

Contents

<i>List of Figures</i>	ix
<i>List of Tables</i>	xi
<i>Acknowledgements</i>	xiii
<i>Notes on Contributors</i>	xiv
1 Introduction: Researching L2 Collocation Knowledge and Development <i>Andy Barfield and Henrik Gyllstad</i>	1
Part I L2 Collocation Learner Corpus Research	
2 Effects of Second Language Immersion on Second Language Collocational Development <i>Nicholas Groom</i>	21
3 Sound Evidence: Phraseological Units in Spoken Corpora <i>Phoebe M. S. Lin and Svenja Adolphs</i>	34
4 Exploring L1 and L2 Writing Development through Collocations: A Corpus-based Look <i>Randi Reppen</i>	49
5 Commentary on Part I: Learner Corpora: A Window onto the L2 Phrasicon <i>Sylviane Granger</i>	60
Part II L2 Collocation Lexicographic and Classroom Materials Research	
6 Towards Collocational Webs for Presenting Collocations in Learners' Dictionaries <i>Susanne Handl</i>	69
7 Japanese Learners' Collocation Dictionary Retrieval Performance <i>Yuri Komuro</i>	86
8 Designing Pedagogic Materials to Improve Awareness and Productive Use of L2 Collocations <i>Jingyi Jiang</i>	99
9 Commentary on Part II: Exploring Materials for the Study of L2 Collocations <i>Hilary Nesi</i>	114

Part III L2 Collocation Knowledge Assessment Research

- | | | |
|----|--|-----|
| 10 | Evaluating a New Test of <i>Whole</i> English Collocations
<i>Robert Lee Revier</i> | 125 |
| 11 | Toward an Assessment of Learners' Receptive and
Productive Syntagmatic Knowledge
<i>June Eyckmans</i> | 139 |
| 12 | Designing and Evaluating Tests of Receptive Collocation
Knowledge: <i>COLLEX</i> and <i>COLLMATCH</i>
<i>Henrik Gyllstad</i> | 153 |
| 13 | Commentary on Part III: Developing and Validating
Tests of L2 Collocation Knowledge
<i>John Shillaw</i> | 171 |

Part IV L2 Collocation Learner Process and Practice Research

- | | | |
|----|---|-----|
| 14 | Collocation Learning through an 'AWARE' Approach: Learner
Perspectives and Learning Process
<i>Yang Ying and Marnie O'Neill</i> | 181 |
| 15 | Learning Collocations through Attention-Drawing
Techniques: A Qualitative and Quantitative Analysis
<i>Elke Peters</i> | 194 |
| 16 | Following Individuals' L2 Collocation Development
over Time
<i>Andy Barfield</i> | 208 |
| 17 | Commentary on Part IV: Processes in the Development
of L2 Collocational Knowledge – A Challenge for
Language Learners, Researchers and Teachers
<i>Birgit Henriksen and Lars Stenius Stæhr</i> | 224 |
| 18 | Conclusion: Navigating L2 Collocation Research
<i>Alison Wray</i> | 232 |

	<i>References</i>	245
--	-------------------	-----

	<i>Index</i>	266
--	--------------	-----

1

Introduction: Researching L2 Collocation Knowledge and Development

Andy Barfield and Henrik Gyllstad

The collocation gap in second language acquisition research

For anyone learning or teaching a second language, collocation is undoubtedly one of the most fascinating (and at times frustrating) challenges that they will face. Equally, for those interested in researching second language (L2) collocation knowledge and development, the challenges are both fascinating and frustrating, but for different reasons. Although several wide-ranging volumes of research in L2 vocabulary acquisition have been published in the last 15 years or so (Arnaud and Béjoint, 1992; Coady and Huckin, 1997; Schmitt and McCarthy, 1997; Read, 2000; Schmitt, 2000; Nation, 2001; Schmitt, 2004; Bogaards and Laufer, 2004; Daller et al., 2007; Fitzpatrick and Barfield, 2009), they have rarely included dedicated studies of L2 collocation knowledge and development. In fact, in the last decade, only five book-length publications in English stand out for the more specific focus that they take on L2 collocation knowledge and use (Cowie, 1998c; Lewis, 2000; Nesselhauf, 2005; Schmitt, 2004; Meunier and Granger, 2008). The first situates collocation within the broader field of phraseology and provides a far-ranging exposition of corpus-based studies, some of which are collocation-focused. *Teaching Collocation*, edited by Lewis, is also multi-authored and is directed towards the pedagogic treatment of collocations in the classroom. Nesselhauf's solo-authored volume provides an in-depth analysis of the Verb + Noun collocations in a corpus of essays written by advanced German L1 learners of English. The two other edited volumes (Schmitt, 2004) and (Meunier and Granger, 2008) go beyond collocation itself by taking a generally wider view of the formulaic and phraseological patterning of language. Until now, then,

2 Introduction

there has been no single volume of work focused solely on researching L2 collocation knowledge and development within different local contexts.

To address this gap, *Researching Collocations in Another Language* is an international collection of L2 collocation studies that, for the first time, brings together dedicated research from Asia, Europe, and North America in the following four areas:

- using learner corpora to identify patterns of L2 collocation use (Part I, Chapters 2–5)
- developing appropriate L2 collocation lexicographic and classroom materials (Part II, Chapters 6–9)
- investigating how learners' L2 collocation knowledge can be assessed (Part III, Chapters 10–13)
- exploring the processes and practices by which learners develop their L2 collocation knowledge and use (Part IV, Chapters 14–17).

Each part includes three research chapters and a critical commentary. Written by experts in the respective field (Part I: Sylviane Granger; Part II: Hilary Nesi; Part III: John Shillaw; Part IV: Birgit Henriksen and Lars Stenius Stæhr), the commentary chapters identify and take up issues of interest across each set of research studies and constructively re-frame them within a broader critical view. While Alison Wray looks back, in Chapter 18, the closing chapter, at the whole collection to draw out further connections and potential contradictions in researching collocations in another language, it is our wish, in this opening chapter, to lay out the general contours for the work that follows in the rest of this book. We will first consider differing interpretations of the concept of collocation and how these lead into varying research priorities. We will then highlight some of the major issues that previous research has addressed in the four distinct areas of focus of this book, and outline the particular research studies and commentaries in each of the four parts of *Researching Collocations in Another Language*.

Two major conceptual underpinnings of L2 collocation research

Research on collocation has commonly been carried out within two different but sometimes somewhat overlapping traditions, which we can refer to as the *frequency-based* and the *phraseological* traditions. In the former, frequency and statistics are intrinsic ingredients in the

analysis of textual instantiations of collocation. In the phraseological tradition, work on collocation is guided by syntactic and semantic analyses, largely inspired by Russian and continental European work on phraseology.

The frequency-based view of collocation

In the frequency-based tradition, collocations are, in general, seen as units consisting of co-occurring words within a certain distance of each other, and a distinction is often made between frequently and infrequently co-occurring words. Pioneering work within this tradition was carried out by Firth (1952/3, 1956, 1957a, 1957b),¹ Halliday (1961, 1966) and Sinclair (1966, 1987a, 1987b, 1991; Sinclair et al., 1970). Firth was concerned with theorizing how meaning was produced at 'mutually congruent series of levels' (Firth, 1957a: 176) within language (context of situation, collocation, syntax, phonology, and phonetics). Although each level of the system is interdependent with the others, Firth was careful to distinguish 'colligation' (Firth, 1956: 113; 1957a: 181–3) within the syntactic level from collocation. Arguing that one of the meanings of 'night' is established through its collocability with the word 'dark', Firth suggested that part of the meaning of a word could be established by collocation. He summed this up in his famous exclamation 'You shall know a word by the company it keeps' (1957a: 179), and was dismissive of an essentialist semantic view where individual words have intrinsic core meanings. Rather, collocation was for Firth a central dimension in understanding how meaning and functional value are created through use: 'The distribution of the collocations in larger texts will probably provide a basis for functional values or meanings for words of all types' (Firth, 1952/3: 23). Firth saw collocations as sequences of co-occurring words, where the length of sequences varied greatly from two words up to 15. He envisioned different types of collocations such as 'habitual', 'more restricted technical', 'unique', and 'a-normal' (Firth, 1957b), but did not specifically define them or distinguish them clearly from one another.

