

# Decomposition and delexicalisation in learners' collocational (mis)behaviour

Dr Gill Philip

CILTA

University of Bologna

*g.philip.polidoro@gmail.com*

## 1. Introduction

This paper presents some of the theoretical considerations which have arisen from the author's learner corpus research over the past couple of years (Philip, 2005a, 2005b, 2006, 2007, in press). In particular, it highlights a mismatch between the language of native speakers and that of learners, not simply at the textual level, but at a more abstract, conceptual level. The findings of this extended piece of research suggest that it is time to shift the focus of investigation away from learner errors *per se*, and concentrate on what barriers seem to be preventing learners from achieving fluency in a second language. This means not only that the causes of errors, often dismissed as interlanguage, have to be investigated more thoroughly, but also that the notion of *avoidance* has to be taken into consideration. While it is acknowledged that most empirical linguists are uncomfortable with the idea of measuring avoidance, it will be shown during the course of this paper that focusing on error alone is an inadequate means of identifying problems in language acquisition. Learner error must be considered as part of a trio which also includes avoidance, and calquing from the L1 which happens to match L2 patterns (which I will call here *fortuitous accuracy*) both of which can be measured contrastively: avoidance is determined by its low frequency (or absence) in relation to equivalent native speaker texts in the L2 (*cf.* Martelli, 2006), while fortuitous accuracy can be identified through the comparison of back-translated forms in a general reference corpus in the students' L1<sup>1</sup>. The trio of error, avoidance and fortuitous accuracy paints a much more detailed picture of the strategies that advancing learners use in their language production than can be achieved by studying error alone, as is shown in the discussion below. By examining both what is and is not present, and comparing this to L1 and L2 native norms, it is possible to get inside one of the major problems affecting language learning at higher levels: collocations in conventionalised abstract and figurative language.

Some collocations are easier to learn than others. Many noun+noun and adjective+noun collocates reflect linguistically what can be observed in the real world, and so their lexical co-occurrence seems natural and relatively unproblematic. However, as language moves away from the concrete and observable towards the abstract and, apparently, arbitrary, the likelihood of a collocate to be deemed 'logical' by a learner diminishes, as does its accurate recycling in free production<sup>2</sup>. Although there is a growing body of research based on learner-corpus data, it is somewhat limited by its focus on error alone. In particular, while it is possible to state, for example, that "the total number of collocational errors is 105. Of these 0 are adverb + verb collocations, 6 are collocations in which the entire expression is involved, 50 are adjective + noun collocations, 45 and verb + noun collocations." (Martelli, 2006: 1008), such a calculation of frequency is based on the total number of erroneous collocations, not the total number of collocations, whether good or bad. It can also be misleading to imply that, e.g. adverb+verb collocations are relatively unproblematic because they are rarely misused, because it may be that adverbs are simply not being used; if so, absence is as significant as error.

---

<sup>1</sup> The Corpus di Italiano Scritto (CORIS) held at the University of Bologna: <http://corpora.dslo.unibo.it/>

<sup>2</sup> Free production refers here to any written or spoken production, and should be contrasted with production which arises from tasks designed to elicit particular language forms, e.g. gap-fill, transformation or translation exercises (*cf.* Nesselhauf 2003).

What most researchers are in agreement over is that errors occur when one or other of the collocates has a restricted collocational range. Such restrictions are deemed to be arbitrary (Lewis, 1993, 1997, Thornbury, 2002), with the result that they are difficult to learn and recycle well. But what has not yet been investigated satisfactorily is whether collocation errors are indeed mislearned or misremembered collocations, or if they are something else entirely – calqued or invented forms – with the ‘arbitrary’ collocations being avoided instead. This is clearly an enormous area for research, and will only be touched upon in this paper. What will be investigated here is the interplay between conceptual space and the lexical collocations which fix and are fixed by it, and how mismatches between the L1 and L2 underlie error, avoidance and fortuitous accuracy.

