JUSTYNA LEŚNIEWSKA

Psycholinguistic Accounts of Collocation

Introduction

In the simplest possible description of language production, words – the building blocks of language – are combined into utterances according to what the speaker intends to say (semantic restrictions) and the way in which things can be said in a language (syntactic restrictions). There is, however, a pervasive phenomenon which makes this simple system much more complicated – the phenomenon of collocation, or the tendency of some words to typically combine with others in a way which is not determined exclusively by semantic and syntactic restrictions. For several years now, collocations have been attracting more and more attention in applied linguistics and in second language acquisition research.

There are many different approaches to defining and classifying collocations; however, they are not the focus of this paper. It is sufficient here to say that all types of words combinations can be seen as collocations of differing levels of "fixedness," ranging from relatively open combinations such as "have children" to idioms such as "drive a hard bargain." There are degrees of "fixedness" and they involve various additional criteria, such as syntactic regularity, restrictedness, compositionality, and institutionalization. However, a more common use of the term "collocation" is the one in which reference is made to somewhere in the

middle of this continuum of fixity, that is, to word combinations which are not entirely free and not completely fixed either, such as "inevitable consequence" or "glaringly obvious." This is the sense in which the term *collocations* is used in foreign language teaching.

Interest in patterns of lexical co-occurrence comes from various disciplines: linguistics (e.g. Firth 1957; Halliday 1966; Sinclair 1991; Hoey 2005), lexicography (e.g. Aisenstadt 1981; Cowie 1981), corpus linguistics (e.g. Altenberg 1993; Kjellmer 1990; Sinclair 1991; Partington 1998; Hoey 2005), language processing (e.g. Pawley and Syder 1983; Peters 1983; Wray 2002a), discourse analysis (e.g. Ferguson 1976; Wray 2002a), second language acquisition research (e.g. Hakuta 1976), second language pedagogy (e.g. Alexander 1984, 1988; Nattinger and DeCarrico 1992), language assessment (e.g. Schmitt 1998), studies of learner language (e.g. Howarth 1996; Altenberg and Granger 2001; Nesselhauf 2003). Phraseology is sometimes mentioned as a separate field of study (Cowie 1998). However, although the pervasiveness of collocations in language is widely recognized, and although collocations merit special attention in foreign language teaching, relatively little is known about the psychological reality behind the use of word combinations. The aim of this paper is to overview the available findings concerning the psycholinguistic aspect of collocational knowledge and their implications for foreign language acquisition.

Collocational Competence

Benson, Benson, and Ilson define collocation on the example of the expression *commit murder* (1986: 253, emphasis added):

> *To commit murder* differs from free combinations in two ways. Firstly, the synonymy of the verb is restricted. In this instance the only synonym seems to be *to perpetrate*. Secondly, and more importantly, the combination *to commit murder* is used frequently; **it springs readily to mind; it is psychologically salient**; it is a "fixed phrase" in English.

This description of collocation moves from the purely linguistic criteria to the psycholinguistic dimension. What is the mechanism, then,

that makes certain phrases "spring readily to mind"? A number of the-
oretical frameworks account for this phenomenon in various terms,
though what makes them different is mostly the degree of emphasis
placed on collocation rather than the basic understanding of the phe-
nomenon.

A strong view of the importance of collocational competence is ad-
vocated by Ellis (2001), who argues that language users store chunks of
language in long-term memory and acquire the experience of how likely
particular items are to co-occur. A crucial role is played by associations
between items which are observed to appear in the vicinity of each other.
Language users are able to break up the chunks according to the gram-
mar rules of the language, but can produce and comprehend them with-
out reference to those rules. A lot of learning can also be accounted for
in terms of learning by association, as a result of encountering certain
word combinations.

Hoey (2005) goes even further and argues that lexical patterns are
responsible for the structure of language, and that grammar is merely
an outcome of the pervasiveness of collocation. Collocation, in turn,
is presented as a psychological concept. The recurrent co-occurrence
of words is explained by means of the psychological concept of prim-
ing. Hoey writes that "we can only account for collocation if we as-
sume that every word is mentally primed for collocational use. As a
word is acquired through encounters with it in speech and writing, it
becomes cumulatively loaded with the contexts and co-texts in which
it is encountered, and our knowledge of it includes the fact that it
co-occurs with certain other words in certain in certain kinds of con-
text" (2005: 8). Grammar emerges from the recurrent patters of word
combinations.