Building on the sometimes rather vague writing of Firth, Halliday took a slightly different conceptual view of collocation as a syntagmatic association of lexical items that could be quantified textually in terms of their probability of occurrence at a certain distance from one another (1961: 276). Halliday posited, alongside the paradigmatic category of 'set', the syntagmatic category of 'collocation' for understanding lexis in language. The interaction of these two axes allows analysis of 'a very simple set of relations into which enter a large number of items'

(Halliday, 1966: 153). According to Halliday, collocation restricts the co-occurrence of particular lexical items and may allow for prediction of items that co-occur 'with a probability greater than chance' (Halliday, 1966: 156). He used 'lexical item' to mean a lexeme in all its derivative forms. Halliday also introduced the terms 'node', 'collocate', and 'span' to refer to the item under study (*node*), the co-occurring item (*collocate*), and the specified environment in which the node and the collocate may co-occur (*span*), respectively. These terms have proven fundamental to the operationalization of collocation and have served as indispensable tools for subsequent research.

Sinclair's innovative and far-reaching contributions to the work on collocation originate from his attempts to solve some of the practical problems concomitant with a Firthian view of collocation. He applied in practice Firth's original ideas to the Office of Scientific and Technical Information (OSTI) project (Krishnamurthy, 2004), and later also to one of the largest (for its time) and most ambitious research projects in computational lexicography ever carried out: the COBUILD project (Carter, 1998: 167). On the one hand, Sinclair expanded on Halliday's notion of probability of co-occurrence within a certain distance by calculating that a span of ± 4 , that is, four locations (number of orthographic words) to the left and to the right, respectively, of the node, constitutes the optimal environment within which 95 per cent of that node's collocational influence occurs (Jones and Sinclair, 1974: 21). On the other, the COBUILD project revealed that the most frequent of words of English tend to be collocated in delexical senses rather than in a full lexical sense so that they 'function as elements of structure' (Renouf, 1987: 177). Collocation itself was now becoming more clearly understood as a level of language use or 'lexical realisation of the situational context' (Moon, 1987: 92) – as Firth had originally claimed. For example, the differing textual collocates of 'skate' – 'ice', 'roller,' and 'winter' for sporting activity, and 'fish', 'ray', 'shark', and 'water' for fish (Moon, 1987: 91–2) – uncovered the distinct contextually bound meanings of the item. Another major insight from the COBUILD project was that the different senses of lexical items had such constrained typical phrasal patternings that few frequent words could be thought 'to have a residue of patterning that can be used independently' (Sinclair, 1987b: 158).

Two principles of interpretation were proposed by Sinclair (1987a: 318–19; 1991: 109–21) for how meaning is produced in text: *the open-choice principle* and *the idiom principle*. The former envisages language text as the result of a very large number of complex choices to do with individual lexical items (the 'slot-and-filler' model). Texts are then

seen as a number of slots that are filled item by item from a lexicon, if various local constraints are satisfied. The latter principle – the idiom principle – is an important complement to the open-choice principle. One of its central claims holds that ‘a language user has available to him or her a large number of semi-preconstructed phrases that constitute single choices, even though they might appear to be analysable into segments’ (Sinclair, 1991: 110). Sinclair’s view is that the idiom principle takes precedence over the open-choice principle and that text production is constantly constrained by the collocational restrictions that words (and multiword phrases) carry in relation to other words (and multiword phrases).

A fairly recent extension of the pioneering frequency-based work outlined above can be seen in ‘lexical bundle analysis’. Lexical bundles are loosely defined as ‘recurrent expressions, regardless of their idiomaticity, and regardless of their structural status’ (Biber et al., 1999a: 990). Used predominantly in corpus-based analyses of recurrent sequences of words widely distributed across texts in specific registers, the lexical bundle approach allows researchers to search for identical instantiations of *n*-word bundles (e.g. 2-word, 3-word, 4-word, 5-word bundles, and so on). Lexical bundles are seen by some as having a pre-fabricated or formulaic status (see, for example, Biber and Barbieri, 2007: 265). They are also considered to display non-idiomatic or transparent meanings (Biber and Barbieri, 2007: 269) and to occur more often in spoken than written discourse.

The phraseological view of collocation

In contrast to the frequency-based tradition, the common ground for those working within the phraseological tradition lies in the treatment of collocation as a word combination, displaying various degrees of fixedness, and in the preoccupation with collocation typology, that is, the decontextualized classification of collocations. The approach is heavily influenced by work carried out primarily in Russia in the 1940s (Cowie, 1998a, 1998b), but a close connection between a phraseological view of collocation and lexicography and pedagogy dates back further at least to the work of Bally² in Geneva and Palmer³ in Tokyo in the early twentieth century. Bally worked on developing various phrase-based (Hausmann, 1979: 189; Siepmann, 2006: 9) French learning materials for foreign students at Geneva University in 1909, whereas Palmer collaborated with A. S. Hornby (Smith, 1998; Cowie, 1999; Smith, 1999) to categorize and meticulously codify the lists of collocations they had drawn up for the development of appropriate English language

materials in the pre-war period in Japan. Their innovative treatment of multiword units saw the light of day in general-purpose English dictionaries for learners (Hornby et al., 1942, 1948) and is detailed by Cowie (1998b: 210–13; 1999: 52–81).

Other more theoretical views of the variability of multiword units emerged from the work of Vinogradov and Amosova in the Soviet Union (Cowie, 1998b: 213–18) and were taken up in the West from the 1970s onwards, in particular with regard to ‘collocation restriction’ (Aisenstadt, 1979; 1981). Seminal work within the phraseological tradition can be found in Aisenstadt (1979, 1981), Cowie (1981, 1988, 1991, 1994, 1998c), Benson (1989), Howarth (1996), Benson et al. (1986a, 1986b, 1997), Cowie and Howarth (1996), Howarth (1996, 1998a, 1998b), Mel’čuk (1998) and Nesselhauf (2005). The phraseological school has been rather less concerned with frequencies and statistical significance per se, and rather more interested in word combinations, their degree of opacity, and the commutability (also called substitutability) of the word elements in these combinations. Frequency-based approaches alone do not suffice: ‘... phraseological significance means something more than what any computer algorithm can reveal’ (Howarth, 1998a: 27).

Cowie (1981, 1988, 1991, 1994, 1998a) has argued that collocations are associations of two or more lexemes (or roots) occurring in a specific range of grammatical constructions. Restricted collocations have been further defined by Cowie as ‘word-combinations in which one element (usually the verb) [has] a technical sense, or a long-established figurative sense which [has] lost most of its analogical force’ (Cowie, 1991: 102). On the whole, Cowie (1981, 1998a), followed by Howarth (1996, 1998a, 1998b) and Nesselhauf (2005), presents the case for a scalar analysis of word combination categories, ranging in the form of a continuum from transparent, freely recombinable collocations at one end to unmotivated and formally invariable idioms at the other. A straightforward illustration of this can be seen in Howarth’s (1998b: 164) examples of ‘free combination’ (*blow a trumpet*), ‘restricted collocation’ (*blow a fuse*), ‘figurative idioms’ (*blow your own trumpet*), and ‘pure idioms’ (*blow the gaff*). An interesting feature of the scalar view is that, while collocations are in most cases lexically variable, they are also characterized by arbitrary limitations of choice at one or more points. Cowie exemplifies this with combinations like *cut one’s throat* and *slash one’s wrist*, which are appropriate in English, but where substitution of the verb creates infelicitous variants such as **slash one’s throat* and *?cut one’s wrist* (see also Howarth, 1996; 1998a; 1998b).

