Classroom Concordancing in the 21st Century: The New Generation

GILL PHILIP
University of Bologna, Italy

ABSTRACT

Data-driven learning has come of age; yet although most teachers have by now heard of classroom concordancing in its various forms — hand-out, hands-on, or simply handy (as pre-digested data presented in the form of kibbitzers) — the relative paucity of published studies suggests that its actual use remains the preserve of the few. On the other hand, research centered around the analysis of learner corpora is flourishing. While some of the findings from these studies find their way into language reference works, most are slow to filter down into classroom materials. This paper discusses how the findings from an ongoing learner corpus study are being integrated into teaching using corpora in the classroom.

BACKGROUND

It is now 19 years since Tim Johns first used the term "data-driven learning" (DDL) (Johns 1991a), and 22 years since he first wrote about "classroom concordancing" (Johns 1988). So irrespective of whether one considers adulthood to start at eighteen or at twenty-one, it can safely be said that the use of corpora in classroom teaching has come of age. Coming of age tends to be viewed as a good thing: it brings with it recognition of one's status as an autonomous individual, a degree of respect appropriate to one's age and achievements, and great expectations for the years ahead.

Thanks to the concerted efforts of a small but influential group of scholars, there are very few language teachers now who are not familiar with corpora, at least in name. Yet despite the availability of clear instructions on how to prepare DDL materials (see especially Tribble & Jones 1991; see also Johns 1991a, 1991b), and empirical evidence of their pedagogical usefulness (Boulton, forthcoming a), the conclusive proof that DDL has not been entirely successful can be found in its

conspicuous absence in foreign language textbooks. While I do not intend to dwell on this point it is worth stressing that one of the factors contributing to this state of (non)affairs is time: classroom concordancing is a time-consuming business, and time as a commodity is in ever-shorter supply.

Over in the other camp, the view is somewhat rosier. Learner corpus studies are flourishing, and more and more is being discovered about the linguistic problems that learners really have, as opposed to the ones their teachers intuitively believe they have. Researchers have not only taken advantage of such pre-compiled learner corpora as ICLE (Granger 2003), but have also embraced the idea that they can compile corpora of their own from the texts produced by their own students, and thus conduct research that has direct relevance to their pedagogical context. Yet although the analyses are revealing and their implications for teaching pressing, their conversion into practical applications for use in the classroom is negligible.

What happens once a learner corpus study has been completed? Some findings inevitably feed into dictionaries and other published language resources, though the ways in which they do so are not always clear. But what about the rest? There seems to be a disconnect between research findings and practical applications, probably due to the fact that most corpus linguists are not classroom teachers and most classroom teachers are not corpus linguists. However, those of us who carry out both activities have a duty to bring our research full circle and prevent or remedy the very problems that we identify in our data. This paper illustrates how learner corpus research can combine with DDL, and other methods in applied linguistics. It reports the reconnection of data findings and language pedagogy with reference to an ongoing investigation into collocation and conceptualization, as revealed in a learner corpus (Philip 2005, 2006, 2007, 2008, forthcoming, in press).

COLLOCATION AND CONCEPTUALIZATION: NON-LITERAL MEANINGS IN THE L2

All research is triggered by something unusual, surprising, or otherwise eye-catching. This research began after coming across a fairly wide range of strange lexicalisations in advanced students' coursework assignments (Examples 1-4).

(1) The author's ability in manage such a long and rich period of time without loosing any relevant *knot* is surprising.

(2) However, a gap already exists between this availability of resources and the software which allows people to use them: virtualization is *raising* as the enabling technology to *climb over* the matter.

(3) Cullen *exposed* an account of causality which had been successfully applied in medicine and did not refer to necessity at all.

(4) He changed the format of his pedagogical works and chose more palatable essays instead of a *cold* and dry philosophical treatise.

Considered individually, the highlighted lexical items are best described as collocation errors. However, viewed collectively, a different (though related) explanation emerged: these are conceptual errors. Simply put, there is no problem with the ideas being expressed except that they are not realized in English through the lexicalization chosen by the learners. Examples 1-4 are typical of the kinds of oddities that teachers find extremely difficult to explain to their students. Because they do not violate any language rules, they are not mistakes as such; and yet they are barely acceptable because they make use of the L2 in unusual, unconventional ways.

It is probably true to say that most unexplainable errors like these are caused by calquing from the L2. It is also true that greater focus on form, especially in phrases, is required by both teacher and learner if similar mistakes are not to occur in the future. But there is a greater issue at stake; one which pervades language and yet is systematically overlooked in the foreign language syllabus: metaphor.

