

Part 1

# Formulaic Language in Acquisition and Pedagogy



## Chapter 2

# The Development of Collocation Use in Academic Texts by Advanced L2 Learners: A Multiple Case Study Approach

Jie Li and Norbert Schmitt  
*University of Nottingham*

### Introduction

It is now generally agreed that the native-like use of collocations (word combinations such as *heavy smoker*, *make a speech*, *bitterly cold*) is an important element of proficient language use (e.g. Sinclair, 1991; Wray, 2002). However, researchers have found that L2 learners rely heavily on creativity so as to produce expressions which are simply not used by native speakers (Pawley & Syder, 1983; Wray, 2002). Skehan (1998) and Foster (2001) claim that non-native speakers, unlike native speakers, generate a great proportion of their language from rules instead of lexicalized routines. Native speakers use conventional expressions to convey meaning, while learners often express meaning with unidiomatic combinations of words.

A number of studies (Granger, 1998; Howarth, 1998; Nesselhauf, 2003) have shown that even advanced L2 learners often experience problems with collocations in written English. For example, Granger (1998) used a corpus-based approach to look at *-ly intensifier + adjective* collocations automatically extracted from advanced French learners' academic essays and similar essays written by native English students. She found that one type of intensifier, that is, 'boosters' (e.g. *deeply*, *strongly*, *highly*) were underused by French learners compared with the frequency (i.e. the number of types and tokens) of those used by natives. She then concluded that advanced French learners of English did use collocations, whereas they tended to underuse native-like expressions but overuse those unidiomatic word pairs which have direct L1 translation equivalents.

Using a frequency-based statistic approach, Lorenz (1999) also investigated *intensifier-adjective* collocations in "expository-argumentative" texts

produced by advanced German learners and native British students. By calculating association measures of collocations and type-token ratios, he found that advanced German learners of English had smaller repertoires of collocations (as measured by type-token ratio) and overused a limited number of high frequency collocations (as measured by t-score and MI).

Building on Lorenz's statistical approach, one of Siyanova and Schmitt's (2008) three studies used corpus-based frequency data and mutual information statistics (MI) to investigate adjective-noun collocations in advanced Russian learners' and native university students' written English. By consulting the BNC for counting frequency and calculating the MI value of each collocation, they found that only 45 per cent of the collocations used by advanced Russian learners in their writing texts were appropriate (i.e. frequent and strongly associated English word combinations).

Following a phraseological approach, Howarth (1998) focused on restricted verb-noun collocations (e.g. *make a claim*, *reach a conclusion*) identified from native and advanced non-native academic written corpora. Based on the norms established in native speaker writing, he reported that advanced non-native MA students employed about 50 per cent fewer restricted collocations than natives. He also found that approximately 6 per cent of collocations produced by advanced learners are non-conventional. Based on Howarth's analysis, it seems that among the three collocational groups (i.e. restricted collocations, free collocations, and idioms), restricted collocations are most problematic for advanced non-native learners.

Another more comprehensive study which explored advanced German speaking learners' verb-noun collocation (e.g. *take a break*, *shake one's head*) is that of Nesselhauf (2003). Like Howarth, she also adopted a phraseological approach and classified collocations into three groups, namely, free combinations (e.g. *want a car*), collocations (e.g. *take a picture*) and idioms (e.g. *sweeten the pill*). She found that learners made the greatest proportion of errors with collocations (79 per cent), followed by free combinations (23 per cent) and idioms (23 per cent).

As can be seen, the existing studies all used a corpus-based native versus Non-native comparison to investigate learners' collocation use and identify gaps between these two populations. In general, three main approaches have been employed to define and identify collocations in written texts. One has studied all word combinations of a particular grammatical form (e.g. -ly amplifier + adjective), regardless of whether they are 'idiomatic' or not (Granger, 1998). A second is the so-called phraseological approach, represented by Howarth (1998) and Nesselhauf (2003, 2005). Borrowing the Russian School's definition and classification of phraseology (Cowie,

1998), collocations are typically identified according to two defining criteria: semantic opacity – the degree to which words are used with their ‘dictionary’ meanings, and fixedness – the degree to which elements of a phrase can be substituted. A final approach uses occurrence frequency of word combinations within the investigated corpus as identification criteria. Thus, Lorenz (1999) and Siyanova and Schmitt (2008) compared word pairs in non-native and native equivalent corpora and used statistical ‘association measures’ to identify which pairs were characteristically idiomatic.

Although different approaches have been employed to identify and define collocations, they point to the same conclusion: L2 learners have difficulties with collocation use in their language production. However, existing studies have largely been descriptive in nature, and tend to focus on one-off compositions produced by learners. Little research has focused on an empirical analysis of L2 learners’ collocations over time, which could inform about how collocational knowledge develops. A small number of longitudinal studies have been undertaken to investigate the role of formulaic language improvement in young L2 learners’ language acquisition (e.g. Wong-Fillmore, 1976; Huang & Hatch, 1978). Apart from Adolphs and Durow’s (2004) longitudinal study of two L2 postgraduates’ three-word formulaic language improvement in spoken English, few studies have done the same for advanced L2 learners’ improvement of formulaic language.

The only truly reliable way to identify patterns of development in the use of collocations by L2 learners is to conduct longitudinal studies of the same learners over time. This study attempts to do this using a multiple case-study approach. The purpose is to provide a rich and detailed description of several individual learners’ use of collocations over a period of one academic year. We are also interested in how the individual results combine into group results. As our goal is descriptive, we begin with no formal research questions. However, the following general questions helped to focus the investigation:

1. Will advanced Chinese L2 learners improve their collocation use in academic writing assignments over a one-year study abroad postgraduate programme? Are the collocations used by Chinese students similar to those used by published authors?
2. Are the statistical measures of collocations we use valid for the investigation of collocation improvement over an academic year? To what extent we can put our faith in the statistical results of group patterns of collocation development?

## Methodology

### Participants

The four participants were female Chinese postgraduates, on a one-year MA programme in English Language Teaching (here after ELT) in the School of Education at the University of Nottingham. All of them were English majors from China, with ages ranging from 26 to 29. Their English language learning experiences in China were similar in that their exposure to target language was mainly from non-native teacher-dominated classroom instruction, which was generally grammar-based and input-poor. The participants had similar career plans and expectations, namely returning to China to start teaching in colleges or universities. Overall, the four participants were advanced English language learners, who received similar English language training in China, and were exposed to the same L2 environment at a British university. Details of the individual participants are shown in Table 2.1 as follows:

**Table 2.1** Participant's Personal Details

Participant	Age	Education background	Teaching experience	IELTS/TOEFL score
LH	29	Technical College	5 years	6.5, Writing: 6.0
TT	26	Bachelor's Degree	4 years	640 (TOEFL), Writing: 5.5
WL	27	Technical College	5 years	6.5, Writing: 6.0
YJ	27	Bachelor's Degree	5 years	6.5, Writing: 6.0

Since the scores of IELTS and TOEFL are not directly comparable, it is impossible to compare TT's English language competence to the other three participants. Nevertheless, similar to the other members of the participant group, she is a proficient English language user on the basis of her TOEFL marks.