## **2. Data**

The corpus used in the current study is approximately 60 000 words in size, and was compiled by the author from the written work produced by top-end B2–C1 students of English at the University of Bologna’s language centre. The students were nearly all native speakers of Italian, and none of them were following degree courses in linguistics or foreign languages, making them neither expert or trainee-expert linguists. This point is worth stressing because students of language and literature might be better equipped to tackle the sorts of phraseological and figurative language described in the analyses that follow. The texts fall into three distinct groups – those produced as the result of a structured classroom activity (such as the *Phrasebuilder*, Philip, 2007), homework commentaries written in response to a stimulus (typically a newspaper article or current affairs topic), and assessed essays. The language discussed in this paper is drawn from students’ critical commentaries and essays, in which no specific restrictions were imposed on the language content.

The corpus texts were all submitted for correction, so were all read ‘manually’, and in fact most of the examples discussed in this paper were initially identified during this phase. Given that the corpus was compiled from course assignments, it features several clusters of related texts which make use of similar language and ideas, helping to counteract the limitations of a small data set, i.e. a lack of repetition of lexis, with subsequent lack of identifiable patterns, which makes it difficult to arrive at generalisations of any sort. As in most corpus-based analyses of figurative language (*cf.* Charteris-Black, 2004, Partington, 2003), concordancing was used as a secondary option, mainly to search for additional examples or to verify hypotheses formed from the analysis of the manually-identified examples. Concordances were the main source of data regarding the frequency and use of accurate collocations in the corpus, with which the inaccurate could be compared and contrasted. Concordances from general reference corpora were also used to ascertain native speaker norms (L1 and L2) which could serve as a benchmark for measuring the learners’ production.

## **3. Delexicalisation, decomposition, relexicalisation**

The teaching of collocation is gradually becoming part of the mainstream language learning curriculum, but is often presented as a problematic issue, mainly because there is a general perception of collocations as being arbitrary and, in a practical sense, inexplicable. However, by focusing too closely on the apparent problems of L2 collocations, it is easy to lose sight of the very obvious fact that students of foreign languages use collocations every day in their L1, and that these collocations are equally arbitrary and inexplicable.

For native speakers, the vast majority of everyday language is phraseological and involves a greater or lesser degree of delexicalisation. In very general terms, the idiom principle (Sinclair, 1991) governs the meaning of conventionalised language, shifting the semantic focus away from the individual words present in a given stretch, and prioritising the phraseological and pragmatic meaning of the utterance as a whole. It hardly needs to be pointed out to the native speaker that *highly original* expresses the extent of originality rather than its height. Delexicalisation colludes in

masking the arbitrariness of many collocations because it detracts one's attention from the fully salient meanings (Louw, 2000) of the collocates while allowing them to be reactivated for rhetorical purposes (cf. Philip, 2003), so although a *glaring error* may well "stare you in the face just as it is" (Firth, 1957: 182), it is the use of the citation here that resuscitates the fully salient lexical meaning of *glaring*, whose role in the collocation is simply to indicate extent. So language which is normally delexical can be relexicalised, with the non-compositional being broken down into its component parts. This can only happen under particular cotextual or contextual conditions, though. Or rather, this is the native-speaker view of language: the learner's view is a little different.

While the everyday language of the native speaker is delexicalised to a considerable extent, the same language when learned as a L2 is perceived as being fully salient. This is partially the result of a pedagogical methodology which prioritises exposure to the concrete and literal, making learners more inclined to favour literal, compositional meaning in both comprehension and production. Collocations, therefore, are initially seen as compositional combinations of words rather than as a phenomenon of co-selection. As a result, L2 collocations which correspond to equivalent forms in the student's L1 are deemed 'normal' and thus effectively remain delexical in the mind of the learner. For example, the Italian student who encounters expressions such as *highly qualified*, *highly specialised*, *highly (im)probable*, and *highly toxic*, will assimilate them to the corresponding Italian forms *altamente qualificato*, *altamente specializzato*, *altamente (im)probabile*, and *altamente tossico*, without giving the meaning of *highly* so much as a second thought. On the other hand, L2 collocations which deviate from patterns in the L1 are perceived by the learner as 'strange'; they are treated analytically rather than delexically, which means that the learner interprets the meaning of each part in its fully salient sense. Remaining with *highly*, in the equally conventional collocation *highly original* (which does not correspond directly to the student's L1 equivalent: *\*altamente originale*<sup>3</sup>) the collocation will seem strange, in precisely the same way that a collocation error sounds odd to the native speaker.