One of the corollaries of offering beginners only concrete, referential language is that later on, even competent learners have difficulty coping with the non-literal, phraseological and/or noncompositional meanings of a word they already "know." Philip (2007; in press) finds evidence to suggest that this is due to a strong one-to-one mapping between L1 and L2 lexical items during the early stages of language learning. The polysemy of core vocabulary items — kinship terms, colors, parts of the body — is rarely if ever commented upon in learners' textbooks. As a result, the learner is ill-equipped to deal with the non-salient, textual meanings which are encountered during progress through the (upper-) intermediate and advanced levels and, crucially, lacks awareness of the semantic flexibility of those basic lexical items (Philip 2007). This leads almost inevitably to calquing, as from the time that they are matched until such a time as that connection is forcibly broken, the L2 lexical item will be used in the same patterns as its L1 equivalent. Why? Because in the pairing process, the knowledge of how the L1 lexical item is used in text — the phrases it participates in, the prepositions it requires, the collocations it forms — is

not severed from its form. As the meaning of the L2 lexical item has not yet been ascertained through textual use but only through concrete reference ("this is a table," "that's my book," "my hands are cold"), the "missing" information is supplied by virtue of equivalence (e.g. "So 'hand' means mano; that means I can use 'hand' when I would use 'hand', mano."). In other words, all the primed patterns (c.f. Hoey 2005) of the L1 lexical item mano are transferred wholesale to the L2 lexical item, and what seemed to be a simple cross-linguistic pairing of two lexical items reveals itself to be a complex web of meanings paired to as-yet-untranslated prefabricated structures.

That L1-L2 mapping is a largely successful strategy belies the fact that "persistent calquing actually prevents students from acquiring a sense of the word's conceptual range in the L2, negatively affecting textual fluency and cohesiveness" (Philip, in press: 13). The problem facing teachers is how to prevent the extensive phraseological associations brought from the L1 from crystallizing around the L2 lexical item, or at the very least, how to remedy the problem once it manifests itself. It is a problem that goes beyond the right or wrong use of collocates, yet if meaning is created and fixed by collocation (c.f. Firth 1957), and conceptual ranges are defined and delimited by collocation, attention to how frequent words collocate is probably a good place to start. And this is where the corpus comes in.

REMEDIES AND PREVENTATIVE CURES FOR THE ADVANCED LEARNER

This section reports the use of corpora in tasks aimed at heightening learners' awareness of the conceptual range of frequently misused lexical items using collocate lists and concordance data. Section 3.1 is an extended description of an exercise created to teach the use of intensifying adverbs; Section 3.2 is a shorter description of the use of corpora in L1 and L2 to compare the collocational (and hence, conceptual) range of a frequently misused metaphor: "birth."

Slightly embarrassing: Mis- and under-use of intensifiers

There are of course more important things in the world than intensifiers, but their neglect in textbooks, reference works and both mono- and bilingual dictionaries (Philip 2008) has resulted in noticeable differences in their use by non-native speakers, compared to natives (Granger 1998, Lorenz 1999, Martelli 2006). My own learner corpus revealed that my previous students had rarely used any intensifiers other than "very," "really," and "quite" (Philip 2008), and on the few

occasions that they used lexical intensifiers (e.g. "highly," "closely") these were in calqued collocations and usually (though not always) unacceptable. The situation has not changed: with every new group of advanced students, the problem recurs and shows no sign of improving without direct intervention.

```
always somebody who strongly believes that living
   need for e-skills strongly binds the educational
o the Bible and are  strongly convinced that the
   most of people is strongly influenced by the
    e study period were strongly influenced by the
   esearch in Italy is strongly limited by lack of
s. The rioters also  strongly recommended Chinese
   a Cabrera were also strongly related to parish
```

Figure 1. Italian advanced learners' use of the intensifier "strongly."

Figure 1 gives an idea of the extent of the problem: of all eight occurrences of "strongly" in the corpus, none could be said to be the best choice of collocate, and several are downright wrong ("strongly limited," "strongly related"). The same problems repeat themselves across the board with other lexical intensifiers such as closely, deeply, heavily, highly, and "profoundly." Having identified the problem and its likely sources (Philip 2006, 2007, forthcoming 2010), the next step is to remedy it.

