11

How can a corpus be used to explore patterns?

Susan Hunston

1 Types of patterns

The term "pattern" is often used in corpus linguistics with a variety of meanings. Its core definition might be "an observed regularity", but the nature of that regularity is not fixed.

When considering this variation, a number of distinctions might be made. These can be summarised as:

a. The level of abstraction. "Pattern" may refer to specific words occurring together – something akin to "phraseology" – or it may refer to abstract categories such as grammatical classes or semantic categories.
b. The scale of the phenomenon. "Pattern" may be used to describe local regularity, that is, recurring sequences of words or categories, or a global regularity, that is, the greater or lesser frequency of a feature in specific populations of texts.
c. The nature of the corpus. "Pattern" may be used to describe the language in an individual text, or in the texts produced by an individual writer or in a general corpus of texts.
d. The source of the phenomenon. In those cases where a "pattern" is observed in a general corpus, it may be assumed that the pattern is the consequence of constraints imposed by the language itself (such as *start* followed by a to-infinitive verb) or that it is a consequence of people repeating the same ideas (such as *start* occurring in a clause beginning with the conjunction *before*).

These distinctions can be used to discuss a number of examples.

Repetition in the language of individual speakers may be observed without recourse to a corpus (see distinction c). Political speeches have long been known to incorporate this kind of regularity. One instance is the speeches of former US President Barack Obama, which contain many instances of patterning, such as those in examples 1 and 2.

DOI: 10.4324/9780367076399-11

(1)
I've gone to some of the best schools in America and lived in one of the world's poorest nations.
(2)
The church contains in full the kindness and cruelty, the fierce intelligence and the shocking ignorance, the struggles and successes, the love and yes, the bitterness and bias that make up the black experience in America.

In example 1 we find repetition of the superlative and its co-text: *some of the best schools in America; one of the world's poorest nations*. In example 2 there is repetition of binomial opposites: *the kindness and cruelty; the intelligence and the ignorance; the struggles and successes; the love and the bitterness*. It is also notable that although Obama here creates patterns of repetition within his own speech, he also draws on patterns that are widely used in English, such as "quantity + *of the* + superlative adjective + plural-noun + *in* place" (*some of the best schools in America; one of the richest women in Germany*) or "*the* + group of people + *experience* (*the black experience; the visitor experience; the US experience*).

Turning to patterns identified in a corpus, the word "pattern" may be used to describe the co-occurrence of individual words (see distinction a). A researcher may talk about the patterns of usage of a word, meaning the phrases that word occurs in or the collocates it attracts. For example, the word *cake*, when it is used figuratively, might be observed to occur in the following phrases:

- *the icing on the cake;*
- *a piece of cake;*
- *have one's cake and eat it.*

These might be described as the "patterns" of the figurative word *cake*.

The word "pattern" can also refer to more abstract categories of language (see distinction a). For example, Lei and Liu (2018) use the term "collocation patterns" to refer to identified combinations of word classes. They specify 11 such patterns, including "adjective + noun" and "verb + noun". Thus there is a finite number of patterns and each one is precisely identified as a pair of word classes. Similarly, Lehmann (2018) refers to "four variant patterns" that are used with the verb *provide*. These are specified as "the *with*-pattern", "the *to*-pattern", "the *for*-pattern" and the "npnp pattern". Examples from the British National Corpus are "*providing* them with a route out of poverty", "*provide* exit visas to intellectuals", "*providing* a single nest site for a group of four birds" and "we did *provide* them facilities". In this example the "pattern" is not simply the co-occurrence of two words (such as *provide* and *with*), but the sequence of more abstract elements: verb + noun phrase + prepositional phrase.

Hanks (2013) specifies semantic categories as well as grammatical ones. For example, one of the patterns noted for the verb *agree* in the Pattern Dictionary of English Verbs (https://pdev.org.uk) is:

- [Human | Institution] **agree** to [Activity = Future | Proposition]

This notation specifies that in an example such as example 3, the subject (*he*) must be a noun phrase indicating a human being or human institution, and the to-infinitive clause (*to stand as Vice President*) necessarily indicates a future activity.

(3)
…he's agreed to stand as Vice President.

[BNC]

In the examples used so far, what has been described as a pattern might be said to be an integral part of how English works. Alternative explanations can be found for patterns, however (see distinction d). For example, the wordform *decide* is often preceded by phrases that suggest either rights and obligations (*a woman's right to decide; have the opportunity to decide; it is up to each local authority to decide; they will now have to decide; has 28 days to decide*) or possibility, ability and futurity (*a debate which could decide; they just cannot decide; the office will decide; has yet to decide; is planning to decide*). This has more to do with real-world conditions in which decisions are taken than with restrictions on the verb: we can or cannot decide, we have an obligation to decide or the right to make a decision rests with us. Hunston (2008) refers to patterns occurring as the result of what people often say as "semantic sequences".