### The learner corpus

The learner corpus consisted of 36 academic writing assignments (including eight essays and one dissertation for each participant) written over a period of one academic year (i.e. three terms). Since the four participants are all from the one-year MA programme in ELT, their writing requirements were the same except for the coursework for elective modules. This MA course is

comprised of two core modules: Applied Linguistics and Syllabus Design & Methodology; four elective modules and a final dissertation requirement. Each module requires a 3,000 word essay for coursework apart from core modules, which require 6,000 words (i.e. two essays of 3,000 words each). The word count requirement for the dissertation in Term 3 is 12,000 words. Overall, each participant is required to produce 12,000 words in each of three terms over the course of 12 months.

Developing the corpus involved collecting each text in electronic form, cleaning it (i.e. removing unnecessary parts: titles, headers, footers, captions, and reference list), and categorizing it according to the term it was written. The resulting corpus contained 149,587 running words (tokens) and 7,259 types, which was divided into three subcorpora: Term 1 – 50,376 running tokens, Term 2 – 48,530, and Term 3 – 50,681. The three subcorpora are, therefore, directly comparable in terms of text length and text style.

### **The BNC academic written corpus**

The academic written sub-set of the BNC World Edition (2000) was used as the ‘proficient writer’ comparison corpus. It consists of 501 texts totaling over 16 million words, selected from books and journal articles in the six disciplines proposed by Lee (2001): humanities/arts, medicine, natural science, politics/law/education, technical/engineering, and social science.

### **Procedure**

All adjective-noun combinations were extracted from the learner corpus in the following manner. The corpus was searched for the 187 nouns from Sublist 1 of Coxhead’s (2000) Academic Word List (AWL), and those selected which were used by at least one participant over time (at least in two of the three terms). Then WordSmith 5.0 was used to locate adjacent adjective collocates for each of these recurring academic nouns. Collocations were excluded from analysis if they included one of the following constituents: hyphenated adjectives (e.g. *corpus-based approach*), pronouns, possessives, determiners, numbers/ordinals, adjectives to signify nationalities (e.g. *Chinese, English*), and terminology (e.g. *Lexical Approach, Universal Grammar*). This selection procedure produced 41 nouns, leading to 147 different adjective-noun collocation types in Term 1, 95 in Term 2, and 107 in Term 3, a total of 494 collocation tokens and 299 collocation types.

The number of collocate types (i.e. different collocations) and tokens (i.e. occurrences of each type) produced by each participant across terms was counted and recorded. For example, the node noun *role* was used with different adjectives by participant WL in her academic texts across three terms, that is, *important* in Term 1, *central* in Term 2, and *key*, *potential*, *critical*, *significant* in Term 3.

Based on these frequency counts, the type-token ratio (TTR) of each collocation type was calculated for all of the four participants. The t-score and MI value for each collocation type was also calculated. Since it is claimed that low-frequency collocations jeopardize the reliability of all association measures (Manning & Schütze, 1999; Evert & Krenn, 2001), all the extracted collocations with less than four occurrences in the BNC academic corpus were excluded from MI and t-score calculations.

Various researchers define their cut-off points differently: Manning and Schütze (1999) suggest a minimum of three occurrences; Stubbs (2001) five occurrences; Church and Hanks (1990) five occurrences. We used a cut-off point of four to include as large a set of learner collocations as possible.

To explore the four participants' collocational development pattern over the period of 12 months, the TTR, t-score, and MI values for each participant were then averaged within each term and these averages compared across terms. Finally, to explore the development of the strongly-associated collocations preferred by expert writers, the adjective-noun collocations were ranked into different bands according to their MI values.

## Results

The value of case-studies is the elicitation and analysis of rich data, and so we will report both the group results and the results of each individual participant.

### **Participants' overall collocation use**

#### ***Group***

494 adjective-noun collocation tokens were identified from the learner corpus, made up of 299 types. Of these 299 types, over 40 per cent can be considered frequent and strongly-associated (which we will term *robust* in this section), at least according to the criteria of appearing four or more



**Table 2.2** Participant Group's Overall Collocation Use

Adjective-noun collocations	Tokens (total)	Types (total)	Tokens (Term 1)	Types (Term 1)	Tokens (Term 2)	Types (Term 2)	Tokens (Term 3)	Types (Term 3)
F<4	143	128	65	62	35	33	43	38
F≥4 & MI>3 & t-score>2	283	123	112	67	81	42	90	54
Total	494	299	198	147	142	95	154	107
% of robust collocations	57.3	41.1	56.6	45.6	57.0	44.2	58.4	50.5

times in the 16 million-word BNC academic subcorpus and having association figures of MI>3 and t-score>2 (Table 2.2). On the other hand, there was a similar percentage of rarely-occurring combinations, that is, appearing less than four times in the BNC subcorpus. However, in terms of instances of use (tokens), the participants used the robust collocations considerably more often than the infrequent ones: TTR of 0.43 for robust collocations in comparison to 0.90 for infrequent ones. On average, each robust collocation type occurred more than twice in the 36 academic writing texts.

The table also shows how the participants' use of robust collocations developed over the academic year. Although the number of types used declined after Term 1, the percentage of robust types used remained essentially the same from Term 1 to Term 2, and then increased slightly by the time the dissertation was written up in Term 3. In terms of tokens, the number of robust collocations used over the three terms showed a similar pattern as that of types, while the percentage of robust tokens only ranged from 56.6 per cent to 58.4 per cent during the academic year, and so was relatively stable. Overall, after the one-year exposure to an English academic environment, it appears that the participant group as a whole displayed little if any improvement in the number of robust adjective-noun collocation types or tokens produced in their academic writing.

