Collocation in beginner learner writing: A longitudinal study

CrossMark

Anna Siyanova-Chanturia*

School of Linguistics and Applied Language Studies, Victoria University of Wellington, Kelburn Parade, Wellington 6140, New Zealand

A R T I C L E   I N F O

A B S T R A C T

The fact that second-language (L2) learners have problems with collocation is widely attested. However, few studies have investigated the development of L2 collocational knowledge longitudinally. Fewer still have employed more than a handful of participants. In addition, almost all studies to date have looked at advanced learners of English. Other L2 proficiencies and backgrounds have, by and large, been disregarded. The lacuna left by this paucity motivated the present study. Thirty-six Chinese beginner learners of Italian wrote a composition at the beginning (Level 1), in the middle (Level 2), and at the end (Level 3) of an intensive course. A small corpus of L2 Italian was compiled with the aim of investigating N + Adj combinations. Analyses revealed that the number of combinations at the beginning and at the end of the course was comparable; the number of higher frequency items, however, was greater in Level 3 compositions than Level 1 compositions. Importantly, analyses showed an increase in strongly associated collocations in Level 3 writings compared to Level 1 writings. Thus, Level 3 compositions contained not only more higher frequency combinations, but also more strongly associated collocations than did Level 1 compositions. Taken together, the study provides new insights into the development of L2 collocational competence.

© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

Recent years have seen a surge in work on collocation in learner corpora. This is hardly surprising given that the appropriate use of collocation is now widely considered to be one of the key prerequisites for proficient language use (Cowie, 1998; Sinclair, 1991; Wray, 2002). Collocations — and other types of multi-word expressions — are ubiquitous in language; they have been found to comprise anywhere from 20 to over 50 percent of spoken and written native-speaker discourse (Biber, Johansson, Leech, Conrad, & Finegan, 1999; Erman & Warren, 2000; Foster, 2001; Howarth, 1998; Sorhus, 1977). Biber et al. (1999) reported that multi-word speech constituted 28% of the spoken and 20% of the written discourse analysed. Erman and Warren (2000) and Howarth (1998) estimated that multi-word speech (collocations, idioms, multi-word verbs, etc.) amounted to 52.3% and 40% respectively of the written discourse they looked at. According to Pollio, Barlow, Fine, and Pollio (1977) and Glucksberg (1989), four multi-word expressions — of one kind or another — are produced by a native speaker in every minute of spoken discourse. Collocation is what makes native speakers' speech idiomatic, fluent and natural. It is also what often renders second language (L2) learners' speech awkward, unnatural and even odd. Indeed, it has been established that L2 learners have problems with collocation in their written and spoken language (Granger, 1998; Howarth, 1998; Nesselhauf, 2003, 2005). Some have argued that L2 learners rely on creativity and make "overliberal assumptions about

---

* Tel.: +64 (0) 463 5922.
  E-mail address: anna.siyanova@vuw.ac.nz.

the collocational equivalence of semantically similar items" (Wray, 2002: 201–202). Others have suggested that collocations present a major problem for L2 learners because collocational competence does not develop in parallel with general vocabulary knowledge (Bahns & Eldaw, 1993). As a result, learners often underuse or misuse native-like expressions and use atypical word combinations instead.

2. Literature review

By far the most common method of investigation of L2 collocational patterns has been to compare word combinations extracted from a written learner corpus with those found in a comparable native speaker (L1) reference corpus, with the aim of evaluating learner collocations qualitatively (correctness and appropriateness) and/or quantitatively (number of L1 combinations vs. L2 combinations). Many of the studies employing this method have focused on learners of L2 English from the same first language (L1) background: Chinese (Chen & Baker, 2010), French (Gilquin, 2007; Granger, 1998; Henderson & Barr, 2010), German (Lorenz, 1999), Hebrew (Laufer & Waldman, 2011; Levitzky-Aviad & Laufer, 2013), Swedish (Bartning, Forsberg, & Hancock, 2009; Forsberg, 2010), and Russian (Siyanova & Schmitt, 2008). Fewer studies have compared the patterns of collocation usage in learners from a number of different L1 backgrounds. For example, Waibel (2008) and Alejo Gonzalez (2010) compared the use of phrasal verbs in argumentative essays written by German versus Italian, and Spanish versus Swedish learners, respectively; while Cross and Papp (2008) looked at Chinese versus German versus Greek learners' use of V + N collocations in writing (also see Paquot, 2010, for a small-scale analysis of collocation use across learner groups representing ten L1 backgrounds). However, it is not only L2 written production that has been of interest to researchers in the context of collocation learning and use; there have also been studies, albeit relatively few, into collocation use in learner speech. For example, Crossley and Salsbury (2011) recorded casual conversations to investigate learners' use of lexical bundles, while De Cock (2004) focussed on two-to-six word recurrent strings in a spoken corpus of L2 English. In another study, Gotz and Schilk (2011) compared three-word lexical bundles in English as a foreign language (EFL) versus English as a second language (ESL) varieties of spoken English. Finally, Bartning et al. (2009) and Forsberg (2010) analysed conventional sequences in advanced L2 spoken French − a rare example of the target L2 being non-English.

Clearly, the studies into L2 collocational knowledge differ in a number of methodological ways. Some have focussed on L2 written production, while others analysed L2 speech. Many compared L2 learners of the same mother tongue against an L1 reference corpus, whereas others compared L2 learners of two or even three different L1 backgrounds. In addition, researchers have focussed on a variety of multi-word expressions: Premodifier + N collocations (Durrant & Schmitt, 2009; Siyanova & Schmitt, 2008), V + N collocations (Altenberg & Granger, 2001; Cross & Papp, 2008), Intensifier + Adj collocations (Granger, 1998), lexical bundles (Chen & Baker, 2010; Crossley & Salsbury, 2011), phrasal verbs (Alejo Gonzalez, 2010; Waibel, 2008), lexical verbs (Granger & Paquot, 2009), and conventional sequences (Bartning et al., 2009; Forsberg, 2010). Learner corpus studies, however, also differ vastly in the way in which the corpus data were extracted and analysed. For example, some have used statistical measures, such as frequency, *t*-score and mutual information to identify and compare collocations in L2 and L1 corpora (Durrant & Schmitt, 2009; Granger & Bestgen, 2014; Groom, 2009; Lorenz, 1999; Siyanova & Schmitt, 2008); others have employed an automated frequency-driven approach (Chen & Baker, 2010), or a frequency index (Crossley & Salsbury, 2011). A different approach altogether has been to use L1 judgements and collocation dictionaries to evaluate the formulaicity, appropriateness, and correctness of L2 collocations (Bartning et al., 2009; Forsberg, 2010; Howarth, 1996, 1998; Foster, 2001; Laufer & Waldman, 2011; Nesselhauf, 2005).

Because learner corpus studies markedly differ in the methods of data extraction and analysis, they correspondingly vary in the formal definition of collocation adopted. Some have relied on native speaker judgements of 'formulaicity' in their definition of collocation (Foster, 2001). Others have defined collocation as word combinations that are necessarily restricted in some way (Howarth, 1998; Nesselhauf, 2005). However, L1 intuitions would inevitably play a role in this approach, too. Other researchers have adopted a purely 'frequency-based' approach in which intuitions and native-speaker judgements play no role (Durrant & Schmitt, 2009). The frequency-based tradition defines collocation as "the relationship a lexical item has with items that appear with greater than random probability in its (textual) context" (Hoey, 1991: 7). That is, two words can be considered a collocation if they are found together more often than their individual frequencies would predict (Jones & Sinclair, 1974; Manning & Schutze, 1999; also see Hoey, 2005; Sinclair, 2004; Stubbs, 1995).

