Your participation is greatly/highly appreciated

Amanda Edmonds, Aarnes Gudmestad

PRE-PRINT VERSION
Edmonds, A., & Gudmestad, A. (2014). Your participation is *greatly/highly* appreciated : Amplifier collocations in L2 English. *Canadian Modern Language Review, 70*(1), 76-102.

Your participation is *greatly/highly* appreciated:

Amplifier collocations in L2 English

Amanda Edmonds & Aarnes Gudmestad

Abstract

The current study sets out to investigate collocational knowledge for a set of 13 English amplifiers among native and nonnative speakers of English, by providing a partial replication of one of the projects reported on in Granger (1998). The project combines both phraseological and distributional approaches to research into formulaic language to examine whether natives and nonnatives demonstrate similar patterns of saliency and agreement in their judgments of adverb-adjective collocations. A total of 55 English native speakers and 120 Francophone learners of English (first-year university students, third-year university students, and Master's students) completed two tasks targeting such collocations. Our quantitative analysis reveals that Master's students and native speakers performed similarly on the different tasks, and that both groups differed significantly from the first- and third-year university learners. This pattern holds for all analyses of salience and for all but one analysis of agreement. We interpret these findings as evidence of development toward nativelike patterns with respect to the collocations under study.

Keywords: amplifiers, collocations, formulaic language, salience

Résumé

La présente étude a pour objectif d'examiner les connaissances de collocations associées à 13 intensificateurs anglais pour un groupe de locuteurs natifs ainsi qu'un groupe de locuteurs non-natifs d'anglais, à l'aide d'une étude qui prend pour modèle l'une des expérimentations présentées par Granger dans son article de 1998. Pour ce faire, nous avons recours à deux types d'approche de l'étude des séquences formulaïques (approches phraséologiques et distributionnelles) afin de déterminer si des locuteurs natifs et non-natifs jugent des collocations adverbe-adjectif de manière semblable selon des critères de saillance et d'accord. Nous avons sollicité la participation de 55 anglophones et de 120 apprenants francophones de l'anglais à trois niveaux: des étudiants en première année d'université, des étudiants en troisième année d'université et des étudiants en Master. Deux expérimentations ont été menées et l'analyse quantitative des résultats révèle que les anglophones et les étudiants en Master se sont comportés de manière semblable. Les différences entre les natifs et les étudiants en Master d'un côté, et les étudiants en première et en troisième année de l'autre, se

sont avérées significatives. Ces résultats semblent indiquer un développement dans les connaissances collocationnelles chez les non-natifs qui mène vers une maîtrise proche de celle des natifs.

Mots clés: intensificateurs, collocations, séquences formulaïques, saillance

For several decades now, researchers interested in questions concerning formulaic language[1] have pointed out the importance of such phenomena in linguistic analyses in general and for analyses specific to the field of second language acquisition (SLA) in particular (see discussions in the following recent collected volumes: Barfield & Gyllstad, 2009; Corrigan, Moravcsik, Ouali, & Wheatley, 2009; Granger & Meunier, 2008; Meunier & Granger, 2008). Over this same period, two important developments in the field of linguistics have taken place, each of which has supported this renewed interest in formulaic language. On the one hand, corpus linguistics has come into its own, providing the tools necessary for the examination of lexical patterns in large data sets (Kucera & Francis, 1967; Paquot & Granger, 2012; Sinclair, 1991). On the other hand, the importance of such patterns has led to numerous theoretical innovations that have drawn attention to the importance of lexis (versus grammar), although to varying degrees and within different frameworks (Fillmore, 1988; Jackendoff, 2002; Sinclair, 1991). In the realm of SLA, it was Pawley and Syder's landmark (1983) article that clearly established the importance of formulaic language. As they pointed out, nonnative speakers (NNSs) are confronted with difficulties of both nativelike selection ("the ability of the native speaker routinely to convey his meaning by an expression that is not only grammatical but also nativelike," p. 191) and nativelike fluency ("the native speaker's ability to produce fluent stretches of spontaneous connected discourse," p. 191). It is the first of these issues that interests us in the current project.

Numerous researchers have investigated the difficulties associated with selecting a string that is not only grammatical but also nativelike in a second language (L2) (i.e., nativelike selection). Results from such studies have generally shown that mastery of

formulaic language is acquired late in L2 acquisition. In the current article, we present results from a research project in which we investigated a subset of formulaic language – collocations – in native and nonnative English. Our project was conceived of as a partial replication of Granger (1998), a study which examined (among other things) amplifier-adjective collocations. We intend to build upon the insights offered in Granger's article to contribute to discussions concerning the saliency of collocations for NNSs and agreement with a native norm across several proficiency levels.

Collocations

What are they?

Although most researchers would probably agree with Tutin and Grossmann (2002), who defined *collocation* as a "cooccurrence lexicale privilégiée de deux éléments linguistiques entretenant une relation syntaxique" [preferential lexical co-occurrence of two linguistic elements between which there exists a syntactic relationship](p. 8), the nature of the relationship in question continues to be debated, as highlighted in Nesselhauf's (2005) description of collocation as "some kind of syntagmatic relationship of words" (p. 11). Recent work on collocations and on formulaic phenomena in general has highlighted two approaches to phraseology, which Granger and Paquot (2008) referred to as phraseological (or traditional) approaches as opposed to distributional ones. Phraseological approaches to formulaic language rely on traditionally linguistic criteria (syntactic, semantic, and/or pragmatic) to identify the strings of interest. In the case of collocations, Nesselhauf's (2003) criterion of "arbitrary restriction on substitutability" (p. 225) is an example of one such linguistically based criterion. If authors working within such an approach agree on the importance of linguistic criteria, disagreements remain concerning which are able to best draw the line

between idioms (i.e., syntactically and/or semantically noncompositional strings) and collocations, on the one hand, and between collocations and free expressions, on the other. A small sample of strings identified as collocations in studies adopting a predominantly phraseological approach is provided below:

- *pay attention* (Howarth, 1998)
- *profondément enraciné* "deeply rooted" (Granger, 1998)
- *long time* (Siyanova & Schmitt, 2008)
- *take a picture* (Nesselhauf, 2003)

Distributional approaches, on the other hand, define formulaic sequences (including collocations) with respect to the frequency with which a string is used and/or to the strength of the statistical co-occurrence between the different members of the string. Statistical co-occurrence measures, which include Mutual Information (MI) scores[2] and *t*-scores,[3] compare the actual frequency with which two or more elements co-occur in a given corpus to their expected frequency of co-occurrence given the size of the corpus and the frequencies of the individual words. Those authors who rely exclusively on frequency or statistical co-occurrence criteria for the identification of formulaic language may claim to be interested in *lexical bundles*, *recurrent word units*, or *collocations*. Biber, Conrad, and Cortes' (2004) work on lexical bundles, which relies exclusively on an observed frequency criterion of 40 occurrences per million words for the identification of such sequences, is representative of this trend. Some of the lexical bundles identified in their project include *the end of the*, *I mean I don't*, and *want to do is* (p. 381).

Finally, many authors attempt to reconcile the two approaches, incorporating both linguistic- and frequency-based information into their definitions and identification criteria (e.g., Durrant & Schmitt, 2009; Shin & Nation, 2008; Wolter & Gyllstad, 2011). Thus, for Gyllstad (2009), for example, a collocation is a sequence of at least two words that are

"frequently co-occurring" and "where one of the words is used in a figurative, delexical, or technical sense" (p. 155), thus combining both distributional and phraseological approaches to the study of collocations. Examples cited in Gyllstad's study include *run a business* (p. 157) and *catch a cold* (p. 158).