### *Individual participants*

Although the group results showed little improvement, an analysis of the individual participants' results shows quite varied behaviors. Table 2.3 shows that although all participants used about 40 different collocation types in Term 1, LH and WL used steadily fewer over the year, while TT dropped in

**Table 2.3** Participants' Use of Robust Collocations over Three Terms

Participant	Adjective-noun collocation	Term 1 (tokens)	Term 1 (types)	Term 2 (tokens)	Term 2 (types)	Term 3 (tokens)	Term 3 (types)
LH	F $\geq$ 4 & MI $>$ 3 & t-score $>$ 2	20	17	16	13	9	8
	Total	44	40	36	32	32	27
	% of robust collocations	45.5	42.5	44.4	40.6	28.1	29.6
TT	F $\geq$ 4 & MI $>$ 3 & t-score $>$ 2	18	17	18	12	22	19
	Total	43	41	32	22	38	33
	% of robust collocations	41.9	41.5	56.3	54.5	57.9	57.6
WL	F $\geq$ 4 & MI $>$ 3 & t-score $>$ 2	36	19	20	13	17	11
	Total	57	39	29	21	25	17
	% of robust collocations	63.2	48.7	69.0	61.9	68.0	64.7
YJ	F $\geq$ 4 & MI $>$ 3 & t-score $>$ 2	38	28	27	19	42	27
	Total	54	43	45	37	59	42
	% of robust collocations	70.4	65.1	60.0	51.4	71.2	64.3

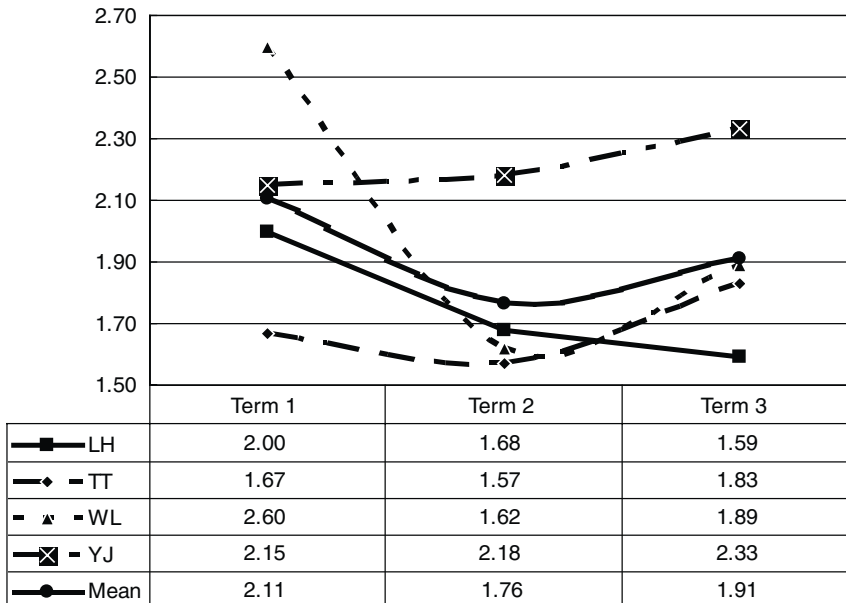
Term 2, but recovered somewhat in Term 3. YJ remained relatively stable in the number of types she used through the year. Regarding the number of collocation tokens used by the four participants across three terms, it displays a rather similar development trend to that of collocation types apart from that of YJ. She tended to use more collocation tokens (59 in Term 3 compared with 54 in Term 1) by the end of the academic year, largely due to frequent repetition (with TTR 0.71 in Term 3, and 0.80 in Term 1). It is also interesting to note the percentage of robust collocations used. In conjunction with her reduced diversity of collocation types/tokens, LH also dropped in the percentage of robust collocations, both types and tokens. Overall, her mastery of these collocations seems to have deteriorated over the year. Conversely, although WL declined in the number of collocation types/tokens used, her percentage of robust collocation types increased over the year, while her percentage of robust collocation tokens reached the peak in Term 2, then dropped slightly in Term 3. She therefore used

fewer types over time, but a greater percentage of those types were similar to those used by proficient English writers. YJ had a dip in Term 2, but ended up in Term 3 essentially where she began in Term 1, both in number of types and percentage of robust types/tokens. Thus, her usage of collocations was relatively stable over the year. TT also had a dip in numbers of types/tokens produced in Term 2, but her percentage of robust collocations (both types and tokens) steadily increased over the three terms. Overall, her figures indicate gradual improvement in collocation mastery.

### Development in the diversity of adjective-noun collocations produced

The average number of adjective types used to describe academic nouns provides a general measure of the diversity of adjective-noun collocations produced. The group mean result in Table 2.4 exhibits U-shaped behavior, with the Term 3 figure not recovering to the Term 1 figure. However, this group average does not show the substantial differences between the individual participants. In fact, the group profile only serves to disguise the very

**Table 2.4** Average Number of Adjective Types per Noun across Three Terms



real differences in the participants' individual development of adjective variation.

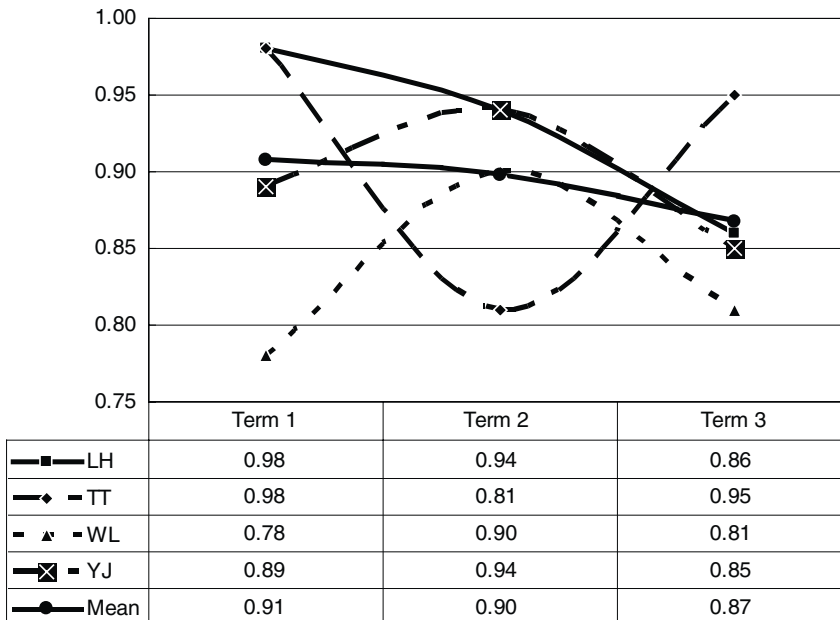
LH started with an average of two adjective types in the first term of her academic year, followed by a consistent decrease from 1.68 to 1.59 in the next two terms afterward. This continuous decline in the mean number of adjective types over the course of three terms indicates that LH used less diverse adjective-noun collocations (about 20 per cent less) by the end of her study abroad programme. In contrast, participant YJ showed an opposite developmental trend in the use of adjectives to describe academic node nouns in her academic writing tasks over the course of three terms. At the beginning of the academic year, an average of 2.15 adjective types were used. This figure rose to 2.18 and 2.33 respectively in the following two terms. The steady increase suggests that YJ used slightly more diverse (approximately 8.4 per cent more) adjective-noun collocations by the end of her MA course.