A somewhat different typology of collocation within the phraseological tradition has been developed by Benson et al. (1986b, 1997) in

distinguishing between grammatical and lexical collocations (cf. Firth's distinction between collocation and colligation). In a grammatical collocation, a dominant word (noun, adjective or verb) is combined with a preposition or grammatical structure. Eight major types of grammatical collocations (G1–G8) are identified, such as Noun + Preposition, Noun + *to* + Infinitive, Adjective + Preposition, with each consisting of a varying number of subtypes. Type G8, for example, comprises no fewer than 19 different English verb patterns. The seven main categories of lexical collocation represent word combinations consisting of nouns, adjectives, verbs, and adverbs only, but without any function words (Verb + Noun, Adj + Noun, Noun + Verb, Adj + Adv and Verb + Adv).

In stark contrast to the Benson, Benson and Ilson dictionary, radically different forms of phraseological dictionary are currently being developed, where collocation entries are first organized by key themes or concepts (see, for example, Siepmann, 2005, 2006, 2008; Pecman, 2008) and the user is then guided from conceptual frames to phrase-based collocational encodings. The first generation of such electronic 'onomasiological' collocation dictionaries, once completed, will probably enable interactivity within and between entries through bilingual displays, hyperlinked access points, and semantic query functions (Pecman, 2008) – and may well help to reconcile some of the differences between the two traditions presented here.

Summing up

In presenting the evolution of two major conceptual views of collocation, it has been our intention to summarize some of the major issues that have concerned researchers working within these two traditions. The picture we have painted is by necessity somewhat partial and incomplete,⁴ but it serves, we hope, to illustrate how different conceptualizations can lead into quite distinct research agendas. Against this notably complicated set of background issues, we will now introduce the L2 collocation research in this volume.

L2 collocation research in this volume

Each of the sections that follow in this introduction is separated into two parts: *Setting the scene*, and *Research and commentary*. In *Setting the scene*, we give a general overview of previous research in each area and draw out various questions that L2 collocation research has tried to address. In *Research and commentary*, we introduce the three research

studies and the respective commentary chapter in each of the four parts of this book.

Part I: L2 collocation learner corpus research

Setting the scene

Since the creation of the first computer-readable, written English corpus – the Brown corpus (Kucera and Francis, 1967) – over 40 years ago, we have witnessed, over the last two decades, the development of a number of different ‘learner corpora’. These corpora, either ready-made for generic use or especially compiled for a specific study (aka do-it-yourself corpora), are collections of cross-sectional or longitudinal data, in the form of L2 users’ written or spoken language productions. One of the more well-known computerized learner corpora is the International Corpus of Learner English (ICLE) (see Granger, 1998a, 2003), which now enables researchers to investigate the language use of advanced students of English coming from as many as 16 L1 backgrounds (Granger et al., 2009). Despite technological advances in corpus analysis, the number of corpus-based studies devoted to L2 collocation research has been fairly modest until now. In particular, research investigating spoken L2 collocation use (including the phonological realization of collocation) is conspicuously absent (see, though, De Cock, 2000, 2004), but a handful of L2 collocation corpus studies have investigated written L2 collocation production (Dechert and Lennon, 1989; Zhang, 1993; Chi, Wong and Wong, 1994; Howarth, 1996, 1998a, 1998b; Granger, 1998b; Gitsaki, 1999; Nesselhauf, 2005).

A common methodological denominator of these studies is the adoption of a typological and more phraseologically oriented approach to collocation. Zhang (1993) analysed English essays written by 30 non-native speaker and 30 native speaker first-year university students in terms of their use of no fewer than 66 types of lexical and grammatical collocations. Similarly, Gitsaki (1999), who looked at English essays written by Greek L1 secondary school students, classified her subjects’ collocations into 37 types. On the whole, the close attention to a precise lexicographic description of phraseological units and their varying degree of fixedness and idiomaticity has largely been extended into L2 collocation research with learner corpora. Yet, given the dominance of frequency-based studies in L1 corpus studies, it is perhaps surprising that this kind of phraseologically oriented approach has been so dominant in L2 collocation corpus studies to date.

A second point of interest in previous studies is the object of analysis. Some studies (Zhang, 1993; Gitsaki, 1999) have cast a very wide net indeed by using a highly comprehensive classification system for collocations. This may be problematic, especially in comparisons of relatively small data sets of native speaker and learner data, because some types of collocations may occur only once or twice in a data set, if at all. Other studies (Howarth, 1996, 1998a, 1998b; Granger, 1998b; Nesselhauf, 2005) have concentrated their analysis on one or, at most, two types of collocation. Granger (1998b) analysed French L1 learners' use of English Adv + Adj combinations, whereas both Howarth (1996) and Nesselhauf (2005) aimed at mapping out L2 learners' use of restricted Verb + Noun collocations. Howarth examined essays written by postgraduates coming from a large number of L1 backgrounds, and Nesselhauf analysed essays written by advanced undergraduate German L1 learners of English. The assumption behind the focus on restricted collocations is that it is precisely these word combinations that form the very large repository of phrases between free combinations and idioms, and which pose the greatest collocation challenge to learners. The advantage of limiting the focus to fewer types of collocation is that it is much more likely that a data set taken from a corpus – even a smaller one – will contain sufficient tokens for a clear comparative analysis.

Research and commentary

The three studies in this section make use of methods and techniques not commonly seen in mainstream L2 collocation research. In Chapter 2 Groom uses a lexical bundle analysis, and a more traditional node and collocate analysis, on corpus data from Swedish undergraduate learners of English to examine whether increased, immersion-based L2 exposure leads to significant improvements in terms of the number of correctly produced collocations. In Chapter 3, Lin and Adolphs explore uncharted waters by investigating whether holistic storage of collocations can be phonologically analysed. This chapter provides a rare example of L2 collocation research based on a corpus of spoken language. In Chapter 4, Reppen applies a lexical bundle analysis to the writing of primary school children across a whole school year. By using a corpus of written English essays from young L1 English and L1 Navajo users, Reppen explores variation by age and language group in the production of 3-word bundles. The results point to a number of interesting language and sociocultural questions worthy of further investigation. All three studies break with the phraseologically oriented approach to L2 collocation prevalent in

previous studies. Granger's commentary in Chapter 5 highlights how conflicting results from phraseological studies may emerge as an effect of different operationalizations of collocation, and how previously perceived weaknesses in learner output can be reinterpreted in a different light. Granger particularly welcomes research involving spoken corpora, and she also emphasizes the need for small-scale exploratory research to try out and evaluate new research methodologies.

Part II: L2 collocation lexicographic and classroom materials research

Setting the scene

In addition to the *The BBI Combinatory Dictionary* (Benson, Benson and Ilson, 1986; 1997) mentioned earlier, there are two other English print⁵ monolingual collocation dictionaries commonly available to researchers, teachers, and learners: Kozłowska and Dzierżanowska's (1982) *Selected English Collocations* (SEC), later republished as *The Dictionary of Selected Collocations* (DOSC) (Hill and Lewis, 1997), and *The Oxford Collocations Dictionary for Students of English* (OCDSE) (2002). The microstructure of entries varies in these dictionaries as regards the display of possible collocate ranges for headwords and the degree of syntactic and semantic information. This variation bears direct relationship to the conceptual view of collocation taken. The BBI comes from within the phraseological/word combination tradition, while the SEC/DOSC and the OCDSE adhere to the lexical frequency/co-occurrence view of collocation. '(C)omplied without any sort of computer assistance whatsoever' (Knowles, 1993: 300), SEC/DOSC organizes entries for adverbs and nouns by listing possible lexical collocates in separate subentries (Cowie, 1998b: 222–4) for each of its 3200 headwords. Unlike the SEC/DOSC (and the BBI), the OCDSE draws on corpus data – the British National Corpus (BNC; Oxford University, 2005) – for the selection of collocates under its 9000 or so headwords. It targets collocations that should be useful for upper-intermediate+ learners dealing with a moderately formal register. The OCDSE divides polysemous words into different senses. It also groups collocates for each sense into distinct word classes and typographically separates sets of similar collocates within each word class. Many of the collocate groupings also have an example sentence to illustrate the contextualized use of a specific collocation.

