"...AND I DROPPED MY JAW WITH FEAR"

The role of corpora in teaching phraseology

Gill Philip
University of Bologna – Italy

Keywords: collocation, phraseology, web as corpus, learner autonomy, language transfer.

1. Introduction

How can the effects of corpora on the language learning process be effectively assessed? This is an old question which is, however, just as important now as it was ten or twenty years ago. Does corpus use aid students in a measurable way? How does corpus use affect students' subsequent language use? Is there a marked qualitative difference between the work of students who make use of corpora in their studies and those who do not? Does corpus use necessarily lead to an improvement in students' language production?

This research reports an ongoing study into phraseological production in advanced learner writing (Philip 2005a; 2005b; 2006/forthcoming); in particular, it investigates corpus use and non-use in a semi-structured writing task. Unlike other studies in the EFL literature, this research examines multiple versions of a single base text, making it possible to observe what effect the use of corpora has on the production of phraseological units in text, rather than in gap-fill sentences or other, more traditional types of elicitation experiment, such as those found in Deignan et al. (1997 and Boers (2000). The rationale behind the present approach is influenced by the use of parallel corpora in translation studies in which multiple versions of single texts – in two or more languages – can be visualised and examined as KWIC concordances.

The data presented here has been gathered from five groups of advanced learners between March 2005 and June 2006. The learners, following advanced (C1) general English courses at the University of Bologna. Four of the groups have used corpora during the course of their studies; of these four groups of corpus users, two have used corpora extensively, and have been taught how to carry out advanced searches (multiple nodes, wildcards, and node plus tag). The fifth group, having had no exposure to language corpora, serves as a control.

2. Phrase building with and without corpora

The "phrase-building" task is an exercise in guided creative writing. It consists of a fixed sequence of key-words and collocations which form the basic skeleton of the story, and these fixed elements are presented sequentially on flash cards or as a slide presentation, making it often necessary for students to revise and reformulate their story in real time. The task is timed to occupy the best part of a two-hour lesson, with students writing their texts directly onto the computer. This allows them to edit and re-work their texts more easily than if they were writing by hand, and they can keep several windows open with on-line dictionaries, Internet search engines and corpora at their fingertips.

The keywords provided are all fairly common words, such as *wine* + *party, nerves, footsteps* +

stairs, and they are all relatively conventional collocations. However they need not be interpreted as being close collocates (node +/-1), with the result that they can effectively be used compositionally (linked by the context alone) or phraseologically (in fixed collocations, idioms and other turns of phrase), as the writer sees fit. The students' task is to write their version of the story using all the words provided, making as few changes as possible to the order of the keywords presented and to the word form given (plurals should remain plural, tenses of verbs should not be modified).

3. Discussion

What the data collected so far shows is that there is a qualitative difference between the texts produced by corpus users and those produced by non-users. However this difference is difficult to measure statistically because of the nature of the task: it is not an exam or test, and there are no right and wrong answers. The qualitative difference lies in the overall cohesiveness of the corpus-users' texts, because they appear to pay as much attention to the structural and grammatical elements of the phraseology as they do to the more salient words. Proficient users of advanced learner dictionaries also achieve this level of accuracy, though they are constrained by the limitations of their dictionary. The Google-dependent students – those students who had been given corpus training but felt more comfortable with the Internet's rough-and-ready alternative – showed much lower degree of phraseological accuracy than either corpus users or proficient dictionary users. This group was the one which used most interlanguage, and which appeared most reluctant to check the meaning of unusual or unfamiliar words. The control group did not produce less accurate texts than the corpus users, but they were less adventurous in their choice of phraseology, tending on the whole to favour the compositional use of the keywords over the phraseological, non-compositional possibilities; here again, the proficient dictionary users were able to produce more complex language chunks than those who relied predominantly on their existing language knowledge and more rudimentary dictionary skills.

What has emerged from the data is that those students who have had very limited exposure to corpora, and those who are unable to carry out more than the most simple of searches, show no better language skills than those who simply use a good learner dictionary. This may be partially due to the fact that English dictionaries for advanced learners are now very comprehensive repositories of information, most of which is derived from corpus analysis, but it also indicates that training students to search a corpus well is possibly as important as teaching them to read concordances.