As pointed out by Paquot and Granger (2012), the studies into collocational knowledge in learner corpora have yielded a wealth of interesting findings. However, methodological and conceptual differences, such as different types of corpora used and definitions of collocation adopted, various approaches to data extraction and analysis, a wide range of collocations and other kinds of multi-word expressions investigated, and the type of L2 learner make it difficult to directly compare many of the studies in question. Despite this heterogeneity in both data and methods, nevertheless, a number of common findings can be identified. Arguably, the most commonly reported finding is under-, over- or misuse of target L2 items compared to L1 reference data (Altenberg & Granger, 2001; Chen & Baker, 2010; De Cock, 2004; Durrant & Schmitt, 2009; Gilquin, 2007; Granger, 1998; Granger & Paquot, 2009; Henderson & Barr, 2010; Howarth, 1998; Laufer & Waldman, 2011; Nesselhauf, 2005; Paquot, 2010). For example, L2 learners in Durrant and Schmitt (2009) were found to underuse less frequent, strongly associated collocations, identified using mutual information (*densely populated*), but their use of frequent collocations, identified using frequency and *t*-score (*hard work*), was found to be comparable to that of L1 speakers. Patterns of underuse in L2 writing have also been observed within a specific domain. For example, Granger and Paquot (2009) investigated the use of lexical verbs in English for Academic Purposes writing of advanced EFL learners from a range of L1

backgrounds. Granger and Paquot (2009) reported that learners used lexical verbs in phraseological patterns that were not found in L1 writing. Specifically, learners underused 'academic verbs' (*include, report*) and used 'conversational verbs' (*think, like*) instead. When learners did use academic verbs, they restricted themselves to a limited range, which was in contrast with L1 writing. In addition, patterns of overuse relative to L1 data have also been found across a wide range of L1 backgrounds. In her analysis of academic vocabulary in learner writing, Paquot (2010) reports that some phrasal units under investigation were significantly overused by the learners from all L1 backgrounds analysed (Italian, German, Dutch, Czech, Spanish, Swedish, Finnish, French, Polish, and Russian). For example, the chunk *I think* was used around 24 times more frequently in Swedish essays than in L1 writing. The least amount of overuse was reported in Polish EFL learners, who, nevertheless, used the chunk about three times as often as L1 writers.

Another commonly cited finding is a negative influence of learners' L1 (Altenberg & Granger, 2001; Gilquin, 2007; Granger & Paquot, 2009; Henderson & Barr, 2010; Lorenz, 1999). For example, Lorenz (1999) found that L1→L2 transfer was the most frequent strategy that resulted in the erroneous use of adjective intensifiers (*highly estimated*) by German school and university learners of English. In her analysis of collocations with the verb *make* (*make progress*) in advanced EFL learner writing, Gilquin (2007) reports that learners did not make many errors but underused *make*-collocations, when compared to L1 writing, and limited themselves to those that had a direct equivalent in their L1 (French). Interestingly, L1 positive influence has also been observed. Alejo Gonzalez (2010) and Waibel (2008) analysed the use of phrasal verbs (*point out, sum up*) by advanced Spanish and Swedish, and Italian and German EFL learners, respectively. They found that the learners whose L1 belonged to the same language family as English (Swedish and German) were more likely to produce phrasal verbs and made fewer mistakes using them than the learners from Romance L1 backgrounds (Spanish and Italian).

Further, a number of studies have highlighted the role of immersion-based L2 exposure (Groom, 2009; Siyanova & Schmitt, 2008; Waibel, 2008). For example, in the study with German and Italian EFL learners mentioned above, Waibel (2008) focused on the use of phrasal verbs in L2 writing. Employing a range of quantitative and qualitative techniques, she found that the time spent in an L2 country was a significant predictor of the number of phrasal units produced. A similar observation was made in Groom (2009), who investigated the use of two-to-five word lexical bundles in the writing of two groups of Swedish learners of English: one group with less than one month in an English L1 environment and one with 12 months or more in an English-speaking environment. Having conducted *t*-score and MI analyses, Groom (2009) found that collocational usage and the time spent abroad were positively correlated, in that the texts in the 12+ months group were characterised by a higher degree of collocational accuracy. Despite this finding, however, Groom (2009: 33) concluded that "it remains clear that the process of L2 collocational development is likely to be a slow and occasionally painful one quite irrespective of the linguistic environment in which the learner happens to be immersed".

Finally, an interesting, if somewhat counterintuitive, finding has been reported with regard to the role of L2 proficiency in collocation production. Laufer and Waldman (2011) and Nesselhauf (2005) found that more proficient EFL learners (or those with more years of L2 learning) showed a similar proportion of erroneous collocations as less proficient learners (or those with fewer years of L2 learning), or, indeed, produced more deviant collocations than their less proficient counterparts.[1] For example, employing a combination of collocation dictionaries, corpora and native-speaker judgements, Nesselhauf (2005) found that a third of L2 V + N collocations produced by German EFL learners in writing were erroneous. The most common mistake was a wrong choice of verb. Generally, verbs were misused in restricted collocations (*make a decision*), but also were a source of errors in free combinations (*make a cake*). Interestingly, the learners who studied English between ten and Seventeen years showed a similar proportion of erroneous collocations as those who studied English between five and ten years. This finding goes hand in hand with Bahns and Eldaw's (1993) proposition that collocational competence may not develop in parallel with general vocabulary knowledge. As Laufer and Waldman (2011: 664) argue, "the development of collocation use is slow and uneven", which is why higher proficiency does not necessarily entail higher accuracy with collocations.

Most of the studies reviewed above are based on the data collected at a single point in time. Only a handful of studies — all focussing on a small number of learners — have investigated the development of L2 collocational knowledge longitudinally. In one of the early studies, Li and Schmitt (2009) followed one Chinese ESL learner over the course of one year with the aim of investigating the development of lexical phrases in her written production (eight essays and one MA thesis). The authors report that, over the course of one year during which the learner lived and studied in the UK, the learner acquired 166 new lexical phrases, improved in the degree of their appropriate usage, and gained confidence using them. However, the learner was also found to over-rely on a limited range of phrases, which appeared to be, in the words of Nesselhauf (2005: 69), her "collocational teddy bears".

In a similar study, Li and Schmitt (2010) analysed Adj + N collocations in essays and theses written by four Chinese ESL learners while they were enrolled in a MA programme at a British university. It was found that, over the course of one year, these learners showed little, if any, change in their production of collocations. Analyses also revealed a high degree of variability among the learners, with some students producing more collocations at the end of the one-year course and others producing fewer.

---

[1] We direct the reader to Thewissen (2013) for an interesting discussion of the link between language proficiency and error production (albeit not in the context of collocation development).

In another longitudinal study focussing on L2 spoken production, Crossley and Salsbury (2011) followed six ESL learners from four L1 backgrounds (Spanish, Korean, Japanese, and Arabic) during the period of one year. The analysis of L2 use of two-word lexical bundles in casual conversation recordings showed that the amount of time spent in an English-speaking country had a significant influence on the learners' acquisition and use of lexical bundles. It was also found that with time, the learners began to use lexical bundles at frequencies that approximated those of L1 speakers.

More recently, Yuldashev, Fernandez, and Thorne (2013) analysed three L2 Spanish learners' use of contiguous and dis-contiguous instances of the chunk *es/que* ("[it] is/that") in a corpus of out-of-class instant messaging and blogs (the students were enrolled in a Spanish advanced placement course in a U.S. high school). These authors found that, over the course of 38 weeks, the learners were able to use the chunk in a fixed and increasingly schematic fashion (although some individual differences were also present). Although informative with respect to the use of *es/que* in L2 Spanish, the study's findings may not be generalisable to the use of other frequent chunks and constructions.

Lastly, Li, Eskildsen, and Cadierno (2014) traced one ESL learner's development and use of English motion constructions (*go to Mexico, go down*) over the course of three and a half years. Employing a longitudinal audio—video database of classroom interaction, it was shown that over time, the learner's use of English motion constructions became increasingly productive and varied "with emergent patterns building on previous experience" (p. 612). The study is important in that it looked at an L2 learner who gradually progressed through a range of proficiency levels − from beginner to high intermediate − which allowed for a detailed picture of key developmental patterns to emerge. However, similar to the longitudinal studies reviewed above, Li et al. (2014) is a case study that focused on a very limited number of participants and whose findings may not be easily generalisable to other L2 populations, contexts and environments.