The three print dictionaries mentioned above have been positively received (Gold, 1988; Herbst, 1988; Cueto, 1998; Appleby, 2000; Howarth, 2000; Klotz, 2003), but, in contrast to the wealth of

user-related investigations for general-purpose dictionaries (for comprehensive overviews, see Béjoint, 2000; Cowie, 1999; Hartmann, 2001), there have been few dedicated studies of monolingual collocation dictionary L2 users. Benson (1989) reported how, under controlled conditions, advanced learners (Russian teachers of English) can improve their use of collocation with the aid of a collocation dictionary (the BBI). Studies by Béjoint (1981) and Bogaards (1990) pointed to variation in users' look-up strategies for multiword combinations in general-purpose dictionaries, whereas Atkins and Varantola (1997) discovered that non-native speakers often looked for collocation information in dictionaries to be reassured about their L2 collocation knowledge. An investigation by Frankenberg-Garcia (2005) required her fourth-year translation majors to translate a newspaper text. Although only 16 per cent of their 146 look-ups were concerned with finding a suitable collocate, students rated their collocation searches as highly successful and helpful.

As Rundell (1999) observes, identifying suitable collocations and understanding collocation restrictions is one of the most important productive needs of learners. Yet, the overall picture we have of what learners do, and how they do it and why, when they use lexicographic collocation resources for their production, can be described as extremely sketchy at best. The same may be said for our understanding of how learners use pedagogic materials specifically designed for developing their L2 collocation ability. Following Lewis's (Lewis, 1993, 1997) groundbreaking 'populist' (Thornbury, 1998: 13) works on a lexical approach to language learning and teaching collocation (Lewis, 2000), learner resources aimed at the global ELT market have been produced (for example, McCarthy and O'Dell, 2005; for a review, see Pulverness, 2007). However, research has yet to explore how learners use and evaluate these kinds of materials in local contexts.

Research and commentary

The three studies in Part II directly address the telling gaps just noted. In Chapter 6, Handl explores a new multi-dimensional classification for elaborating the criteria by which collocations can be selected for advanced learners' dictionaries. This leads to an alternative micro-structure for print and electronic collocation dictionaries that helps learners identify more easily the collocations they need. Komuro examines in Chapter 7 what problems learners encounter in using specific OCDSE entries in order to identify improvements for the design of more user-friendly collocation entry structures in the future. In the

final study of this section, Jiang (Chapter 8) takes a critical look at how English textbooks in China present vocabulary to learners. Noting an almost complete lack of focus on collocation, Jiang reports on a series of quantitative and qualitative interventions informing the design, use, and evaluation of pilot collocation classroom materials with Chinese university students. In her commentary on these three studies (Chapter 9), Nesi considers the wider implications of the detailed description of collocational relations that these chapters present. Nesi also discusses the methodological implications of these studies for future research, as well as the insightful critiques of existing practices that they offer for further improving lexicographic and pedagogic materials.

Part III: L2 collocation knowledge assessment research

Setting the scene

Assessment of learners' acquisition, knowledge and use of L2 collocations forms an essential part of furthering our understanding of how learners cope with these challenges. Broadly speaking, previous studies fall into those that focus on the development process of an assessment instrument itself (Schmitt, 1998, 1999; Bonk, 2000; Gyllstad, 2007), and those that develop assessment instruments predominantly for the purpose of a particular research project (Biskup, 1992; Bahns and Eldaw, 1993; Farghal and Obiedat, 1995; Granger, 1998b; Gitsaki, 1999; Mochizuki, 2002; Barfield, 2006; Keshavarz and Salimi, 2007; Laufer and Girsai, 2008).

Earlier studies tended to use a relatively small number of items in the assessment of L2 collocation knowledge. Farghal and Obiedat (1995) tested Arabic L1 learners' knowledge of 11 Noun + Noun collocation-pairs and 10 Adjectival + Noun collocation-pairs, and Bahns and Eldaw (1993) examined German L1 learners' knowledge of 15 English lexical Verb + Noun collocations, whereas Gitsaki (1999), in one of her questionnaires, investigated Greek L1 learners' knowledge of just ten English collocations. With so few items, drawing well-founded and generalizable conclusions is difficult, and the content validity as well as the reliability may be compromised. In recognition of these threats, more recent studies have used a larger number of items: Bonk (2000): 50; Mochizuki (2002): 70; Barfield (2006): 120; Gyllstad (2007): 50 and 100; Keshavarz and Salimi (2007): 50.

There has also been great variation in item selection for different assessment measures. Gitsaki (1999) drew on English textbooks used in Greek

schools for her test items, and Bahns and Eldaw (1993) took target collocations from textbooks and dictionaries. Farghal and Obiedat (1995), on the other hand, used collocations from the domains of food, clothes, colours, and weather. In some cases, researchers have not reported how their items were chosen (Biskup, 1992; Keshavarz and Salimi, 2007), which hinders proper evaluation of their findings. Several studies have used either corpora or corpus-based frequency lists to sample items for assessment (for example, Mochizuki, 2002; Gyllstad, 2007). A growing trend in current research is to cross-check selected items against information available in mega-corpora like the BNC (Oxford University, 2005) or the Bank of English (HarperCollins, 2007).

With the exception of Gitsaki (1999), Bonk (2000), and Keshavarz and Salimi (2007), most previous studies have involved assessment of lexical collocations. One of the more popular ways of assessing knowledge of lexical collocations has been L1–L2 translation (Biskup, 1992; Bahns and Eldaw, 1993; Farghal and Obiedat, 1995; Laufer and Girsai,⁶ 2008), either through the translation of sentences, or isolated items. Some researchers have used short, decontextualized prompts in a ‘stimulus-response’ manner (Schmitt, 1998, 1999; Barfield, 2009). Other measures, including assessment of either grammatical or lexical collocations, have involved L2 sentence cloze items (Bahns and Eldaw, 1993; Farghal and Obiedat, 1995; Gitsaki, 1999; Bonk, 2000), and discrete receptive tasks of different kinds (Granger, 1998b; Bonk, 2000; Mochizuki, 2002; Gyllstad, 2007).

A common question in previous research has been whether L2 collocation knowledge develops alongside general L2 proficiency. Gitsaki (1999), Bonk (2000), and Gyllstad (2007) all claim a positive correlation, whereas Howarth (1996, 1998a, 1998b) and Barfield (2006) did not find support for such a position. Other more disparate findings include Farghal and Obiedat’s (1995) assertion that paraphrasing as a type of lexical simplification strategy is used extensively among learners to compensate for a lack of well-developed collocation knowledge, Howarth’s (1996, 1998a, 1998b) finding that learners’ use of infelicitous collocations often involves blends of two or more acceptable natively-like collocations from overlapping collocation clusters, and Granger’s (1998b) claim that learners have a weak sense of salience and that they overuse certain word elements in combinations as ‘safe bets’.

Research and commentary

The three studies in Part III are all concerned with defining and operationalizing the construct they are assessing, and creating inherently

reliable and valid measures. In Chapter 10, Revier reports on the initial development process of a discrete test of productive collocation knowledge called CONTRIX. This is designed to assess learners' knowledge of *whole* collocations, where potentially required function words such as determiners and prepositions are considered as important as the lexical elements. Revier also takes great pains to separate the degree of transparency of a collocation as a factor in the assessment. In Chapter 11, Eyckmans traces the development of a reliable, corpus-based test of receptive phrasal knowledge called DISCO. Combining techniques both from the frequency-based and the phraseological traditions, Eyckmans sets up a longitudinal, experimental design where receptive L2 collocation knowledge is mapped in relation to spoken L2 proficiency. Gyllstad reports, in Chapter 12, on the development of two tests of receptive collocation knowledge, called COLLEX and COLLMATCH. The research is focused on arriving at reliable and valid measures of receptive collocation knowledge as a construct, and investigating the relationship between this construct and vocabulary size. In his commentary in Chapter 13, Shillaw draws out a number of common themes in the three studies, including how to define collocation knowledge as a construct and operationalize such a construct in a test. He also looks at the underlying assumptions of the tests featured and addresses the different validation techniques used.