The final example of how the word "pattern" is used takes us beyond concordance lines and refers to the distribution of a word or structure across a corpus or between corpora (see distinction b). For example, Laws and Ryder (2018) refer to 'the effect of context formality on the usage patterns of derivatives between different speech registers'. Here, 'usage patterns' means the relative frequency of a particular feature (derivatives) in sub-corpora comprising registers. In this chapter, however, it is local patterns, specifically those observed using concordance lines, that are the focus of discussion.

It is clear from this introductory discussion that the term "pattern" is a problematic one and is used in at least two senses. In a general sense it means any regularity in phrasing or in frequency that has been observed. In a technical sense a pattern can be a defined set of entities incorporating aspects of phrasing and/or frequency. This ambiguity as to whether we are looking at a general thing here or a specific one carries over into the examination of concordance lines, which is the topic of the next section.

2 Reading concordance lines

In this section we explore local patterns in a corpus, identified using the "keyword in context" (KWIC) method, otherwise known as the set of concordance lines. Essentially, concordance lines consist of a node word or phrase with a small amount of context (measured in characters) to the left and the right (see Chapters 9, 10 and 14, this volume, for more on concordances). In most concordancing programmes the amount of context can be increased on demand. Showing more context requires wrapping the lines, which gives a more "normal" reading experience but reduces the visual impact of any repetition. Most programs, too, allow the concordance lines to be manipulated in ways that make repetition visually salient. These techniques range from being able to sort lines so that the word(s) before or after the node occur in alphabetical order to using colour to identify word class.

For example, if concordance lines are obtained with the wordform *view* as the node word (the examples that follow are from the Bank of English corpus), sorting those lines so that the words immediately to the left of the node are in alphabetical order draws attention to the frequency of *from* and *of* in that position. These prepositions occur in the phrases *obscured / hidden from view* and *point of view*. Left-sorting also brings together, and so makes noticeable, those lines where *view* is preceded by *the* or *this*. *The view* focuses attention on the phrase *the view that*:

```
things that they didn't need. But the view that design is dead portrays the
  medieval scholars. Yet this was the view that all major Christian theologians
 eventually takes place. However, the view from the Moroccan side of the wall
      from the applicant. It took the view that the risk of his association
```

The frequency of *This view…* suggests that *view* is often used to refer back and to summarise a segment of the preceding discourse; this is labelled "encapsulation" in Sinclair's (1990/2004) terminology:

```
 the success of their treatment. This view of clients as experts on matters
 in Capital and Class (1986). 10 This view was expounded by RIIA prominents
biological and chemical weapons. This view was then supported by a senior
```

Sorting the lines to the right of *view* additionally draws attention to phrases where *view* is followed by *of*, and we might note that *view of* is often preceded by an adjective:

```
the success of their treatment. This view of clients as experts on matters
  Black himself did not have a clear view of the age at which a child could
  historians by and large take a dim view of my grandfather's role and
      which reflected a market-frame view of the state's responsibility to
```

Expanding the concordance lines allows ambiguities to be resolved. For example, expanding the lines where *view* is followed by *that* shows that there are two distinct patterns – *view* + appositive clause (example 4) and *view* + relative clause (example 5):

(4)
But the view that design is dead portrays the same lack of perspective for which the 1980s themselves are so famous. (Appositive clause)
(5)
Yet this was the view that all major Christian theologians insisted on – and many still do today. (Relative clause)

[BoE]

What should be obvious is that the concordancing programmes only find and organise the data. Interpretation is a human activity. We now consider what skills are needed to find patterns in concordance lines. As noted earlier, observing patterns involves identifying similarity and forming notional categories. It also involves ignoring distractors, that is, separating what might be important from what is unlikely to be so. (Note, though, that it is impossible to be precise about what is and is not patterned – what is overlooked by one observer might be noticed by another.) These points will be illustrated with a small set of concordance lines from the Bank of English obtained by searching for the lemma REACT. The result of this search gives examples of all wordforms in the lemma: *react, reacted, reacting, reacts*. Twenty random lines have been selected and then sorted so that the words following the node word are in alphabetical order.

```
1      could not believe the way Vieira reacted after he was dismissed. The
2   at all. When asked today how they'd react if the White House sent them a ne
3    step, which will enable viewers to react immediately to what they have see
4       two-thirds of the radical pairs reacting (in a field of typically only
5   any more, I don't know how he would react. Is there any point in making
6    growth because stock markets could react." Mr Visco said stock markets in
7       police officer at Selhurst Park reacted similarly to the Cantona inciden
8       mail, in New York, Adrian Clark reacted to Simon Hoggart's discussion of
9  market has come, and how people will react to it. The best seats and places
10 strength of a substance and the body reacts to fight off any diseases which
11      from the air and induce them to react to form harmless gases. Last
12           is the poster!" Herzen was reacting to a swelling trade in images. T
13    efforts you may find the magician reacting too early or late. Also bear in
14   conference was to see how he would react when asked questions by journalis
15 protect. How is management likely to react when a group threat- ens to quit?
16     twenty-year-old son felt free to react with such ferocity indicates that
17 eposition sulfur and nitrogen oxides react with atmospheric water vapor to
18 above such common tasks, refusing to react with the molecular masses. <p> Bu
19         Commentators and crowd alike reacted with astonishment when Lara
20       during the investigation. They reacted with anger and said: `The findin
```