Both TT and WL experienced a decrease in the number of adjective types from Term 1 to Term 2, and a substantial increase from Term 2 to Term 3, although their developmental profile is very different. WL's employment of adjective types dropped sharply from 2.6 to 1.62 (about 37.7 per cent less), followed by a substantial rise of nearly 17 per cent from 1.62 to 1.89. This left her using less diversity of adjective-noun collocations over the course of the year. On the other hand, TT initially experienced a slight decline of approximately 6 per cent from 1.67 to 1.57, followed by a rise of 16.6 per cent from 1.57 to 1.83. Unlike WL, by the end of the academic year, TT used more various adjective-noun collocations (9.6 per cent more) compared with those used in Term 1.

### **Changes in the repetition of adjective-noun collocation**

TTR value can provide indication of the repetition frequency of collocation use. Table 2.5 shows the TTR value of target adjective-noun collocations for each participant and for the participant group as a whole. The TTR pattern for the group shows a rather stable and subtle decline from 0.91 to 0.87 over the academic year. This steady drop in TTR value over time suggests slightly more repetition of collocation by the four participants as a whole. However, as the decrease is less than 4.5 per cent, it is probably not particularly meaningful.

Of more interest is the individual behavior, which again varies substantially among the participants. The only participant with a profile which in

**Table 2.5** Type-Token Ratios of Adjective-Noun Collocations across Three Terms

any way resembles the group profile is LH, and even here the rate of decrease is much more extreme than the group profile. She started with a TTR of 0.98 in Term 1, dropping to 0.94 and 0.86 in Terms 2 and 3 respectively. This steady decline (nearly 12 per cent decrease) in the TTR suggests that LH tended to repeat collocations more often at the end of the academic year.

Unlike LH, participant TT's TTR value underwent a noticeable fluctuation over the course of three terms. Her TTR value began at 0.98 in Term 1, followed by a considerable drop to 0.81 in Term 2, and ending up nearly where she started at 0.95 by the end of her MA course. It is difficult to say what caused the drop in Term 2, other than to note that it was not based on a single aberrant paper, as TT submitted four papers in this term, as did all the participants. Participants WL and YJ share a similar trend, both experiencing a rise of TTR in Term 2, and a drop afterward in Term 3. Overall, WL showed slightly more repetition of collocations than YJ throughout the year.

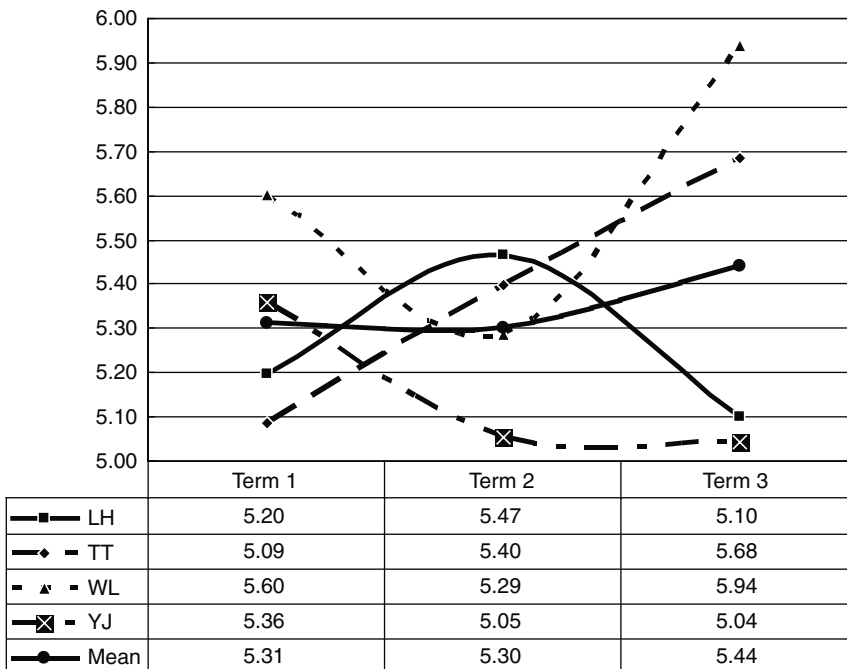
### Development of high-frequency and typical collocation use

We have seen changes in the participants' diversity and repetition of adjective-noun collocation use, and now focus on their production of the type of collocations frequently used by native professional writers in their academic publications, as measured by the t-score statistic and the BNC Academic reference corpus (Table 2.6).

The group result indicates no change in t-score from Term 1 to Term 2 (5.30), and then a slight improvement to 5.44 in Term 3. This suggests that the four participants as a group used more frequent/typical adjective-noun collocations in their dissertations than in their earlier assignments.

However, in this case, this profile accurately represents none of the individual participants' profiles. LH's average t-score in Term 1 was 5.20, which rose to 5.47 and then dropped to 5.10 in Term 3. Thus, over the year, there was no improvement in LH's higher-frequency collocation use. It suggests that LH did not use more native-like adjective-noun collocations which are commonly used by professional expert writers in academic texts. WL's

**Table 2.6** T-scores of Adjective-Noun Collocations across Three Terms



developmental trend is almost a mirror image of LH's with t-score averages of 5.60, 5.29 and 5.94 for Terms 1, 2, and 3 respectively. It seems that by the end of the MA programme, when WL wrote up her high-stakes dissertation, she tended to use collocations with higher frequency levels, compared with those used in her earlier assignments.

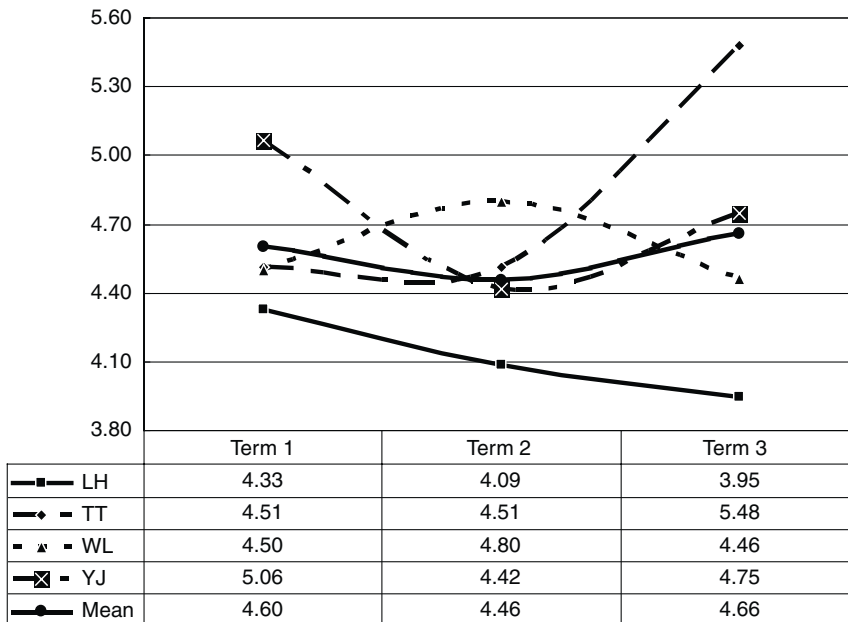
YJ produced a declining profile, dropping from an initial t-score of 5.36 to 5.04/5.05, which indicated a tendency to use adjective-noun collocations which were less frequent and typical by the end of the academic year. Finally, TT produced the type of profile which one might expect given the rich linguistic environment, consistently rising throughout the three terms. She started with an average t-score of 5.09, which thereafter rose to 5.40 and 5.68 in the following two terms. This steady increase suggests that the collocations which occurred in TT's academic writing assignments over the course of the 12-month postgraduate programme were, generally speaking, increasingly more typical of proficient writers.