"From collocates to textual meaning" was the task administered to the seven students who were present at that day's Academic Writing in English course, a highly specialized course for academics and aspiring academics requiring advanced writing skills in order to get their work published. The task was programmed to take about an hour in class. Students were supplied with a worksheet which both outlined the task and provided space for note-taking, and each had a PC with internet connection to work at. The focus was explicitly on "closely," "deeply," "highly," "slightly," "widely," and "largely," although "lightly," "profoundly" and "broadly" were also suggested as optional extras, time-permitting. Using the WordBanks Online corpus sampler, each member of the class had to call up the collocations listings (t-score or MI, or both) for each of the adverbs in turn, and take notes. The idea was to identify some broadly-defined semantic groups associated with each of the adverbs as indicated by the most statistically significant collocates. Once these notes had been taken, students were encouraged to share their findings with the rest of the class members and, with my help, decide what kind of relationships each adverb suggested, and why.

how much they had managed to complete, how effective their collocate groupings were, and how well they had been able to complete the usage "rules."

Tables 1-6 condense the information from the collocates listings, providing a category name (on the left) and some examples of the collocates which were assigned to that category. Beneath each table is a brief commentary on these categories and a résumé of the responses noted on the students' worksheets.

Table 1. *Summary of collocates: "closely"*

Category	Example collocates
Proximity	aligned; gauged; linked; tied; knit; woven;
Resemblance	resemble; resembled; identified
Correspondence	correspond; paralleled; matching; related; associated;
Focus	examined; inspect; look; observe; scrutinised; studying
Collaboration	cooperating; involved; operate; together; work

"Closely" always implies nearness, but different kinds of nearness exist. "Proximity" in Table 1 refers to physical space (vertical/horizontal), though these terms may also be used metaphorically to refer to social relationships (and grouped with "collaboration"). "Resemblance" is both similarity of aspect and of function; "focus" is an entirely separate grouping, connected to the literal meaning of closely, and can be glossed as "examining in detail". The categories that students recognized were: accuracy; the and contact; similarities and comparison; relation; focus/looking.

"Deeply" has a less complex range of collocates than "closely." In fact, although four categories are indicated in Table 2, the first, "underground," features collocates which are normally metaphorical in combination with "deeply," i.e. they also belong to the "belief/ideology" category. The categories share a relationship between something that is manifested on the surface and its origins or nucleus below.

The vast majority of the collocates refer to emotions of all descriptions, closely followed by beliefs. In metaphorical terms, "deeply" refers to embodied experience and core beliefs, convictions and emotions. The more intensely something is felt or believed, the "deeper" it is said to be.

Students recognized the following categories: involvement or "deeply"; inclusion; relation. However, four of the seven students only annotated categories after the class discussion, copying the words I had

used to explain the abstract notions (distance, intensity, focus, accuracy). These were all students who annotated only the first collocates on the listing and had not managed to identify and semantic groupings.

Table 2. *Summary of collocates: deeply*

Category	Example collocates
Underground	buried; embedded; entrenched; ingrained; rooted
Life functions	breathe; exhaled; sighed; drank; asleep;
Emotions	angered; concerned; confused; depressing; disappointed; embarrassed; honoured; impressed; offended; shocked
Belief/ideology	communism; committed; religious; subversive

"Highly" was problematic for students, as the t-score collocates listing revealed a distinct lack of content words. Shifting to the MI calculation allowed students to locate groupings, but only one student managed to identify a semantic group: skills and qualifications.

The reasons for the difficulty is that there is indeed a very wide range of collocate groupings associated with highly. The best defined of these appear in Table 3, but the most effective description for students was to understand that, unlike "closely" and "deeply" which are metaphorically related to the collocates, *highly* is metaphorically related to the graphic representation of intensity on a bar chart or line graph (the vertical axis indicates the percentage or proportion of whatever is expressed by the collocate.) Viewed from this perspective, the meaning of "highly" becomes much clearer, its function having been metaphorically motivated by reference to the image of a bar chart (scrawled hastily on the board).

Table 3. *Summary of collocates: "highly"*

Category	Example collocates
Likelihood	improbable; likely; unlikely; unusual
Importance	placed; significant; visible
Concentration	charged; concentrated; enriched; radioactive; toxic
Evaluation	attractive; controversial; desirable; infectious; suspicious
Competence	effective; motivated; organized; productive; professional
Complexity	complex; detailed; developed; sophisticated; technical

"Slightly" was surprisingly interesting to examine because it turns out that it has a distinct semantic preference for emotions and states of affairs which are viewed as undesirable, i.e. have a negative semantic association (Hoey 2005). This comes through to a certain extent in the

collocates listed in Table 4, but is marked enough to justify the gloss of "slightly" to be not merely "a little bit," but "only a little bit."