The current project: Amplifier-adjective collocations

Although the current study can be considered to present a predominantly phraseological approach to the study of collocations, our analysis and interpretation will combine aspects of both types of approaches, as issues of frequency and collocational strength will be taken into consideration in the analysis. The choice of a phraseological approach was in large part determined by the design of our study, which is a partial replication of Granger (1998). In her study, which will be presented in more detail in the following section, Granger examined amplifier-adjective collocations such as *fully aware* and *blissfully ignorant*. These same combinations are also the target in the current project. Following Granger, we adopted Quirk, Greenbaum, Leech, and Svartvik's (1985) definition of amplifiers as adverbs that express an increase in intensification starting from a given norm. Quirk et al. further divided amplifiers into *maximizers* and *boosters*: maximizers refer to those amplifiers denoting "the upper extreme of the scale" (e.g., *absolutely*, *fully*, *perfectly*, *utterly*), whereas boosters designate amplifiers denoting "a high degree, a high point on the scale" (p. 591), for example, *extremely*, *heavily*, *highly*, and *blissfully*.

NNSs and collocations

General findings

Using a translation task and a cloze task, Bahns and Eldaw (1993) analyzed data collected from German-speaking learners of English. Their results led them to conclude that "a knowledge of collocations is essential to full communicative mastery of English" (p. 109) and that, moreover, collocations should be taught in language classrooms. Since that time, research into L2 learners' knowledge (and use) of L2 collocations has grown (see Paquot & Granger, 2012, for an overview of studies based on learner corpora). Numerous authors have found that NNSs' collocation knowledge lags both behind that of native speakers (NSs) (e.g., Erman, 2009; Fan, 2009; Farghal & Obiedat, 1995) and behind their general L2 vocabulary knowledge (Arnaud & Savignon, 1997; Bahns & Eldaw, 1993; Gyllstad, 2009). Despite the robustness of this finding, recent research examining Swedish-speaking learners of French has found that highly advanced NNSs who have been living in the target-language community are indistinguishable from NSs when their use of lexical collocations is examined (Bartning, Forsberg, & Hancock, 2009; Forsberg, 2010).[4] In the case of Forsberg (2010), the author examined semi-structured interviews lasting from 15 to 20 minutes between a Swedish L1 (first language)-French L2 speaker and a NS of French. Data from four groups of NNSs were included: beginners, high school students, advanced university students, and very advanced L2 users (who had been living in France for between 4.5 and 11 years). Forsberg consistently found significant differences between her NSs in comparison to the beginners, the high school students, and the advanced university students. Interestingly, no such differences were found between the very advanced L2 users and the NSs with respect to use of lexical collocations.

In several recent studies, the attention has turned to the distributional characteristics of collocations used and recognized by NNSs (Durrant & Doherty, 2010; Ellis, Simpson-Vlach, & Maynard, 2008; Siyanova & Schmitt, 2008; Siyanova-Chanturia, Conklin, & van Heuven, 2011). In these studies, both NNSs and NSs were found to show sensitivity to distributional characteristics of collocations in their processing of such strings. Ellis et al., for example,

presented several experiments in an attempt to determine whether corpus-derived formulaic strings were psycholinguistically valid. Several corpora of academic English were consulted, and all three-, four-, and five-word sequences occurring at least 10.9 times per million words were extracted. For each string, frequency and MI scores were calculated from the corpora. The first task asked 11 NNSs and 11 NSs to judge whether or not a string was English, whereas 6 NNSs and 6 NSs were recruited to read aloud the same test strings in Task 2. In both cases, reaction times were recorded and used as the dependent variable in multiple regression analyses. Results showed that speed of NS responses in the grammaticality judgment task and in the read-aloud task was significantly predicted by the MI score (the more strongly the string cohered, the faster the NSs responded). A significant predictor for the learner responses on both tasks, on the other hand, was frequency (higher frequency strings were reacted to more quickly). According to Ellis et al., their results provide evidence showing "that formulaic expressions can be identified statistically from corpora of usage, and that native speakers and advanced ESL learners have become sensitive from their usage histories to these expressions so that they process them preferentially. But native speakers and learners are sensitive to different determinants of fluency — learners to *n*-gram frequency, fluent natives to MI" (p. 389). Thus, both NNSs and NSs are argued to be sensitive to distributional patterns in the L2, albeit to different ones.

Amplifier-adjective collocations

There are at least two studies that precede our own that have looked at amplifier-adjective collocations in L2 English: Lorenz (1999) and Granger (1998). In his study, Lorenz set out to examine adjective intensification in native and nonnative English in all of its forms (including, but not limited to, amplifier-adjective collocations). The author based his study on

two native and two nonnative (German L1) corpora of written argumentative writing in English. Several conclusions come out of this project, of which we will mention but two. First, Lorenz finds a significant overuse of high frequency intensifiers (such as *very*) among his NNSs. Second, after having calculated the MI scores and frequencies for each intensifier-adjective pair, Lorenz was able to conclude that learners' combinations had lower MI scores than did those used by NSs, implying less cohesion in the sequences produced by the NNSs.

Granger's (1998) article is the starting point for the current project, as we have set out to replicate one portion of the study she reported. In her article, Granger examined amplification in native and nonnative (French L1) English using a written corpus. Following this examination, she administered a word combination task to 56 NSs of English and 56 French L1-English L2 speakers. The purpose of this task was "to extract introspective data on collocations" (p. 152). To do so, Granger selected 11 amplifiers and 15 adjectives, and asked participants to select which combinations (amplifier-adjective) were acceptable in English. The presentation of the task is shown in Figure 1 (see description in footnote 5 of Granger's article).

readily significant reliable ill different essential aware miserable
available clear happy difficult ignorant impossible cold important
bitterly significant reliable ill different essential aware miserable
available clear happy difficult ignorant impossible cold important

Figure 1. Example experimental items from word combination task (Granger, 1998)

Participants were instructed to circle all adjectives that collocated with the amplifier in question and to place an asterisk beside any adjective they felt was more frequently associated with the amplifier than the others (what we will call "best combinations"). The results reported by Granger concerned only those combinations that received an asterisk, combinations that Granger argued were "particularly salient in the subjects' minds" (p. 152). The analysis provided by the author examined the total number of asterisked pairs (token analysis), as well as which pairs were selected by the two groups (type analysis). The token

analysis revealed that NNSs significantly underselected best combinations when compared to NSs: As a group, NNSs attributed 280 asterisks, whereas the NSs placed 384. With respect to types of combinations asterisked, the NNSs selected a greater number of types than the NSs, indicating less agreement within the NNS group than within the NS group with respect to what constitutes the best combinations. On the basis of these results, Granger (1998) concluded that her "learners' sense of salience [was] not only weak, but also partly misguided" (p. 152).

Taking these results as a starting point, our study set out to further explore issues of collocational saliency and agreement among NSs and NNSs of English, using a collocation judgment questionnaire inspired by Granger's (1998) word combination task. However, it must be stated that we have introduced three important differences to the task, to respond to some of the limitations in Granger's study. First, we have added a fill-in-the-blank production task to help in the interpretation of the collocation judgment results. Second, we have carried out a quantitative analysis of the judgment data that took into consideration the entire set of responses. This approach contrasts with Granger's type analysis, which relied on a small number of examples. Third, we have collected data from NNSs at three proficiency levels, which will allow us to draw conclusions concerning proficiency and collocation judgments. Proficiency was not addressed in Granger's project, as her 56 NNSs belonged to an undifferentiated group.[5] Two research questions guided this project:

1. Do NSs and NNSs reveal similar patterns of salience in their judgments of amplifier-adjective sequences?
2. Do NSs and NNSs agree with a NS norm in their judgments and their use of amplifier-adjective sequences?

