They make the point that learner agency “can enable the learner to surpass instructional intervention and become a better, self-regulated learner”. In such a way learners can attain *depth* of knowledge, i.e. building an integrated lexicon on a particular topic or word, such as its collocations, semantic prosody and sub-senses, as well as *breadth* of knowledge, i.e. a concentration on a linear increase in vocabulary size.

Corpus-based activities are also a prime way to promote constructivist learning, which has been variously defined in the literature (see Cobb, 2006). As well as being associated with sociocultural theory offering the learner support in the form of “scaffolding”, proponents of this approach see it as also enabling learners to create their own learning through linking new knowledge to their existing knowledge. As Widmann *et al.* (2011: 168) point out “the more possible starting points a corpus offers for exploitation, the more likely it is that there exists an appropriate starting point for a

specific learner”. It is this concept which is behind many of the recent search engine interfaces for grammar instruction specifically designed with learners in mind. Corpus-based programs such as the Chemnitz Internet Grammar allow students to work either deductively accessing grammar rules, or inductively through searches for grammar patterns, thereby accommodating different learning styles (Schmied, 2006).

This presentation will thus showcase SLA theories underpinning corpus-based pedagogy, i.e. ‘noticing’ hypothesis, sociocultural and constructivist theories of learning, illustrated with accounts in the literature, as reported in Flowerdew (2008, 2012a, 2012b, in press, 2012)

References

- Cobb, T. (2006). ‘Constructivism.’ In: Brown, K. (ed.) *Encyclopedia of Language and Linguistics*, 2nd ed. vol. 3. *Foundations of Linguistics*. Oxford: Elsevier. 85-88.
- Flowerdew, L. (2008). ‘Corpus linguistics for academic literacies mediated through discussion activities.’ In: Belcher, D., & Hirvela, A. (eds). *The Oral/Literate Connection: Perspectives on L2 Speaking, Writing and Other Media Interactions*. Ann Arbor, MI: University of Michigan Press. 268-287.
- Flowerdew, L. (2012a). *Corpora and Language Education*. London: Continuum.
- Flowerdew, L. (2012b). ‘Exploiting a corpus of business letters from a phraseological, functional perspective.’ *ReCALL*, 24: 152-168.
- Flowerdew, L. (in press, 2012). ‘Corpora in the classroom: An applied linguistic perspective.’ In Hyland, K., Chau, M.H., & Handford, M. (eds). *Corpus Applications in Applied Linguistics*. London: Continuum.
- O’Keeffe, A., McCarthy, M. & Carter, R. (2007). *From Corpus to Classroom*. Cambridge: Cambridge University Press.
- Schmied, J. (2006). ‘Corpus linguistics and grammar learning: tutor versus learner perspectives.’ In Braun, S., Kohn, K. & Mukherjee, J. (eds). *Corpus Technology and Language Pedagogy*. Frankfurt am Main: Peter Lang. 87-106.
- Vygotsky, L. (1986). *Thought and Language*. Cambridge, MA: The MIT Press.
- Widmann, J., Kohn, K., & Ziai, R. (2011), ‘The SACODEYL search tool – exploiting corpora for language learning purposes.’ In: Frankenberg-Garcia, A., Flowerdew, L. & Aston, G. (eds). *New Trends in Corpora and Language Learning*. London: Continuum. 167-178.