4 Productivity in Word-Formation

4.1 THE OPEN-ENDEDNESS OF THE LEXICON

One of the goals of morphological theorising is to account for the ways in which speakers both understand and form, not only 'real' words that occur in their language, but also potential words that are not instantiated in use in utterances. While it is true that a large percentage of 'real' words listed in dictionaries (such as *pear* and *pair*) are, and in fact must be, memorised, it is equally true, and of great theoretical interest, that countless words used in conversation (and to a lesser extent in writing) are new, made up on the spur of the moment. So, morphology has to throw light not only on the structure of established words like *pair,* but also on that of freshly coined **neologisms** like *spam* (unsolicited electronic mail).

The consensus appears to be that the words of a language are listable in a way in which sentences are not (see Section 13.3.2). The meanings of many words (e.g., *pear* and *pair*) must be listed in the lexicon because there is nothing about their sounds or morphological structure that would enable one to work out their meaning. In this respect morphology differs from syntax. Syntax cannot be restricted to cataloguing only those sentences that occur in some **corpus** (i.e., a body of texts), since language is vast and no list of sentences, no matter how long, could exhaust the set of possible well-formed sentences. Typically, speakers do not merely recycle sentences memorised from previous conversations. Rather, they tend to construct fresh sentences to suit the occasion.

However, by and large, people do not routinely make up new words each time they speak. Nonetheless, the lexicon cannot be seen as a *static* list. No dictionary, however large, (not even the complete *Oxford English Dictionary*, including all its supplements) can list every word in the English language. Why is this so?

Until recently, word-formation rules have tended to be seen as being largely passive in the sense that they are basically used to analyse existing words rather than to create new ones. It is significant in this connection that, whereas reasonably comprehensive dictionaries and wordlists for dozens of languages exist, there are no equivalent, all-encompassing sentence lists for any language. Lists of sentences such as those found in phrase-books for foreigners make no pretence of being exhaustive.

The verdict on whether or not morphology and the lexicon deal with what is effectively a closed list of words will hinge, to some extent, on our attitude

to nonce-words (like *uncomplicatedness*), created by an individual, which do not catch on in the speech community. The more such words are recognised as part of the language, the bigger and more open-ended will be our lexicon. But, perhaps, it will depend to an even greater extent on which forms we recognise as compound words and, hence, part of the province of morphology and the lexicon, and which forms we treat as phrases and, hence, the domain of syntax. Since both compounds and phrases are made up of words, determining which combinations of words are compounds and which combinations are syntactic phrases is not always straightforward.

For instance, consider *Lakeside Grammar School Former Pupils*. We have here the compound nouns *lakeside* and *grammar school*. Is the larger unit *Lakeside Grammar School* also a compound noun? It is even more unclear whether the whole string *Lakeside Grammar School Former Pupils* is also a compound noun. At first blush, we might decide that it is not a compound noun but a syntactic phrase. But we might change our minds if we discovered that *Lakeside Grammar School Former Pupils* is the name of a rugby team that plays in the local league. If we recognise strings like *Lakeside Grammar School Former Pupils* as compounds, we will have a very open-ended lexicon. (See Chapter 13 for further discussion.)

Recall also that in the last chapter we mentioned another aspect of productivity in word-formation. Although no word is infinite in length, in principle there is no upper limit to the length of forms that may function as bases from which new words are formed. In many languages, if not all, we cannot state categorically what the longest possible word is: words can be made longer, if the need arises, by the repetition of an affix or the addition of yet another affix as we saw in [3.24].

Finally, it is also possible to add to the lexicon of a language indefinitely by pillaging the vocabulary of other languages. This is called **borrowing**. English has borrowed from a great number of other languages with which it has been in contact. There are Latin **loanwords** like *port* (from *portus* 'harbour'); French loanwords like *omelette*; Italian loanwords like *fresco,* and so on. In this book we will not deal with the expansion of the vocabulary by borrowing, for which see, for instance, Katamba (2005). We will only be concerned with productive word-creation using the internal resources of a language (see Section 4.2.1.2).

The upshot of this discussion is that morphology is productive. There is no limit to the number of potential words in a language. Therefore, a lexicon that merely attempted to list the words of a language in some corpus would be woefully inadequate.