### **Development of strongly-associated collocation use**

Since MI value is known to emphasize a rather different set of collocations from t-score (Schmitt, in press), a similar analysis was carried out using the MI statistic. It highlights collocations which are typically not very frequent, but which are strongly associated when they do occur (e.g. *tectonic plates*). The group averages (Table 2.7) show a very shallow U-shaped profile, which can probably be best interpreted as no meaningful change across the different terms. But again, the group averages do not accurately represent any of the individual profiles.

LH's collocation use showed a continuous decline in MI values from 4.33 to 3.95 over the course of three terms. This consistent decrease suggests that participant LH tended to use adjective-noun collocations with less association strength in her dissertation, compared with those used in her writing tasks completed in Term 1. By contrast, TT's collocation use displayed an opposite developmental direction. Her MI averages increased over time (4.51, 4.51, 5.48), which indicates TT's use of adjective-noun collocation by the end of her study abroad programme was more native-like, since such strongly-connected collocations characterized the professional writers' academic texts in the BNC sub-corpus.

Although participants WL and YJ underwent completely different developmental trends over the year, they both ended up with lower MI scores in comparison with their initial levels in Term 1. Despite the fluctuations which took place within the length of 12-month postgraduate programme,

**Table 2.7** MI Scores of the Adjective-Noun Collocations across Three Terms

both WL and YJ showed little growth in the employment of adjective-noun collocations with stronger association strength. This suggests that the collocations used by participants WL and YJ did not become more expert-writer-like after the one-year exposure to the academic target language environment.

### **Differences in the distribution of collocations according to MI banding**

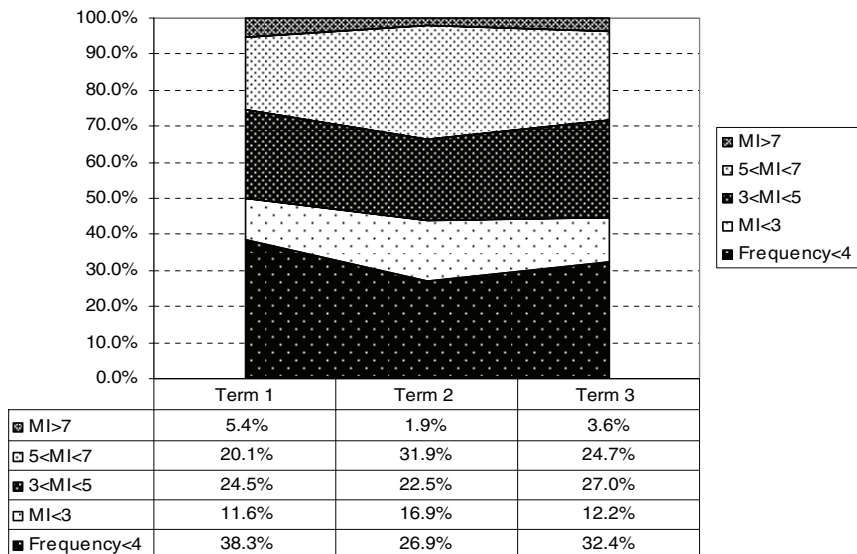
In order to investigate the participant group's collocation patterns in terms of the distribution of strength of association over time, we classified all the adjective-noun collocations used by the four Chinese MA students into five bands on the basis of their MI score values and raw frequency counts obtained from the reference BNC sub-corpus. The MI statistic tends to highlight collocations which are not frequent, but which are highly associated, and are thus likely to be very salient to native speakers (and perhaps proficient non-natives as well). It is thus useful to explore whether the participants began using more of the higher MI collocations, as these may be particularly important in providing a sense of native-likeness to written



compositions (Durrant & Schmitt, in press). We use the term ‘non-associated’ to represent those two-word combinations which are either unattested or with raw frequency of below four in the BNC academic texts. (We recognize that these very infrequent combinations may well be associated, but use this terminology in order to clearly differentiate these combinations from our other categories.) ‘Weak-strength collocations’ are those which occur more than four times in the BNC reference corpus with a MI score of less than 3. ‘Moderate-strength collocations’ have MI scores of  $3 < MI < 5$ , ‘stronger collocations’ have MI strengths of  $5 < MI < 7$ , and ‘extremely strong collocations’ have MI over 7. (All with frequencies of four or more in the BNC sub-corpus.)

When we look at the four participants as a group, the developmental pattern measured by MI values across three terms is shown in Table 2.8. The combined percentage of ‘non-associated’ and ‘weak-strength’ collocation types decreased from nearly 50 per cent in Term 1 to 44.6 per cent in Term 3. This subtle drop indicates that the four participants as a whole used a somewhat lower percentage of less native-like lower-strength collocation types in their academic writing. On the other end of the scale, the participants used only small percentages of ‘extremely strong’ collocation types, and the amount of usage remained about the same over the year. The ratios for both ‘moderate-strength’ and ‘stronger’ collocation categories