Which of these lines might be grouped together to illustrate the "same pattern"? Looking only at what follows the node word, we might observe the following:

- a full stop (lines 5 and 6);
- the word *to* (lines 8, 9, 10, 11, 12);
- the word *when* (lines 14 and 15);
- the word *with* (lines 16, 17, 18, 19, 20).

Looking at the same evidence but in a more linguistically informed way we might express this as follows:

- REACT is followed by a subordinating conjunction (*after* in line 1, *if* in line 2 and *when* in lines 14 and 15); in each case the subordinate clause indicates an event that is the stimulus of the reaction;
- REACT is followed by a preposition (*to* in lines 8, 9 and 12, *with* in lines 16–20);
- REACT is followed by an adverb (lines 3 and 7); in each case the adverb is followed by the preposition *to*;
- REACT is followed by a to-infinitive clause (lines 10 and 11); in each case the clause indicates a reason for the reaction;
- In none of the lines is REACT followed by a noun phrase, an object. In other words, it is an intransitive verb.

A further observation might be that *react with* in lines 17 and 18 works differently from *react/reacted with* in lines 16, 19 and 20. The question that might prompt line 17 or 18 is something like "what does the object/substance (not) react with?", whereas the question prompting the other lines is something like "how did the person/people react?". Putting all this together suggests that a maximum of eight different patterns might be identified in these lines:

1. REACT followed by a subordinate clause indicating stimulus

```
1      could not believe the way Vieira reacted after he was dismissed. The
2   at all. When asked today how they'd react if the White House sent them a ne
14   conference was to see how he would react when asked questions by journalis
15 protect. How is management likely to react when a group threat- ens to quit?
```

2. REACT followed by the preposition *to*

```
8       mail, in New York, Adrian Clark reacted to Simon Hoggart's discussion of
9  market has come, and how people will react to it. The best seats and places
12           is the poster!" Herzen was reacting to a swelling trade in images. T
```

3. REACT followed by an adverb and then by the preposition *to*

```
3    step, which will enable viewers to react immediately to what they have see
7       police officer at Selhurst Park reacted similarly to the Cantona inciden
```

4. REACT followed by a to-infinitive clause indicating consequence

```
3    step, which will enable viewers to react immediately to what they have see
7       police officer at Selhurst Park reacted similarly to the Cantona inciden
```

5. REACT followed by the preposition *with* answering the question 'how?'

```
10 strength of a substance and the body reacts to fight off any diseases which
11      from the air and induce them to react to form harmless gases. Last
```

6. REACT followed by the preposition *with* answering the question 'what?'

```
16    twenty-year-old son felt free to react with such ferocity indicates that
19        Commentators and crowd alike reacted with astonishment when Lara
20      during the investigation. They reacted with anger and said: `The findin
```

7. REACT followed by a full stop

```
17 eposition sulfur and nitrogen oxides react with atmospheric water vapor to
18 above such common tasks, refusing to react with the molecular masses. <p> Bu
```

8. Other lines

```
5    any more, I don't know how he would react. Is there any point in making
6     growth because stock markets could react." Mr Visco said stock markets in
```

Some observers might wish to amalgamate some of these groups. For example, it might be argued that group 3 is simply a variant of group 2 – that the presence or absence of the adverb does not affect the pattern of "REACT + *to* + noun". It is possible, too, to join group 6 with groups 2 and 3 because in each case the prepositional phrase is obligatory. Others might argue that groups 1, 4 and 7 should be conflated because in each case REACT is the end of a clause. Still others would want to add group 5 to those because, it could be argued, in those lines the prepositional phrase beginning with *with* adds only peripheral information. Adding line 13 from group 8, this would yield only two groups:

A. REACT coming at the possible end of a clause

```
1      could not believe the way Vieira reacted after he was dismissed. The
2   at all. When asked today how they'd react if the White House sent them a ne
5   any more, I don't know how he would react. Is there any point in making
6    growth because stock markets could react." Mr Visco said stock markets in
10 strength of a substance and the body reacts to fight off any diseases which
11      from the air and induce them to react to form harmless gases. Last
13    efforts you may find the magician reacting too early or late. Also bear in
14   conference was to see how he would react when asked questions by journalis
15 protect. How is management likely to react when a group threat- ens to quit?
16     twenty-year-old son felt free to react with such ferocity indicates that
19         Commentators and crowd alike reacted with astonishment when Lara
20       during the investigation. They reacted with anger and said: `The findin
```

B. REACT followed by the preposition *to* or *with* as a necessary part of the clause

```
3    step, which will enable viewers to react immediately to what they have see
7       police officer at Selhurst Park reacted similarly to the Cantona inciden
8       mail, in New York, Adrian Clark reacted to Simon Hoggart's discussion of
9  market has come, and how people will react to it. The best seats and places
12           is the poster!" Herzen was reacting to a swelling trade in images. T
17 eposition sulfur and nitrogen oxides react with atmospheric water vapor to
18 above such common tasks, refusing to react with the molecular masses. <p> Bu
```

There are, of course, intermediate positions – it is possible to make three or four groups here as well as eight or two. The point is that no one grouping is absolutely right or wrong; all the groupings use formal information (that is, information based on the form of words) but also linguistic interpretation (distinguishing between the preposition *to* and the to-infinitive, for example, or between the two uses of *with*). A smaller number of groups tends to give a limited amount of information – the division into groups A and B, for example, tells us very little except that REACT may occur with *to* and *with* or may not. On the other hand, division in many groups runs the danger of masking genuine similarities – placing the lines with adverbs into a different group from those without tends to hide the importance of the link between REACT and *to*.

In addition, of course, quite different groups can be made if different aspects of the concordance lines are brought into account. For example, all the lines in which the subject of REACT is non-intentional (not a thinking human being or animal) can be grouped together, giving a set comprising these lines (in line 18 the subject is *Gold*):

```
4       two-thirds of the radical pairs reacting (in a field of typically only
6    growth because stock markets could react." Mr Visco said stock markets in
10 strength of a substance and the body reacts to fight off any diseases which
11      from the air and induce them to react to form harmless gases. Last
18 above such common tasks, refusing to react with the molecular masses. <p> Bu
```

It is noticeable that this set includes all the lines where REACT is followed by a to-infinitive clause and the only line where it is followed by *with* giving essential information.

To summarise: Observing patterns in concordance lines essentially involves grouping those lines together. In most examples, several alternative groupings could be proposed, each highlighting different kinds of information. There is no objectively correct

grouping, although some will be more useful for particular purposes than others. Although the presence of individual words may provide help in grouping, usually a wider context and more interpretation are needed to form groups (that is, to identify patterns) that might be thought to be appropriate.

3 Grammar patterns

In this section of the chapter the discussion will be restricted to a particular kind of pattern, described by Francis (1993) and Hunston and Francis (2000) as grammar patterns. (Note that each individual pattern is referred to as a grammar pattern, while the concept is referred to as pattern grammar.) This is a technical use of the word "pattern" and describes a specific phenomenon: the restrictions on what is found to occur as the complementation of verbs, nouns and adjectives. It is sometimes referred to as an extension of Firth's (1968) notion of "colligation", or co-occurrence of grammatical categories. The concept of pattern grammar ignores lexical collocation and the kind of patterning that is the product of what speakers often say and instead describes only the co-occurrence of specific words with classes of word, phrase and clause. In an example such as *they have to decide how best to collect waste*, the repetition in the corpus of "obligation + *decide how to*" is ignored but the co-occurrence of the lemma DECIDE and a clause beginning with a question word is given the nomenclature **V wh** (verb followed by a clause beginning with a *wh-* word). In the examples where REACT is followed by a clause (*how they'd react if the White House sent them…*), the sequence "REACT + if-clause" is not treated as a pattern, whereas the sequence "REACT + *to* + noun phrase" (*he reacted to Simon Hoggart's discussion…*) is coded as the pattern **V *to* n** (verb followed by a prepositional phrase beginning with *to*). The two examples of *view* followed by the word *that* are treated differently (see examples 4 and 5 above). Where *that* introduces a relative clause, it is ignored, but where *that* introduces an appositive clause, the example is coded as the pattern **N that** (noun followed by that clause). The distinction being made here is between patterns that are used with a whole word class and those that constrain the specific words that they are used with. For example, a modal of obligation (as in *have to decide*) is used with any member of the word class "verb", but a *wh-* clause (as in *decide how best to collect waste*) complements only certain verbs. A relative clause is used with any member of the word class "noun", but appositive *that* clauses are used with only certain nouns (*announcement, claim, hint, intuition, observation, suggestion, theory, view,* etc). There is an important notion of dependency here: The *wh-* clause is dependent on the verb DECIDE and the appositive that-clause is dependent on the noun *view*. The if-clause following *react* (*how they'd react if…*), however, is not dependent on the verb REACT. Only the dependent patterns are termed grammar patterns. The number of such patterns is large but is limited by the fact that in English there are only a certain number of phrase and clause types (noun phrase, verb phrase, that-clause, to-infinitive clause, *wh-* clause, etc.) and only a certain number of prepositions (*at, in, on, with,* etc). Altogether there are about 200 grammar patterns: about 100 relating to verbs and 50 each for nouns and adjectives.