Part IV: L2 collocation learner process and practice research

Setting the scene

As Willis notes (2003: 219), 'Many learners ... are not consciously aware of collocation, or of the importance of fixed phrases.' For this reason, advocates of specific approaches to collocation learning (for example, Sinclair and Renouf, 1988; Baigent, 1999; Lewis, 2000) emphasize noticing language chunks as a crucial initial process in dealing with comprehended input, but also point out that noticing alone is not sufficient. Learners need to be guided to notice 'similarities, differences, restrictions and examples arbitrarily blocked by usage' (Michael Lewis, 2000: 184) if they are to transform input into uptake; such hypothesizing should be constantly encouraged as learners record and experiment with collocation examples and continually readjust their changing awareness and use of collocation. This, the argument goes, will help them build systematic L2 collocation awareness and knowledge.

System building in L2 vocabulary acquisition is often seen in terms of lexical network building (Meara, 1990, 1992, 1996; Vermeer, 2001; Read, 2004; Meara and Wolter, 2004; Wilks and Meara, 2007). How such network building *may* apply to L2 collocation development is illustrated in an experimental study by Haastrup and Henriksen (2000). They identify the three major phases of noticing, analysing, and integrating whereby learners incorporate new words into their L2 lexical networks. Haastrup and Henriksen's position is that the analysing phase is where learners can create syntagmatic links between L2 words, which arguably suggests that learners tend to analyse and segment, before resynthesizing collocations as they become part of their developing L2 lexical networks. Haastrup and Henriksen further characterize such network building as 'a very slow process' (2000: 235). As to potential theoretical reasons for this slowness, Wolter (2006) suggests, particularly with regard to non-equivalent L2-L1 collocations, that successful acquisition of such multiword units first requires changes in learners' conceptual worlds, which makes learners less adept at mastering non-equivalent L2-L1 collocations. These different claims about collocations raise fundamental questions about whether collocations are stored holistically or not, as has been argued for formulaic sequences (see, for example, Schmitt and Carter, 2004: 4–6). Although the jury is still out on resolving such crucial issues, there is now some specific empirical evidence that collocations enable greater speed of processing (Ellis, 2008: 6) and that non-native speakers process collocations less quickly than native speakers do (Sivanova and Schmitt, 2008).

Although such theoretical positions have been advanced, very little research has been completed into learners' actual practices of L2 collocation development (Coxhead, 2008: 159). At a general level, Hunt and Beglar (2005: 35) observe: 'Although little is known about the acquisition of L2 collocational knowledge, ultimately most EFL learners probably develop it through processing large quantities of written input in which most of the vocabulary is known.' Hunt and Beglar further suggest that such rich exposure can be enhanced through conscious attention to, and recycling of, frequent collocations and the collocates of words that learners already know well. A few studies shed more specific light on learner practices and processes of L2 collocation development. Yang and Hendricks (2004) explored how collocation-awareness-raising (CAR) tasks helped postgraduate students improve their collocation use in drafting and redrafting essays. Their students also reported that the CAR approach made them more aware of collocations in their out-of-class reading and that they were keen to continue

developing their use of collocation in other language tasks. In a smaller scale pilot study, Coxhead (2008) found that students could become sensitized to noticing and retrieving multiword combinations from reading texts, but their motivation to reproduce newly retrieved phrases was hindered by a sense of risk and fear of negative evaluation within the particular institutional setting. Coxhead's investigation reminds us that, in addition to metacognitive awareness and active involvement, other influential factors clustering around learners' changing L2 collocation processes and practices need to be considered too, such as previous learning experiences, awareness of collocation, institutional and sociocultural contexts of exposure and use, not to mention learners' short-term and longer-term goals in learning English.

Research and commentary

In Chapter 14, Yang and O'Neill investigate how a group of adult EFL learners develop their collocation learning over a five-month period on an intensive English course. The authors interviewed the students at the beginning and end of the language programme, and also analysed the students' reflective learning journals. This qualitative longitudinal investigation illuminates different problems that learners face, the great variety of strategies they use, and important adjustments they may make, in becoming more collocationally aware and proficient. The research by Peters in Chapter 15 explores the effects of an attention-drawing technique on the recall of individual words and collocations with two groups of advanced EFL learners. Finding no significant quantitative difference between groups in their recall of individual vocabulary items and collocations, Peters takes a detailed qualitative look at the students' strategy use and their perceptions of the task and test in order to probe further learners' positions and decision making about collocation learning. In his narrative reconstruction of key processes in learners' emerging L2 collocation awareness and development, Barfield (Chapter 16) follows learners' decision making in great depth over a nine-month period. Examining how learners interpret their collocation practices in relation to their changing sense of who they are in terms of their own understandings of the world, this study identifies both lexical and sociocultural reorganization as overarching characteristics of learners' L2 collocation development. In their commentary on these three studies, Henriksen and Stenius Stæhr address central processes in the development of L2 collocational knowledge. They also discuss the challenges that learners experience in acquiring such knowledge, as well as the challenges that researchers

face in identifying and interpreting the processes at play. Henriksen and Stenius Stæhr conclude by briefly considering different pedagogic implications for teachers wishing to help their learners develop L2 collocational knowledge.

Closing remarks

We began this chapter by noting the collocation gap in L2 acquisition research, and we then presented two major traditions in collocation research, the frequency-based tradition and the phraseological-typological school. In the second half of the chapter, we set the scene for each of the four areas that this volume covers. Here, we briefly reviewed previous L2 collocation research in each area, before we highlighted the particular focus of the research studies and respective commentary chapter in each part of this volume. This was all done by way of giving a general introduction and re-interpretation of the field, where we could signpost important background issues and highlight some of the major conceptual positions, practices and priorities that run through the L2 collocation research that follows. How these become further developed and re-constructed is at the heart of the research chapters and commentaries, as well as the final conclusion chapter by Alison Wray, in *Researching Collocations in Another Language: Multiple Interpretations*. We hope that many different interesting spaces for further developments in practice and theory will continue to be produced in the interplay between the work that researchers do and the multiple interpretations that others make of the research done.

Notes

1. We have dated Firth's publications here according to the indication given by F. R. Palmer (Palmer, 1957) and Firth (1957c).
2. Bally differentiated between 'les associations libres et occasionnelles, les séries phraséologiques ou groupements usuels et les unités indissolubles' in developing these materials (Hausmann (1979: 189). He did not however use the term 'collocation'.
3. Harold Palmer taught at University College London from 1915 to 1922 before taking a specially created position as linguistic adviser to the Japanese Department (= Ministry) of Education in 1922 (Cowie, 1999: 4–5; Smith, 1999: 57–67), where the Institute for Research in English Teaching (IRET) was established in 1923. In 1933, Palmer published the *Second Interim Report on English Collocations* (Palmer, 1933b); this was also the year that he and A. S. Hornby started working together on collocations (Smith, 1999: 131). Palmer (1931, 1933a, 1933b, 1934) coined the term collocation, noting, for example,

in a 1934 report for the IRET (Palmer, 1934: 20): 'In 1930 ... we presented in mimeographed form a rough draft of a collection of collocations (culled for the most part from Saito's *Idiomological Dictionary*).'

4. For a more comprehensive review of how the phenomenon of collocation has been treated in the literature, see Nesselhauf (2004, 2005).
5. We have not included the *COBUILD English Collocations on CD-ROM* (1995) as it is no longer being produced and so will probably be unavailable for future students, teachers, and researchers.
6. This study also includes L2–L1 translation.