Another approach is that which sees collocational knowledge as
merely one of the sub-components of word knowledge. For each lexi-
cal item, the language user has to store information on its pronuncia-
tion, meaning, spelling, etc., and also on the collocations that the word
usually appears in. One such list of the components of word knowledge
was proposed by Richards (1976; see also Meara's (1996b) discussion of
Richard's framework). I.S.P. Nation (1990: 31) presents a similar approach
in stating that knowing a word involves knowing its form (spoken, writ-
ten), position (grammatical patterns, collocations), function (frequency,

appropriateness), meaning (concept, associations) – all of these in both the receptive and productive mode.

The third approach is based on the distinction between two kinds of mechanisms in language production. Pawley and Syder (1983) argue that language users store both individual lexical items and hundreds of thousands of preconstructed phrases (so that a particular lexical item is stored many times, as part of various chunks). Language users resort to grammatical rules and to chunked information interchangeably.

Similar, "dual" distinctions have been made by Sinclair (1991), Skehan (1998), and by Wray (2002a, 2002b). Sinclair has introduced the distinction between the open choice principle and the idiom principle. The open-choice principle

> is a way of seeing language text as the result of a very large number of complex choices. At each point where a unit is completed (a word or a phrase or a clause), a large range of choice opens up, and the only restraint is grammaticalness. . . . It is often called a "slot-and-filler" model. . . . At each slot, virtually any word can occur.   (1991: 109)

The open choice principle operates, therefore, like traditional grammar-centred models of language: there are a number of syntagmatic choices available for each slot along the paradigm. On the other hand, the idiom principle "is 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 analyzable into segments" (1991: 110).

A similar distinction has been made by Skehan (1998) between two modes of processing available to language users. The first mode operates on the level of grammatical rules, which make it possible to generate novel utterances by putting individual words together; for example when meanings have to be expressed with precision or creativity. The second mode is based on memorized multi-word items, which can be quickly retrieved. This makes it possible for the speaker to communicate fluently under normal time constraints. According to Skehan, the two modes mentioned above – syntactic rules versus retrieval of multi-word items – are used interchangeably in language production. Native speakers of a language can flexibly combine them according to the demands of the situation.

Wray (2002a, 2002b) takes this concept even further, suggesting that "formulaic processing is the default," and that "construction out of, and reduction into, smaller units by rule occurs only as necessary" (2002b: 119). This fact can also be seen as an explanation for the existence of irregularity in language: "if we only create and understand utterances by applying rules to words and morphemes, it is difficult to see why irregularity should be tolerated, let alone why an item or construction should progress from regular, to marked, to antiquated, to a fossilized historical relic" (Wray 2002b: 118).

Pawley's and Syder's claim that language users have access to both individual lexical items and to entire memorized chunks is given as the explanation for what they call "two puzzles for linguistic theory," namely, that the language production of native speakers is characterized by "native-like selection" and "native-like fluency." The first characteristic refers to the ability to produce phrases which are the natural form of expression for native speakers, instead of odd collocations: "the ability of the native speaker routinely to convey his meaning by an expression that is not only grammatical but also native-like; what is puzzling about this is how he selects a sentence that is natural and idiomatic from among the range of grammatically correct paraphrases, many of which are non-native-like or highly marked usages" (1983: 191). The other characteristic, fluency, is the ability to produce fluent stretches of connected discourse: "there is a puzzle in here that human capacities for encoding novel speech in advance or while speaking appear to be severely limited, yet speakers commonly produce fluent multi-clause utterances which exceed these limits" (1983: 191). Comparing recordings of spontaneous speech, Pawley and Syder note the typical speed and prosodic patterns of utterances, and observe that some expressions – e.g. "I don't need anyone to tell me what to do" – are pause free and more promptly delivered that the normal rate for language production would allow.

The two phenomena – native-like selection and fluency – provide the evidence for the fact that "fluent and idiomatic control of a language rests to a considerable extent on knowledge of a body of 'sentence stems' which are 'institutionalized' or 'lexicalized'" (Pawley and Syder 1983: 191). Pawley and Syder argue that clause-length units are partly or wholly fixed. This minimizes "clause-internal encoding" to be done and frees more memory needed for other tasks in communication. The abil-

ity to recall larger chunks from memory does not mean that the chunks are not analyzable into segments. In other words, the function of grammar which makes it possible to construct an infinite number of new sentences on the basis of a finite set of rules – most influentially put forward in Chomsky's generative grammar – is not excluded as a principle of language production, rather, both mechanisms are available to the speaker. As Pawley and Syder say, native speakers "do *not* exercise the creative potential of syntactic rules to anything like their full extent, and that, indeed, if they did so they would not be accepted as exhibiting native-like control of the language" (1983: 193).