Although an ability to use corpora with some skill and flexibility is without doubt an aid to the production of high-quality text, it is not the corpus alone that makes this so. Adventurous corpus users are motivated students: they make a considerable effort to find the language that they need and are prepared to try various permutations in their searches with before they settle on the one that they are most satisfied with. They appear to actively avoid translating directly from their L1, especially when the language is figurative. In other words, they are archetypally good learners who add corpora to their repertoire of language learning strategies.

References

Boers F. (2000) Metaphor awareness and vocabulary retention. *Applied Linguistics* vol. 21, 553-571.

Deignan A., D. Gabrys & A. Solska (1997) Teaching English metaphors using cross-linguistic awareness-raising activities. *ELT Journal* vol. 51, 352-360.

Philip G. (2005a) From Concept to Wording and Back Again: Features of learners' production of figurative language. In J. Barnden, S. Glasbey, M. Lee, A. Wallington & L. Zhang (eds) *Proceedings of the*

Third Interdisciplinary Workshop on Corpus-Based Approaches to Figurative Language (held in conjunction with Corpus Linguistics 2005). Birmingham: University Press, Birmingham, 46-53.

Philip G. (2005b). Figurative language and the advanced learner. *Research News: The Newsletter of the IATEFL Research* SIG. vol 16, 16-23.

Philip G. (2006/forthcoming) Metaphor, the dictionary, and the advanced learner. *Proc. of EURALEX 2006.*

7th Conference on

Teaching and Language Corpora

1st - 4th July 2006, Paris

TaLC 2006

CLILLAC

UFR EILA

UNIVERSITÉ PARIS 7 - DENIS DIDEROT

BnF

Table of Contents

Plenary Speakers **1**

Des machines pour enseigner les langues 2
Georges Antoniadis

Raising teachers' awareness to corpora 4
Ana Frankenberg-Garcia

WebCorp Linguist's Search Engine – the next order of magnitude 7
Antoinette Renouf

Key Words and Key Sections 9
Mike Scott

Main Conference **10**

Discovering language through corpora 11
Francesca Bianchi, Elena Manca

Collex-Biz 14
Kristin Blanpain, Liesbet Heyvaert, An Laffut, Nele Olivier, Dirk Speelman

Bringing corpora to the masses 16
Alex Boulton

From Data-Driven Learning to Needs-Driven Corpora 19
Sabine Braun

How to teach with corpora: 23
Yvonne Breyer

Designing a Learner Translator Corpus for Training Purposes 25
Sara Castagnoli, Dragoş Ciobanu, Kerstin Kunz, Natalie Kübler , Alexandra Volanschi

Lexique de l'émotion en classe de FLE: une expérimentation basée sur corpus 27
Cristelle Cavalla, Agnès Tutin

Language learning as discourse analysis: playing games in a corpus of French journalistic discourse 30
Angela Chambers

'However, there are several problems': Using EAP corpora to teach linking adverbs of contrast 33
Maggie Charles

Teaching language variation 36
Isabella Chiari

Des corpus oraux en didactique du Français Langue Etrangère : pourquoi et comment? 39
Jeanne-Marie Debaisieux, Virginie André

"The use of prepositions by Spanish learners of English at University level: main problems" 42
Mª Belén Díez Bedmar, Antonio Vicente Casas Pedrosa

Corpora, grammar and peer teaching 45
Maria Estling Vannestål, Hans Lindquist

Ways of Studying and Comparing Collocations 47
Jarmila Fictumová

ENTRAD, an English Spanish parallel corpus created for the teaching of translation 50
Celia Florén

Use of signalling nouns across native English speaker and learner corpora 53
John Flowerdew

Corpus-based evaluation of pedagogical materials 55
Costas Gabrielatos

Highly polysemous words in Foreign Language Teaching: 58
Gaëtanelle Gilquin

The International Corpus of Creative English [ICCE] 61
Peter John Hassall

Translation Corpus Aligner, version 2 65
Knut Hofland, Øystein Reigem

Combining Ethnographic and Learner corpora 67
Juan Pablo Jiménez-Caicedo, Meg Gebhard, Amy Rivera