Table 4. Summary of collocates: "slightly"

Category	Example collocates
Change of state	altered; coarser; larger; paler; thicker; thinner
Deterioration	marred; rumpled; shabby; wrinkled
Chemical property	acidic; alkaline; radioactive
Bent	angled; askew; inclined; bowed; bulging; hunched
Evaluation	bemused; crazed; incongruous; misguided; risqué

The students found "slightly" even more problematic than "highly," but the main reason for its inclusion was that the Italian equivalent, "leggermente," can translate as either "lightly" or "slightly," and the inappropriateness of "lightly" has to be pointed out. My suspicion is that they will continue to avoid using this word because it is attitudinally loaded, and this makes it difficult to use effectively.

"Largely" and "widely" are a different matter, however. "Large/ly" and "largo/amente" are false friends, and to make matters more complicated, "largely" has two near synonyms, "widely" and "broadly" (not discussed here). Both "largely" and "widely" intuitively refer to horizontal space, but they do so in different ways, "widely" referring to a horizontal spread from a central point outwards, much like the ripples caused by a stone thrown into water, and "largely" indicating a widening in the sphere of reference of a definition. Table 5 summarizes the main collocate groupings for "largely," which can be compared with those for "widely" (Table 6). It is immediately evident that the adverbs relate to very different fields, though all but one of the students failed to recognize them (probably because of a lack of time to finish the task).

Table 5. Summary of collocates: "largely"

Category	Example collocates
Cause - effect	caused; determine; due; fuelled; result; stems; thanks
Importance	futile; ineffective; irrelevant; superficial; uneventful
Acknowledgement	ignored; overlooked; unknown; unnoticed; unsung
Preservation	decline; remains; survived; unchanged; untouched
Content	composed; consists; drawn; formed
Restriction	confined; dominated; imposed; responsible; restricted

One of the prominent features of "largely," visible to a certain extent in Table 5, is that it displays a strong preference for adjectives with

---

negative prefixes ("ineffective," "irrelevant," "uneventful," "unchanged," etc.). So although it can be glossed as "for the most part," it has a function similar to that of "slightly," only in reverse: while "slightly" indicates only minimal presence of the undesirable feature, "largely" ensures that its presence is not absolute.

Table 6. Summary of collocates: "widely"

Category	Example collocates
Variety	differing; varied/varies/vary/varying
Distribution	circulated; dispersed; distributed; scattered; spaced
Commonality	adopted; available; practised; prescribed; used
Knowledge	discussed; disseminated; known; read; reported; seen
Belief	accepted; assumed; held; interpreted; regarded; viewed
Commendation	acclaimed; praised; respected; supported; welcomed
Condemnation	blamed; condemned; criticised; ignored; reviled

The distributional meaning of "widely" is only rarely literal ("variety" contains only two lemmas). Instead, it collocates with groupings which concern information and practice, plus achievement (both positive and negative), possibly related to the former categories in particular. The one student who was able to distinguish between these two adverbs unaided noted that "largely" was associated with negative adjectives (relating to information) while "widely" meant "being distanced, spaced."

It can be seen from the summary above that none of the students achieved a satisfactory degree of detail, and in future versions of the task the number of collocates would have to be limited, or else a number of categories would have to be suggested in advance, especially because all of the students required assistance in making generalizations about the meanings and sphere of reference of the adverbs chosen for the task. However, the approach combining hands-on use of corpora with teacher-led synthesis and picture-drawing on the board, worked to my satisfaction. The focus on collocates made the task feasible in an hour, which would not have been the case using concordances. In fact, three of the students did not manage to complete the task because they were evidently flipping between collocation and concordance modes to examine how the collocates were used syntactically. This being the case, the teacher-led synthesis allowed all the students to complete the worksheet, which formed part of their course notes, without penalizing those who had been side-tracked by their corpus exploration. That (some) learners show a willingness to use corpora autonomously for their own ends is encouraging, but the details of

syntax and phraseology must come after the general range of reference of the lexical items has been ascertained, especially in a task like this one where the underlying problem is conceptually motivated. Collocating provides a time-effective means of fixing the meaning and reference of confusing lexical items while limiting data overload. This makes it very suitable for classroom tasks where time is at a premium and a balance has to be struck between the needs of the class as a whole and those of individual learners.

I wasn't born yesterday: From collocate to concept

Fortunately, not all classroom use of corpora is so intensive. Brief glimpses of collocational patternings can be an effective way of illustrating why some errors are unacceptable, without resorting to overly abstract descriptions or long-winded examples.

The BIRTH AS BEGINNING metaphor is over- and mis-used by Italian learners of English (Philip 2006, 2007, forthcoming). The problem is not that Italian does not use the same metaphor to refer to these concepts, but that it also uses the BIRTH AS BEGINNING metaphor to refer to emotional states, beliefs, problems, and disagreements too. In other words, its conceptual range is considerably wider than that of its English equivalent.