Although Granger does not define *salience*, the interpretation of her data indicates that every match provided on her word combination task was considered indicative of a salient

match for the participant in question, with those combinations receiving an asterisk being interpreted as particularly salient. In our own project, we follow this implicit definition of salience, considering that the identification of any combination as a good match indicates that the match in question was salient to the participant.[6]

Methodology

In what follows, we will begin by describing our participants. We will then explain the selection process employed to choose which amplifiers and adjectives to test. Finally, we will describe the design and analyses associated with our two tasks.

Participants

Native English participants were recruited from among undergraduate students at a university in England. Nonnative participants were all French NSs studying for an English degree (undergraduate or graduate) at a university located in France. NNSs were either first-year English majors, third-year English majors, or enrolled in an English Master's program. No participant completed both tasks, and there were a total of 89 participants for Task 1 and 86 for Task 2. A short background questionnaire was included with each task; information collected concerning age, time abroad in an English-speaking country, and time studying English is presented in Table 1.

Table 1. Participant details

Task	*n*	Age (*M*)	Time abroad[a] (weeks)	Time studying English (years)
Collocation judgment				
NSs	26	19.2	—	—
Master	12	26.5	94	13
Third year	30	23.6	9.3	10
First year	21	19.3	9.2	8.7
Fill-in-the-blank				
NSs	29	20	—	—
Master	11	22.4	24.3	11

Third year	23	20	9.6	9.4
First year	23	19	9.7	8.6

[a]In an English speaking country.

Amplifier and adjective selection

Our research project targeted 13 amplifiers and 18 adjectives. The targeted items were taken from two studies addressing amplifier-adjective collocations in English: Granger (1998) and Kennedy (2003). With respect to the amplifiers, 10 of the 11 items included in Granger's word combination task were included in our experiment; the remaining three amplifiers (*extremely, greatly, terribly*) were selected from among those examined by Kennedy. As for the adjectives, 14 of the original 15 from Granger's study were retained, with four additional ones (*appreciated*, *criticized*, *devastated*, *sorry*) being included in our project. Thus, all amplifiers and adjectives from Granger's article, with the exception of *bitterly* and *cold*, were included in our project, and seven items from Kennedy's study completed our target list. The seven items selected from Kennedy's study were retained because they met two criteria. First, we wanted to include items that would be familiar to intermediate-level NNSs (judgments as to potential familiarity were reached in agreement with English professors who taught the NNSs in question). Second, a separate goal of this project (which will not be addressed in the current article) was to examine semantic prosody, which guided the selection of lexical items.

Task 1: Collocation judgment questionnaire

Task 1 constituted a partial replication of Granger's (1998) word combination task. The presentation was the same, and the participants were asked both to select all acceptable combinations and to indicate the best combination for each adverb. The instructions read:

> In what follows, you have one adverb followed by a set of adjectives.
>
> Circle all adjectives that you believe naturally follow the adverb in English.

Finally, put an asterisk next to the adjective (or adjectives) that you think is most commonly associated with the adverb in question.[7]

Two versions of this task were compiled; for each version, the order of presentation for both the adverbs and the adjectives was changed. Figure 2 provides two sample items.

perfectly essential significant difficult ill different aware sorry available appreciated clear ignorant important criticized devastated miserable reliable happy impossible
readily essential significant difficult ill different aware sorry available appreciated clear ignorant important criticized devastated miserable reliable happy impossible

Figure 2. Example items from collocation judgment task

Several analyses of the data collected from this task were undertaken. With respect to research question 1, we first compared the number of combinations selected by NSs and NNSs, looking at both best combinations (those that had received an asterisk) and acceptable combinations (all combinations that were circled). We then conducted a type analysis, to determine the total number of combinations per amplifier per group, a measure that gave us an indication of within-group salience and agreement. For research question 2, we looked to individual patterns in an attempt to determine to what extent each participant's responses agreed with two NS norms. Whereas Granger (1998) addresses agreement as part of her type analysis, citing a small number of examples to show that her learners asterisked combinations that were not selected by the native group (see Granger, Table 7.7, p. 153), we chose a different strategy insofar as we set out to quantify the idea of agreement, a step that allowed us to examine the full data set (essentially determining overall agreement instead of looking item by item, combination by combination) and to perform statistical analyses on our data. To quantify the collocation judgments, we first needed to determine what responses were expected. In other words, we needed to establish an answer key for both best combinations and acceptable combinations against which the responses of each of our participants would be compared. To design these answer keys, we turned to principles associated with distributional

approaches to formulaic language – that is, frequency and collocational strength scores. As mentioned in the literature review, there is some recent evidence that NNSs are sensitive to frequency in their L2, whereas NSs are more sensitive to collocational strength. For this reason, we created two answer keys for each of best combinations and acceptable combinations: one based on observed frequency, and the other on *t*-scores.[8] Frequencies and *t*-scores were noted for all amplifier-adjective combinations on the basis of the British National Corpus (BNC). To construct the "best combination answer keys" for each amplifier, the adjective with which it occurred most frequently and the adjective with which it received the highest *t*-score were considered the best combinations on the basis of frequency and collocational strength, respectively. For acceptable combinations, however, a cut-off had to be adopted. In the case of frequency, we considered that any sequence occurring more than 10 times in the BNC would be considered acceptable. As for *t*-scores, the commonly accepted cut-off of 2 was used (Hunston, 2002). As a result, 32 amplifier-adjective combinations were considered acceptable using our measure of frequency, whereas 48 such combinations were deemed acceptable using *t*-scores. The answer keys are provided in the Appendix, in Tables A1 and A2.

Task 2: Fill-in-the-blank production task

To complement the collocation judgment data, a production task was created. In creating this task, we queried the BNC and the Corpus of Contemporary American English[9] for example strings involving the best combinations for each of the 13 amplifiers according to both frequency and *t*-scores. For 9 of the 13 amplifiers, the same combination was considered "best" on the basis of the two measures, meaning that only one string targeted each of these adverbs in Task 2. For *blissfully*, *perfectly*, and *absolutely*, however, best combinations differed on the basis of the two measures, and thus two strings targeting these adverbs were

included. Finally, although two strings targeting *utterly* should have been included, because of a design oversight only the string involving the most frequent combination (*utterly different*) was presented. This gave us a total of 16 strings, which were presented in two randomized orders, and in which the targeted adverb was replaced by a blank (see examples in Table 2). The example strings retained were randomly selected from among the occurrences found in the two corpora, and the final list contained an equal number of British and American examples. The following instructions were provided to participants:

> Each of the following sentences has been taken from a corpus of authentic English and one of the original words has been replaced by a blank. Read each sentence and provide *one* word that you find appropriate and natural in order to fill in the blank.

Table 2. *Examples of items from fill-in-the-blank task*

Targeted Amplifier	Experimental Item
perfectly	From the tone of his voice, it was _____ clear that his decision was not open to negotiation.
perfectly	Despite lavish celebrity weddings, a multitude of dating websites and stacks of self-help books about finding your soul mate, singles are a growing segment of the population – and increasingly say they are _____ happy with their singlehood, thank you very much.
readily	Since they were mass-produced, they are _____ available in the market today.

The data collected using this task were subjected to an analysis in which every occurrence of each of the 13 targeted amplifiers was examined. These patterns will be brought to bear when the Task 1 results with respect to research question 2 are presented.