3. Present study

The rationale behind the present investigation stemmed from a number of gaps in learner corpus research. First, the level of L2 proficiency targeted has for most part been upper intermediate or advanced. The notable exceptions are Laufer and Waldman (2011) and Forsberg (2010), who, employing a cross-sectional design, also looked at basic and beginner EFL learners, respectively; and Li et al. (2014) who followed one ESL learner's progression from beginner to high intermediate. Indeed, researchers have called for a greater emphasis on a range of L2 proficiencies (Granger & Bestgen, 2014). With this in mind, the present investigation focussed on complete beginner learners, that is, learners with no prior knowledge of L2. By focussing on beginner, rather than advanced, learners, we hoped to obtain new insights into how collocations are learnt and used in L2 writing. For example, unlike advanced learners, beginners will have much smaller repertoires of single words and phrases.

Second, as follows from the above review, few studies have investigated learner collocation production from a longitudinal perspective. Although recent years have seen a few such studies (Crossley & Salsbury, 2011; Li et al., 2014; Li & Schmitt, 2009, 2010; Yuldashev et al., 2013), longitudinal studies employing more than a handful of learners are rare, such that Paquot and Granger (2012) and Laufer and Waldman (2011) explicitly call for more studies looking at the patterns of phraseological development over a period of time. If one is to investigate L2 collocation developmental patterns, the only reliable way of doing so is by analysing longitudinal data collected from the same group of learners. Data collected at one point in time are limited in what they can tell us about the development of L2 collocation knowledge.

Finally, most learner-corpus based studies have investigated the production of collocation by L2 learners of English. Learners from other L2 backgrounds have been largely disregarded (but see Bartning et al. (2009) and Forsberg (2010) for corpus-based work on advanced L2 French, and Yuldashev et al. (2013) for corpus-based work on advanced L2 Spanish). Given this, the present study aimed to focus on L2 learners of Italian.[2] In the context of the Italian language, one recent study is of relevance. Spina (in press) investigated the use of general (N + Adj: *scuola elementare* "elementary school"; V + N: *fare una domanda* "ask a question") and academic (N + Adj: *prospettiva teorica* "theoretical perspective"; V + N: *affrontare il tema* "address the issue"; N + Prep + N: *ambito di studio* "field of study") multi-word expressions in L1 and advanced L2 essays. It was found that, compared to L1 speakers, L2 learners tended to overuse general expressions within the highest frequency ranks; they also exhibited a limited ability to vary their lexical choices, when compared to the native speaker data. Further, the analysis showed that L2 learners' repertoire of academic expressions was only half as rich as that of native speakers (and, indeed, many frequent academic expressions were never used in any of the L2 essays analysed). These patterns of over- and underuse of multi-word speech in L2 Italian offer support to earlier studies on L2 English (Durrant & Schmitt, 2009; Granger & Paquot, 2009; Paquot, 2010).

Although we do not believe that learning Italian collocations is fundamentally different from learning English (French, German, or Spanish) collocations, we do believe that it is important for a wider range of L2s to be represented alongside English, in the field of second language acquisition, in general, and in learner corpus research, in particular (also see Ortega [2009] for a similar stance).

---

2 A number of researchers have contributed to our better understanding of the acquisition and use of collocations and other phrasal elements in L1 and L2 Italian. For example, see Siyanova-Chanturia and Spina (2015), Spina (2010, 2014, in press), as well as *Dizionario Combinatorio Italiano* (Lo Cascio, 2013) for work on Italian collocations. The latter is a two-volume combinatory dictionary of Italian. It contains around 220,000 Italian word combinations and is aimed at advanced L2 learners and native speakers of Italian.

3.1. Research question

The main research question of the study concerned the quality and quantity of N + Adj combinations produced by beginner learners of L2 Italian over a period of time:

> Do Chinese learners of L2 Italian improve in their use of appropriate N + Adj combinations over time? Specifically, will the compositions written at the end of a language course contain a greater number of frequent and strongly associated collocations, and fewer instances of infrequent word combinations, than the compositions written at the beginning of the course?

We hypothesised that more experience with the target language should lead to a more native-like, idiomatic output. We expected more advanced compositions to contain a greater number of frequent and statistically significant collocations, as defined in the 'frequency-based' tradition (Hoey, 2005; Sinclair 2004; Stubbs, 1995, see Method), and fewer very low frequency word combinations, than less advanced compositions.

3.2. Method

The present investigation focused on Chinese non-native writers' use of L2 collocation, as defined in the 'frequency-based' tradition (Hoey, 2005; Sinclair, 2004; Stubbs, 1995). That is, we define collocation as "the relationship a lexical item has with items that appear with greater than random probability in its (textual) context" (Hoey, 1991: 7), where two words can be considered a collocation if they are found together more often than their individual frequencies would predict (Jones & Sinclair, 1974). Because we believe collocation to be, by and large, a statistical, probabilistic phenomenon, we adopted a frequency-based approach as one being least subjective and not dependent on native-speaker judgements, intuitions, or inter-rater agreement. In our approach to collocation enquiry, we followed Siyanova and Schmitt (2008), Durrant and Schmitt (2009), Lorenz (1999), and others who have used a similar frequency-based approach.

3.2.1. Participants

For the purpose of the present investigation, we used short compositions written by 36 Chinese learners of Italian (21 females, age range 19—23). These learners came to Italy in order to study the Italian language. The learners came from across Mainland China and can be said to be of comparable socio-economic background. At the time of data collection, the learners were full-time students enrolled in an intensive Italian language programme at a university in Italy. The students had around six hours of Italian from Monday to Friday, on top of being exposed to Italian on a daily basis. None of the students knew Italian prior to commencing the course. Thus, the students were homogeneous in terms of their age, socio-economic status, L1 (Mandarin speakers), L2 level (beginners), time spent in the L2 country (all students had arrived in Italy prior to the start of the course). In addition, during the course, the students were taught by the same group of Italian language teachers and, hence, experienced a comparable teaching style. In terms of the knowledge of other foreign languages, all learners had studied English at school. No English language proficiency data were gathered, but it can be assumed that the learners had comparable English language proficiency, as all had studied English as a compulsory subject for a similar period of time.

The duration of the intensive course was around five months. Three exams were incorporated into the curriculum; each was made of a series of vocabulary and grammar tests and a short composition (150—200 words). The three exams were taken at a roughly equal interval, approximately seven, 14 and 21 weeks into the course. Of interest to us is the comparison between the data collected at the beginning of the course (Level 1, week seven) and the data collected at the end of the course (Level 3, week 21) in order to compare the greatest developmental gains (see Crossley & Salsbury, 2011, for a similar approach).

3.2.2. Learner corpus

The compositions were written in class, under time pressure with no dictionaries allowed. A small learner corpus of L2 Italian was compiled on the basis of the three 150—200 word compositions written by each student. The topics of the compositions varied between Levels 1—3, but all can be said to be comparable, in terms of the genre and register, short narratives, such as, hobby, Italy versus China, holidays, travel, and childhood.[3] The learners were free to choose one topic from a few provided for Level 1, 2, and 3 compositions. The topics were equally distributed within the three levels, but differed across the three levels, as, arguably, the learners would have found it repetitive to have to write a composition on the same topic two or three times. The final corpus (108 compositions or 17,912 words in total) was made of three sub-corpora: Level 1 compositions (n. = 36; 5282 words), Level 2 compositions (n. = 36; 6263 words), and Level 3 compositions (n. = 36; 6367 words). Undoubtedly, the corpus is small and any generalisations to other learner populations can be tentative at best. However, the unique features of the Italian corpus — such as its longitudinal character and the homogeneous nature of the learners who contributed to it — make it, despite its modest size, a well-suited tool for studying the development of collocational knowledge in beginner learner writing. As pointed out by Paquot and Granger (2012: 132),

---

[3] The topics varied between Levels 1, 2, and 3. In an ideal world, one would use the same topic across and within the three levels. But, in a language classroom, this is hardly possible. Importantly, all topics were simple, drawing on the learners' basic vocabulary. Given that these were beginner learners writing short narratives, it is unlikely that the choice of the topic affected the findings reported in the study.

"small corpora may also be of considerable value". Regardless of the potential worth of small learner corpora, the small size is, perhaps, to be expected if the focus of one's investigation is on beginner learners who would not be able to produce longer compositions.

The longitudinal studies conducted thus far have been characterised by a very limited number of informants. In this respect, it was hoped that following a relatively large group of L2 learners during an intensive course would allow us to gain new insights into the development of learner collocational knowledge.