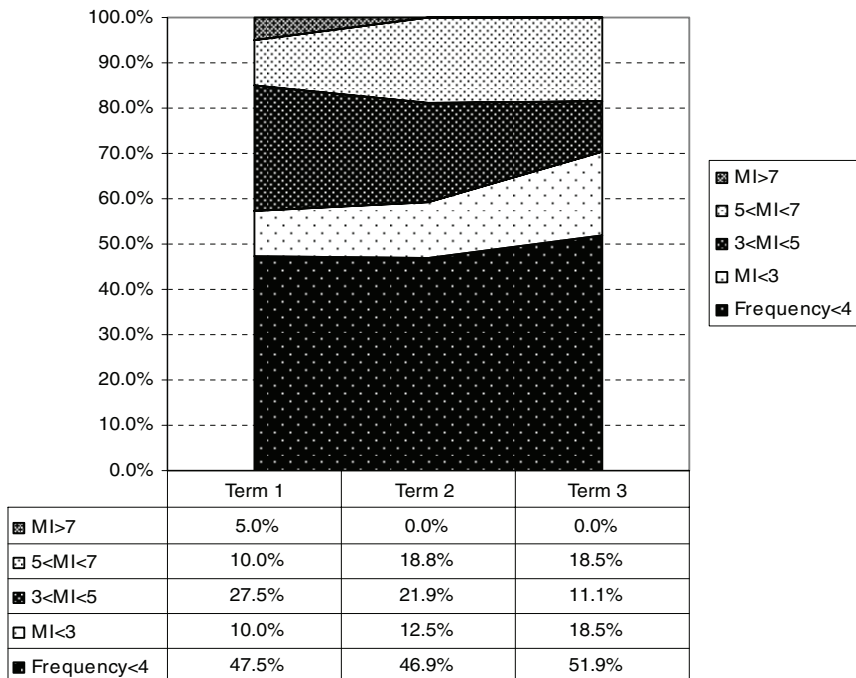
**Table 2.8** Participant Group’s MI Distribution across Three Terms (Collocation Types)

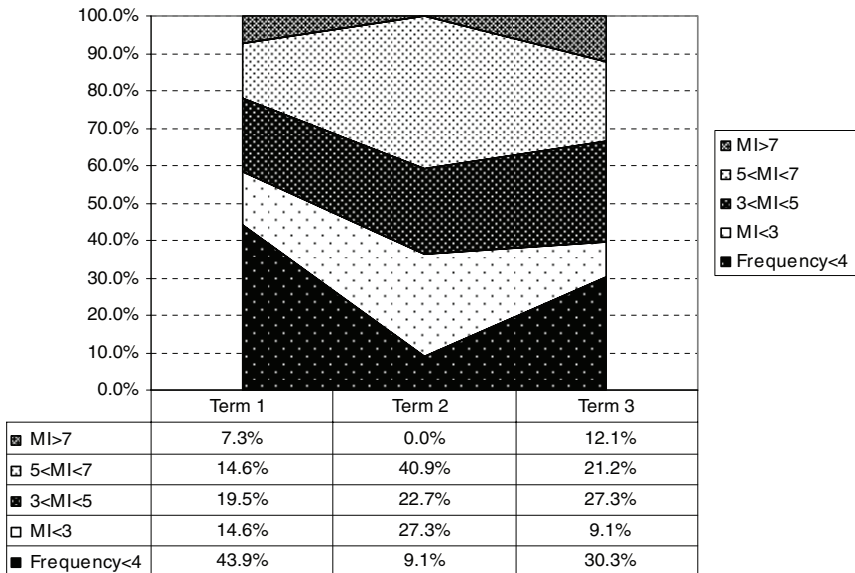


rose slightly from 24.5 per cent to 27 per cent, and from 20.1 per cent to 24.7 per cent respectively. Overall, the group produced slightly less non-associated combinations, and slightly more weak/moderate/stronger strength collocations, although the amount of the most strongly associated collocations (extremely strong collocations) decreased slightly.

This group summary can be compared against the individual results. Table 2.9 shows that LH's combined use of 'non-associated' and 'weak-strength' collocation types has increased over the course of the academic year, from a total of 57.5 per cent in Term 1 to 70.4 per cent in Term 3. The consistently rising figures indicate that LH used increasingly larger proportions of less strongly associated collocation types over the course of the year. As for those collocations with an MI value of above 3, the drop in percentage of 'moderate-strength' collocation types is largely offset by the increase in the 'stronger' collocation types, which would suggest some shift towards the use of more strongly associated collocations. However, working against this conclusion is the disappearance of all collocations with an MI over 7.

**Table 2.9** LH's MI Distribution across Three Terms (Collocation Types)

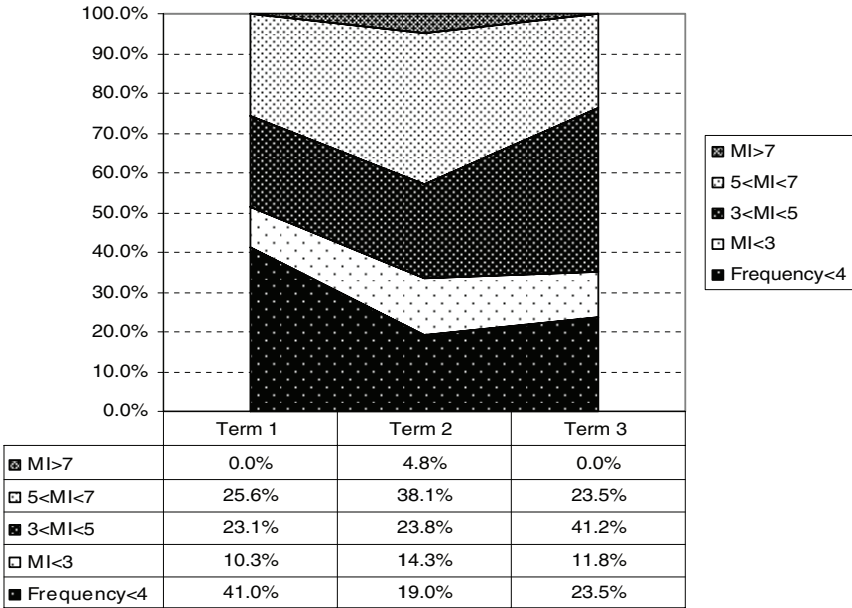


**Table 2.10** TT's MI Distribution across Three Terms (Collocation Types)

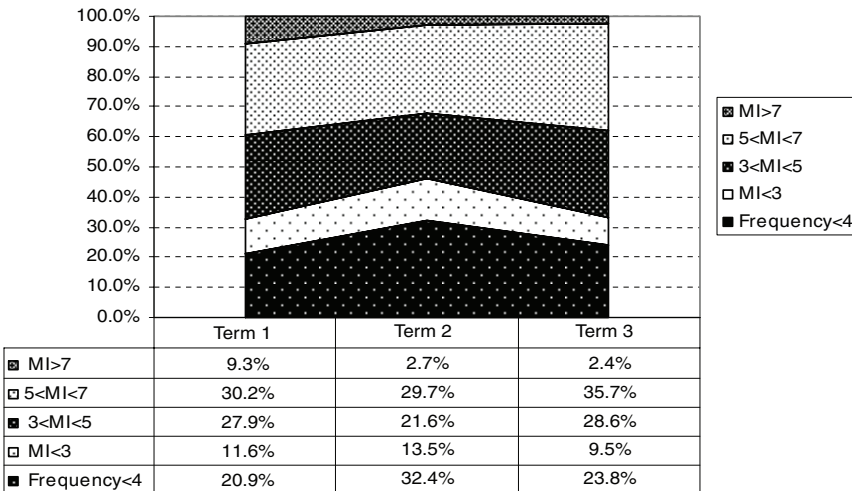
TT's MI distribution profile is completely different from that of LH. As shown in Table 2.10, TT used a smaller percentage of 'non-associated' and 'weak-strength' collocation types in her course assignments written up in Terms 2 (36.4 per cent) and 3 (39.4 per cent) compared to Term 1 (58.5 per cent). Over the same period, she used a larger proportion of the 'extremely strong' adjective-noun collocation types. The most noticeable feature of the graph is the great increase in 'stronger collocations' in Term 2, and then equally dramatic decreases in Term 3. Overall, TT increased in her percentage of collocations with the scores of 3 or above from Term 1 to Term 2, and then remained relatively stable in this regard from Term 2 to Term 3, with just the distribution among the three strongest bands varying. A similar overall summary also describes WL's profile (Table 2.11), although the increase of collocation use in the 3<MI<5 band is noticeable in her profile.