When language chunks are broken down into their component parts, the phraseological meaning can be distorted if too much attention is paid to salient meaning. In particular, decomposition can distort a learner's perception of markedness. The strangeness of a collocation – its apparent unnaturalness – can trigger various reactions, including error and avoidance, although over-use may also occur. In the case of *highly original*, errors are likely to emerge in the choice of adverb used to collocate with *original*, either through use of a calqued form, such as *\*absolutely original*, or through the creation of a novel collocation based on a synonym of *highly*, (e.g. *greatly* or *vastly*). While the former strategy can be equated with the learner trying to use an expression that feels less marked than the L2 does, the latter shows that the learner has activated the semantic range of the word in its fully salient sense and has therefore over-lexicalised the collocate. Avoidance, identified by absence of the collocation in a corpus of learner English, occurs when a learner shies away from using the alien collocation *highly original*, preferring a teddy-bear adverb such as *very* or *really*, or indeed eliminating the adverb altogether so as to circumvent the problem<sup>4</sup>. Over-use (identified using the same methods as those for avoidance) may occur, squeezing out other contending collocates such as *extremely* or *entirely*.

The section 4 takes examples drawn from the author's learner corpus to illustrate the points made above.

#### 4. Case studies

Polysemy is a major source of problems for language learners. Very little lexis undergoes explicit recycling in the language teaching syllabus, and that which does attract attention is presented as being somehow unusual or special. Cases in point include idioms and other fixed

---

<sup>3</sup> Not attested in CORIS.

<sup>4</sup> A general tendency to avoid using adverbs may well account for the absence of errors reported for adverb+verb collocations in Martelli 2006.

phrases, and phrasal verbs. The extended meanings of core vocabulary items, such as parts of the body, colours, and kin, are rarely if ever commented upon. Thus the student who learns the words *head*, *green*, and *father* in the elementary stages of language learning, is ill-equipped to cope with their less salient textual meanings which s/he will encounter at upper-intermediate and advanced levels: ‘top’, ‘commander’, ‘intelligence’, ‘side of a coin’; ‘ecologist’, ‘vegetables’, ‘lawn’, ‘envy’; ‘inventor’, ‘mentor’, ‘god’. For so many learners, words are metaphorically ‘ticked off’ as they are learned; they are assigned to a L1 equivalent, initially as an aid to comprehension. But the inexperienced concept of translation equivalence effectively maps the newly-acquired L2 word onto L1 patterns, and this has negative repercussions on the learning of sub-senses, figurative extensions and phraseological meaning. As Philip states, “Students generally encounter words in their literal sense first, match them to a translation equivalent in their L1, and from then on, unless instructed otherwise, use the word in calqued forms of the L1 phraseology. The relative success of this strategy effectively masks the underlying problem, which is more serious than simply getting collocations wrong. 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, 2007: 13).

#### 4.1 Beyond the literal: *born*

As a first example of the problem, let us look at *born* and its nominalised variant *the birth of...* in the non-core, metaphorical sense of birth as initiation or beginning: “When an idea or organization is born, it comes into existence. If something is born of a particular emotion or activity, it exists as a result of that emotion or activity.” (COBUILD 3<sup>rd</sup> ed.). The metaphor is also used in Italian although a comparison semantic preferences of *birth* and its Italian equivalent, *nascita* (Figure 1) shows that the Italian metaphor extends into the emotional sphere but the English does not.

<b>BIRTH</b>	<b>NASCITA</b>
nations; businesses	nazioni; aziende
political movements	movimenti politici
organisations	organizzazioni
social trends	tendenze socioculturali
academic disciplines	discipline accademiche
[emotions]	emozioni
[misunderstandings]	equivoci
[difficulties]	difficoltà
[trouble]	guai
[problems]	problemi

Figure 1: Comparison of semantic preferences of *birth* and *nascita*<sup>5</sup>

This imperfect correspondence between the two languages is reflected in the students’ use of the metaphor (see lines 2-3 in Figure 2).