Acknowledgement

We would like to thank Elke Peters, Jingyi Jiang, Marnie O'Neill, Randi Reppen, Robert Revier, Sylviane Granger, and Ying Yang for their helpful feedback on earlier drafts of this chapter.

Index

- Adjective + Noun collocations
see collocation types
- Adolphs, S. 9, 34–48, 61, 62, 209,
 238, 239
- assessment
see L2 collocation knowledge
 assessment research
- AWARE approach 183–92, 226, 231
see also collocation awareness
 raising (CAR) approach
- Bachman, L. 132, 135, 154,
 161, 169
- Bahns, J. 12, 13, 125, 128, 153, 157,
 182, 209, 217
- Bally, C. 5, 17
- Bank of English 13, 26, 31, 115, 143
- Barfield, A. 1, 12, 13, 16, 60, 141,
 142, 153, 169, 182, 208–23, 225,
 226, 228, 230, 231, 236, 239
- BBI Combinatory Dictionary of
 English 6–7, 10, 11, 69–70, 87
- BBI Dictionary of English Word
 Combinations 6, 7, 10, 87, 196
- Benson, M. 6, 10, 11, 196
- Biber, D. 5, 22, 23, 34, 49, 50, 53,
 57, 64, 241
- Bonk, W. 12, 13, 125, 128,
 153, 209
- British National Corpus (BNC) 10,
 13, 80, 85, 86, 115
 test item selection 129–30, 143–4,
 159–60, 169, 172–3
see also corpus analysis; learner
 corpora
- chunks/chunking 14, 37–8, 49,
 61–2, 143, 225, 238–9
see also collocations; word
 combinations
- cloze tests 13, 80, 82–3, 84, 129,
 140, 141, 144, 147, 149,
 150, 172
- clusters 13, 22, 49, 105, 109–10,
 111, 120, 222
see also word clusters
- COBUILD 4, 18, 101, 143,
 173, 196
- Collocate Matching Test
 (COLLMATCH) 14, 153–70,
 173–5, 177
- Collocating Lexis Test (COLLEX) 14,
 153–70, 171, 174–5, 177
- colligation 3, 7
- collocates 4, 9, 10, 11, 15, 25, 125
 grouping in learner
 dictionaries 10, 11, 71–3,
 75–8, 80, 87–98, 115, 116,
 119, 121
 selection for test items 169–70,
 181, 193–4, 207: *see also*
 collocation range; collocation
 span; node; node and collocate
 analysis; set
- collocation
 and colligation 3, 7
 and mutual expectancy 70, 85
 as independent construct 127–9,
 171–2, 239–41
 as level of language 3, 4
 as word property 22, 24–6,
 125–6, 154, 159, 169,
 237–9
- conceptualizing ~ 2–7, 10, 17,
 22, 73–5, 127–8, 141–4, 154–5,
 237–41
- defining ~ 3, 6, 13–14, 70–1, 75,
 126–7, 141–2, 154–7, 238
- frequency-based view 2, 3–5, 6, 8,
 10, 14, 17, 34, 70, 241–2
- operationalizing ~ 151–3, 186,
 187–9: *see also* tests
- phraseological view 2, 3, 5–7, 8,
 9–10, 14, 17, 154
- psychological reality 62, 127, 134,
 135, 172, 173, 174–5

- terminology 3–7, 17–18, 22, 63–4, 154, 158, 242
- typological view of
 - collocation 8, 17, 70
- collocation awareness raising (CAR)
 - approach 15–16, 182, 206–7, 225–6, 228, 231
 - see also* AWARE approach
- collocation competence/phrasal
 - competence/syntagmatic competence 102, 141, 139–5, 172, 181–2, 184
 - see also* L2 collocation knowledge; phrasal knowledge
- collocation dictionaries 6, 7, 10–12, 111, 188
- existing displays/
 - microstructures 10–11, 71–3, 93, 94–5, 96–7, 98, 115
- alternative displays/
 - microstructures 11, 73, 76–8, 80, 82, 84, 116, 119
- user research 11–12, 69–85, 86–98, 114–21
- see also* electronic dictionaries; general purpose dictionaries; learners: look-up processes for collocations; Oxford Collocations Dictionary for Students of English; print dictionaries
- collocation direction 74–5, 76–8, 80–5, 116, 120
- see also* collocation links; collocation partners
- collocation links 15, 74–7, 80, 117, 208
- see also* collocation direction; collocation partners
- collocation partners 70, 74, 75, 76, 77, 78, 80, 82, 83, 84, 116
- see also* collocation direction; collocation links
- collocation patterns 4, 7, 39–41, 54–7, 78, 85, 90, 104, 224, 229, 231, 236, 238–9, 241, 243
- collocation processing
 - and compositionality 128, 155
 - and learner identity 208, 213, 214, 216, 217, 219–21, 224, 226, 228, 231
 - and memorization 187, 191, 202–3, 212, 214, 216, 217, 219
 - and past vocabulary
 - practices 14–16, 110–11, 185, 186, 189, 212–13, 221, 222
 - and quantity-quality shift in lexical knowledge 217–19, 227, 228:
 - see also* L2 lexical networks; mental lexicon; vocabulary depth
 - and time in an English-speaking environment 21, 22, 30–3:
 - see also* language exposure
- breaking collocations down 15, 128, 208, 225, 235, 240
- comprehending/recognizing
 - collocations 14, 96, 105, 141, 155–6, 181–2, 185–6, 194–5, 217, 218–9, 225, 227, 243–4
- noticing collocations 14, 15, 16, 105–7, 109, 112, 113, 117, 121, 183–4, 186, 194–5, 202, 228
- overusing/underusing
 - collocations 13, 21, 23–4, 28, 60, 61, 63, 102, 139, 181, 221
- producing collocations 11, 15, 107, 112, 117, 126, 128–9, 189–91, 194–5, 214–22, 228, 243–4
- recalling collocations 16, 156, 191, 200, 205, 226, 229
- retrieving collocations 16, 105, 110, 111, 112, 187–9, 202–5, 213–5, 228
- translating collocations 11, 13, 89–96, 188, 195, 197, 199–200, 205, 206: *see also* translation, tasks
- see also* learners; learning
 - collocations
- collocation range 10, 70, 74, 75, 76, 116
- see also* collocates; collocation span; node; node and collocate analysis; set
- collocation restrictions 3, 4, 5, 6, 9, 11, 14, 70, 139

- collocation restrictions (*Continued*)
see also collocations; figurative idioms/idiomatic collocations/idiomatic word combinations; free combinations; idioms; restricted collocations; selectional restrictions; word combination
- collocation span 4, 24–5, 26, 29, 142, 143
see also collocates; node; collocation range; set
- collocation types 3, 5, 6–7, 8, 9
 Adjective + Noun collocations 7, 12, 89, 93–5
 Preposition + Noun collocations 89, 95–6
 Verb + Noun collocations 1, 6, 7, 9, 12, 21, 89, 91–3, 99, 126–7, 141–4, 146, 150, 155, 157, 159, 169, 172–3, 176–7, 214, 215, 221
- collocation webs 26, 78–9, 84, 85, 120
- collocations
 as co-occurring sequences 3–4, 6, 22, 24–6, 50, 60, 63, 70, 86, 87, 141, 142, 143, 192, 227
 as holistic units 9, 15, 34, 35, 36, 37, 47, 128, 151, 239–41
 as lexical bundles 5, 9, 22–4, 27–8, 49–59, 60, 62–3, 64, 241
 as phraseological sequences/units 8, 29, 33, 34–48, 61, 63
 constituent words 70, 74, 78, 126–31, 136, 137, 143, 144, 225
 whole ~ 14, 125–37
see also chunks/chunking; free combinations; figurative idioms/idiomatic collocations/idiomatic word combinations; idioms; restricted collocations; word combinations
- concordances 31, 40, 41, 130–1, 143, 151, 188
- Constituent Matrix Test (CONTRIX) 14, 129–37, 173–4, 177
- corpus analysis 1, 4–5, 13, 235, 241–2
- learner ~ 8–10, 21–65, 99–102, 113, 235–6, 242
see also British National Corpus; learner corpora
- Cowie, A. P. 1, 5–6, 10, 11, 17, 21, 34, 63, 71
- depth of word knowledge
see vocabulary depth
- dictionaries
see collocation dictionaries; electronic dictionaries; general purpose dictionaries; learners, look-up processes for collocations; print dictionaries
- Discriminating Collocations Test (DISCO) 14, 140–2, 144, 148–51, 171, 173–7
- Durow, V. 209, 229
- electronic dictionaries 7, 11, 71, 78, 84, 120–1
see also collocation dictionaries; general purpose dictionaries; learners, look-up processes for collocations; print dictionaries
- Eldaw, M. 12, 13, 125, 128, 153, 157, 182, 209, 217
- Farghal, M. 12, 13, 125, 139, 153, 182, 209, 217
- figurative idioms/idiomatic collocations/idiomatic word combinations 6, 8, 9, 70, 72, 73, 74, 86, 87, 104, 116, 141, 143, 144, 146, 150, 173, 176
see also collocations; free combinations; idioms; restricted collocations; selectional restrictions; word combinations
- figured worlds 209, 210, 213, 217
- Firth, J. 3, 5, 7, 69, 70
- formulaic sequences 5, 15, 22, 28, 35–6, 39, 47, 48, 49, 209, 224–5, 229