Although the range of patterns is finite, this does not mean that which word occurs in which pattern is also fixed. There is variation between varieties of English. In British

English, for example, the verb MISTAKE is used in the pattern **V n *for* n** (see example 6), whereas in other varieties it is also used in the pattern **V n *as* n** (see example 7).

(6)
…it might be possible to mistake these species for large gastropods.
[BNC]

(7)
…he might mistake a friendly comment as an invitation…
[ANC]

Words also change the patterns they occur with. Well-known examples in British English might be the verb IMPACT, which used to be used only with *on*, as in example 8, but which is now used also in the pattern **V n**, as in example 9.

(8)
The way we undertake our Census may impact on our ability to understand our future society.
[English Web 2013 sample]

(9)
This evolution ultimately will impact society in an incredibly positive way…
[English Web 2013 sample]

In spite of this variation, it might be said that most patterns remain consistent over long periods of time, making it feasible to specify both patterns and the words used with them. One attempt to do this is described in Hunston and Francis (2000); the results can be found at https://grammar.collinsdictionary.com/grammar-pattern. (This web resource incorporates an updated version of the information found in Francis *et al.* 1996, 1998). In this work, all the words occurring with a given pattern are divided into "meaning groups". For example, the entry for the pattern **ADJ that** lists 115 adjectives divided into 12 groups. Examples include *surprised that…, angry that…, glad that…, certain that…, insistent that…, lucky that…*. In the next section we consider how this information might be relevant to English language teaching and learning.

4 Grammar patterns and language education

Willis (2003: 28–47) discusses grammar patterns in terms of 'the grammar of class', which, along with 'the grammar of structure' and 'the grammar of orientation', comprises the grammar learners need to know. For example, in example 10, knowing the grammar of structure means knowing that the subject *I* precedes the verb *had forgotten* and that the indirect question is formed by the sequence "question phrase + subject + verb + object" (*how much I love this pasta*). Knowing the grammar of orientation means being able to use the past perfect (*had forgotten*) to locate the forgetting before the act of making the pasta. Knowing the grammar of class means being able to use *forget* with a *wh-* clause (*had forgotten how much*), that is, knowing the grammar patterns that are used with the verb FORGET.

(10)
I had forgotten just how much I love this pasta until I made it earlier this week.
[English Web 2013 sample]

It is apparent that some of the information covered under the topic of pattern grammar is routinely covered in any English language teaching syllabus. The use of that-clauses or to-infinitive clauses with some verbs, for example, is a familiar topic in grammar. The term "sentence pattern" is often used to distinguish SVO clauses, SVOO clauses and others, with passing mention that the pattern(s) available will depend on the choice of main verb in the clause. (The pattern grammar equivalents of these observations are the **V that** pattern, the **V to-inf** pattern, the **V n** pattern and the **V n n** pattern.) The value of a systematic listing of lexically dependent patterns, and the lexis on which they are dependent, however, is still being explored. The most obvious learner group targeted by such exploration is learners of English as a foreign language (Hunston 2002; see also Chapters 22 and 23, this volume), but there are also developments in teaching English in contexts where it is a first or additional language (McSorley and Patten 2019). Here I will mention just four aspects of this exploration: awareness, accuracy, complexity and academic English.

It has long been acknowledged that a language teaching syllabus is unlikely to contain a sequenced list of patterns in the way that many contain a sequenced list of tenses, for example. Like collocation, a confidence in using grammar patterns is likely to be fostered by raising **awareness** of patterns. Most recommended activities, then, encourage this awareness. The activities can accompany a text comprehension exercise, can involve activities such as rewriting or can emerge as one of the outcomes of hands-on corpus investigation by students (see Chapters 21, 29 and 30, this volume). But why should learners be made aware of grammar patterns? The next two aspects answer this question.