1 markets. So, new ways of doing business are born. The term e-business stands  
 2 On-line Resolution process, a conflict born on the Internet can be resolved  
 3 a condominium conflicts and discords can be born with others.  
 4 commercial and political relationship. A new-born Scottish science, political economics  
 5 Reflections on beauty grew up with the birth of the Aesthetic, that involves

Figure 2: Birth as beginning.

<sup>5</sup> English data from BNC; Italian data from CORIS. Square brackets indicate that the translation equivalent is unattested in the corpus.

Although dictionary definitions do specify what kinds of things can be *born*, there is no guarantee that a student will check or, having checked, will appreciate the relevance of finding *ideas* and *organisations*, but not *conflict* or *disagreement*, included in the definition. What these collocation errors show is that the students are unaware that their knowledge of *born* – a word that they know very well – is imperfect at a collocational and conceptual level. Not only can *conflict* and *discord* not occur together with *born*, but neither can any other semantically-related word such as *trouble*, *problem*, *anger* or *resentment*: they are semantically incompatible with *born*, though not with *nato*. The fact that they are conventional collocates in the students' L1 means that they are also delexicalised to some extent, meaning that they are treated non-compositionally. As a result of this, there is no reason why decomposition should occur, so the expression is transferred whole into the L2.

#### 4.2 Familiarity breeds contempt: interlanguage at work

Conceptual mismatches do not merely concern the more colourful uses of language that students may produce (cf. Philip, 2005a, in press), although these are undeniably more interesting to study. A less eye-catching example of interlanguage found in the written production of Italian native speakers is the over-use of the structure *in a ... way*, which operates in tandem with a corresponding under-use of simple adverbs. For the Italian speaker, the equivalent *in modo ...* is unmarked and used in free variation with the corresponding adverbial of manner. In English, however, where both forms are possible, *in a ... way* is typically used to emphasise the manner in which something is done. The concordances in Figure 3 illustrate the problem nicely: while ex. 5 uses the structure correctly, all the others could – and probably should – be substituted with simple adverbs (*aggressively*, *badly*, *differently*, *persuasively*, *suspiciously*, *unworthily*).

1 from the baseline but in a very aggressive way, putting a lot of pressure on the  
 2 In fact Agassi was playing in a very bad way, it seemed he had lost his usual game  
 3 to organize the calendar in a different way, if necessary, I indicate monday,  
 4 believes that things went in a different way. He opened a trial to establish if  
 5 emperor's son he was educated in a military way too and he played an important role  
 6 present with the past in a very persuading way. Gould's main claim is that the two  
 7 and one of her friend dies in a suspicious way, so she thinks that these facts are  
 8 that she conquered it, in a not worthy way, as many people today do. So, is or

Figure 3: *in a ... way*

The remaining occurrences of *way* in the corpus (Figure 4) further reflect the calquing going on.

1 spend their final part of life in the best way they can. I can add that in country  
 2 , of being totally unaware of the normal way of being, unaware of how an apple could  
 3 is "the" choice, meaning that every other way will bring you to a deadline, but in  
 4 natural law and then he left it on his own way. But they hold that God can suspend  
 5 vices that big towns offer. At the same way old people can find the silence the  
 6 the faith. He organized meeting in the same way but there were less people and a little

Figure 4: *way*

Although there are a couple of occurrences of fortuitous accuracy, where the English and Italian structures match (ex. 1 and 6), the others can be traced back to their L1 origins; and all reveal that *way* has been learned as the one-to-one equivalent of *modo*<sup>6</sup>. This reliance on equivalence at word level is encouraged in the early phases of language learning, and it is therefore difficult for students to adopt different strategies later on. The importance of phraseological meaning, and phraseological equivalence, is simply not given enough attention when students are learning how to learn a language. Once students are aware there is a problem, they are more likely to resort to avoidance (in this case, the avoidance of lexical adverbs altogether) than they are to

<sup>6</sup> The Italian equivalents, respectively, are: *nel miglior modo possibile*, *il solito modo di essere*, *ogni altro modo*, *a modo suo*, *allo stesso modo*, *nello stesso modo*.

tackle the problem head-on. Thus errors creep into the student's written and spoken production, many remaining uncorrected because the error is difficult to pin down: while grammar or orthography can be clearly right or wrong, 'naturalness' is often a matter of degree and personal preference which can be difficult to explain.