To Pawley and Syder, "native-like selection" is an element of communicative competence. Many of the examples they give refer to the kind of language choices which one would normally call "choosing the right expression for the right situation." For example, it is more usual to say *I'm so glad to see you* than *to see you gladdens me so*. There are of course degrees of unnaturalness of particular expressions, and it is perfectly possible for a speaker to use a less natural expression on purpose. It must be noted that the "naturalness" of certain expressions that Pawley and Syder are preoccupied with can only be properly judged given the situational context, as the use of the right expression is determined by linguistic convention. This does not, however, change the basic fact that part of the native speaker's command of language is knowing which usages are more typical and which are marked, or unusual.

By discussing at length various examples, Pawley and Syder prove that native-like selection is not a matter which can be explained in purely grammatical terms (as it could, for instance, if it were true that the most "natural" of alternative expressions also happened to be always the shortest one, or the one with a specific word order). The point they make is that there is no sense in stretching the rules of grammar by indefinitely making them subtler and subtler to account for choices which are not a question of grammar – they show a different principle at work.

There are two issues which lie at the heart of the model produced by Pawley and Syder. To clarify the picture, let us separate them:

1) The pragmatic aspect: "lexicalised sentence stems" have a particular pragmatic role. The stress is on the importance of the functional aspects of some word combinations, their pragmatic non-

compositionality and their specialized roles in discourse. As Weinert puts it, "certain language sequences have conventionalized meanings which are used in certain predictable situations" (1995: 196).

2) The psycholinguistic aspect: "memorized sentences" and "lexicalised sentence stems" are to some extent processed as wholes, as ready-made chunks of language. A more extensive discussion of the psycholinguistic processes behind the notion of chunking seems to be in order at this point.

Chunks and Formulas

What then, is the nature of the process of chunking which enables the "idiom principle"? As I.S.P. Nation illustrates (2001: 319), chunking occurs at all levels of language. Complex words are usually processed as wholes, not as combinations of individual morphemes. Morphemes, in turn, are processed as units, not as sequences of individual phonemes. Chunking means that smaller units can be grouped into larger wholes, but wholes can also be later analyzed into segments. As Pawley and Syder (1983) demonstrate, the main advantage of chunking is reduced processing time, and, therefore, faster language comprehension and production. The disadvantage of chunking is that it takes up storage space – there are, obviously, many more chunks than individual components. Nation explains the economics of chunking by drawing on an analogy with word building. Research evidence suggests that high-frequency complex words, such as *unable*, are stored as wholes, whereas low-frequency ones, such as *unambiguousness*, are re-created by rules when they are needed (see Aitchison 1987). High-frequency items, then, are stored as chunks, reducing processing time, since they occur often enough to make up a large proportion of the overall language produced. Low-frequency items, on the other hand, do not "deserve" separate storage space.

Chunking is based on the concept of the storage of complex units in memory, which brings the internal organization of the mental lexicon into the discussion. Although many hypotheses have been put forward concerning the structure of the mental lexicon (see Aitchison 1987;

Singleton 1999), it is not entirely clear how units larger than individual words are stored and processed in the mental lexicon. Carter notes that the storage in memory of such items remains an "interesting factor to investigate," and particularly the question "whether items are stored singly or as whole composite units" (1998: 75). The notion of a chunk implies that a multi-word combination is stored just like an individual word, or, at least, that the constituent items of the chunk are recalled and produced in a certain linear order. As Weinert (1995: 197) suggests, there is no reason to believe that all kinds of formulaic chunks are stored in the same way. Evidence which suggests that multi-word units are stored as wholes comes mostly from research on the processing of figurative language, including idioms. Studies which compare the speed of processing of phrases which can have both a literal and a figurative meaning, and which are put in contexts which are ambiguous, suggest that subjects take longer to process the literal meaning.