Results

Research question 1: Salience

Any combination selected by a participant was considered to be a salient combination for that individual, with items receiving asterisks taken to be particularly salient. To examine patterns of saliency in the data, we first conducted a token analysis followed by a type analysis, mirroring the analysis presented in Granger (1998).

Token analysis

For the token analysis, we calculated the average number of best combinations circled by each participant by group,[10] as well as the average number of acceptable combinations per participant per group.[11] The results are presented in Figure 3.

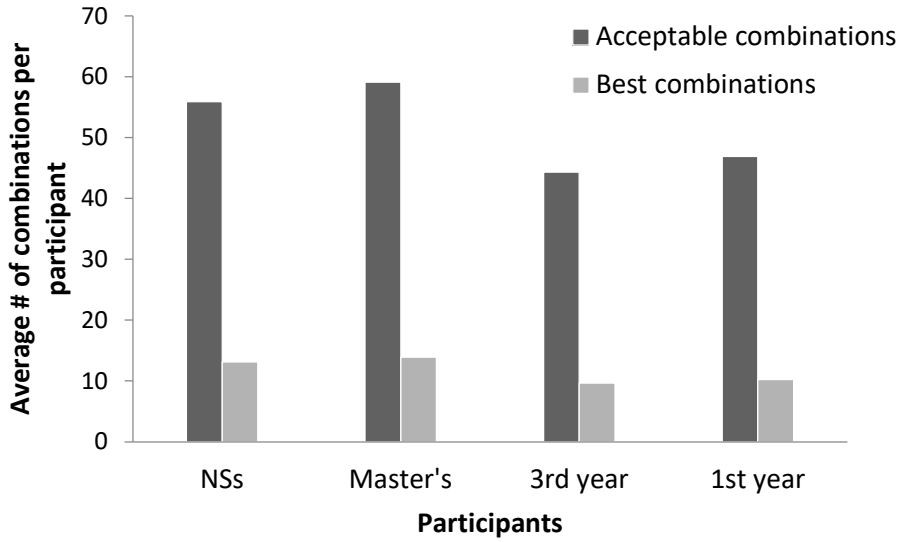

Figure 3. Average number of acceptable and best combinations selected per participant per group

One-way analyses of variance (ANOVA) were carried out for best combinations and for acceptable combinations. For best combinations, the one-way ANOVA was significant (Welch's $F(3) = 26.277$, $p < .001$). The post-hoc analysis for this and for all subsequent ANOVAs was carried out using four Gabriel's post-hoc tests (with a Bonferroni protection of $\alpha = .0125$); the Gabriel test was chosen because of the different sample sizes in our four groups. The post-hoc tests showed the NSs and Master's groups providing on average more asterisked responses than the first- and third-year participants. For the total number of

acceptable combinations, however, no significant differences among the groups were found (Welch's $F(3) = 1.035$, $p = .389$).

Type analysis

We next examined the number of combinations per adverb per group for both best combinations and all acceptable combinations. With respect to best combinations, as a group the NSs and the Master's students selected on average 5 and 4 adjectives, respectively, of the 18 possible as best combinations for each of the 13 adverbs. The first- and third-year students showed less within-group agreement as to what constitutes a best combination, with averages of 6 adjectives per adverb for the third-year students and 7 adjectives per adverb for the first-year group. Table 3 presents the adjectives selected as best combinations for the adverbs *seriously* and *fully* in each of the four groups.

Table 3. Best combinations selected for *seriously* and *fully*

	Group			
Adverb	NSs (*n* = 26)	Master's (*n* = 12)	Third year (*n* = 30)	First year (*n* = 21)
seriously	ill (18) important (2) appreciated (1) devastated (1) difficult (1)	ill (8) criticized (1) difficult (1) ignorant (1)	ill (14) devastated (4) difficult (3) different (1) important (1) reliable (1)	ill (4) devastated (4) criticized (3) difficult (2) sorry (2) different (1) important (1) impossible (1) reliable (1)
fully	aware (20) appreciated (2) aware (1)	aware (5) appreciated (3) clear (1) devastated (1) ignorant (1)	aware (13) reliable (3) different (2) clear (1) devastated (1) impossible (1) significant (1)	aware (8) different (3) devastated (2) appreciated (1) available (1) essential (1) happy (1) impossible (1) reliable (1)

Note. The numbers provided in parentheses give the number of participants in each group that asterisked the adjective in question.

When all acceptable combinations were considered, results revealed that the first- and third-year groups tended to circle a greater number of adjectives for each adverb than did the Master's group and the NS group, which is evident in Table 4.

Table 4. Number of different adjectives (ADJ) circled per adverb (ADV) per group

Group							
NSs		Master's		3rd year		1st year	
ADV	# of ADJ	ADV	# of ADJ	ADV	# of ADJ	ADV	# of ADJ
blissfully	7	blissfully	8	blissfully	12	blissfully	10
readily	8	vitally	8	vitally	14	heavily	12
vitally	9	seriously	12	highly	15	perfectly	15
fully	12	greatly	13	terribly	15	vitally	15
greatly	14	heavily	13	absolutely	16	fully	16
highly	14	perfectly	13	fully	16	greatly	16
perfectly	14	readily	13	greatly	16	readily	16
absolutely	16	highly	15	heavily	16	terribly	16
heavily	16	terribly	15	perfectly	16	absolutely	17
seriously	17	utterly	15	readily	16	extremely	17
terribly	17	absolutely	16	extremely	18	highly	17
utterly	17	fully	16	seriously	18	seriously	17
extremely	18	extremely	18	utterly	18	utterly	17
Average	13.8	Average	13.5	Average	15.8	Average	15.5

For the first- and third-year groups, 9 of the 13 adverbs were associated with 16 to 18 of the adjectives tested, indicating a diffuse idea of acceptable combinations and a weak sense of within-group salience. Master's students, on the other hand, clearly patterned with the NS participants, and both groups showed a sense of saliency that was more restricted and more consensual than did the first- and third-year groups.

Research question 2: Agreement

Agreement with a NS norm was examined through three different analyses. First, we examined agreement on best combinations (i.e., those that received an asterisk), comparing each participant's results to our two answer keys, the first one developed with respect to frequency and the second with reference to *t*-scores in the BNC. We then turned to all acceptable combinations (meaning all combinations that were circled), which were analyzed

in two ways. We first compared what percentage of each participant's circled combinations matched the expected acceptable responses on the two answer keys. For example, for the observed frequency answer key, we calculated how many of the 32 acceptable combinations had also been circled by each participant. Although this measure provides an indication of agreement on acceptable combinations, it does not take into consideration the number of combinations circled that did not belong to the answer key. More concretely, this measure makes no difference between a participant who circled all 32 of the acceptable combinations and no others and a participant who circled the same 32 and an additional 100 others. To take into consideration the number of off-answer key responses provided, in our second analysis of the acceptable combination data we divided the number of responses matching the answer key by the number of other responses provided by each participant (acceptable combinations vs. unacceptable combinations).

Best combinations

The asterisked responses provided by each participant were compared against the two best combination answer keys, and percentages were calculated representing how many of a participant's best responses matched the best combinations as determined by *t*-scores and frequency. The results by group are presented in Figure 4.