Perhaps the most obvious relevance of grammar patterns for learners is that their correct use (that is, the use that aligns with that of expert speakers of English) increases **accuracy** and reduces error. Learners sometimes over-generalise patterns, for example, using SUGGEST in the same way as RECOMMEND (e.g. "he suggested me to go") or using the verb EMPHASISE with *on* ("they emphasised on the importance…") by analogy with the noun EMPHASIS ("they place emphasis on the importance…"). Learners may sometimes simply be uncertain which preposition to use with a given noun or adjective, or may find a set of patterns difficult to manage. Examples such as "Students are easy to use on-line dictionaries" suggest a confusion between "On-line dictionaries are easy to use" and "Students find it easy to use on-line dictionaries". It should be noted, however, that pattern alone does not guarantee accuracy. Alqarni (2019) has investigated patterns in learner English and notes that examples such as that shown in example 11 are inaccurate in terms of standard English.

(11)
I agree to ban smoking in restaurants.
[ICLE-Chinese]

In this example, the pattern **V to-inf** is correct, but the meaning is not: In Hanks's (2013) terms, the implicature of *agree to ban smoking* is that the writer themselves will execute

the ban, whereas the more likely intended implicature is that someone else, such as the government, will ban smoking and the writer simply agrees with this proposal. Alqarni (2019) draws on Hanks's extensive work on verb patterns to explain the inaccuracy in usage here.

Awareness of pattern does more than increase accuracy, however. It also expands language repertoire and increases the **complexity** of language use. Many grammar patterns extend a basic clause to link ideas. For example, example 12 uses the pattern ***it* v-link ADJ *of* n to-inf** to connect the idea "it was good of you" with the idea "you let me know".

(12)
It was good of you to let me know.

[English Web 2013 sample]

The patterns that involve prepositional phrases require an idea to be expressed as a noun phrase and can require some sophisticated manipulation of language, requiring learners to be flexible and raising awareness of what can be expressed in a noun. For example, the pattern **v-link ADJ *about* n** links adjectives that express emotion (*happy, angry, sad,* etc.) with the cause of the emotion. That cause may be a simple entity or it may be an event. Rewriting exercises can be used to practise the pattern. Table 11.1 gives some examples. The prompt column shows what is given as a prompt to learners. The response column shows what learners are expected to write. In example 1 the task is easy because the noun phrase *the terrible mess* is present in the prompt itself. Example 2 is slightly more difficult because the verb *meet* has to be changed to *meeting* in order to be used following a preposition. In example 3, the verb *were sold* is changed to the noun *sale* when the response is given. Similarly in example 4, the clause *the work was delayed* is changed to the noun phrase *a delay in the work*. In example 5 the rewriting makes use of *the fact that* to produce the noun phrase.

A question that often arises in discussions about grammar patterns is the relevance of this work to **academic English**, both in the context of teaching English for academic purposes in higher education institutions and in the context of increasing command of academic vocabulary among learners in secondary schools, whether they are monolingual or bilingual speakers of English. Grammar patterns are lexis specific rather than register

Table 11.1 Rewriting with the pattern **v-link ADJ *about* n**

Example	*Prompt*	*Response*
1	There was a terrible mess. Robin was cheerful.	Robin was cheerful about the terrible mess.
2	I wanted to meet some friends. Ann was not enthusiastic.	Ann was not enthusiastic about meeting friends.
3	Some paintings were sold. Jenny was unhappy.	Jenny was unhappy about the sale of the paintings.
4	The work was delayed. The residents were furious.	The residents were furious about a delay in the work.
5	I wanted to leave. Dave was bitter.	Dave was bitter about the fact that I wanted to leave.

specific, so there are no patterns which are exclusive to academic English. On the other hand, as the patterns associated with given vocabulary items are an important part of how those items are used, there is increasing interest in recording the patterns of key academic vocabulary. Green (2019), for example, argues for the enrichment of vocabulary lists with pattern information and has built a database of academic vocabulary patterns to support this. McSorley and Patten (2019) have worked with secondary school teachers to change the way academic vocabulary is taught to include patterns as part of that instruction (see Chapters 24 and 28, this volume, for more on vocabulary lists).