Table 2.12 displays YJ's changes of collocation use measured by MI values over the three terms. If we look at the combined 'non-associated' and 'weak-strength' percentages, we find that although there was a peak in Term 2, the Term 3 figure (33.3 per cent) was essentially the same as at Term 1 (32.5 per cent). Likewise, the moderate-strength percentages remained nearly

**Table 2.11** WL's MI Distribution across Three Terms (Collocation Types)



**Table 2.12** YJ's MI Distribution across Three Terms (Collocation Types)



the same in Terms 1 and 3. The increase in the percentage of 'stronger collocations' (+5.5 percentage points) is mostly accounted for by the decrease in 'extremely strong collocations' (-6.9 percentage points). Overall, there was a spike in the percentage of 'non-associated' and 'weak-strength' collocations in Term 2, but by Term 3 this had been rectified, and YJ ended up the academic year largely where she started in Term 1.

## Discussion

The multiple case-study approach used in this study produced a rich set of data, which was analysed in a number of ways: the number of types and tokens produced, the diversity of adjectives used with each academic noun, the amount of repetition of each adjective-noun collocation, how the collocations produced compared to those in the BNC academic reference corpus according to t-score (largely frequency-based) and MI score (largely based on strength of association), and the degree of strength of association according to a five-band MI rating scale. Overall, the participant group used fewer adjective-noun collocations (both types and tokens) over the course of the academic year, although the percentage of 'robust' collocations (types and tokens) increased slightly. The diversity of collocation (average number of adjective types per academic noun) decreased across the year, during which the collocations were repeated slightly more often in later academic writing tasks. This indicates that the group of four Chinese postgraduates as a whole demonstrated a tendency to use a somewhat smaller group of collocations more repetitively by the end of the 12-month MA programme. In terms of how 'native-like' the collocations were, the group produced a modest increase in t-score over the year, while the MI scores remained relatively static. When the strength of association was explored by MI banding, the group profile was largely similar in Terms 1 and 3. In sum, the statistical approach used in this study was able to show relatively little substantial change in the production of adjective-noun collocations over the course of an academic year. The MI bandings showed some improvement from Term 1 to Term 2 in the decreased use of MI<3 collocations, but this plateaued out and there was no further improvement at Term 3.

However, for most of the measures, the group results fail to adequately represent the individual participants. LH showed fairly consistent decreases in the ability to use collocations on all of the measures. Conversely, TT produced

improved figures on most of the indices. YJ showed fluctuations through the year, but largely ended up in Term 3 near where she started in Term 1, with some indices slightly up and others slightly down. WL produced a mixed set of results, with fewer types produced, but a great percentage of robust collocations. Her diversity of collocations decreased over the year, although so did the amount of repetition, albeit slightly. She used a higher percentage of collocations with higher t-scores, but those same collocations were about the same at the beginning and end of the year in terms of MI score.

Indeed, perhaps the most interesting result of the study is the demonstration that the group figures painted a misleading picture of all the participants. The group results showed little real change, yet the study had one declining student, one improving student, and one with mixed results. There was also one student who ended up nearly where she started, but even here the group figures disguised the amount of variation in YJ's results. Learners typically have a great deal of variation in their acquisition and use results, and this is particularly true of vocabulary (Meara, 1996). The small number of participants in this study also makes it hard for the variation to be 'evened out'. Nevertheless, these results provide a warning to researchers of formulaic language to be careful about generalizing individual behavior from group averages. It may be that the acquisition and use of formulaic language is so idiosyncratic that group averages will have difficulty in making useful statements about any of the individuals within the group. In other words, the group scores may serve only to average out all of the variation inherent in the group, and thus provide a misleadingly 'smooth' representation of what might be quite different behaviors for all of the participants involved.

This is clearly seen in this study, where the participants were very similar to each other. They were all Chinese, female, and of a similar age. They went through the same school system, and attended mainly the same MA courses during their stay in Nottingham. They (with the exception of TT who took the TOEFL) had the same IELTS overall score (6.5) and the same writing score (6.0). Thus we might expect that they would develop the adjective-noun collocations in similar ways. Yet in spite of this, they usually ended up with four quite different profiles in each of the measures. If such a small and homogeneous group as this demonstrates such varying behavior, then it appears that researchers will need to be cautious in their approach to group data in formulaic language research. This point has already been acknowledged by Howarth (1996), who argues that researchers

can lose opportunities for identifying significant differences among learners' processing mechanisms by extracting an average performance from a corpus of various non-native writers. He also goes one step further, and claims that non-native language proficiency is best researched by means of small-scale manual analysis (such as carried out in this study).

This raises the interesting issue of how much variation is inherent in the acquisition and use of formulaic language by second language learners, compared to the amount of consistency among learners. On all of our measures there was variation, but without firm benchmarks, it is difficult to interpret those variations. For example, the average t-scores of the collocations moved between 5.00 and 6.00 for our participants. But is this just normal fluctuation and not meaningful? Perhaps it takes 2–3 full points movement to indicate a truly meaningful change in collocation behavior? Without such benchmarks, researchers can describe the variation, but it is difficult to know if the variation represents real change. Unfortunately, to our knowledge, there are no established benchmarks for t-score, MI, or types and tokens which show the degree of change which indicates real improvement or decline for those measures. Indeed, for t-score and MI, the advice seems to be that they are best used for *ranking* collocations against each other, rather than providing absolute measures of strength of association (Stubbs, 1995; Manning & Schütze, 1999). Most of the measures in this study have been widely used in the study of formulaic language, but it has largely been *descriptive* up until now. If we are to use them to inform about the *acquisition* of formulaic language, the field will either need to somehow establish benchmarks to work against, or to move to alternative methodologies. One of these methodologies that worked well in a single case study of the acquisition of formulaic language was expert rating panels (Li & Schmitt, in press).