3.2.3. Procedure

For the purpose of the present investigation, we decided to focus on one type of word combinations − N + Adj. Unlike English, Italian adjectives normally follow the noun they modify (*piazza grande* "main/big square"*, vino rosso* "red wine"). Adj + N (corresponding to N + Adj in Italian) combinations have been found to be relatively frequent pairings in L2 learner writing (Durrant & Schmitt, 2009; Siyanova & Schmitt, 2008). Given the small size of the corpus, it was hoped that by focussing on this type of combinations we would have a large enough data set for further analyses.

First, all N + Adj pairings were manually extracted from the three sub-corpora. We aimed to be as inclusive as possible and did not apply any further criteria, other than Word 1 being a noun and Word 2 − an adjective. All combinations were extracted by a trained research assistant (a native speaker of Italian) not otherwise involved in the study, who read the essays a few times to verify the extracted items and to ensure no mistakes were made during the extraction process. Second, using our reference corpus of L1 Italian − La Repubblica (as well as two additional reference corpora PAISÀ and Google, more on this below) − we obtained frequency information for Word 1 and Word 2 within each combination, as well as its phrasal frequency. La Repubblica corpus[4] is a collection of 380 million words that contains all the articles published by the Italian national daily newspaper La Repubblica between 1985 and 2000 (Baroni et al., 2004). La Repubblica is the largest and most representative Italian corpus currently available. It includes the following categories: culture, society, sport, weather, education, news, science, politics, economics, and church. Although we used La Repubblica as our main reference corpus in the study (it was used to obtain phrasal frequencies, Word 1 and Word 2 frequencies), we decided to use a second reference corpus − the PAISÀ corpus[5] (Baroni, 2010) − to validate La Repubblica frequencies. By validating La Repubblica against the PAISÀ corpus, we wanted to ascertain the representativeness of the relative frequencies of learner combinations (obtained from La Repubblica) in the Italian language. The PAISÀ corpus (250 million tokens) is a collection of authentic contemporary Italian texts of various genres. The corpus, collected between 2009 and 2012, contains 380,000 documents from about 1000 different websites. La Repubblica frequencies were compared against PAISÀ frequencies; the two were found to correlate strongly, as shown by a Spearman correlation test ($r = .885$, $p < 0.01$). We further obtained Google frequencies for the learner items and correlated these frequencies with those of La Repubblica and PAISÀ. Both corpora frequencies correlated strongly with Google frequencies (La Repubblica vs. Google: $r = .785$, $p < 0.01$; PAISÀ vs. Google: $r = .787$, $p < 0.01$). Although the web is not a perfect corpus, many researchers have argued that the large amount of data available outweighs potential problems (Keller & Lapata, 2003).[6] The high correlations between the three corpora allowed us to conclude that the frequencies obtained from La Repubblica can be said to be representative of the Italian language in general, rather than specific to a particular corpus, category, or genre.

With La Repubblica frequency information in hand, we calculated mutual information (MI) for each pairing. MI is a standard measure used to show how strongly associated the words are and how likely it is that they will co-occur. For example, the combination *excruciating pain* will have a very high MI score, because whenever we come across *excruciating*, we are very likely to see or hear *pain* following it. The word combination *blue balloon* will have a much lower MI score, since many things can be blue, and balloons, in general, can be of different colours. Therefore, MI is good at discovering 'interesting collocations', as it is not so strongly linked with raw frequency as other association measures, such as, for example, *t*-score, which correlates strongly with raw frequency. In addition, according to Hunston (2002: 73), MI is not dependent on the size of the corpus, and is thus good for both larger and smaller corpora. The higher the MI score, the more strongly associated the words within a phrase are. According to Hunston (2002) and Stubbs (1995), the score of 3 or above is indicative of a significant collocation threshold. A word combination can be considered a collocation − in other words, its constituent words appear together more often than would be predicted by chance alone − if the MI score equals, or is above, 3. Conversely, a word combination is not a collocation if the MI score is below 3. Thus, the present investigation will use the MI score of 3 as a threshold for identifying significant collocations.

However, as noted by Durrant and Schmitt (2009), merely dividing learner combinations into 'collocations' and 'non-collocations' is problematic, as this dichotomous method may disguise differences between combinations that just about pass the threshold and much stronger collocations. Thus, following the example of Durrant and Schmitt (2009), learner combinations will also be classified across a scale of collocational strength. Specifically, we will look at learner collocations along a MI score continuum starting with the minimum MI score of 3 (i.e., 3–3.99, 4–4.99, 5–5.99, 6–6.99, 7–7.99, and above 8). This approach is in line with Manning and Schutze (1999) and Evert and Krenn (2001), who argue that association measures, such as MI, are best used for ranking combinations on a collocational strength continuum. In addition, by looking at a range of

---

[4] La Repubblica corpus can be found here http://dev.sslmit.unibo.it/corpora/corpus.php?path=&name=Repubblica.

[5] More information about PAISÀ is available here http://www.corpusitaliano.it/en/index.html.

[6] See a special issue of *Computational Linguistics*, 2003, volume 29, issue 3 dedicated to the web as a corpus.

Table 1
Summary of N + Adj combinations extracted from the learner corpus.

	N. of words	N. of students/compositions	Time of writing	N + Adj combinations (tokens)	N + Adj combinations (types)
Level 1	5282	36	Week 7	167	117
Level 2	6263	36	Week 14	165	128
Level 3	6367	36	Week 21	183	145

collocational strength, one can get a more fine-grained view of the data (Durrant & Schmitt, 2009), which is particularly important given that the focus of the present study is on (possibly, rather small) developmental differences in L2 collocation use in writing.

3.3. Analysis and results

3.3.1. Total number of N + Adj combinations produced

The extraction procedure resulted in a total of 515 phrasal tokens and 390 phrasal types extracted from the 108 learner compositions (Table 1). Overall, the numbers can be said to be comparable across the levels (it is noteworthy that Level 2 and Level 3 compositions were slightly longer than Level 1 compositions, which resulted in slighter larger Level 2 and Level 3 sub-corpora).

To address the question of idiomaticity and appropriateness of L2 combinations, two analyses were performed employing a range of thresholds: (1) frequency and (2) frequency and MI.

3.3.2. Frequency analysis

Drawing on previous research (Durrant & Schmitt, 2009; Siyanova & Schmitt, 2008) and using La Repubblica raw frequency information, we split each of the three sets (Level 1, 2, and 3) of learner pairings into five frequency bands: 0 (failed to appear in La Repubblica), 1–5, 6–20, 21–100, and >100 occurrences in La Repubblica (Fig. 1).

What we can see from Fig. 1 is that Level 1 compositions contained by far the largest number of zero frequency items (those that failed to appear in the reference corpus, e.g., *strada confortevole* "comfortable street", *cibo difficile* "difficult food"). Further, the two medium frequency bands (6–20 and 21–100) suggest that more items in these frequency ranges (*prezzo pieno* "full price", *città moderne* "modern cities") were present in Level 3 compositions than Level 1.

Perhaps, a clearer picture emerges if we collapse the five frequency bands into two: 0–5 occurrences (zero and very low frequency items) and ≥6 occurrences (medium–low, medium–high and high frequency items). As can be seen from Fig. 2, Level 1 compositions contained the highest proportion of zero and very low-frequency items, while Level 3 compositions contained the lowest proportion of such items. Correspondingly, Level 1 compositions contained the lowest proportion of the medium–high frequency combinations, while Level 3 compositions contained by far the highest.

Having obtained the numbers of 'relatively frequent' (≥6) and 'relatively infrequent' (0–5) word combinations, we calculated whether or not the difference between these figures was statistically significant. Our main goal was to see if the difference was significant for the same L2 writers at the beginning (Level 1) versus at the end (Level 3) of the course. A chi-square test of independence (with no correction for continuity because the sample size is greater than 20, Gries, 2013) showed a small but significant difference in Level 1 versus Level 3 use of 'frequent' and 'infrequent' combinations[7]: χ^2 (1, $N = 262$) = 4.35, $p < 0.05$, Cramer's $V = 0.13$, suggesting that Level 3 compositions were more likely than Level 1 compositions to contain more frequent N + Adj pairings and, correspondingly, fewer infrequent word combinations.