5 Grammar patterns and language theory

The concept of patterns raises questions about language theory, that is, what account of language structure best fits observed use of language, on the one hand, and empirical evidence for how language is stored in the brain, on the other. The patterning of language is a topic dealt with under a wide variety of headings – "formulaic language", "chunks", "lexical bundles", etc. – all of which make the following observations (see Chapters 15 and 17, this volume):

- Language production depends on stocks of ready-made language. From Pawley and Syder's (1983) seminal article onwards, the contribution of formulaic phrases to the ability to speak at normal speed has been noted. This may be a trivial observation that has nothing to do with the "real" structure of language and that is a consequence only of real-time memory capacity and the tendency of speakers to repeat what they have heard. It may, on the other hand, be a profound observation about the nature of language.
- Language processing (comprehension) depends on "chunking" i.e. deciding what unit of language makes the most sense. The unit may be a word ("apple") or a phrase ("apple of one"s eye") or an instantiated pattern ("divide the apple in two"). Sinclair (1991) noted that the default position is to interpret the phrase, or unit of meaning, rather than the individual words. For example, the hearer understands "apple of one"s eye" as a single unit, not as a combination of the name of a fruit and the organ of sight. Research by Mauranen (2019) seeks to establish empirically whether the chunks used by language processors (individuals hearing a spoken utterance) match intuitive sequence chunking.
- Lexis and grammar are closely interlinked. Descriptions of language need to acknowledge the dependency of one on the other. At the least, dictionaries need to record grammatical information about the words in them, and grammar reference books need to record information about the lexical restrictions on the features described.

One of the questions raised by these observations is how the apparently systematic structures of grammar arise from the patterning of individual words. Proponents of emergent grammar (e.g. Hopper 2011) see grammatical systems as constantly in flux and dependent upon an ongoing amassing of language instances. This is compatible with the proposal that pattern recognition arises from the experience of individual instances.

Another question is how language is stored in the brain, or what the learnt units of language look like. One answer is the proposal of constructions, a concept arising from cognitive linguistics (Goldberg 2006; Hoffmann and Trousdale 2013; Hilpert 2014; Patten and Perek 2019). A construction is a unit of any size which matches form and function or meaning. A construction can be a word (*apple*), a phrase (*apple of one's eye*), a structure such as the ditransitive (*give someone an apple*) or an abstraction such as subject–verb inversion that has the meaning of "interrogative" (*did you eat the apple*). Constructions, properly speaking, exist in the brain, but the term is often used for regularities observed in corpora, whether or not there is empirical evidence that speakers have them stored (see Chapters 22 and 23, this volume).

Construction grammar is appealing to corpus linguists because it is entirely compatible with, and indeed utilises, observations of patterning in language. It offers a useful vocabulary that can cut through the tangle of terminology ("formulaic language", "chunk", "pattern", etc.). Any kind of regularity in language can be labelled a construction. This is both a strength and a weakness: "Construction" is a neutral term for "observed regularity" that does not draw on any one model of grammar, but because anything can be described as a construction, the term tends to lose meaning. Because there are no unit types, there is no hierarchy of unit types. This again is both a strength and a weakness: It allows the description to move away from the restrictions of group, phrase and clause, but it means that observations of different levels of abstraction are treated alike. There is a hierarchy of constructions, but the levels in the hierarchy are not clearly described. This is where pattern grammar can be of assistance, because there are obvious similarities between one level of construction, exemplified by the ditransitive, and its "pattern" counterpart **V n n**. Hunston and Su (2019) and Hunston (2019) suggest that the most useful alignment of pattern and construction would label as a "construction" not the pattern itself, but the pattern along with a sub-set of the words occurring with it. For example, one construction might be proposed that consists of the pattern **ADJ that** where the adjective is an emotion word (*angry that…* etc.), another that consists of the same pattern but this time with adjectives associated with speaking (*insistent that…* etc.) and another where the adjective relates to a judgement of luck or correctness (*lucky that…, right that…* etc.). This offers the possibility of building a comprehensive "constructicon" (Perek and Patten 2019) comprising all constructions in English at a given level of abstraction based on the records of grammar patterns and their node words. This would represent a concrete example of cooperation between corpus linguistics and cognitive linguistics.

Further reading

Hunston, S. and Perek, F. (eds) (2019) *Constructions in Applied Linguistics*, special issue of the *International Journal of Corpus Linguistics* 24(3). (This journal issue explores the applicability of patterns to discourse analysis and language teaching, the development of resources such as constructicon and the connections between patterns and construction grammar. It comprises papers by U. Römer, S. Gries, N. Groom, S. Hunston and F. Perek and A. Patten.)

Hunston, S. and Francis, G. (2000) *Pattern Grammar: A Corpus-Driven Approach to the Lexical Grammar of English*, Amsterdam: Benjamins. (This book outlines the connections between Grammar Patterns, meaning, discourse, and approaches to grammar.)

McSorley, E. and Patten, A. (2019) 'Addressing the Vocabulary Gap Using the Pattern Grammar Approach', *Impact* 6xx. (This paper discusses how attention to patterns can improve the teaching of academic vocabulary.)