The problem of mis-use has to be pointed out to learners, but this need not take more than a couple of minutes. I had already classified the collocates for a conference paper, and the file remains permanently on my teaching USB key.

Table 7. *Comparison of collocate groupings:* "Nascita" *and* "birth"

Nascita	**Birth**
Nazioni	Nations
Movimenti politici	Political movements
Aziende	Businesses
Organizzazioni	Organizations
Tendenze socioculturali	Social trends
Discipline accademiche	Academic disciplines
Difficoltà	-
Guai	-
Equivoci	-
Problemi	-

Working down the list, is immediately obvious that there are close correspondences, i.e. the metaphor mappings are similar in both

languages. However, it is equally evident that the range of the Italian metaphor extends beyond that of the English one: there simply are no English categories corresponding to "*difficoltà*" (difficulties), "*equivoci*" (misunderstandings), "*guai*" (trouble), or "*problemi*" (problems). If learners are made aware that "born" is never used in English to refer to this broad semantic area, they ought not to transfer the metaphor inappropriately. And should they want to know what metaphor to use instead, they can consult the collocates for "difficulty," "misunderstanding," "problem," and "trouble," to get an instant(ish) answer.

NON-REFERENTIAL MEANING FOR BEGINNERS

If the problems surrounding learners' inability or unwillingness to deviate from one-to-one equivalence are to be resolved rather than merely remedied, intervening at the higher levels where non-literal meanings become common is probably too late to be truly effective. Intervention is necessary much earlier on – as early on as possible. In this subsection I briefly report methods I have been using with beginners and post beginners (A1-A2) which serve to heighten their awareness of how vocabulary works. The activities reported here are not formalized tasks (although they could be formalized): instead they represent an approach to vocabulary acquisition and revision which is used consistently throughout the course. Students' overall response to the approach, along with their ability to grasp non-core meanings, and their appreciation of contextual meaning and collocation in determining meaning are only discussed on the basis of observation: the effects of teaching metaphor awareness are only likely to become visible at upper-intermediate level and above, i.e. several years after initial exposure to figurative thinking. Given semester-long courses, patchy attendance and discontinuity from one academic session to the next, tracing its effectiveness longitudinally has been unfeasible.

It's hardly rocket science: Collocates delimit meaning

Most foreign language learners at tertiary level come to the classroom with rather fixed ideas about language and how to learn it, having already had some experience of language learning at secondary school or even earlier. Italian learners place particular emphasis on the learning of grammar and underestimate the importance of acquiring vocabulary. Further to this, their view of what vocabulary is and how it works is extremely simplistic, based on a fully compositional model (i.e. "open choice," Sinclair 1991) in which L2 words are translations of L1 words.

Although collocations have taken on more prominence in beginner level coursebooks over the past ten years or so, low-level learners generally fail to appreciate their value unless they are made consciously aware of collocations in their L1 – a matter left entirely to the teacher to introduce.

Explaining collocation is a first step in explaining the more unsettling prospect that words do not have meanings – meanings have words. This notion takes time to sink in but is easy to demonstrate; and the demonstration helps learners develop a number of skills related to vocabulary learning and reading comprehension: notation of vocabulary which incorporates its immediate collocates (and translation of the whole); the value of example sentences; how to recognize a word's syntactic role; how to use context to disambiguate, and so on.

At its heart, the task is a guessing game which hones learners' awareness of how words change their meanings when they change their collocates. It is a simple, board-based activity using words already present in the learners' active core vocabulary. A single word is written on the board – e.g. "book" – and learners are asked what it means. The salient meaning for most learners will be the concrete noun referring to reading matter: "the book," "a book," so the article, or any adjectives or compounds they suggest are added to the word on the board to fix that meaning ("book shop," "new text-book," etc.). The words fixing the noun meaning are subsequently erased, and learners are asked to remember the verb form, "to book," and suggest phrases and formulae which fix the meaning "to reserve." Then "book" is transformed into "booking," which students know as the present participle in progressive verb constructions, but have not encountered as a noun. They are invited to suggest what "a booking" could possibly mean (a) in a restaurant or on a train, and, moving into more abstract, "figurative thinking" (Littlemore & Low 2006), (b) in a football match.