### 4.3 Delexical language and markedness: (de)lexical adverbs

The over-use of *in a ... way* helps to disguise a more general problem that these students have with adverbs. While sequencing and discourse adverbials are present as expected in their writing, lexical adverbs (such as *highly*, cf. 3) occur infrequently and with varying degrees of success. There is a distinct preference in the writing of borderline B2-C1 students to use *very* and *really* to the virtual exclusion of any other adverb, making it difficult to gather enough data on the errors to get a clear picture of the problems which underlie adverb use. C1 students make more use of conventional collocations, and in so doing make visible some of the processes that are going on in their learning.

The least problematic of adverbs are those used compositionally, especially when the ensuing collocations are also acceptable in the students' L1. Compositional use allows for the full meaning of the words to be exploited, and errors are rare because the language itself tends to be literal. Adverb+verb/ adjective collocations in this group include *teach privately*, *study obsessively*, *try intentionally* and *intrinsically fair*. More formulaic collocations such as *officially deny*, *state publicly*, *calculate precisely*, *participate actively*, *drop dramatically*, and *strongly suggest* may well have been penned compositionally by the students, but their closeness to the mother tongue makes it impossible to tell whether they have been learned, acquired or if they are just lucky guesses. Other collocations are easier to identify as having been learned, because of the restricted collocability of one or other collocate. In the data examined, only the following three could be identified unequivocally as being learned, not calqued: *eat heartily*, *fail miserably*, *furnish sparsely*. The presence of so few collocations of this type is partly due to the lack of lexical adverbs in general, as well as to the substantial overlap between English and Italian collocational patterns, which allows students to fall back on familiar expressions rather than these more idiomatic ones.

A1 And there is always somebody who strongly believes that living in the countryside  
A2 in any case, the need for e-skills strongly binds the educational system to the  
A3 interpretation to the Bible and are strongly convinced that the world is less than  
A4 the author, the most of people is strongly influenced by the interest and the  
A5 during the study period were strongly influenced by the geography of the  
A6 Scientific research in Italy is strongly limited by lack of resources and by a too  
A7 companies. The rioters also strongly recommended Chinese people to boycott  
A8 isolation in La Cabrera were also strongly related to parish distances from the  
A9 from migration matrices, suggest strongly that the 'outlier' status of this parish  
B1 to have this reputation, to be so deeply appreciated and celebrated all over the  
B2 and I think it concerns me very deeply as I'm a biotechnologist working in  
B3 going to be the same for lecturers, deeply attached to the traditional means of  
B4 a dead-lock. Economical problems deeply influence Italian society which is ageing  
B5 is, far to be totally objective, is deeply influenced by several personal aspects. In  
B6 students the possibility of going deeply into any discipline in a more creative and  
B7 everywhere and making you feel deeply upset. >G05STi-----//NEWS HEADLINES AIRPORT  
C1 the aforementioned issues? It is highly ambiguous and dubious proposition. In fact,  
C2 La Cabrera (province of Leon), an highly isolated mountainous region of North-  
C3 Finally I can say that I think very highly of her otherwise we wouldn't have been  
C4 Results from the latest trials are highly positive and let hope that the benefits  
C5 of his group. The Byrds are highly responsible for the diffusion of the song  
C6 an inherently uncertain thing and highly speculative reasoning related to their  
D1 on kidneys and stomach, which are heavily affected by what we eat. Obesity pills  
D2 be easy as Italian universities are heavily split between tradition and innovation. In  
D3 their behaviour and they changes heavily their lives in certain periods. The most  
E1 Newton's physics was analyzed and widely developed. Moreover, following Newton's  
E2 The blasphemy in Down Under: It is widely known that successful modernization has  
E3 SP04----//2. The use of computer is widely widespread and it's becoming relevant as a