Some evidence of chunking comes from studies of formulaic language. Although the exact criteria for defining formulaic utterances vary, and sometimes the phenomenon is described under different names (such as "prefabricated" or "ready-made language," "routines," "chunks"), the focus in studies of formulaic language is always on strings of language which seem to be acquired and/or processed as wholes, and are often unanalyzed. In studies of language acquisition, their existence in learner language as chunks which are "produced or recalled as a whole" is noted on the basis of various criteria, among which the most common are: phonological coherence ("fluent, non-hesitant encoding" – Weinert 1995: 182), the fact that the sequence contains lexemes which do not appear elsewhere in learner output, the lack of productive uses of the rules underlying a sequence, the idiosyncratic or inappropriate use of sequences, situational dependence, frequency and invariance in form. Entirely fixed strings (*How are you?*) are often distinguished from sequences with open slots, sometimes called prefabricated patterns (*Can you? . . .*), in which part of the chunk is fixed and one or more elements can be exchanged or modified. It is not difficult to predict that the precise classification of a string of words as formulaic or not in a particular learner's production may be problematic; indeed, research studies dealing with formulaic language have to cope with serious methodological problems in this respect. Formulas may include virtually any type of word com-

bination, as any type of word combination may be used by a particular subject as a chunk. Formulaic language may thus include both "free" and "restricted" word combinations, and combinations of varying degree of pragmatic specialization.

The presence of formulaic speech has been observed in child L1 and L2 acquisition. Formulas seem to sometimes play a role in adult L2 acquisition, but research findings in this area are often inconclusive or contradictory. One thing seems clear – namely, that both in first and second language acquisition, chunks are used as wholes *before* they become analyzed into parts. Whether strings of words can acquire formulaic status through frequent production is an entirely different question, which so far has not been thoroughly researched.

Weinert postulates the view of language as a "formulaic-creative continuum" and quotes Langacker (1987: 46):

> Speakers do not necessarily forget the forms they already know once the rule is extracted, nor does the rule preclude their learning additional forms as established units. . . . Out of this sea of particularity speakers extract whatever generalizations they can. Most of these are of limited scope, and some forms cannot be assimilated to any general patterns at all.

According to Weinert, neurological evidence supports "the psychological reality of formulaic language in terms of storage and production" (1995: 185). She quotes studies in psycholinguistics and neurology which suggest that rote and rule processes may interact. However, the nature of such interaction is far from simple:

> The view of language as a formulaic-creative continuum suggests that the units of knowledge and production may vary, including fixed formulas, mini-grammars, and general rules. Language may be represented in dual or even multiple form, analyzed at various levels. (Weinert 1995: 198)

In many psycholinguistic accounts collocations are often discussed together with multi-word units, and their acquisition and processing is often discussed alongside the more fixed elements of language, for example, Yorio (1989: 55) uses the term "conventionalized language" to refer to "idioms, formulas, collocations, etc." Most sources usually imply that what is said about formulaic language is somehow relevant to collo-

cations in general, but the topic is left without further elaboration. However, findings on the use of formulaic language in L2 use are clearly not directly applicable to collocations, since the latter category consists of word combinations which would not normally be produced as unanalyzed wholes, rarely have a specific pragmatic or situational function, and are unlikely to play a role in language acquisition. However, since collocations are not produced entirely by rule, their storage and processing must involve chunking to some extent. What is needed is a theoretical model that can account for links of variable strength holding between words.

The Mental Lexicon and Collocations

It is a widely held view that the mental lexicon is a complex system of interrelated items, although the precise models of lexical storage and processing vary (see Aitchison 1987; Singleton 1999). The particular items are connected by links of varying strength. Each node is connected to many other nodes, which in turn have their own networks of connections. Any node which is activated in turn activates those points with which it is most strongly connected. As Aitchison (1987: 226) puts it in very illustrative terms:

> Finding a word in the mental lexicon can be envisaged as following a path through this complex network, with some networks being stronger than others. For well-known common words, the paths are well-worn, and it is easy to travel fast. But for words used only occasionally, the paths are narrow and dimly lit. Meanwhile, new tunnels are perpetually being dug.

Aitchison reviews several word-association studies and notes that such experiments produce evidence for mostly four types of connection between nodes: co-ordination, collocation, superordination, and synonymy. Collocational links are almost as common as co-ordinate links, and their strength depends on the degree of fixedness of a particular collocation. Although Aitchison does not discuss this type of links in greater detail, the idea that collocations are enabled by links of varying strength holding between the individual items in the mental lexicon seems a very

attractive way of conceptualizing the storage of word combinations. Collocations could be seen as having links which are weaker than those in multi-word items.