Admittedly, "book" is not a particularly interesting word for this task, though it serves its purpose as a first foray into collocation. Once students have begun to notice that meaning is realized through word combinations, more metaphorically active words, e.g. body parts, can be brought in. The aim, however, is not to overload learners with new vocabulary and they are explicitly told that what they learn will not be examined in their end of course tests: the aim is to open their minds. Figurative thinking tasks have the potential to switch on learners' ability to deduce meaning from context, while at the same time reflecting on how their L1 lexicalizes the same meanings. And once they demonstrate competence in the board-based task, they are ready for real collocates in a corpus.

Bite-sized classroom collocating

Sometimes there seems no way of escaping rules and vocabulary lists. Explaining to low-level learners that the manner adverb derived from "hard," not "hardly" (even though "hardly" also exists...), is a case in point. One group struggled to accept this despite having been provided with translations for the terms: the presence of "hard" in "hardly" proved too much of a distraction. What this group needed was an idea of the collocational profile of each of the two words, which would give a general idea of their different meanings. It barely seemed worthwhile spending much time on just a couple of words, and it occurred to me that some "hit-and-run" DDL might be effective.

Having instant access to an on-line corpus in the classroom, paired with a projector screen linked up to an array of multimedia equipment is one of the best things about teaching in a university. In under a minute, the collocates of "hard" (WordBanks Online) were listed on the screen and students were recognizing most of them. When read out as collocations ("work hard," "find it hard to," "hard disk," "hard currency," etc.). In doing this, it became evident that "hard" was both an adjective and an adverb; and the class also saw directly how many different nuances of meaning expressed with "hard."

Table 7. Top 10 lexical collocates (t-score) for "hard" and "hardly"

Hard	Hardly
work	surprising
find	believe
working	anything
worked	likely
disk	knew
currency	anybody
line	noticed
hard	matters
believe	need
hit	seemed
imagine	see

The same procedure was carried out with "hardly," and the learners were immediately struck by the different fields of reference (Table 7). The translation/ explanation that they had been given only a few minutes earlier clicked into place, comprehension dawned, and we moved on to the next exercise.

HIT-AND-RUN SOLUTIONS TO REAL-TIME LANGUAGE QUERIES

The approaches and tasks described in the Section 3 serve a formative purpose, actively engaging low-level learners with a view of language in which meaning emerges from words when they combine forces. It is still with the higher levels that hit-and-run DDL comes into its own. Upper Intermediate and advanced learners demand not only rules, but *explanations and verifiable facts*, when working out the differences between lexical items or grammatical structures. This has always been the case, but two things have changed in recent years: the technology and the learners themselves. While corpus linguists revel in the technological advances that offer ever-increasing speed, availability, accessibility and user-friendliness, I have yet to see comments in print regarding the changing profile of learners, who have grown up expecting the information they require to appear instantly, usually via the internet. Traditional DDL is not dead, but it is a child of its time, developed as an educational tool for the learners of its time. Those learners have changed, and if DDL is to survive, it too must change.

Hit-and-run DDL is a response to learners' demands for faster, better, cheaper and personalized responses to their language queries. These learners are the Information Generation and have grown up surrounded with facts, data and information on every topic under the sun. They are very good at filtering though this data themselves to find what they want, and they expect their teachers to be able to do the same, only better (faster and more accurately), since they are the teachers, after all. These students, or rather *consumers of education*, expect a service from their teachers, and are quick to complain if they feel they are not getting value for money.

The short subsections below illustrate a few more examples of hit-and-run DDL used to respond in real time to classroom queries.

I never used to worry about it

Sometimes the grammar presented in learner textbooks is a little suspect. Often, it is inadequate, but can pass. Other times it just sounds wrong. Unfortunately, teaching the same structures year after year can inure the teacher to textbook incongruities and EFL-ese. But corpora come to the rescue.

The negative and interrogative of the "*used to inf.*" structure which indicates past habit are, according to most textbooks, "didn't use to" and "did X use to?" respectively. The first time I taught this structure, that negative form sounded distinctly odd, a feeling confirmed by there being a total of five occurrences in the 56 million word WordBanks

Online corpus sampler. The form that sounded more natural to my ear, "never used to," was the most frequent (more than the forty concordance lines returned, at least), and other permutations, including "didn't used to" (35 occurrences), were also more common than the book version.

It is advisable to tread carefully when criticizing the coursebook in front of students, not least because learners are more apt to believe what they see in print than the ad lib statements of their teacher. But when they too saw the raw data from the corpus projected on the screen, not only were they convinced, but their smugness at having better information than their peers was also palpable. EFLese was nipped in the bud, and faith in their teacher reinforced.

Accurate and satisfactory, albeit cursory

There are always times when an innocent question sends the teacher into existential crisis. "Can I use *albeit* instead of *although* in that context?" is a case in point. There are a few possible responses to any query like this one: (a) "In general, yes/no." (b) "have you checked your dictionary?" (c) I'll let you know next lesson, and (d) "Well, let's have a look!" This fourth option is the hit-and-run DDL response.