References

Alqarni, A. (2019) *Applying Corpus Pattern Analysis to Learner Corpora: Investigating the Pedagogical Potential of the Pattern Dictionary of English Verbs*, unpublished PhD thesis, University of Birmingham.

Firth, J. (1968) 'A Synopsis of Linguistic Theory, 1930–55', in F. R. Palmer (ed.) *Selected Papers of J. R. Firth (1952–59)*, London: Longmans, pp. 168–205.

Francis, G. (1993) 'A Corpus-Driven Approach to Grammar: Principles, Methods and Examples', in M. Baker, G. Francis and E. Tognini-Bonelli (eds) *Text and Technology: In Honour of John Sinclair*, Amsterdam: Benjamins, pp. 137–56.

Francis, G., Hunston, S. and Manning, E. (1996) *Collins Cobuild Grammar Patterns 1: Verbs*, London: HarperCollins.

Francis, G., Hunston, S. and Manning, E. (1998) *Collins Cobuild Grammar Patterns 2: Nouns and Adjectives*, London: HarperCollins.

Goldberg, A. (2006) *Constructions at Work: The Nature of Generalization in Language*, New York: Oxford University Press.

Green, C. (2019) 'Enriching the Academic Wordlist and Secondary Vocabulary Lists with Lexicogrammar: Toward a Pattern Grammar of Academic Vocabulary', *System* 87: 102158.

Hanks, P. (2013) *Lexical Analysis: Norms and Exploitations*, Cambridge: MIT Press.

Hilpert, M. (2014) *Construction Grammar and its Application to English*, Edinburgh: Edinburgh University Press.

Hoffmann, T. and Trousdale, G. (eds) (2013) *Construction Grammar Handbook*, Oxford: Oxford University Press.

Hopper, P. (2011) 'Emergent Grammar', in J. P. Gee and M. Handford (eds) *The Routledge Handbook of Discourse Analysis*, London: Routledge.

Hunston, S. (2002) *Corpora in Applied Linguistics*, Cambridge: Cambridge University Press.

Hunston, S. (2008) 'Starting with the Small Words: Patterns, Lexis and Semantic Sequences', *International Journal of Corpus Linguistics* 13(3): 271–95.

Hunston, S. (2019) 'Patterns, Constructions, and Applied Linguistics', *International Journal of Corpus Linguistics* 24(3): 324–53.

Hunston, S. and Francis, G. (2000) *Pattern Grammar: A Corpus-driven Approach to the Lexical Grammar of English*, Amsterdam: Benjamins.

Hunston, S. and Su, H. (2019) 'Pattern, Construction and Local Grammar: The Case of Evaluation', *Applied Linguistics* 40(4): 567–93.

Laws, J. and Ryder, C. (2018) 'Register Variation in Spoken British English: The Case of Verb-forming Suffixation', *International Journal of Corpus Linguistics* 23(1): 1–27.

Lehmann, H. (2018) 'Lexical Preference and Variation in the Complementation of *Provide*', *International Journal of Corpus Linguistics* 23(1): 55–84.

Lei, L. and Liu, D. (2018) 'The Academic English Collocation List: A Corpus-Driven Study', *International Journal of Corpus Linguistics* 23(2): 216–43.

Mauranen, A. (2019) 'Repeated Sequences: What Can Corpora Tell us about Units of Processing?' Paper given at the CAAD conference, Universitat Jaume 1, November 2019.

McSorley, E. and Patten, A. (2019) 'Addressing the Vocabulary Gap Using the Pattern Grammar Approach', *Impact* 6. https://impact.chartered.college/article/addressing-the-vocabulary-gap-pattern-grammar-approach/.

Patten, A. and Perek, F. (2019) 'Pedagogic Applications of the English Constructicon', in H. Boas (ed.) *Pedagogic Construction Grammar: Data, Methods and Applications*. Berlin: de Gruyter.

Pawley, A. and Syder, F. (1983) 'Two Puzzles for Linguistic Theory: Nativelike Selection and Nativelike Fluency', in J. Richards and R. Schmidt (eds) *Language and Communication*, London: Longman, pp. 191–226.

Perek, F. and Patten, A. (2019) 'Towards an English Constructicon Using Patterns and Frames', *International Journal of Corpus Linguistics* 24(3): 354–84.
Sinclair, J. (1991) *Corpus Concordance Collocation*, Oxford: Oxford University Press.
Sinclair, J. (1990/2004) 'Trust the Text', in J. Sinclair *Trust the Text: Language, Corpus and Discourse*, London: Routledge, pp. 9–23.
Willis, D. (2003) *Rules, Patterns and Words: Grammar and Lexis in English Language Teaching*, Cambridge: Cambridge University Press.