Figure 5: Most frequent (de)lexical adverbs

The area where most errors occur is with conventionalised collocations which are abstract and delexical. It is useful to remember at this point that only linguists (and not even all of them) freely admit that delexicalisation exists, so the average foreign language learner majoring in computer science, or history, or economics, is unlikely to give much thought to such notions as idiomaticity, pragmatic effect, and co-selection. A great many students express surprise when polysemy in their own language is pointed out to them; and when asked why they have used e.g. *strongly* or *heavily* in an English collocation where these words do not normally appear, they are often unable to appreciate that the word is being used in a figurative sense. The same clearly applies to the corresponding adjectives. In the corpus used in this study, *strongly* is the most frequently-occurring (de)lexical adverb, closely followed by *deeply* and *highly*, with *heavily* and *widely* occurring less than half as often (see the concordances displayed in Figure 5).

The use of the adverbs in Figure 5 reflect the Italian patternings of the corresponding delexical adverbs in Italian. This is most obvious when mistakes and oddities such as *strongly limited* (A6), *highly isolated* (C2), and *widely widespread* (E3) are present, but even correct collocations such as *suggest strongly* (A9), *deeply influence* (B4-5), and *highly speculative* (C6) stem from L1 patternings. Simply put, the strategy of word-for-word translation is endemic because in the earlier phases of language learning with its focus on simpler, more compositional language, it has proved successful. It takes an act of will on the student's part to experiment with unfamiliar collocational patterns, and s/he will only do so if s/he believes it absolutely essential. If the student is made aware of the delexical nature of the L1 patterns s/he uses, however, a degree of motivation is created for leaving these patterns behind and seeking something less familiar.

Different collocational patternings reflect back on the perceived semantic range that the words – in their literal sense – seem to cover, and this creates a symbiosis between form and meaning. This is a tricky area of language, because it is the place where the marked and the unmarked converge. If students do not learn to use L2 collocations, and to accept them as unmarked (despite the fact that their word-for-word translation would be marked in their L1), they will continue to produce familiar, calqued patternings derived from their L1 (which are correspondingly marked in the L2). This inevitably leads to misunderstandings, both in comprehension and in the meaning that they effectively communicate. It also hampers their long-term linguistic proficiency, as it prevents them from picking up on exploitations in which the collocate is relexicalised because of the semantic set being activated (*cf.* section 3, par 2), not to mention more sophisticated linguistic devices such as allusion, humour, irony, and extended metaphor.

## 5. Containing collocational misbehaviour

The examples discussed in 4 illustrate some of the general problems facing learners in their production of non-literal language. Several implications can be drawn from the research carried out over the course of this project, perhaps the most important being that comparison and translation have to be welcomed back into the classroom. Students will always translate new terms into their own language, and surely it is better to help them do this properly rather than pretend that they only need make comparisons in a monolingual context. There is a real need for students to be encouraged to compare collocational patternings in the L2 with those of their L1 so that they are able to identify possible areas of difficulty for themselves. This not only raises the profile of bilingual and bilingualised dictionaries, but also makes the availability of L1 corpora an important consideration. Learning a foreign language well involves making discoveries about one's L1 habits, and corpora greatly enhance this process of discovery.

A second consideration regards the place of phraseological meaning in the vocabulary acquisition process. Current pedagogical practice is taking collocation far more seriously than it used to, but it is often restricted to the teaching of new vocabulary rather than the teaching of new

meanings of already-learned items. Figurative meanings and delexical forms abound in the language, and when they are used wrongly, the flow of meaning is interrupted, resulting in incomprehension, misunderstanding and, at times, embarrassment or unintended humour. This can be contrasted with the effect of continuing to use *do* instead of a lexical verb such as *carry out*, *prepare* or *arrange*, where meaning is impoverished, not disjointed.