Extemporaneous analysis is (just) possible with forty concordance lines and consists in reading aloud four or five extracts, inviting class members to look for anything interesting, and hoping that you can come up with a satisfactory answer before the students do. With "albeit," the first thing to notice is that its use is more restricted than that of "although," and a dictionary will also indicate that it belongs to a higher register. The immediate response to the student's query, on the basis of the screenful of examples viewed, was that "albeit" is used primarily (no absolutes here) to introduce contrast between attributes or a series of related attributes contrasted with one other: "a^1, a^2, a^3, albeit b" (Examples 5 and 6). So the answer to the student's question was "no", but it was a motivated answer supported by illustrative examples.

(5) His eyes were *dark, well-fringed, kind*, albeit *male*.
    [feminine traits]   [masculinity]
(6) He was willing to *push* for many of those policies, albeit *reluctantly*.
    [force]   [resistance]

It does, however, prove a point

Punctuation does not feature highly in the EFL curriculum. With the exception of full stops and capital letters, the only explicit information that one can guarantee will be present in a general English course is the

use of commas surrounding non-restrictive relative clauses. Yet to be effective, discursive writing requires competence in how to use punctuation appropriately. I have found that zapping concordances up on the projector screen is much more effective than more abstract explanations, and eliminates much of the need for tedious grammar-book exercises to consolidate the theory. Observing the punctuation patterns in a screenful of concordances puts a stop to calqued punctuation (indeed, punctuation is also calqued). Italian learners of English frequently mis-punctuate connectors like "so" and "however"; but not after looking at concordance data.

Recurrent punctuation patterns are laid bare in a list of concordances. The visual impact of seeing rows and rows of "however" physically distanced from the surrounding text by punctuation cannot be underestimated. Italian learners of English frequently mis-punctuate connectors "however" by omitting its surrounding commas; on the other hand, they often insert a comma between "so" and the consequence it introduces. These kinds of punctuation problems can be easily resolved by illustration with concordance data, and doing so takes very little time. In fact, it takes longer for the students to write their notes than it does to call up the concordances and see the regularities.

CONCLUDING REMARKS

The aim of this paper has been to show (i) the importance of bringing learner corpus research full-circle, creating or improving teaching materials so that the errors identified are remedied or avoided altogether and (ii) to demonstrate how corpora and other classroom technology must be used in accordance with the expectations of a new generation of students which is more demanding than the previous one on several counts, and has an insatiable appetite for information.

The shift from concordancing in the classroom to collocating in the classroom may meet with the disapproval of long-term DDL practitioners, but my defense is that it addresses the need for information and detail in a way that suits a new breed of learner, is compatible with the ever-present requirement for speed, lends itself to spontaneous illustration and description of language points in the classroom, and eats far less into the ever-more precious time available for materials preparation and in the classroom itself, than concordance-based worksheets do.

In an ideal world, the use of corpora in teaching would whet students' appetite for linguistic data. So to end this paper on a positive note, I can report that this is indeed the case for a good number of my own upper-intermediate and advanced learners. After being shown concordances and collocates during lessons, in order to answer specific language points raised by the students themselves, they are seduced by the potential empowerment that corpus data can offer them. For a generation which has grown up with customizable everything, from ring tones to playlists to videogames, and even sneakers, it should not strike us as surprising that corpora are being incorporated into the range of information sources available on the internet which students exploit when customizing their language learning. They trust the text because they see that their teachers do. The fact that they ask how they can access corpora outside the classroom is an encouraging sign of learner autonomy. In using corpora in our teaching, we are leading horses to water: it is up to them how much they want to drink.

REFERENCES

Boulton, A. (forthcoming a). Empirical research in data-driven learning – A summary. Web supplement to A. Boulton. Learning outcomes from corpus consultation. In M. Moreno Jaén, F. Serrano Valverde and M. Calzada Pérez (Eds.), Exploring New Paths in Language Pedagogy: Lexis and Corpus-based Language Teaching. London: Equinox.

— (forthcoming b). Data-driven learning: On paper, in practice. In T. Harris & M. Moreno Jaén (Eds.), Corpus Linguistics in Language Teaching. Bern: Peter Lang.

Cambridge ESOL. 2008a. Certificate in Advanced English Handbook for Teachers. Cambridge: Cambridge University Press. Available online: <http://www.cambridgeesol.org/assets/pdf/resources/teacher/cae_hb_deco8.pdf>.