Although Aitchison does not discuss the mental lexicon of the L2 user, it is likely that all kinds of links and associations in the L2 lexicon are weaker than those in the L1 lexicon. Such a possibility is considered by Meara (1996a, 1996b), who argues that the mental lexicon of language learners could be characterized in terms of three general dimensions rather than in terms of the knowledge of individual lexical items. The three properties of the lexicon as a whole are its size, the speed of access to its elements, and the richness of the lexical structure which links the items in the lexicon, that is, "lexical organization." Meara suggests that L2 speakers' vocabularies are not as well structured as the vocabularies of L1 speakers, and that the degree of lexical organization can be studied by means of testing the ability of producing networks of associations linking two vocabulary items. For example, native speakers can easily produce numerous chains of associations linking *sea* and *butterfly*, such as the following (Meara 1996a: 49):

    sea – blue – sky – fly – butterfly
    sea – horse – horse fly – butterfly
    sea – swim – butterfly stroke – butterfly
    sea – sand – sandwich – butter – butterfly
    sea – weed – flower –butterfly

In Meara's words, "the fact that a large number of plausible chains can be found, and the fact that these chains are relatively short, suggests that there is a very high level of interconnection between words in an L1. For L2 speakers, connections seem to be more tenuous (1996a: 49)." Meara does not comment on the fact that there are different kinds of associations at work in the above examples. The main problem with such a measure of the lexical organization of the L2 lexicon seems that the associations are to a varying extent language-specific. The association of *sky* with *fly* is augmented by the rhyme, while the association of *horse* with *horse fly* is based on similarly language-specific knowledge of compound nouns. However, the association of *sea* with *river, holiday, sand* or *blue* can be seen as a reflec-

tion of the knowledge of the world which is independent of a particular language.

Nevertheless, the idea that the L1 and L2 lexicons differ in terms of the strength of the links between particular items seems very plausible. It is well known that the L2 mental lexicon changes its structure with the development of the L2, moving from phonological to semantic organization. This fact accommodates the possibility of gradual qualitative and quantitative changes to the network of links in the lexicon. To a great extent, words do not have meaning in isolation, but derive their meaning from a number of other words in the language, and part of lexical competence is the ability to relate a word to other words. Someone whose mental lexicon is highly structured will have a good command and awareness of the links and associations between various words. The possibility that the collocational links of language learners are weaker than those of native speakers implies that the learners are more likely to put words together according to principles other than collocational.

An important question concerns the extent to which the structure of the L1 lexicon affects the functioning of the L2 lexicon. This question involves the ongoing debate about the extent of separation versus integration of the mental lexicons of a bilingual person. Common sense suggests that they cannot be wholly integrated nor entirely separated, but the precise extent of the interconnection is still uncertain and the issue of much ongoing psycholinguistic research (see Schreuder and Weltens 1993; de Groot and Kroll 1997; Singleton 1997; Garrod and Pickering 1999; Nicol 2001).

This separation/integration issue is the reason why the discrepancies between the L2 mental lexicon of a learner and that of a native speaker can be viewed in two different ways:

1) As caused by the fact that the L2 system is not yet fully developed, the implication being that the L2 mental lexicon is on its way to becoming fully native-like, with the potential of becoming indistinguishable from that of a native speaker.

2) As unavoidable regardless of the level of development, because the competence of a bilingual in one of his languages will always be different from that of a person who speaks only one language. The assumption here is that a mind with more than one language will never

function in any of the languages in the same way as the mind of a monolingual – the simple fact of there being more than one language introduces some qualitative changes into the system. Cook (1992, 2002) introduced the notion of "multi-competence" or "the knowledge of more than one language in the same mind" (2002: 10) in order to emphasize that the language knowledge of bilinguals will always by definition be different with respect to any of the languages known than the knowledge of a monolingual person.

If the language systems of a bilingual can be kept completely separate, than indeed a bilingual operating in, for example, the L1, will be qualitatively equivalent to any monolingual speaker of the L1. However, if complete separation is impossible – which is very likely – a bilingual's use of the L1 or the L2 will never be the same as that of a monolingual speaker of either of the two languages. It is even theoretically possible that the lexicons of native speakers and very advanced learners may differ in terms of the underlying processes of storage, access, or retrieval, without affecting the language production of the speakers in a perceptible way.