- free combinations 6, 9, 70, 73, 146, 170, 176
see also collocations; figurative idioms/idiomatic collocations/idiomatic word combinations; idioms; restricted collocations; selectional restrictions; word combinations
- frequency-based view of
 collocation 2, 3–5, 6, 8, 10, 14, 17, 34, 70, 241–2
see also collocation
- general purpose dictionaries
 collocation displays 6, 71–3, 82, 115, 121
see also collocation dictionaries; electronic dictionaries; learnness: look-up processes for collocations; Oxford Collocations Dictionary for Students of English; print dictionaries
- Gitsaki, C. 8, 9, 12, 13, 125, 128, 153, 182, 208
- Granger, S. 1, 2, 8, 9, 10, 12, 13, 22, 60–5, 141, 208, 233, 241, 242
- Groom, N. 9, 21–33, 60, 61, 62, 63, 64, 238, 241
- Gyllstad, H. 12, 13, 14, 142, 153–70, 171–3, 175, 177, 236, 239
- Halliday, M. A. K. 3–4
- Handl, S. 11, 69–85, 114, 115, 116, 117, 118, 119, 120, 121, 239
- Henriksen, B. 2, 15, 16, 17, 127, 155, 158, 208, 224–31, 235, 242, 244
- holistic processing and storage 9, 15, 34, 35, 36, 37, 39, 47, 62, 128, 151, 229, 239–41
- Hornby, A. S. 5, 6, 17
- Howarth, P. 6, 8, 9, 10, 13, 21, 25, 26, 87, 127, 153, 155, 157, 182, 208
- idiom principle 4–5, 23, 70, 139
compare open choice principle
- idiomaticity 5, 8, 70, 73
compare selectional restrictions
- idioms 6, 9, 70, 73, 74, 170, 176
see also collocations; figurative idioms/idiomatic collocations/idiomatic word combinations; free combinations; restricted collocations
- Jiang, J. 12, 114, 115, 116, 117, 118, 119, 121, 184, 239
- Komuro, Y. 11, 86–98, 114, 115, 116, 117, 118, 119, 121, 239
- language exposure
 and L2 collocational development 9, 15, 16, 21, 22, 30–3, 69, 229
see also collocation processing: time in an English-speaking environment
- learner corpora 1, 8–10, 23–4, 60, 62, 64, 65
- International Corpus of Learner English (ICLE) 8, 26
- development of pedagogic materials 99–113
- spoken language 8, 9, 35
- Chinese Learner English Corpus (CLEC) 99, 100–1, 115
- English and Navajo parallel corpora 9, 50–9
- German Corpus of Learner English (GeCLE) 26, 27
- Nottingham International Corpus of Learner English of Chinese Learners (NICLE-CHN) 35, 38–9, 46
- Uppsala Student English Corpus (USE) 26–32, 60
see also British National Corpus; corpus analysis
- learners
 collocation learning strategies 11, 16, 182–4, 187–9, 191–2, 217, 225, 229, 230
- decision making about
 collocations 16, 96, 97, 208–9, 213, 217
- figured worlds 209, 210, 213, 217

learners (*Continued*)

- identity changes 213, 214, 216, 217, 219–21, 224, 226, 228, 231
- lexical and sociocultural reorganisation 212–22, 229
- look-up processes for
 - collocations 11, 82–4, 85, 91–8, 111, 118–20, 213–4, 219, 222
- see also* collocation dictionaries; electronic dictionaries; general purpose dictionaries, Oxford Collocations Dictionary for Students of English; print dictionaries
- metacognitive awareness 16, 183–4, 189, 191
- noticing 14, 15, 16, 105, 107, 183, 184, 186, 187, 188, 189, 190, 191, 192, 201–3, 206, 219, 225–6, 227
- past vocabulary practices 14–16, 110–11, 185, 186, 189, 212–13, 221, 222
- see also* collocation processing; learning collocations
- learning collocations
 - difficulties 9, 11, 15–16, 29, 73, 96, 97–8, 102, 110, 114, 115, 119, 139, 181–2, 185–6, 190–1, 192, 194–5, 208, 212–13, 214–15, 221, 225–8, 242–4
 - journals/reflections 16, 112, 185–92, 210–11, 230, 231
 - materials 5, 6, 10–12, 103–13, 114–121
 - notes 124–5, 133, 245–7, 247–8, 249, 254, 255, 256, 264, 268–70
 - strategies 11, 13, 16, 182–92, 201–3, 229
- see also* collocation processing; learners
- Lewis, Michael 1, 10, 11, 14, 49, 60, 140, 181
- lexical approach 11, 14, 60, 140, 195
- lexical bundles 5, 9, 22–4, 27–8, 49–59, 60, 62–3, 64, 241
- Lin, M. S. P. 9, 34–48, 61, 62, 238, 239
- longitudinal studies 14, 16, 39, 50–9, 62–3, 141–52, 184–93, 209–23, 226, 230
- L2 collocation development 9, 11, 13, 15–17, 21–33, 39, 49–59, 62–3, 107, 113, 135, 140, 141, 150, 169, 172, 175, 176, 182, 186, 208–223, 224–31, 243
- L2 collocation knowledge 12–17, 21, 23–4, 28, 33, 39, 80, 82, 91, 95, 97, 98, 100–2, 112, 115, 155–7, 158–9, 171, 173, 177, 181, 182, 190–1, 208, 210, 212, 215, 218, 220, 224–31, 239–41
- productive ~ 8, 11, 145, 16, 21, 23, 32, 35, 47, 62, 105–7, 117, 126–9, 137, 172, 173
- receptive ~ 13, 14, 62, 141, 144, 150, 151, 155–7, 159, 172, 173, 176
- relationship to vocabulary size 14, 161, 166–9, 172, 175, 217–19
- see also* collocation competence/ phrasal competence/ syntagmatic competence; collocation processing; phrasal knowledge
- L2 collocation knowledge
 - assessment research 12–14, 125–77, 242
- L2 collocation learner corpus
 - research 8–10, 21–65, 99–102, 113, 235–6, 242
- L2 collocation learner process and practice research 14–17, 181–231
- L2 collocation lexicographic and classroom materials
 - research 10–12, 69–122
- L2 lexical networks 15, 26, 228
- see also* collocation processing, and quantity-quality shift in lexical knowledge; mental lexicon; vocabulary depth
- mental lexicon 43, 46, 73, 151, 182, 186, 215, 216, 276, 277
- see also* collocation processing, and quantity-quality shift in lexical knowledge; L2 lexical networks; vocabulary depth