— 2008b. Certificate of Proficiency in English Handbook for Teachers. Cambridge: Cambridge University Press. Available online: <http://www.cambridgeesol.org/assets/pdf/resources/teacher/cpe_handboo k.pdf>.

Firth, J. R. 1957. A Synopsis of linguistic theory 1930-55. In F. R. Palmer (Ed.), Selected Papers of J. R. Firth 1952-1957 (pp. 168-205). London and Harlow: Longmans.

Granger, S. 1998. Prefabricated patterns in advanced EFL writing: Collocations and lexical phrases. In A. Cowie (Ed.), Phraseology: Theory, Analysis, and Applications (pp. 145-160). Oxford: Oxford University Press.

— 2003. The international corpus of learner English: A new resource for foreign language learning and teaching and second language acquisition research. TESOL Quarterly, 37/3, 538-546.

Hoey, M. 2005. *Lexical Priming. A New Theory of Words and Language.* London and New York: Routledge.

Johns, T. 1988. Whence and whither classroom concordancing? In T. Bongaerts, P. De Haan, S. Lobbe & H. Wekker (Eds.), *Computer Applications in Language Teaching* (pp. 9-27). Dordecht: Foris.

—— 1991a. From printout to handout: Grammar and vocabulary teaching in the context of data-driven learning. *CALL Austria*, 10, 14-34.

—— 1991b. Should you be persuaded: Two samples of data-driven learning materials. *English Language Research Journal*, 4, 1-16.

—— & King, P. (eds.). 1991. *Classroom Concordancing. English Language Research Journal 4.* Birmingham, UK: University of Birmingham, Centre for English Language Studies.

Littlemore, J. & Low, G. 2006. *Figurative Thinking and Foreign Language Learning.* Basingstoke: Palgrave Macmillan.

Lorenz, G. 1999. *Adjective Intensification – Learners versus Native Speakers.* Amsterdam: Rodopi.

Martelli, A. 2006. A corpus based description of English lexical collocations used by Italian advanced learners. In E. Corino et al. (Eds.), *Proceedings of EURALEX XII International Lexicography Congress* (pp. 1005-1011).

Philip, G. 2000. L'uso delle concordanze bilingui nell'insegnamento dei "falsi amici". In R. Rossini Favretti (Ed.), *Linguistica e Informatica. Corpora, Multimedialità e Percorsi di Apprendimento* (pp. 363-373). Rome: Bulzoni.

—— 2005. From concept to wording and back again: Features of learners' production of figurative language. In J. Barnden et al. (Eds.), *Proceedings of the Third Interdisciplinary Workshop on Corpus-Based Approaches to Figurative Language* (pp. 46-53). Birmingham: Birmingham University Press.

—— 2006. Metaphor, the dictionary, and the advanced learner. In E. Corino et al. (Eds.), *Proceedings of EURALEX XII International Lexicography Congress* (pp. 895-905). Alessandria: Edizioni dell'Orso.

—— 2007. Decomposition and delexicalisation in learners' collocational (mis)behaviour. In M. Davies et al. (Eds.), *Proceedings of Corpus Linguistics 2007.* Available online: <http://corpus.bham.ac.uk/corplingproceedings07/paper/170_Paper.pdf>.

—— 2008. Adverb use in EFL student writing: From learner dictionary to text production. In E. Bernal & J. De Cesaris (Eds.), *Proceedings of the XIII Euralex International Congress* (pp 1301-1310). Barcelona: Universitat Pompeu Fabra. Istitut universitari de lingüística aplicada.

—— (in press). "...and I dropped my jaw with fear": The role of corpora in teaching phraseology. In N. Kübler (Ed.), *Corpora, Language, Teaching and Resources.* Frankfurt: Peter Lang. Pre-print available online: <http://amsacta.cib.unibo.it/archive/00002361/>.

—— (forthcoming). "Drugs, traffic, and many other dirty interests": Metaphor and the language learner. In G. Low et al. (Eds.), *Researching and Applying Metaphor in the Real World.* Amsterdam: Benjamins.

Rundell, M. (ed.). 2007. *Macmillan English Dictionary for Advanced Learners (MEDAL)* 2nd ed. Oxford: Macmillan.

Sinclair, J. M. 1991. *Corpus, Concordance, Collocation.* Oxford: Oxford University Press.

Tribble, C. & Jones, G. 1991. *Concordances in the Classroom.* London: Longman.

ACKNOWLEDGEMENTS

GILL PHILIP
UNIVERSITÀ DI BOLOGNA, ITALY.
E-MAIL: <G.PHILIP.POLIDORO@GMAIL.COM>