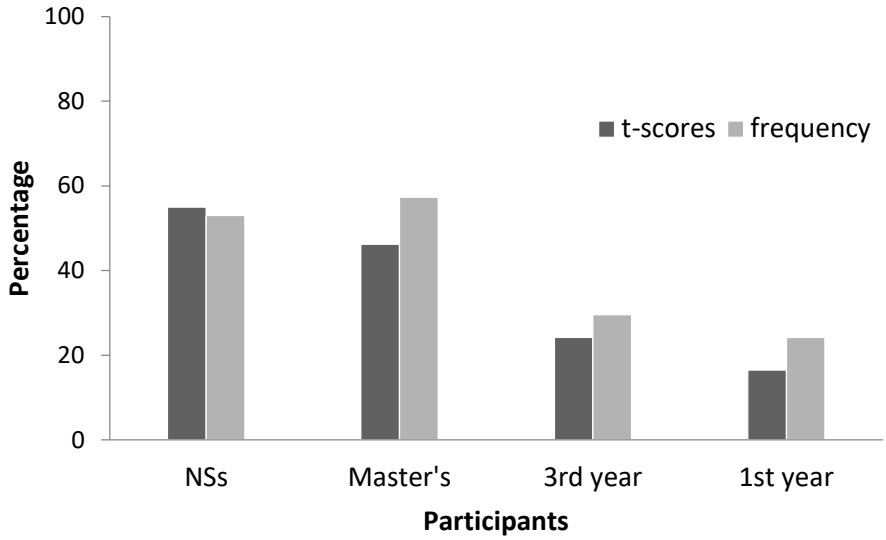

Figure 4. Percentage of best combinations matching the *t*-score and frequency answer keys

As can be seen in this figure, NSs performed similarly, regardless of whether their responses were compared to the *t*-score answer key or to the observed frequency one. The NNSs, on the other hand, consistently performed better when best combinations were defined with respect to frequency as opposed to with respect to collocational strength. As assumptions of normality in the distribution of these percentages were met, we performed one-way ANOVAs on the two data sets (*t*-score and frequency) to examine between-group differences. The results were similar, with significant *F* values being found in the analysis of both the *t*-scores, $F(3) = 50.927$, $p < .001$, and frequency, $F(3) = 31.314$, $p < .001$. In both cases, the post hoc analysis, using Gabriel's test ($p < .0125$), revealed a significant difference between the NSs and the first- and third-year groups and a significant difference between the Master's group and the first- and third-year groups; on average, approximately 50% of the combinations asterisked by the NSs and Master's students were defined as "best combinations" on the basis of frequency or collocational strength, whereas only between 16.5% and 29.6% of the first- and third-year students' asterisked responses were identified as "best combinations" in our NS norms. There was no difference between NSs and the Master's group, and no difference between the first- and third-year students.

All acceptable combinations

We next determined what percentage of those matches appearing on our two "acceptable combinations" answer keys were also circled by each participant. The results are provided in Figure 5, from which it is clear that all participant groups performed better when acceptable combinations were defined as sequences occurring more than 10 times in the BNC (frequency), as opposed to when they were defined with respect to *t*-scores.

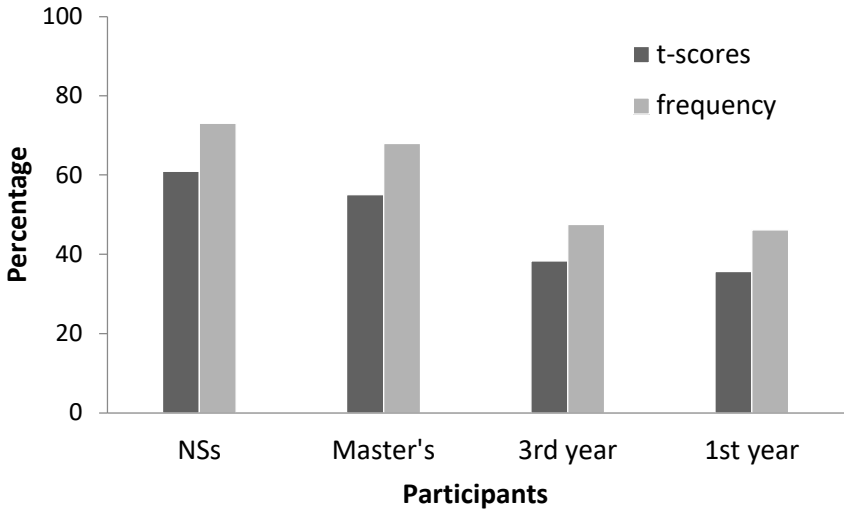

Figure 5. Percentage of acceptable combinations matching the *t*-score and frequency answer keys

We then examined the data for the possibility of between-group differences and, once again, results from the two one-way ANOVAs were similar. Specifically, a significant difference between groups was found when acceptable combinations were defined both with respect to *t*-scores ($F[3] = 18.324$, $p < .001$) and with respect to frequency ($F[3] = 17.423$, $p < .001$). The Gabriel's post-hoc tests ($p < .0125$) revealed similar patterns in both ANOVAs: a significant difference between NSs and the first- and third-year groups and a significant difference between the Master's group and the first- and third-year groups. Thus, the NSs and the Master's students circled a significantly greater number of combinations deemed acceptable on the basis of both *t*-scores and frequency than did the first- and third-year

participants. No differences were found between NSs and Master's students or between first- and third-year students.

In investigating agreement, we also considered the production data from Task 2. When all tokens of the 13 targeted amplifiers were examined, we noted that NSs used these adverbs in a greater variety of combinations (41) than did the Master's students (31 different combinations), the third-year students (18 combinations), or the first-year students (13 combinations). This is certainly in part due to the dominance of the amplifier *par excellence* – *very* – in the responses from the first- and third-year groups. Whereas *very* constitutes 10.3% and 18.3% of the NSs and Master's group's total responses, respectively, the first- and third-year groups provided this amplifier in 35% percent of the contexts. Looking at the 13 targeted adverbs, the nonnative uses matched in large part the native patterns. For the Master's students, only 13% of the combinations they used on the production task were not attested in the native responses. For the third-year group this percentage rose to 22.2%, whereas for the first-year students only 1 of the 13 combinations attested was not also found among the native responses. Some examples of combinations provided by NNSs but not attested among NS responses are

- *highly different* / *utterly significant* Master's
- *extremely criticized* / *seriously criticized* third-year
- *terribly ignorant* first-year

Acceptable combinations vs. unacceptable combinations

In our final analysis, we divided the number of acceptable combinations by the number of off-answer-key responses provided by each participant. The resultant percentage indicates what portion of a participant's total responses corresponded to each of the two answer keys. The results are presented in Figure 6.

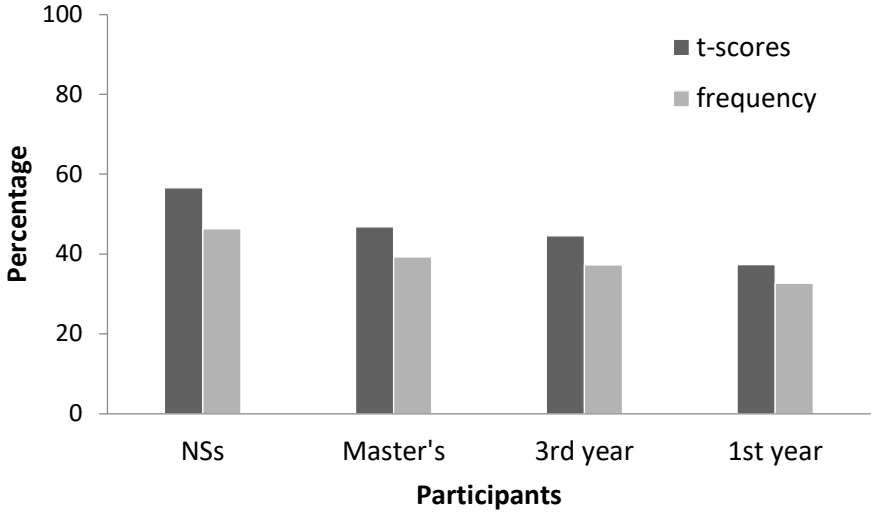

Figure 6. Percentage of acceptable vs. unacceptable combinations matching the *t*-score and frequency answer keys