The errors examined in this paper and those preceding it highlight the importance that formulaic and conventional chunks have at all stages of the language learning process. Much idiomatic usage is surprisingly mundane, but it should never be forgotten that the learner perceives linguistic ordinariness in the light of his/her L1. Communicative effectiveness demands that patterns are adhered to, and if this means that patterns must be granted more space in the curriculum, then the implications of this should be considered, because the increased attention to one aspect of language inevitably means that another one loses out. Most importantly, however, there is a need to revisit collocation from the non-native speaker's point of view, paying particular attention to delexical collocates whose semantic motivation may be difficult to appreciate. Only in this way can highly motivated students become less heavily dependent on teacher corrections, and be confident enough to cast aside their deeply held (and strongly felt) convictions of how well-known words actually collocate.

## References

- Charteris-Black, J. (2004) *Corpus Approaches to Critical Metaphor Analysis*. Basingstoke: Palgrave Macmillan.
- Firth, J.R. (1957) 'A Synopsis of Linguistic Theory, 1930-55'. In F.R. Palmer (ed.) (1968) *Selected papers of J.R. Firth 1952-1957*, pp. 168-205. London and Harlow: Longman.
- Lewis, M. (1993) *The Lexical Approach*. Hove: Thomson Heinle LTP.
- Lewis, M. (1997) *Implementing the Lexical Approach*. Hove: Thomson Heinle LTP.
- Louw, W.E. (2000) 'Some implications of progressive delexicalisation and semantic prosodies for Hallidayan metaphorical modes of expression and Lakoffian "Metaphors we Live By"' privately-distributed version of 'Progressive delexicalization and semantic prosodies as early empirical indicators of the death of metaphors' paper read at the 11th Euro-International Systemic Functional Workshop: Metaphor in systemic functional perspectives, University of Gent (Belgium), 14-17 July 1999.
- Martelli, A. (2006) *A Corpus Based Description of English Lexical Collocations Used by Italian Advanced Learners*. In E. Corino, C. Marengo and C. Onesti (eds) *Proceedings XII EURALEX International Congress*. Alessandria: Edizioni dell'Orso, pp. 1005-1011.
- Nesselhauf, N. (2003) *The Use of Collocations by Advanced Learners of English and Some Implications for Teaching*. *Applied Linguistics* 24 (2), 223-242.
- Partington, A. (2003) *The Linguistics of Political Argument: The spin-doctor and the wolf-pack at the White House*. London and New York: Routledge.
- Philip, G. (2003) *Connotation And Collocation: A Corpus-Based Investigation Of Colour Words In English And Italian*. PhD Thesis. Birmingham: The University of Birmingham. Available on-line from <http://amsacta.cib.unibo.it> (accessed: 30 June 2007)
- 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 and L. Zhang (eds) *Proceedings of the Third Interdisciplinary Workshop on Corpus-Based Approaches to Figurative Language*, pp 46-53. Birmingham: University Press, Birmingham. Available on-line from <http://www.cs.bham.ac.uk/~amw/figLangWorkshop/Philip.pdf> (accessed: 30 June 2007)
- Philip, G. (2005b) 'Figurative language and the advanced learner' *Research News: The Newsletter of the IATEFL Research SIG 16*, 16-19 Preprint available on-line from <http://amsacta.cib.unibo.it> (accessed: 30 June 2007)



- Philip, G. (2006) Metaphor, the dictionary, and the advanced learner. In E. Corino, C. Marelllo and C. Onesti (eds) Proceedings XII EURALEX International Congress. Alessandria: Edizioni dell'Orso, pp. 895-905. Preprint available on-line from <http://amsacta.cib.unibo.it> (accessed: 30 June 2007)
- Philip, G. (2007) "...and I dropped my jaw with fear" – The role of corpora in teaching phraseology. Paper read at the 7th Teaching and Language Corpora Conference (TaLC) 2006. Paris, 1-4 July 2006. Preprint available on-line from <http://amsacta.cib.unibo.it> (accessed: 30 June 2007)
- Philip, G. (in press) 'Metaphor and the language learner.' In L. Cameron and G. Low (eds) Researching and Applying Metaphor - 10 years on. Amsterdam: John Benjamins. Preprint available on-line from <http://amsacta.cib.unibo.it> (accessed: 30 June 2007)
- Sinclair, J.M. (1991) Corpus, Concordance, Collocation. Oxford: Oxford University Press.
- Thornbury, S. (2002) How to Teach Vocabulary. London: Longman.