3.3.3. Frequency and MI analysis

One way to look at learner combinations is in terms of their frequencies in an L1 reference corpus. Another way is to look at them in terms of their association strength. Although it is possible to calculate MI scores for all collocations with a frequency above zero, researchers have noted that the lowest-frequency data present a serious challenge for MI (Evert, 2008; Evert & Krenn, 2001; Manning & Schutze, 1999; Xiao & McEnery, 2006). Therefore, following Siyanova and Schmitt (2008) and Evert and Krenn (2001), we adopted a two-tier threshold based on: (1) corpus frequency and (2) MI. Based on the previous research (Siyanova & Schmitt, 2008), we selected a frequency threshold of 6, in order to exclude the very low frequency items that may present a challenge for MI (also see Durrant & Schmitt, 2009 for a similar procedure). As discussed above, the MI threshold was set at ≥3. Thus, any item appearing ≥6 times in the corpus and having a MI score of ≥3 can be considered a strongly associated collocation in which the two words appear together more often than would be expected by chance alone.

We find that under a third of Level 1 learner combinations meet the two-tier collocation threshold of frequency ≥6 and MI ≥3 (Fig. 3). Again, a different picture emerges if we look at Level 3 data. Based on the joint frequency/MI score criterion, it

---

[7] All chi-square tests were performed using corresponding contingency tables with the raw counts of learner combinations specific to each comparison (Field, 2005; Gries, 2009). In the figures throughout the manuscript, proportions of learner combinations are provided for easier visual inspection and comparison.
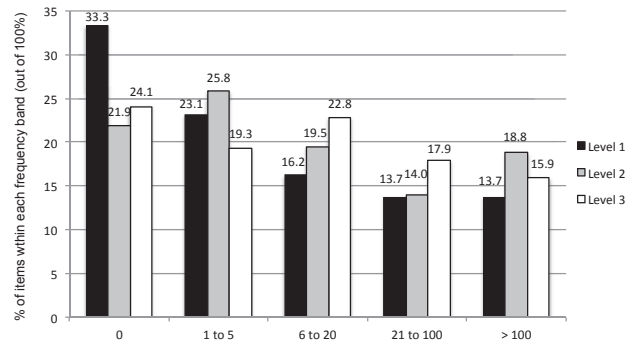
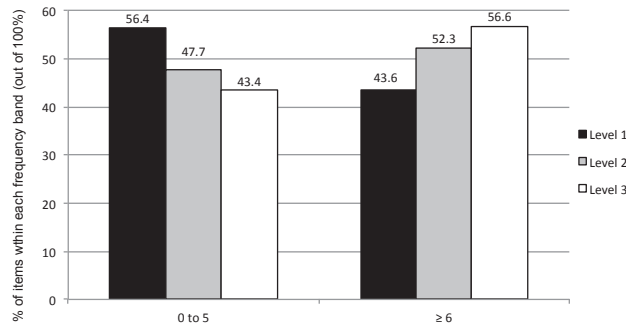**Fig. 1.** Distribution of learner combinations (Types) across the five frequency bands.

Fig. 2. Distribution of learner combinations (Types) across the two collapsed frequency bands.

appears that 41.4% of Level 3 N + Adj combinations can be said to be strongly associated, statistically significant collocations (*spazi verdi* "green spaces", *tempo libero* "free time").

Having identified the collocations with frequency of ≥6 and MI score of ≥3, we calculated whether or not the difference between these figures was significant. As in the earlier analysis, our primary comparison was Level 1 versus Level 3 compositions. A chi-square test of independence (with no correction for continuity because the sample size is greater than 20, Gries, 2013) showed a modest but significant difference in Level 1 versus Level 3 use of collocations that met our two-tier collocation threshold: χ^2 (1, $N = 262$) = 4.91, $p < 0.05$, Cramer's $V = 0.14$, suggesting that Level 3 compositions were more likely than Level 1 writings to contain strongly associated N + Adj collocations.

Learner collocations can be classified across a scale of collocational strength. In the analysis above, we employed the minimum, commonly adopted, threshold of MI ≥3. Keeping the same frequency threshold, we can consider a higher MI
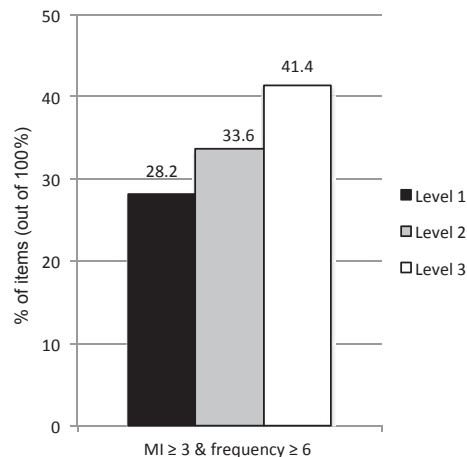

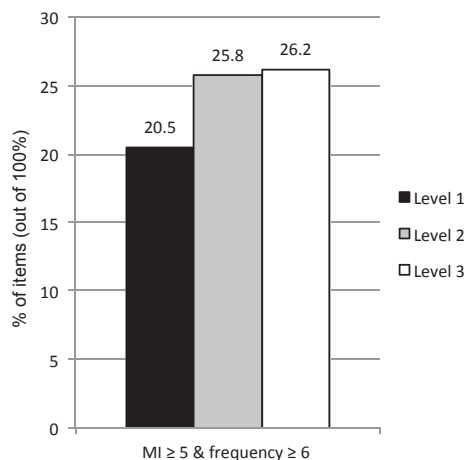**Fig. 3.** Distribution of learner collocations (Types) with mutual information (MI) ≥3 and frequency ≥6.

Fig. 4. Distribution of learner collocations (Types) with mutual information (MI) ≥5 and frequency ≥6.

threshold, for example, MI ≥5 (Fig. 4). A chi-square test of independence showed no significant differences between Level 1 versus Level 3 compositions in terms of the number of collocations that met the higher MI threshold ($p > 0.1$).

Finally, another way of looking at the data is by grouping all significant collocations − meeting the collocation threshold of frequency ≥6 and MI ≥3 − along a MI score continuum. If we look at Fig. 5, it appears that Level 3 compositions contained a larger proportion of significant collocations at different levels of MI scores than Level 1 compositions: 3—3.99 (*giornata splendida* "splendid day"), 4—4.99 (*cibo quotidiano* "daily food"), 5—5.99 (*gente simpatica* "nice people"), 6—6.99 (*tempo pieno* "full time"), 7—7.99 (*musica popolare* "popular music"), and ≥8 (*musica classica* "classical music"). However, a *Z*-test for proportions, performed on the MI bands, showed no significant differences in collocation proportions between Level 1 and Level 3 compositions ($p > 0.1$).

4. General discussion

The present study aimed to investigate the production of N + Adj combinations in learner writing from a longitudinal perspective. Short compositions were written by 36 beginner learners of Italian (L1 Chinese) at the beginning, halfway through, and at the end of an intensive course. Our focus was on the extent to which beginner learners made use of N + Adj combinations (*pelle abbronzata* "tanned skin") as their proficiency increased, and how this increase was reflected in the quality and quantity of the combinations produced. Our main focus was on the possible differences between the combinations produced at the beginning versus at the end of the course. The following findings emerged.