The one-way ANOVAs performed on the *t*-score data set revealed significant differences among our groups, $F(3) = 11.363$, $p < .001$, shown through the Gabriel's post-hoc tests ($p < .0125$) to be attributable to a significant difference between the NSs and the first- and third-year groups: 56.6% of the responses selected by NSs also had high *t*-scores in the BNC, a percentage which drops to 44.5% for the third-year group and to 37.3% for the first-year group. The behaviour of the Master's group was not found to be significantly different from that of any of the other three groups tested. Similar findings resulted from the analysis of the frequency data set, which used a Kruskal–Wallis non-parametric test as these data were not normally distributed. It was found that there were significant differences among the groups, $H(15.422)$, $p < .01$, which three follow-up Mann–Whitney U tests (with a Bonferonni protection of $\alpha = .016$) showed to be due to a significant difference between NSs and the third-year group ($p < .016$) and between NSs and the first-year group ($p < .016$).

Discussion

Although corpus linguistics deserves much of the credit for the renewed interest in formulaic language, we agree with Lorenz (1999, p. 18) when he highlighted that NS

judgments are an important complement to data-driven approaches. His comments echo Granger's (1998) own remark concerning the utility of a task like our collocation judgment questionnaire, which in her own project she judged to be potentially "valuable in providing a clearer notion of what constitutes a significant collocation" (p. 153). Our project took a closer look at NS and NNS judgments on amplifier-adjective collocations to examine notions of salience (research question 1) and agreement (research question 2) with respect to strings that are generally accepted to be late acquired in an L2.

The analyses relevant to the question of salience revealed several important findings. First, no differences in the number of tokens were found among the four groups when all responses were considered. When we limited our analysis to number of best responses (as did Granger, 1998), we found that NSs and the Master's group attributed significantly more asterisks than did the first- and third-year groups. Granger found that her NSs placed more asterisks than did the NNS group as a whole, which she interpreted as evidence of a weak sense of salience on the part of the nonnatives, a conclusion that our results do not completely support. All groups tested in our study selected similar numbers of combinations as acceptable, suggesting that it would be erroneous to speak of a weak sense of salience for acceptable combinations. With respect to best combinations, the more advanced NNSs were indistinguishable from the NSs in the number of combinations asterisked, indicating that if number of best combinations is taken to be an indication of strength of salience, more advanced NNSs may have a sense of salience similar in strength to that of the NSs.

Similar patterns were found when types of combinations were considered. In her own data, Granger (1998) found that "[o]n balance, the learners marked a greater number of types of combinations than the natives, indicating that the learners' sense of salience is not only weak, but also partly misguided" (p. 152). Of our three NNS groups, we also found that the first- and third-year groups, on the whole, selected a greater number of combinations in the

collocation judgment questionnaire, indicating less agreement within each of these groups as to what constitutes a significant collocation (see Tables 3 and 4). The Master's students, on the other hand, once again behaved like the NSs, selecting fewer combinations as a group, a finding that was true both for all acceptable combinations and for best combinations. Thus, using number and type of combinations selected as indications of salience, we conclude that NNSs' sense of salience may be weak and (at least in part) misguided, but that proficiency is an important modulating factor. In other words, the patterns presented here point to development over the three NNS proficiency levels, in that the advanced group performed similarly to the NSs.

By quantifying the concept of agreement (research question 2), we further examined to what extent the NNSs' sense of salience was potentially misguided. In other words, although the rather low level of within-group agreement for the first- and third-year participants seen in the type analysis certainly suggests that on the whole their sense of salience was misguided, it was the analyses associated with our second research question that allowed us to quantify to what extent their judgments were off-target. In the current project, *agreement* was defined with respect to two NS norms, each of which was established by applying one distributional approach principle to data from the BNC: frequency and *t*-scores. Two different NS norms were included because of recent findings suggesting that (at least with respect to processing) NNSs are sensitive to frequency in their L2, whereas NSs are more sensitive to collocational strength. Given this finding, it is reasonable to suspect that these different sensitivities may be visible in a collocation judgment questionnaire, meaning that NNSs may perform better when best and acceptable combinations are defined with respect to frequency than when they are defined with respect to collocational strength. This possibility finds only limited support among the results presented in this article, in particular with respect to best combinations. As was shown in Figure 4, NSs showed similar scores regardless of the way in which best

responses were defined. NNSs, on the other hand, received clearly higher scores when best combinations were defined with respect to frequency in the BNC, a finding that held in all three NNS groups. For the remaining analyses, significant patterns were identical, regardless of whether measures of frequency or collocational strength were used. The question of native and nonnative sensitivity to different distributional patterns in formulaic language would merit additional attention.

The three agreement analyses both confirmed and further developed the findings reported with respect to salience. On the one hand, the Master's students continued to pattern with the NSs. When we examined results from best combinations and from all acceptable combinations, we found that these two groups were indistinct from each other, but significantly different from the first- and third-year groups. Thus, not only do the NSs and the Master's students appear to have a similar sense of salience (defined with respect to number and types of combinations selected), but they also show similar levels of agreement when their responses are compared to two NS norms. On the other hand, the findings change somewhat when we take into consideration all responses circled, effectively factoring in the number of combinations selected that were not part of the NS norm used as a yardstick. In this final analysis, NSs were found to behave significantly differently from the first- and third-year groups in analyses based on both *t*-scores and frequency; the behaviour of the Master's students was not found to significantly differ from that of the NSs or from that of the first- and third-year groups. Thus, whereas the Master's group clearly patterned with the natives as opposed to the first- and third-year groups in the analyses that took only expected responses into consideration, the final analysis (in which the number of other – off-target – responses were factored in), the Master's students' performance was different neither from the NSs nor from the lower proficiency groups. Taken together, the results with respect to agreement provide evidence of development toward nativelike patterns in the selection of best and

acceptable matches, but without a concomitant restriction in the selection of "other" combinations. This seems to indicate that identification of acceptable collocations may precede the rejection of unacceptable ones for NNSs. It is our opinion that findings such as these, highlighting development in collocational knowledge, constitute an important complement to recent research that has convincingly demonstrated developmental patterns in L2 collocation use (e.g., Cobb, 2003, Paquot & Granger, 2012).

Conclusions and future directions

This project has examined questions of salience and agreement with respect to amplifier-adjective collocations in native and nonnative English. We focused predominantly on the data collected via our collocation judgment questionnaire, which provided insight into strings that NSs and NNSs judge to be significant collocations. Results were interpreted with respect to saliency and agreement, demonstrating that the Master's group tended to behave similarly to the NS group, whereas the first- and third-year groups were similar to each other. We have argued that these patterns underscore the importance of the proficiency variable with respect to collocation judgments, a variable that was not taken into account in Granger (1998). These results echo the recent findings by Bartning et al. (2009) and Forsberg (2010), in which very advanced L2 users showed patterns of usage similar to those of NSs for lexical collocations, a finding that goes against numerous studies reporting a consistent gap between NSs and NNSs with respect to collocation use and knowledge. On the basis of our results, it appears that advanced NNSs can also approach nativelike patterns with respect to metalinguistic judgments of collocations.