Data-driving Critical Discourse Analysis 69
Bernhard Kettemann

Le corpus Carmel 73
Olivier Kraif, Marc El-Bèze, Régis Meyer, Claude Richard

Looking for Semi-Frozen Expressions using an Aligned Corpus 76
Kraif Olivier, Tutin Agnès

On the Role of Semantic Preference and Semantic Prosody in the Specialist Language Class 79
Sylvana Krausse:

What can we learn about discourse particles from a corpus to improve teaching materials? 81
Phoenix Lam

Autonomous learning through retrieval in a monolingual corpus 84
Heike van Lawick

Self-mention in EFL argumentative writing 87
Agnieszka Leńko-Szymańska

Teacher-controlled interactions in Singapore English Language classes: A corpus-based approach 89
Yongbing Liu, Huaqing Hong, Shouhui Zhao

Playing on words: A corpus-based methodology in the translation classroom 92
López Rodríguez Clara Inés, Tercedor Sánchez María Isabel

Corpus-based Study of Auto-Correction Learning 94
Hui-Chuan Lu

The Integrated Contrastive Model in Action 95
Ursula Maden-Weinberger

Un logiciel pour l'utilisation de corpus de parole spontanée dans l'enseignement des langues .. 98
Philippe Martin

Three conditions for Verb-Subject order in non-native English: 101
Amaya Mendikoetxea, Cristóbal Lozano

Empowering teachers' and learners' corpus literacy 104
Fanny Meunier, Cédrick Fairon

The Quill at the Gate 107
David Minugh

Umming and erring, analysing disfluencies in an oral learner corpus 108
John Osborne

Learning to learn about language: 111
Ulrike Oster

How can lexical combinatorics awareness improve foreign language learning in specialized fields of knowledge 113
Mojca Pecman

"...and I dropped my jaw with fear" 117
Gill Philip

Corpus research and practice: 120
Ute Römer

Teaching and LearningContrastive Linguistics using an EU Translation Corpus with English, German, French and Spanish 123
Josef Schmied

Corpora and the communication of the deaf 125
Michal Šulc

From Speech Corpora to Textbook Generation 127
Jozsef Szakos, Ulrike Glavitsch, Oliver Hess

Applying some Developments in Corpus Building Technology for Language Teaching and Learning 129
James Thomas, Jan Pomikálek

"They have their words yeah" - the lexis of lectures 131
Paul Thompson

Créer des corpus de données lexicales pour l'enseignement 133
Ophélie Tremblay

Us and them: 135
Chris Tribble, Mike Scott

Reconnaissance automatique de proverbes et applications 138
Olympia Tsaknaki

Breaking the Chains of Rhetorics in Academia: 142
Josta A. van Rij-Heyligers

La base lexicale du français (BLF) 144
Serge Verlinde

Annotation for a learner Translation Corpus 147
Alexandra Volanschi, Philippe Baudrion, Natalie Kübler, Andrei Popescu-Belis, Alina Secara, Martin Thomas, Susan Armstrong,

The Learner's Dictionary and the Sciences: 150
Geoffrey Williams

Hedging in Chinese writing: 153
May Lai-Yin Wong

Post-Colonial Theory, Composition Studies and Text Linguistics 158
Mohamed Zayani

Posters **160**

A Corpus-based Methodology in Grammar Instruction for the Acquisition-poor Environment of Japanese Learners of English 161
Mariko Abe

Bringing Corpora to Chalk-front Teachers 164
Elizabeth Anne, Elena Elizondo,

The MeLLANGE Project 165
Susan Armstrong, Guy Aston, Toni Badia, Silvia Bernardini, Gerhard Budin, Sara Castagnoli, Daniela Ford, Christine Gallois, Tony Hartley, Dragoş Ciobanu, Natalie Kübler, Kerstin Kunz, Mike Lunt, Vaclav Rericha, Notburga Rotheneder, Erich Steiner, Alexandra Volanschi, Alan Wheatley

Discovering Grammar in the Beginning-Level DDL EFL Classroom 168
Kiyomi Chujo, Kathryn Oghigian

A Corpus of Propositional Relations in Texts 171
Andy Cresswell