- metacognitive awareness 16, 183–4, 189, 191
see also noticing
- Moon, R. 4, 34, 35, 36, 38, 49, 61, 110
- multiword units (MWUs) 5, 6, 11, 15, 16, 22, 23, 36, 40, 104, 140, 145, 146, 147, 148, 152, 175, 194, 195, 216
- Nation, P. 1, 100, 102, 125, 156, 158, 159, 160, 194
- native and non-native speaker
 comparisons 8, 9, 15, 23–4, 31, 48, 53, 59, 101–2, 115, 151, 161, 164–5, 166, 175, 181, 208, 225, 240–1
- native-like collocation
 performance 13, 21, 24, 31, 69, 70, 71, 86, 115, 208, 219, 224, 226, 227
- native speaker collocation
 performance 23, 24, 74, 127, 128, 131, 224
- Nesi, H. 2, 12, 114–21, 233
- Nesselhauf, N. 1, 6, 8, 9, 12, 18, 21, 22, 25, 26, 31–2, 33, 60, 61, 65, 127, 141, 142, 143, 144, 153, 157, 208, 209
- node 4, 9, 24–5, 26, 29, 80, 125
see also collocates; collocation range; collocation span; node and collocate analysis; set
- node and collocate analysis 4, 9, 24–6, 28–32, 33, 60
see also collocates; collocation range; collocation span; node, set
- non-native speaker collocation
 performance 11, 13, 39, 60–5, 71, 83, 127, 128, 214, 217, 218, 226
- noticing 14, 15, 16, 105, 107, 183, 184, 186, 187, 188, 189, 190, 191, 192, 201–3, 306, 219, 225–6, 227
see also metacognitive awareness
- Obiedat, H. 12, 13, 125, 139, 153, 182, 209, 217
- O'Neill, M. 16, 225, 226, 228, 230, 231, 239
- open choice principle 4–5
compare idiom principle
- Oxford Advanced Learner's Dictionary (OALD) 71–3, 80, 93, 115, 118, 119, 121, 127
- Oxford Collocations Dictionary for Students of English (OCDSE) 10, 11, 88, 89, 90–1, 96–8, 115, 116, 118, 119, 196, 219
- entry structure 10, 11, 87–9, 96–8, 115, 116, 118, 119
- semantic grouping of collocates 10, 108–9, 110, 114
see also collocation dictionaries; electronic dictionaries; general purpose dictionaries; learners, look-up processes for collocates; print dictionaries
- Palmer, A. 154
- Palmer, H. 5, 17–18, 86
- Pawley, A. 23, 86, 139, 224
- Peters, E. 16, 194–207, 226, 229, 230, 231, 239
- phrasal knowledge 14, 139–41, 144, 145, 149, 172, 176
see also collocation competence/ phrasal competence/ syntagmatic competence; L2 collocation knowledge
- phraseological sequences/units 8, 29, 33, 34–48, 61, 63
- phonological coherence 35, 38, 39, 43–7, 48, 61, 62
- phraseological view of collocation 3, 5–7, 21–2, 34–5, 60–1, 142, 154–5
see also collocation, typological view of collocation
- Preposition + Noun collocations
see collocation types
- print dictionaries
 collocation displays 7, 10–11, 71, 84, 114

- print dictionaries (*Continued*)
see also collocation dictionaries;
 electronic dictionaries;
 general purpose dictionaries;
 learners, look-up processes
 for collocations; Oxford
 Collocations Dictionary for
 Students of English
- productive L2 collocation
 knowledge 8, 11, 145, 16,
 21, 23, 32, 35, 47, 62, 105–7,
 117, 126–9, 137, 172, 173
compare receptive L2 collocation
 knowledge
see also collocation processing
- receptive L2 collocation
 knowledge 13, 14, 62,
 141, 144, 150, 151, 155–7,
 159, 172, 173, 176
compare productive L2
 collocation knowledge
see also collocation processing
- Reppen, R. 9, 22, 49–59, 62,
 63, 238, 241
- restricted collocations 6, 9, 70,
 73–4, 139, 142, 143, 182,
 208, 222
see also collocations; collocation
 restrictions; figurative idioms/
 idiomatic collocations/
 idiomatic word combinations;
 idioms; selectional restrictions;
 word combinations
- Revier, R. 14, 125–38, 153, 155,
 158, 208, 238, 239–41
- Saito, H. 18, 86, 98
- Saito's Idiomatic English-
 Japanese Dictionary 18, 86
- Schmitt, N. 1, 12, 13, 15, 22, 23,
 34, 36, 49, 61, 140, 141, 151,
 153, 160, 228, 229
- selectional restrictions 4, 5, 6, 11,
 70, 73, 74, 155
compare idiomaticity
see also collocation restrictions;
 restricted collocations
- semantic prosody 83, 139
- semantic transparency 5, 6, 14, 38,
 70, 171–2, 173, 225, 239–40, 243
 non-transparent collocations 116,
 127–8, 135–7
 semi-transparent
 collocations 126–8, 130–3,
 135–7
 transparent collocations 116,
 126–8, 130–3, 135–7
- set 3–4, 10, 85, 120
see also collocates; node;
 collocation range; collocation
 span; node and collocate
 analysis
- Shillaw, J. 2, 14, 171–7, 235, 242
- Siepmann, D. 5, 7, 157, 241
- Sinclair, J. 3, 4, 5, 14, 22, 23, 24,
 34, 60, 61, 70, 139
- statistical measures 25, 29, 31,
 42–3, 53–4, 64, 101, 132–7,
 144, 146, 147–9, 199–200,
 201, 204–5
- Stenius Stæhr, L. 2, 26, 17,
 224–31, 235, 242, 244
- Syder, F. H. 23, 86, 139, 224
- tests
 Classical Test Theory (CTT) 160,
 169
 items 12–13, 129–31, 135–7,
 144, 150–1, 157–8, 196–8
 item selection 12–13, 129–30,
 143–4, 159–60, 169, 172,
 173, 177
 item facility (IF) 135–6, 160, 162,
 164–6
 item-total correlation
 (ITC) 135–6, 146, 149,
 160, 162, 164
 reliability 12, 13–14, 132–7, 147,
 148, 149, 153, 161, 166–70,
 172, 174–6
 test format 13, 128–9, 132,
 145–6, 151, 152, 155, 157–9
 validity 12, 13–14, 133–5, 137,
 153, 160, 164–5, 169, 172,
 173–7
- Item Response Theory
 (IRT) 169, 177

- see also* collocation,
operationalizing
- translation
tasks 80, 82, 84, 85, 89–90, 91–6
see also collocation processing,
translating collocations
- transparency
see semantic transparency
- typological view of
collocation 8, 17, 70
see also collocation; phraseological
view of collocation
- Verb + Noun collocations
see collocation types
- vocabulary depth 168, 228
see also collocation processing,
and quantity-quality shift in
lexical knowledge; L2 lexical
networks; mental lexicon
compare vocabulary size/breadth
- Vocabulary Levels Test (VLT) 159,
160, 172, 175
- vocabulary size/breadth 14, 157,
161, 166–8, 172, 228
compare vocabulary depth
- word clusters 105, 109–10, 111,
120, 222
see also clusters
- word combinations 5–9, 10, 11, 13,
16, 22, 34, 75, 84, 86, 87, 89,
102, 104, 115, 128, 129–31,
143, 144, 145, 146–8, 150,
151, 155, 156, 157–9, 168,
169, 170, 172–3, 175, 176,
177, 181–2, 185–6, 190, 192,
195, 212, 213, 215–17, 218,
221, 225, 227, 239–40, 241–2
see also chunks/chunking;
collocations; free
combinations; figurative
idioms/idiomatic collocations/
idiomatic word combinations;
idioms; restricted collocations
- Wray, A. 2, 17, 23, 34, 35, 36, 38,
39, 70, 151, 208, 209, 224,
225, 229, 232–44
- Yang, Y. 15, 16, 181–93, 209, 225,
226, 228, 230, 231, 239