We found that the number of learner combinations extracted from the compositions was comparable across the levels (if the length of the compositions and the size of the corresponding sub-corpora are taken into account). Thus, if we consider the
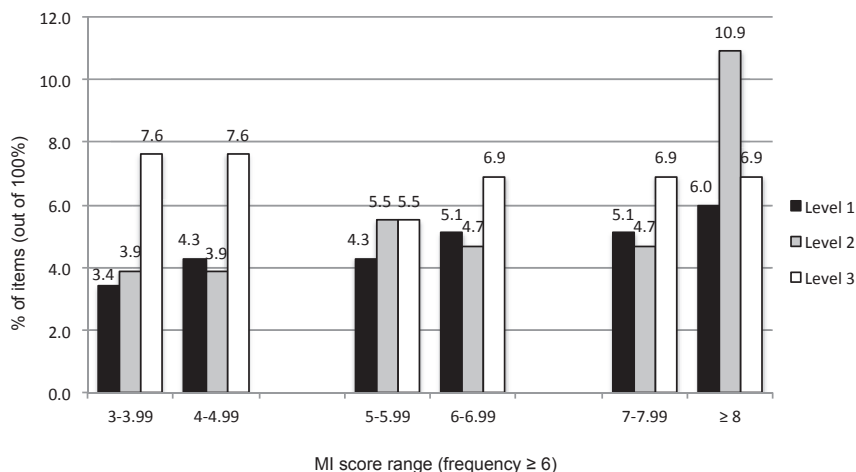

Fig. 5. Distribution of learner collocations (Types) found at different levels of mutual information (MI) scores.

number of all N + Adj combinations produced at the beginning versus at the end of the five-month period, it can be said to have remained about the same.

A more interesting finding emerges if we consider frequency distributions of the learner combinations extracted from the sub-corpora. Our analysis showed that if we consider learner items that appear very infrequently (0—5) and those that appear ≥6 times in the L1 reference corpus, we find a small but significant difference between Level 1 and Level 3 compositions, with the latter containing more higher frequency items (frequency ≥6) and, correspondingly, fewer lower frequency items (frequency 0—5). This suggests that as proficiency increased over the course of five months, the Chinese learners of Italian were more able to produce more frequent, and thus more native-like, word-combinations, while the number of zero and very low frequency, and thus less native-like, items (*studente calmo* "calm student", *spiaggia confortevole* "comfortable beach", *via affascinante* "fascinating street") decreased.

Our further analyses focused on the quality of N + Adj combinations more directly. Specifically, we examined those learner combinations that can be considered statistically significant collocations, according to the frequency-based approach adopted (Hoey, 2005; Sinclair, 2004; Stubbs, 1995). To address the question of how many and what kind of collocations our learners used in their writing, L2 items were analysed using a combined frequency and MI score criterion. Having adopted a two-tier threshold of frequency ≥6 and MI score ≥3, we found that Level 3 compositions contained a significantly higher number of strongly associated N + Adj collocations than did Level 1 writings. This implies that an increase in learner proficiency resulted not only in an overall increase in higher frequency word combinations, but also in those that are more 'phrasal' in nature, that is, statistically significant, strongly associated collocations (*città natale* "birth place", *aria fresca* "fresh air"). This finding echoes that of Durrant and Schmitt (2009), who found that higher proficiency speakers (native speakers of English) made greater use of Premodifier + N collocations with higher MI scores (*densely populated*) compared to lower proficiency speakers (advanced L2 learners). In addition, recent studies have shown that there is a marked improvement in collocation use from intermediate to advanced learners (Granger & Bestgen, 2014). In a cross-sectional study employing Durrant and Schmitt's (2009) paradigm, Granger and Bestgen (2014) analysed intermediate versus advanced EFL learner writings. It was found that the more proficient learners drew on strongly associated bigrams (*immortal souls*) identified using the MI score to a greater extent than the less proficient writers. The findings of the present study are important in that they suggest a noticeable increase in the use of significant, strongly associated collocations (as suggested by MI scores) at the beginner level — the first study to show this. This is remarkable, as one might expect complete beginners to concentrate on grammar and single words rather than word combinations. Despite the fact that it has long been acknowledged that our mental lexicon is made not only of single words, but also of a large number of units above the word level (Jackendoff, 1995; Langacker, 1987; Tomasello, 2003), single word has traditionally been the major unit of vocabulary learning (Laufer, 1989, 1997a, 1997b; Nation, 1990, 2001).

Indeed, our finding that L2 collocation production improved as a result of time spent learning Italian in an L2 environment appears to go against the widely held assumption that L2 learners tend to focus solely on single words. For example, it has been proposed that L2 learners (unlike L1 learners) pay attention to individual words, rather than to combinatorial mechanisms of these words. In other words, learners do not seem to notice collocational relationships (Wray, 2002). As Wray argues:

> The consequence [of focusing on word-sized units in L2 learning] is a failure to value the one property of nativelike input which is most characteristic of the idiomaticity to which the learner ultimately aspires: words do not go together, having first been apart, but, rather, belong together, and do not necessarily need separating. (Wray, 2002: 212)

Our results imply that from the start of the learning process, L2 learners begin to attend to combinatorial linguistic mechanisms. They develop and strengthen lexical connections between words based on frequency of co-occurrence and association strength (Crossley & Salsbury, 2011; Ellis, O'Donnell, & Römer, 2014), as suggested by a decrease in zero and very low frequency items and an increase in strongly associated collocations — an improvement that paralleled an increase in L2 proficiency over the course of five months. The development of collocational competence being a rather slow process (Bahns & Eldaw, 1993; Laufer & Waldman, 2011), reliant on rich and sustained exposure, it is, perhaps, unsurprising that when the MI threshold was increased from the commonly adopted MI ≥3 to a higher, more conservative, MI ≥5, the difference between Level 1 and Level 3 compositions disappeared, a finding further supported by the analysis of learner collocations across a range of MI bands. While Level 3 compositions were superior to Level 1 compositions in terms of frequency distributions of learner items and the number of collocations that met the minimum significant collocation threshold, Level 3 writings appeared comparable to Level 1 writings when much stronger collocations were considered.

Overall, the results of the present study suggest that, despite collocations being a seemingly intractable problem for language teachers and learners (Granger, 1998; Howarth, 1998; Nesselhauf, 2003, 2005), five months might be sufficient for novice learners to *gradually begin* to produce more idiomatic, native-like output. Our findings appear to support those of Crossley and Salsbury (2011), the study that examined beginner learner production of collocation over a period of time and which lends itself to a useful comparison with the present investigation. Crossley and Salsbury (2011) analysed the development of lexical bundles in naturally produced speech of six ESL learners. By focussing on the accuracy of L2 lexical bundles, these researchers analysed how the frequency of learner items changed over the course of one year (the duration of the study) and how L2 frequencies compared against the L1 baseline. In addition, Crossley and Salsbury (2011) investigated whether, as they were becoming more proficient, the six L2 learners were producing lexical bundles that were commonly found in L1 discourse (a question not unlike the one that we set out to answer in the present study). Their analysis showed significant correlations between L2 bundle frequency and the native speaker corpus, with the correlations increasing from *r* = 0.60 in the

first quarter to $r = 0.81$ in the fourth (and final) quarter of learning. This was taken to suggest that lexical bundle accuracy increased as a result of time spent learning English in an L2 environment and thus as a function of increasing English language proficiency.

First, both studies showed that over time, L2 learners begin to produce more frequent, native-like phrases and fewer infrequent, unidiomatic ones. Second, our results, as well as those of Crossley and Salsbury (2011), question the notion that L2 learners focus solely on single words rather than on combinatorial properties of these words (Wray, 2002) — a finding suggestive of the development of depth of L2 lexical knowledge, that is, "knowledge about the strength of connections between words" (Crossley & Salsbury, 2011: 15). These studies are particularly comparable given that both analysed phrasal production of a group of beginner learners over time, during an intensive language course in an immersion-based environment. However, the findings of the two studies also complement each other in a number of important ways — while Crossley and Salsbury (2011) found improvement in the use of lexical bundles in L2 spoken production, we found an increase in significant collocation use in L2 written production. We can thus put forward a strong case for comparable developmental mechanisms involved in L2 spoken discourse (informal register) and L2 written discourse (formal register) of collocations, lexical bundles, and potentially other instances of units above the word level. Further, while Crossley and Salsbury (2011) demonstrated that L2 learners improved in the accuracy of lexical bundles over the course of one year, we observed a significant improvement in collocation production over the course of five months. This suggests that beginner learner collocational knowledge can improve over a relatively short period of time. Finally, although the number of participants in Crossley and Salsbury (2011) was the largest among the longitudinal studies available (six participants), it is, admittedly, rather small. The present study, drawing on the data from 36 beginner learners, lends important credence to the findings reported in Crossley and Salsbury (2011) and contributes to a more complete picture of the development of L2 collocational knowledge — in spoken as well as written production — from a longitudinal perspective.