The present study has attempted to use a statistical approach to investigate a longitudinal learner corpus in order to identify the learners' improvement in collocation use over the course of one academic year. This is different from previous research which used statistical measurements to explore the differences in the collocation use between native and non-native speakers (e.g. Lorenz, 1999; Durrant & Schmitt, in press). Such research has successfully used association measurements of collocation (i.e. t-score and MI) as valid discriminators between the two populations with different proficiency levels, but this study seems to suggest that they are less efficient in identifying improvement in collocation of advanced L2 learners over a relatively short period of time. For example, Lorenz

(1999) and Durrant and Schmitt (in press) found that proficient writers prefer collocations with high MI values but relatively low t-scores, but that the reverse is true for less proficient non-native writers. The present study cannot clearly distinguish Chinese L2 postgraduates' developmental phases of collocation use within the 12-month period of investigation. This is not surprising when we take into account the relatively short-term investigation undertaken. One academic year may simply not be long enough for advanced level Chinese MA students to show meaningful improvement in collocation use.

The present study is not without its limitations. Firstly, the BNC academic sub-corpus is not a parallel reference corpus, which consists of research articles, books, and book sections from all disciplines. Since a number of studies (e.g. Cortes, 2004; Biber, 2006; Hyland, 2008) have shown that there are considerable variations in the frequency of forms and structures across different types of academic writing texts, the use of adjective-noun collocations may vary from discipline to discipline. If a reference native corpus containing the specific texts from students' MA reading list were compiled, it would be a more parallel comparison. In addition, it should be noted that the size of the longitudinal learner corpus is relatively small, consisting of only four participants' writing tasks written within an academic year. This case-study approach has allowed a detailed exploration of individual progress, but a larger-sized longitudinal learner corpus built up over a longer period may yield a more insightful account of L2 learners' collocation development over time.

## References

- Adolphs, S., & Durow, V. (2004). Social-cultural integration and the development of formulaic sequences. In N. Schmitt (Ed.), *Formulaic sequences* (pp. 107–26). Amsterdam: John Benjamins.
- Biber, D. (2006). *University language: A corpus-based study of spoken and written registers*. Amsterdam: John Benjamins.
- Church, K., & Hanks, P. (1990). Word association norms, mutual information and lexicography. *Computational Linguistics*, 16 (1), 22–29.
- Cortes, V. (2004). Lexical bundles in published and student disciplinary writing: Examples from history and biology. *English for Specific Purposes*, 23 (3), 397–423.
- Cowie, A. P. (1998). Phraseological dictionaries: some East-West comparisons. In A. P. Cowie (Ed.), *Phraseology: Theory, analysis and applications* (pp. 145–60). Oxford: Oxford University Press.
- Coxhead, A. (2000). A new academic word list. *TESOL Quarterly*, 34 (2), 213–38.



- Durrant, P., & Schmitt, N. (in press). To what extent do native and nonnative writers make use of collocations? *International Review of Applied Linguistics*.
- Evert, S., & Krenn, B. (2001). *Methods for the qualitative evaluations of lexical association measures*. Paper presented at the 39th Annual Meeting of the Association for Computational Linguistics, Toulouse, France.
- Foster, P. (2001). Rules and routines: A consideration of their role in the task-based language production of native and non-native speakers. In M. Bygate, P. Skehan, & M. Swain (Eds.), *Language tasks: Teaching, learning and testing* (pp. 74–93). Harlow: Longman.
- Granger, S. (1998). Prefabricated patterns in advanced EFL writing: Collocations and formulae. In A. P. Cowie (Ed.) *Phraseology: Theory, analysis, and applications* (pp. 79–100). Oxford: Oxford University Press.
- Howarth, P. (1996). *Phraseology in English academic writing: Some implications for language learning and dictionary making*. Tübingen: Niemeyer.
- Howarth, P. (1998). The phraseology of learners' academic writing'. In A. P. Cowie (Ed.), *Phraseology: Theory, analysis, and applications* (pp. 161–86). Oxford: Oxford University Press.
- Huang, J., & Hatch, E. (1978). A Chinese child's acquisition of English. In E. Hatch (Ed.), *Second language acquisition: A book of readings* (pp. 118–31). Rowley, MA: Newbury House.
- Hyland, K. (2008). Academic clusters: Text patterning in published and postgraduate writing. *International Journal of Applied Linguistics*, 18 (1), 41–62.
- Lee, D. (2001). Genres, registers, text types, domains and styles: Clarifying the concepts and navigating a path through the BNC jungle. *Language Learning & Technology*, 5 (3), 37–72.
- Li, J. & Schmitt, N. (in press). The acquisition of lexical phrases in academic writing: A longitudinal case study. *Journal of Second Language Writing*.
- Lorenz, G. (1999). *Adjective intensification – learners versus native speakers: A corpus study of argumentative writing*. Amsterdam and Atlanta: Rodopi.
- Meara, P. (1996). The classical research in vocabulary acquisition. In G. Anderman & M. Rogers (Eds.), *Words, words, words* (pp. 27–40). Clevedon: Multilingual Matters. Retrieved December 20, 2008, from <http://www.lognostics.co.uk/vlibrary/index.htm>.
- Manning, C. D., & Schütze, H. (1999). *Foundations of statistical natural language processing*. Cambridge, MA: MIT Press.
- Nesselhauf, N. (2003). The use of collocations by advanced learners of English and some implications for teaching. *Applied Linguistics*, 24 (2), 223–42.
- Nesselhauf, N. (2005). *Collocations in learner corpus*. Amsterdam/Philadelphia: John Benjamins.
- Pawley, A., & Syder, F. H. (1983). Two puzzles for linguistic theory: Nativelike selection and nativelike fluency. In J. C. Richards & R. W. Schmidt (Eds.), *Language and communication* (pp. 191–226). New York: Longman.
- Schmitt, N. (in press). *A vocabulary research manual*. Palgrave Press.
- Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford: Oxford University Press.
- Siyanova, A., & Schmitt, N. (2008). L2 learner production and processing of collocation: A multi-study perspective. *The Canadian Modern Language Review*, 64 (3), 429–58.

- Skehan, P. (1998). *A cognitive approach to language learning*. Oxford: Oxford University Press.
- Stubbs, M. (1995). Collocations and semantic profiles: On the cause of the trouble with quantitative studies. *Functions of Language*, 2 (1), 23–55.
- Stubbs, M. (2001). Texts, corpora, and problems of interpretation: A response to Widdowson. *Applied Linguistics*, 22 (2), 149–72.
- Wong-Fillmore, L. (1976). *The second time around: Cognitive and social strategies in second language acquisition*. Doctoral dissertation: Stanford University.
- Wray, A. (2002). *Formulaic language and the lexicon*. Cambridge: Cambridge University Press.