What the above implies for the use of collocations in L2 is that the collocational use of bilinguals may always, by definition, be different from that of monolinguals, and the cross-linguistic influence may manifest itself in both directions, L1 → L2 and L2 → L1. Technically speaking, what is cross-linguistic influence with respect to collocations? Such a notion implies that the collocational links themselves are transferred. If cross-linguistic transfer of collocations does happen, that means that the two L2 words must be activated via a link holding between L1 items. Therefore, collocational cross-linguistic influence seems to fit the model of subordinative or compound rather than coordinate bilingualism, to use Weinreich's (1953) terms, although any model of such kind is bound to be only a simplification (see Singleton 1999: 172ff).

Conclusions

Collocations are a borderline case with respect to chunking vs. rule formation, which means that research findings on formulaic language are

not very helpful in accounting for the existence of collocations in language. It seems likely that collocations are enabled by associative links in the mental lexicon. While fixed multi-word units involve very strong links, the use of collocations involves the activation of links of variable strength. The implication for second language acquisition is that the former category may be easier to learn and use than the latter. If selectional restrictions are best learned when they are strong, the L2 learners would tend to either operate according to the open-choice principle, or to use the word combinations which are fixed. The subtler kinds of collocational restrictions are likely to be more elusive, and more liable to cross-linguistic influence.

REFERENCES

Aisenstadt, Ester. 1981. "Restricted Collocations in English Lexicology and Lexicography." *ITL Review of Applied Linguistics* 53: 53–61.

Aitchison, Jean. 1987. *Words in the Mind: An Introduction to the Mental Lexicon.* Oxford: Blackwell.

Alexander, Richard J. 1984. "Fixed Expressions in English: Reference Books and the Teacher." *ELT Journal* 38: 127–34.

Alexander, Richard J. 1988. "Fixed Expressions, Idioms and Collocations Revisited." *Beyond Words.* Ed. Paul Meara. London: BAAL and CILT.

Altenberg, Bengt. 1993. "Recurrent Verb-Complement Constructions in the London-Lund Corpus." *English Language Corpora: Design, Analysis and Exploitation.* Ed. Jan Aarts, Pieter de Haan, and Nelleke Oostdijk. Amsterdam: Rodopi.

Altenberg, Bengt, and Sylviane Granger. 2001. "The Grammatical and Lexical Patterning of MAKE in Native and Non-Native Student Writing." *Applied Linguistics* 22: 173–94.

Benson, Morton, Evelyn Benson, and Robert Ilson. 1986. *Lexicographic Description of English.* Amsterdam: Benjamins.

Carter, Ron. 1998. *Vocabulary: Applied Linguistic Perspectives.* 2nd ed. London: Routledge.

Cook, Vivian. 1992. "Evidence for Multicompetence." *Language Learning* 42: 557–91.

Cook, Vivian. 2002. "Background to the L2 User." *Portraits of the L2 User.* Ed. Vivian Cook. Clevedon: Multilingual Matters.

Cowie, Anthony P. 1981. "The Treatment of Collocations and Idioms in Learners' Dictionaries." *Applied Linguistics* 2: 223–35.

Cowie, Anthony P., ed. 1998. *Phraseology: Theory, Analysis, and Applications.* Oxford: Clarendon.

De Groot, Anette M.B., and Judith Kroll, eds. 1997. *Tutorials in Bilingualism: Psycholinguistic Perspectives.* Mahwah, NJ: Lawrence Erlbaum.

Ellis, Nick C. 2001. "Memory for Language." *Cognition and Second Language Instruction.* Ed. Peter Robinson. Cambridge: Cambridge UP.

Ferguson, Charles. 1976. "The Structure and Use of Politeness Formulas." *Language and Society* 5: 137–51.

Firth, John R. 1957. *Papers in Linguistics.* Oxford: Oxford UP.

Garrod, Simon, and Martin Pickering, eds. 1999. *Language Processing.* Hove: Psychology P.

Hakuta, Kenji. 1976. "A Case Study of a Japanese Child Learning English." *Language Learning* 26: 321–51.