The findings of the present study can be said to be good news for language teachers and learners. It is, however, imperative that they be considered in the context of the intensive course the learners attended and the immersion-based environment in which they operated. Although no comparison was made with a group of students studying Italian as a foreign (as opposed to second) language, our findings might offer indirect support to earlier studies that highlighted a positive influence of immersion-based L2 exposure on collocational competence (Groom, 2009; Siyanova & Schmitt, 2008; Waibel, 2008).

5. Conclusion

We believe the present study — albeit small-scale — is important in that it is the first to look at collocational knowledge development longitudinally focussing on a relatively large group of beginner learners. Despite the modest length of the intensive course, the Chinese learners of Italian, nevertheless, showed a significant improvement in the use of N + Adj collocations. Not only did the compositions produced at the end of the course contain more higher frequency combinations and, correspondingly, fewer lower frequency ones, they also contained a greater number of statistically significant collocations. Thus, as L2 learners gained (some) control of the Italian language, their production of N + Adj combinations became more native-like, with the learners appearing less reliant on single words to convey meaning and more capable of drawing on a small but gradually improving repertoire of strongly associated lexical items. This suggests that a period of time as short as five months might be sufficient for beginner learners to begin to exhibit more idiomatic, native-like output.

While the findings reported in this study shed light on the development of L2 lexical proficiency over a period of time, there are, nevertheless, reasons why they need to be replicated and complemented. One obvious limitation is the small size of the learner corpus and, correspondingly, a small number of learner items extracted. Undoubtedly, a larger corpus would provide more and richer data. It needs to be pointed out, however, that one can hardly expect complete beginners to be able to produce long compositions. A larger learner corpus is more likely to be achieved by employing a larger participant pool, rather than by increasing the length of individual compositions. Further, due to relatively little data available, we focused on one type of collocation — N + Adj. Given more L2 data, a variety of collocation types could be compared and contrasted (for example, it has been suggested that Adj + N/N + Adj collocations might be easier to learn than V + N collocations, Gitsaki, 1999; Peters, 2015; also see Boers, Demecheleer, Coxhead, & Webb, 2014 for a discussion of potential difficulties associated with V + N collocations). Finally, in the present investigation, we used frequency and the strength of association between the noun and the adjective to identify and qualify learner collocations, as we believe this approach to be least biased and not reliant on one's intuition. Future longitudinal research may want to combine such a frequency-driven approach with one that employs a panel of native-speaker judges to look at different types of learner combinations (e.g., free combinations, restricted collocations, idiomatic expressions, two word compounds).

We are hopeful that future longitudinal research, currently almost non-existent with the exception of a few case studies, will: (1) employ corpora collected from a larger L2 participant pool, (2) look at various types of learner collocations, (3) focus on a wider range of L2 proficiencies and backgrounds, and (4) explore the role of L2 immersion more directly, by comparing collocational development in a second versus foreign language setting. After all, as argued by Li and Schmitt (2010), the only truly reliable way to examine patterns of development in the use of L2 collocations is to conduct longitudinal studies of the same learners over a period of time.

References

Alejo Gonzalez, R. A. (2010). L2 Spanish acquisition of English phrasal verbs: a cognitive linguistic analysis of L1 influence. In M. C. Campoy-Cubillo, B. Belles-Fortuno, & M. L. Gea-Valor (Eds.), *Corpus-based approaches to English language teaching* (pp. 149–166). London, UK: Continuum.
Altenberg, B., & Granger, S. (2001). The grammatical and lexical patterning of "make" in native and non-native student writing. *Applied Linguistics, 22*, 173–194.
Bahns, J., & Eldaw, M. (1993). Should we teach EFL students collocations? *System, 21*, 101–114.
Baroni, M. (2010). *The Paisà project*. Presented at Human Language Technologies for Italian - 2010, Workshop of the Natural Language Processing Group of the AI*IA in conjunction with the 11th AI*IA Symposium on Artificial Intelligence, Brescia, Italy, 1 December 2010 http://www.corpusitaliano.it/en/index.html Accessed 24.08.14..
Baroni, M., Bernardini, S., Comastri, F., Piccioni, L., Volpi, A., Aston, G., et al. (2004). Introducing the "la Repubblica" corpus: A large, annotated, TEI(XML)-compliant corpus of newspaper Italian. In *Proceedings of LREC 2004*. http://dev.sslmit.unibo.it/corpora/corpus.php?path=&name=Repubblica Accessed 24.08.14.
Bartning, I., Forsberg, F., & Hancock, V. (2009). Resources and obstacles in very advanced L2 French. Formulaic language, information structure and morphosyntax. In *EUROSLA yearbook* (Vol. 9, pp. 185–211). Amsterdam: Benjamins.
Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English*. Harlow, UK: Longman.
Boers, F., Demecheleer, M., Coxhead, A., & Webb, S. (2014). Gauging the effects of exercises on verb-noun collocations. *Language Teaching Research, 18*(1), 54–74.
Chen, Y.-H., & Baker, P. (2010). Lexical bundles in L1 and L2 academic writing. *Language Learning & Technology, 14*, 30–49.
Cowie, A. P. (1998). *Phraseology: Theory, analysis, and applications*. Oxford: Oxford University Press.
Crossley, S. A., & Salsbury, T. (2011). The development of lexical bundle accuracy and production in English second language speakers. *International Review of Applied Linguistics in Teaching, 49*, 1–26.
Cross, J., & Papp, S. (2008). Creativity in the use of verb + noun combinations by Chinese learners of English. In G. Gilquin, S. Papp, & M. B. Dıez-Bedmar (Eds.), *Linking up contrastive and learner corpus research* (pp. 57–81). Amsterdam, the Netherlands: Rodopi.
De Cock, S. (2004). Preferred sequences of words in NS and NNS speech. *Belgian Journal of English Language and Literature (BELL), New Series, 2*, 225–246.
Durrant, P., & Schmitt, N. (2009). To what extent do native and non-native writers make use of collocations? *International Review of Applied Linguistics, 47*, 157–177.
Ellis, N. C., O'Donnell, M. B., & Römer, U. (2014). Second language verb-argument constructions are sensitive to form, function, frequency, contingency, and prototypicality. *Linguistic Approaches to Bilingualism, 4*(4), 405–431.
Erman, B., & Warren, B. (2000). The idiom principle and the open-choice principle. *Text, 20*, 29–62.
Evert, S. (2008). Corpora and collocations. In A. Ludeling, & M. Kyto (Eds.), *Corpus linguistics. An international handbook* (pp. 1212–1248). Berlin, Germany: Mouton de Gruyter.
Evert, S., & Krenn, B. (2001, July). *Methods for the qualitative evaluations of lexical association measures*. Paper presented at the 39th Annual Meeting of the Association for Computational Linguistics, Toulouse, France.
Field, A. (2005). *Discovering statistics using SPSS* (2nd ed.). London: Sage.
Forsberg, F. (2010). Using conventional sequences in L2 French. *International Review of Applied Linguistics, 48*(1), 25–50.
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.), *Researching pedagogic tasks: Second language learning, teaching and testing* (pp. 75–93). Harlow, UK: Longman.
Gilquin, G. (2007). To err is not all. What corpus and elicitation can reveal about the use of collocations by learners. *Zeitschrift fur Anglistik und Amerikanistik, 55*, 273–291.
Gitsaki, C. (1999). *Second language lexical acquisition: A study of the development of collocational knowledge*. Bethesda, MD: International Scholars Publications.
Glucksberg, S. (1989). Metaphors in conversation: how are they understood? Why are they used? *Metaphor and Symbolic Activity, 4*, 125–143.
Gotz, S., & Schilk, M. (2011). Formulaic sequences in spoken ENL, ESL, and EFL. In M. Hundt, & J. Mukherjee (Eds.), *Exploring second-language varieties of English and learner Englishes: Bridging a paradigm gap* (pp. 79–100). Amsterdam: John Benjamins.
Granger, S. (1998). Prefabricated patterns in advanced EFL writing: collocations and formulae. In A. Cowie (Ed.), *Phraseology: Theory, analysis, and applications* (pp. 145–160). Oxford, UK: Oxford University Press.
Granger, S., & Bestgen, Y. (2014). The use of collocations by intermediate vs. advanced non-native writers: a bigram-based study. *IRAL, 52*(3), 229–252.
Granger, S., & Paquot, M. (2009). Lexical verbs in academic discourse: a corpus-driven study of learner use. In M. Charles, D. Pecorari, & S. Hunston (Eds.), *Academic writing: At the interface of corpus and discourse* (pp. 193–214). London, New York: Continuum.
Gries, S. (2009). *Quantitative corpus linguistics with R: A practical introduction*. London, New York: Routledge.
Gries, S. (2013). Statistical tests for the analysis of learner corpus data. In A. Díaz-Negrillo, N. Ballier, & P. Thompson (Eds.), *Automatic treatment and analysis of learner corpus data* (pp. 287–309). Amsterdam & Philadelphia: John Benjamins.
Groom, N. (2009). Effects of second language immersion on second language collocational development. In A. Barfield, & H. Gyllstad (Eds.), *Researching collocations in another language* (pp. 21–33). Basingstoke, UK: Palgrave Macmillan.
Henderson, A., & Barr, R. (2010). Comparing indicators of authorial stance in psychology students' writing and published research articles. *Journal of Writing Research, 2*, 245–264.
Hoey, M. (1991). *Patterns of lexis in text*. Oxford: Oxford University Press.
Hoey, M. (2005). *Lexical priming: A new theory of words and language*. London: Routledge.
Howarth, P. (1996). *Phraseology in English academic writing: Some implications for language learning and dictionary making*. Tubingen, Germany: Max Niemeyer Verlag.
Howarth, P. (1998). Phraseology and second language proficiency. *Applied Linguistics, 19*, 24–44.
Hunston, S. (2002). *Corpora in applied linguistics*. Cambridge: Cambridge University Press.
Jackendoff, R. (1995). The boundaries of the lexicon. In M. Everaert, E. van der Linded, A. Schenk, & R. Schreuder (Eds.), *Idioms: Structural and psychological perspectives* (pp. 133–165). Hillside, NJ: Erlbaum.
Jones, S., & Sinclair, J. McH. (1974). English lexical collocations. A study in computational linguistics. *Cahiers de lexicologie, 24*, 15–61.
Keller, F., & Lapata, M. (2003). Using the web to obtain frequencies for unseen bigrams. *Computational Linguistics, 29*(3), 459–484.
Langacker, R. (1987). *Foundations of cognitive grammar* (Vol. 1). Stanford: Stanford University Press.
Laufer, B. (1989). A factor of difficulty in vocabulary learning: deceptive transparency. *AILA Review, 6*, 10–20.
Laufer, B. (1997a). What's in a word that makes it hard or easy? Intralexical factors affecting the difficulty of vocabulary acquisition. In M. McCarthy, & N. Schmitt (Eds.), *Vocabulary description, acquisition and pedagogy* (pp. 140–155). Cambridge University Press.