In future research, we intend to further examine the results from our agreement analyses to determine whether the off-target responses selected by the NNS participants show

certain common characteristics. On the basis of recent (and very robust) findings pointing to the influence (and sometimes the interference) of L1 phraseological patterns on L2 production (see Paquot & Granger, 2012), we hypothesize that many off-target responses will be found to be licit in the NNSs' L1. With respect to formulaic language more generally, we would like to pursue the idea of *saliency*, with the goal of potentially using an operationalization of this concept as an identification criterion. We also believe that it would be fruitful to further pursue the idea that NNSs and NSs may be sensitive to different distributional characteristics of formulaic language, with an aim to determine whether such sensitivities may be visible in tasks using measures other than reaction times (e.g., Durrant & Schmitt, 2009). If this proves to be the case, such a finding would provide important evidence for theories of L2 language learning, in particular with the potential to support frequency-based or usage-based approaches. Finally, we believe that the collocation judgment questionnaire presented in this project may be applied to many other types of formulaic sequences, and that the results may be of interest to both researchers and to language teachers.

Correspondence should be addressed to *Amanda Edmonds, Université de Pau et des Pays de l'Adour, U.F.R. Lettres, Avenue du Doyen Poplawski, BP 1160, 64013 PAU Cedex. Email: amanda.edmonds@univ-pau.fr*

Start Appendix

Appendix

Table A1. Best combinations according to observed frequency and *t*-scores

	Best Combinations (Observed Frequency)		**Best Combinations (*t*-scores)**	
Adverb	Adjective	Observed Frequency	Adjective	*t*-score

extremely	difficult	488	difficult	8.7
readily	available	426	available	8.9
fully	aware	239	aware	8
seriously	ill	227	ill	8.8
vitally	important	191	important	9.2
highly	significant	156	significant	7.2
terribly	sorry	67	sorry	7.6
greatly	appreciated	66	appreciated	8.2
heavily	criticized	54	criticized	7.4
absolutely	clear	149	essential	7.4
perfectly	clear	117	happy	7.1
utterly	different	27	miserable	6.4
blissfully	happy	11	ignorant	12.9

Table A2. Acceptable combinations according to observed frequency and *t*-scores

	Acceptable Combinations			
Adverb	Adjective	Observed Frequency	Adjective	*t*-score
extremely	difficult	488	difficult	8.7
	important	279	important	7.4
	happy	24	happy	4.7
	sorry	18	sorry	4.3
	ill	12	ill	4.4
	significant	12	significant	3.6
	clear	11	clear	2.9
			reliable	3.9
readily	available	426	available	8.9
	appreciated	14	appreciated	6.1
fully	aware	239	aware	8
	appreciated	62	appreciated	7.4
	clear	12	clear	2.7
seriously	ill	227	ill	8.8
vitally	important	191	important	9.2
			aware	4.9
			significant	4.1
highly	significant	156	significant	7.2
	important	38	important	4.2
			reliable	4.1
terribly	sorry	67	sorry	7.6
	important	55	important	6.3
	difficult	27	difficult	5.7
			ill	4.9
			happy	3
			clear	2.3

greatly	appreciated	66	appreciated	8.2
	different	14	different	3.3
heavily	criticized	54	criticized	7.4
absolutely	clear	149	clear	7
	essential	122	essential	7.4
	impossible	19	impossible	4.9
	devastated	18	devastated	6.6
			reliable	3.8
			happy	2.1
perfectly	clear	117	clear	6.8
	happy	97	happy	7.1
	aware	17	aware	4.6
			reliable	4.3
utterly	different	27	different	5.1
	impossible	11	miserable	6.4
			impossible	5.2
			devastated	5.1
			ignorant	5
			happy	2.9
			clear	2.3
blissfully	happy	11	happy	9.8
			ignorant	12.9

End Appendix

Figure 1. Sample of experimental items from word combination task (Granger, 1998).
Figure 2. Sample of items from collocation judgment task.
Figure 3. Average number of acceptable and best combinations selected per participant per group.
Figure 4. Percentage of best combinations matching the *t*-score and frequency answer keys.
Figure 5. Percentage of acceptable combinations matching the *t*-score and frequency answer keys.
Figure 6. Percentage of acceptable vs. unacceptable combinations matching the *t*-score and frequency answer keys.

Notes

1. We will use this term to cover the ensemble of phraseological phenomena.

2. The equation used to calculate MI scores is the following: $\log_2$ (observed freq/expected freq).

3. The equation used to calculate *t*-scores is the following: $\log_2$ (observed freq - expected freq) / √ expected freq.

4. Although the authors refer to these strings as lexical conventional sequences, the description given of the phrasal/denotative lexical conventional sequences matches what many other authors would refer to as collocations. Examples include *avoir envie* "to feel like," *faire du sport* "practice a sport," and *poser une question* "ask a question" (Forsberg, 2010, p. 35).

5. An anonymous reviewer points out that the participants in Granger's (1998) study were identified as advanced French-speaking learners of English. This is true for the corpus analysis that constitutes the main of Granger's article; however, it is not clear that the 56 participants in the word combination have a similar profile.

6. The question of salience can also be linked to issues of input, awareness, and noticing (Carroll, 2006; Schmidt, 1990). However, these issues will not be further explored in the current article.

7. In Granger's (1998) project, the description of the instructions states that if participants "felt that one adjective was more frequently associated with the amplifier than all the others, they were requested to mark it with an asterisk" (p. 152).

8. MI scores and *t*-scores are both measures of collocational strength. However, according to Gries (2010), MI scores have a tendency to return high scores for low-frequency or technical terms. He argues that a *t*-score "provides a better measure of the non-randomness of the co-occurrence" (p. 283). Although we report only results from the *t*-score analysis in this article, we conducted the same analyses using MI scores, and the results from the *t*-score and MI score analyses were non-distinct.

9. Although in the current article we are presenting data from the British native group, we have also collected data from American English speakers, which is why examples from both corpora were used in creating this task.

10. The number of best combination responses ranged from 12 to 14 for the NSs, 12 to 23 for the Master's students, 1 to 13 for the third-year students, and 7 to 13 for the first-year students.

11. The number of acceptable combination responses per participant ranged from 13 to 99 for the NSs, 25 to 104 for Master's students, 11 to 91 for the third-year students, and 21 to 76 for the first-year students.

References

<edb>Arnaud, P., & Savignon, S. (1997). Rare words, complex lexical units and the advanced learner. In J. Coady & T. Huckin (Eds.), *Second language vocabulary acquisition: A rationale for pedagogy* (pp. 157–173). Cambridge: Cambridge University Press.</edb>

<jrn>Bahns, J., & Eldaw, M. (1993). Should we teach EFL students collocations? *System*, *21*(1), 101–114. http://dx.doi.org/10.1016/0346-251X(93)90010-E</jrn>

<edb>Barfield, A. & Gyllstad, H. (Eds.). (2009). *Research collocations in another language: Multiple interpretations*. London: Palgrave Macmillan. http://dx.doi.org/10.1057/9780230245327.</edb>

<jrn>Bartning, I., Forsberg, F., & Hancock, V. (2009). Resources and obstacles in very advanced L2 French: Formulaic language, information structure and morphosyntax. *Eurosla Yearbook*, *9*(1), 185–211. http://dx.doi.org/10.1075/eurosla.9.10bar</jrn>

<jrn>Biber, D., Conrad, S., & Cortes, V. (2004). *If you look at…*: Lexical bundles in university teaching and textbooks. *Applied Linguistics*, *25*(3), 371–405. http://dx.doi.org/10.1093/applin/25.3.371</jrn>