Halliday, M.A.K. 1966. "Lexis as a Linguistic Level." *In Memory of J.R. Firth.* Ed. C.E. Bazell, J.C. Catford, M.A.K. Halliday, and R.H. Robins. London: Longman.

Hoey, Michael. 2005. *Lexical Priming: A New Theory of Words and Language.* London: Routledge.

Howarth, Peter A. 1996. *Phraseology in English Academic Writing: Some Implications for Language Learning and Dictionary Making.* Tübingen: Niemeyer.

Kjellmer, Göran. 1990. "Patterns of Collocability." *Theory and Practice in Corpus Linguistics.* Ed. Jan Aarts and Willem Meijs. Amsterdam: Rodopi.

Langacker, Ronald W. 1987. *Foundations of Cognitive Grammar. Volume 1: Theoretical Prerequisites.* Stanford: Stanford UP.

Meara, Paul. 1996a. "The Dimensions of Lexical Competence." *Performance and Competence in Second Language Acquisition.* Ed. Gillian Brown, Kirsten Malmkjaer, and John Williams. Cambridge: Cambridge UP.

Meara, Paul. 1996b. "The Vocabulary Knowledge Framework." *Vocabulary Acquisition Research Group: Virtual Library.* 15 Dec. 2005 <http://www.swan.ac. uk/cals/calsres/vlibrary/pm96d.htm>.

Nation, I.S.P. 1990. *Teaching and Learning Vocabulary.* New York: Newbury House.

Nation, I.S.P. 2001. *Learning Vocabulary in Another Language.* Cambridge: Cambridge UP.

Nattinger, James, and Jeanette DeCarrico. 1992. *Lexical Phrases and Language Teaching.* Oxford: Oxford UP.

Nesselhauf, Nadja. 2003. "The Use of Collocations by Advanced Learners of English and Some Implications for Teaching." *Applied Linguistics* 24: 223–42.

Nicol, Janet, ed. 2001. *One Mind, Two Languages: Bilingual Language Processing.* Oxford: Blackwell.

Partington, Alan. 1998. *Patterns and Meanings: Using Corpora for English Language Research and Teaching.* Amsterdam: Benjamins.

Pawley, Andrew, and Frances Hodgetts Syder. 1983. "Two Puzzles for Linguistic Theory: Nativelike Selection and Nativelike Fluency." *Language and Communication.* Ed. Jack C. Richards and Richard W. Schmidt. London: Longman.

Peters, Ann M. 1983. *The Units of Language Acquisition.* Cambridge: Cambridge UP.

Richards, Jack C. 1976. "The Role of Vocabulary Teaching." *TESOL Quaterly* 10: 77–89.

Schmitt, Norbert. 1998. "Measuring Collocational Knowledge: Key Issues and an Experimental Assessment Procedure." *ITL Review of Applied Linguistics* 119/120: 27–47.

Schreuder, Robert, and Bert Weltens, eds. 1993. *The Bilingual Lexicon.* Amsterdam: Benjamins.

Sinclair, John. 1991. *Corpus, Concordance, Collocation.* Oxford: Oxford UP.

Singleton, David. 1997. "Cross-Linguistic Aspects of the Mental Lexicon." *Trends in Linguistics: Language History and Linguistic Modelling.* Studies and Monographs 101. Ed. Raymond Hickey and Stanisław Puppel. Berlin: Mouton de Gruyter.

Singleton, David. 1999. *Exploring the Second Language Mental Lexicon.* Cambridge: Cambridge UP.

Skehan, Peter. 1998. *A Cognitive Approach to Language Learning.* Oxford: Oxford UP.

Weinert, Regina. 1995. "The Role of Formulaic Language in Second Language Acquisition: A Review." *Applied Linguistics* 16: 180–205.

Weinreich, Uriel. 1953. *Languages in Contact.* New York: Linguistic Circle of New York.

Wray, Alison. 2002a. *Formulaic Language and the Lexicon.* Cambridge: Cambridge UP.

Wray, Alison. 2002b. "Formulaic Language in Computer-Supported Communication: Theory Meets Reality." *Language Awareness* 11: 114–31.

Yorio, Carlos A. 1989. "Idiomaticity as an Indicator of Second Language Proficiency." *Bilingualism Across the Lifespan: Aspects of Acquisition, Maturity and Loss.* Ed. Kenneth Hyltenstam and Loraine K. Obler. Cambridge: Cambridge UP.