Laufer, B. (1997b). The lexical plight in second language reading: words you don't know, words you think you know and words you can't guess. In J. Coady, & T. Huckin (Eds.), *Second language vocabulary acquisition: A rationale for pedagogy* (pp. 20—34). Cambridge University Press.

Laufer, B., & Waldman, T. (2011). Verb-noun collocations in second language writing: a corpus analysis of learners' English. *Language Learning, 61*, 647—672.

Levitzky-Aviad, T., & Laufer, B. (2013). Lexical properties in the writing of L2 learners over eight years of study: single words and collocations. *Eurosla Monographs Series, 2*, 127—148.

Li, P., Eskildsen, S., & Cadierno, T. (2014). Tracing an L2 learner's motion constructions over time: a usage-based classroom investigation. *The Modern Language Journal, 98*(2), 612—628.

Li, J., & Schmitt, N. (2009). The acquisition of lexical phrases in academic writing: a longitudinal case study. *Journal of Second Language Writing, 18*, 85—102.

Li, J., & Schmitt, N. (2010). The development of collocation use in academic texts by advanced L2 learners: a multiple case study approach. In D. Wood (Ed.), *Perspectives on formulaic language: Acquisition and communication* (pp. 2—46). London, UK: Continuum.

Lo Cascio, V. (2013). *Dizionario combinatorio italiano*. Amsterdam, Philadelphia: John Benjamins.

Lorenz, G. (1999). *Adjective intensification — Learners versus native speakers. A corpus study of argumentative writing*. Amsterdam, the Netherlands: Rodopi.

Manning, C. D., & Schutze, H. (1999). *Foundations of statistical natural language processing*. Cambridge, MA: MIT Press.

Nation, I. S. P. (1990). *Teaching and learning vocabulary*. Rowley, MA: Newbury House.

Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.

Nesselhauf, N. (2003). The use of collocations by advanced learners of English and some implications for teaching. *Applied Linguistics, 24*, 223—242.

Nesselhauf, N. (2005). *Collocations in a learner corpus*. Amsterdam, the Netherlands: John Benjamins.

Ortega, L. (2009). *Understanding second language acquisition*. London: Hodder.

Paquot, M. (2010). *Academic vocabulary in learner writing: From extraction to analysis*. London, New York: Continuum.

Paquot, M., & Granger, S. (2012). Formulaic language in learner corpora. *Annual Review of Applied Linguistics, 32*, 130—149.

Peters, E. (2015). The learning burden of collocations: the role of interlexical and intralexical factors. *Language Teaching Research*. http://dx.doi.org/10.1177/1362168814568131.

Pollio, H., Barlow, J., Fine, H., & Pollio, M. (1977). *Psychology and the poetics of growth: Figurative language in psychology, psychotherapy, and education*. Lawrence Erlbaum.

Sinclair, J. M. (1991). *Corpus, concordance, collocation*. Oxford: Oxford University Press.

Sinclair, J. McH. (2004). *Trust the text: Language, corpus and discourse*. London: Routledge.

Siyanova-Chanturia, A., & Spina, S. (2015). Investigation of native speaker and second language learner intuition of collocation frequency. *Language Learning, 65*(3), 533—562.

Siyanova, A., & Schmitt, N. (2008). L2 learner production and processing of collocation: a multi-study perspective. *Canadian Modern Language Review, 64*(3), 429—458.

Sorhus, H. (1977). To hear ourselves — implications for teaching English as a second language. *ELT Journal, 31*, 211—221.

Spina, S. (2010). The dictionary of Italian collocations: design and integration in an online learning environment. In N. Calzolari, K. Choukri, B. Maegaard, J. Mariani, J. Odjik, S. Piperidis, et al. (Eds.), *Proceedings of the Seventh Conference on International Language Resources and Evaluation* (pp. 3202—3208). European Language Resources Association.

Spina, S. (2014). Il Perugia Corpus: Una risorsa di riferimento per l'italiano. Composizione, annotazione e valutazione. In R. Basili, A. Lenci, & B. Magnini (Eds.), *Proceedings of the First Italian Conference on Computational Linguistics CLiC-it 2014* (pp. 354—359). Pisa: Pisa University Press.

Spina, S. Phraseology in academic L2 discourse: the use of multi-words units in a CMC university context. In E. Castello, K. Ackerley, & F. Coccetta (Eds.), Studies in learner corpus linguistics research and applications for foreign language teaching and assessment. in press, Peter Lang

Stubbs, M. (1995). Collocations and semantic profiles: on the cause of the trouble with quantitative methods. *Functions of language, 2*(1), 1—33.

Thewissen, J. (2013). Capturing L2 accuracy developmental patterns: insights from an error-tagged EFL learner corpus. *The Modern Language Journal, 97*, 77—101.

Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.

Waibel, B. (2008). *Phrasal verbs: German and Italian learners of English compared*. Saarbrucken, Germany: VDM.

Wray, A. (2002). *Formulaic language and the lexicon*. Cambridge: Cambridge University Press.

Xiao, R., & McEnery, T. (2006). Collocation, semantic prosody, and near synonymy: a cross-linguistic perspective. *Applied Linguistics, 27*(1), 103—129.

Yuldashev, A., Fernandez, J., & Thorne, S. L. (2013). Second language learners' contiguous and discontiguous multi-word unit use over time. *The Modern Language Journal, 97*, 31—45.