<conf>Carroll, S.E., Grantham O'Brien, M., Shea, C., & Archibald, J. (Eds.) (2006). Salience, awareness and SLA. In M. Grantham O'Brien, C. Shea, & J. Archibald (eds.), *Proceedings of the 8th Generative Approaches to Second Language Acquisition Conference: The Banff Conference* (pp. 17–24). Somerville, MA: Cascadille Proceedings Project.</conf>

<jrn>Cobb, T. (2003). Analyzing late interlanguage with learner corpora: Quebec replications of three European studies. *Canadian Modern Language Review*, *59*(3), 393–423. http://dx.doi.org/10.3138/cmlr.59.3.393</jrn>

<edb>Corrigan, R., Moravcsik, E.A., Ouali, H., & Wheatley, K.M. (Eds.) (2009). *Formulaic language* (Vols. 1 & 2). Amsterdam: Benjamins.</edb>

<ths>Durrant, P., & Doherty, A. (2010). Are high-frequency collocations psychologically real? Investigating the thesis of collocational priming. *Corpus Linguistics & Linguistic Theory, 6,* 125–155.</ths>

<jrn>Durrant, P., & Schmitt, N. (2009). To what extent do native and non-native writers make use of collocations? *International Review of Applied Linguistics in Language Teaching*, *47*(2), 157–177. http://dx.doi.org/10.1515/iral.2009.007</jrn>

<jrn>Ellis, N.C., Simpson-Vlach, R., & Maynard, C. (2008). Formulaic language in native and second language speakers: Psycholinguistics, corpus linguistics, and TESOL. *TESOL Quarterly*, *42*, 375–396.</jrn>

<edb>Erman, B. (2009). Beyond the single word: Collocations in the writings of native speakers and first term university students of English. In R. Bowen, M. Mobärg, & S. Ohlander (Eds.), *Corpora and discourse – and stuff. Papers in honour of Karin Aijmer* (pp. 23–36). Gothenburg: University of Gothenburg.</edb>

<jrn>Fan, M. (2009). An exploratory study of collocational use by ESL students: A task based approach. *System*, *37*(1), 110–123. http://dx.doi.org/10.1016/j.system.2008.06.004</jrn>

<jrn>Farghal, M., & Obiedat, H. (1995). Collocations: A neglected variable in EFL. *International Review of Applied Linguistics in Language Teaching*, *33*(4), 315–331. http://dx.doi.org/10.1515/iral.1995.33.4.315</jrn>

<jrn>Fillmore, C. (1988). The mechanisms of construction grammar. *Berkeley Linguistics Society, 14,* 35–55.</jrn>

<jrn>Forsberg, F. (2010). Using conventional sequences in L2 French. *International Review of Applied Linguistics in Language Teaching*, *48*(1), 25–51. http://dx.doi.org/10.1515/iral.2010.002</jrn>

<edb>Granger, S. (1998). Prefabricated patterns in advanced EFL writing: Collocations and formulae. In A.P. Cowie (Ed.), *Phraseology: Theory, analysis, and applications* (pp. 145–160). Oxford: Clarendon.</edb>

<edb>Granger, S., & Meunier, F. (Eds.). (2008). *Phraseology: An interdisciplinary perspective*. Amsterdam: Benjamins.</edb>

<edb>Granger, S., & Paquot, M. (2008). Disentangling the phraseological web. In S. Granger & F. Meunier (Eds.), *Phraseology: An interdisciplinary perspective* (pp. 27–50). Amsterdam: Benjamins.</edb>

<edb>Gries, S.T. (2010). Useful statistics for corpus linguistics. In S. Aquilino & A. Moises (Eds.), *A mosaic of corpus linguistics: Selected approaches* (pp. 269–291). Frankfurt am Main: Peter Lang.</edb>

<edb>Gyllstad, H. (2009). Designing and evaluating tests of receptive collocation knowledge: COLLEX and COLLMATCH. In A. Barfield & H. Gyllstad (Eds.), *Researching collocations in another language – multiple interpretations* (pp. 153–170). Palgrave Macmillan.</edb>

<jrn>Howarth, P. (1998). Phraseology and second language proficiency. *Applied Linguistics*, *19*(1), 24–44. http://dx.doi.org/10.1093/applin/19.1.24</jrn>

<bok>Hunston, S. (2002). *Corpora in applied linguistics*. Cambridge: Cambridge University Press. http://dx.doi.org/10.1017/CBO9781139524773.</bok>

<bok>Jackendoff, R. (2002). *Foundations of language: Brain, meaning, grammar, evolution*. Oxford: Oxford University Press.</bok>

<jrn>Kennedy, G. (2003). Amplifier collocations in the British National Corpus: Implications for English language teaching. *TESOL Quarterly*, *37*(3), 467–487. http://dx.doi.org/10.2307/3588400</jrn>

<bok>Kucera, H., & Francis, W.N. (1967). *Computational analysis of present-day American English*. Providence, RI: Brown University Press.</bok>

<bok>Lorenz, G. (1999). *Adjective intensification – learners versus native speakers*. Amsterdam: Rodopi.</bok>

<edb>Meunier, F. & Granger, S. (Eds.). (2008). *Phraseology in foreign language learning and teaching*. Amsterdam: Benjamins.</edb>

<jrn>Nesselhauf, N. (2003). The use of collocations by advanced learners of English and some implications for teaching. *Applied Linguistics*, *24*(2), 223–242. http://dx.doi.org/10.1093/applin/24.2.223</jrn>

<bok>Nesselhauf, N. (2005). *Collocations in a learner corpus*. Amsterdam: Benjamins.</bok>

<jrn>Paquot, M., & Granger, S. (2012). Formulaic language in learner corpora. *Annual Review of Applied Linguistics*, *32*, 130–149. http://dx.doi.org/10.1017/S0267190512000098</jrn>

<edb>Pawley, A.K., & 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. 191–225). London: Longman.</edb>

<bok>Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1985). *A comprehensive grammar of the English language*. London: Longman.</bok>

<jrn>Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, *11*(2), 129–158. http://dx.doi.org/10.1093/applin/11.2.129</jrn>

<jrn>Shin, D., & Nation, P. (2008). Beyond single words: The most frequent collocations in spoken English. *ELT Journal*, *62*(4), 339–348. http://dx.doi.org/10.1093/elt/ccm091</jrn>

<bok>Sinclair, J. (1991). *Corpus, concordance, and collocation*. Oxford: Oxford University Press.</bok>

<jrn>Siyanova, A., & Schmitt, N. (2008). L2 learner production and processing of collocation: A multi-study perspective. *Canadian Modern Language Review*, *64*(3), 429–458. http://dx.doi.org/10.3138/cmlr.64.3.429</jrn>

<jrn>Siyanova-Chanturia, A., Conklin, K., & van Heuven, W.J.B. (2011). Seeing a phrase “time and again” matters: The role of phrasal frequency in the processing of multiword sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(3), 776–784. http://dx.doi.org/10.1037/a0022531</jrn>

<jrn>Tutin, A., & Grossmann, F. (2002). Collocations régulières et irrégulières : esquisse de typologie du phénomène collocatif. *Revue Française de Linguistique Appliquée*, *7*(1), 7–26.</jrn>

<jrn>Wolter, B., & Gyllstad, H. (2011). Collocational links in the L2 mental lexicon and the influence of L1 intralexical knowledge. *Applied Linguistics*, *32*(4), 430–449. http://dx.doi.org/10.1093/applin/amr011</jrn>