4.1.1 What is Productivity?

But what exactly is productivity? We will provisionally view productivity simply in terms of **generality**. The more general a word-formation process is,

the more productive it will be assumed to be. There are two key points requiring elucidation:

(i) Productivity is a matter of degree. It is *not* a dichotomy, with some word-formation processes being productive and others being unproductive. Probably no process is so general that it affects, without exception, *all* the bases to which it could potentially apply. The reality is that some processes are *relatively* more general than others.
(ii) Productivity is subject to the dimension of time. A process which is very general during one historical period may become less so at a subsequent period. Conversely, a new process entering a language may initially affect a tiny fraction of eligible inputs before eventually applying more widely.

Exercise
Study the following data:

[4.1]

chartist	morbid	worker
communist	tepid	painter
racist	timid	swimmer
pianist	splendid	dancer
anarchist	horrid	jogger

(i) Divide the above words into their constituent morphs.
(ii) List all the suffixes.
(iii) State the meaning of the morphemes represented in the data.
(iv) Find five more words that are formed using each of the suffixes that you have identified.
(v) State the word-class of the base to which each suffix is added.
(vi) To what class does the resulting new word belong?
(vii) Was it equally easy to find more words which contain the different suffixes? If not, comment on any problems that you encountered.

The words in [4.1] contain the roots listed in [4.2]:

[4.2]

Noun	Adj/Verb	Verb
chart	morb-	work
commun(e)	tep-	paint
rac(e)	tim-	swim
pian(o)	splend-	danc(e)
anarch(y)	horr-	jog

The status of the bound roots in the second column is supported by additional examples such as *timorous*, *splendour*, *horrify*, etc. There are three suffixes present: *-ist, -id* and *-er*. All of them are derivational.

The suffix *-ist* may be added to noun bases to form other nouns (typically they are noun bases which can also take the derivational suffix *-ism,* such as *chart-ism, communism*, etc.). It can also be added to noun bases to form adjectives (e.g., $race_{N} \rightarrow racist_{Adj}$). In view of this we can distinguish at least three meanings which are realised by *-ist*:

(i) N → N *-ist*: meaning 'advocate of' (as in *anarchist, communist,* etc.);
(ii) N → N *-ist*: meaning 'practitioner of' (as in *pianist, violinist,* etc.);
(iii) N → Adj *-ist*: meaning 'advocating X' (as in *racist, sexist,* etc.).

The suffix *-id* is also of Latin origin and entered English via French. It is added to bound adjectival bases with the meaning of 'having the quality specified by the verb'. In Latin *-id* was used to derive **attributive** adjectives from verbs, for example, *timidus* from *timere* 'to fear' gives us English *timid* (i.e., attributing to an individual the quality of being afraid) and *tepidus* from *tepere* 'to be warm' gives us English *tepid* (i.e., attributing the quality of being warm).

Finally, the native Germanic suffix *-er* is suffixed to verbs to create **agentive** nouns (with the meaning 'someone who does whatever is designated by the verb').

One would expect you to have had some difficulty in finding five more adjectives which contain the derivational suffix *-id*. The morpheme *-id* is at the unproductive end of English morphology. It is frozen. It is no longer used actively (if it ever was) to produce new words. The words containing it, such as *frigid*, *lurid*, *turbid*, *turgid*, etc., could simply be listed in the lexicon.

At the other extreme, agentive nouns containing the suffix *-er,* are numerous and can be added to indefinitely. Most verbs can have a noun formed from them in this way. So, no reasonable case could be made for listing all agentive nouns ending in *-er* in the dictionary. Rather, what is needed is a rule stating that, by suffixing *-er,* an agentive noun can be derived from virtually any verb.

General derivational processes that apply more or less across the board (such as the formation of agentive nouns by using the *-er* suffix) and historical relics (like the formation of attributive adjectives using *-id*) excite little theoretical interest in discussions of productivity. What is fascinating is the grey area in between occupied by morphemes like *-ist*.

By suffixing *-ist* we can form a very large number of nouns with the meaning 'advocate of, follower of, supporter of or practitioner of whatever is designated by the input noun'. But we do not have a *carte blanche* to use it with absolutely any noun. There are unexplainable gaps. For example, a follower of the prophet Mohammed is not a **Mohammedist* though a follower of Buddha is a *Buddhist* and an adherent to Calvin's approach to Christianity is a *Calvinist*. And, note also that a *piano* is played by a *pianist,* a *guitar* by a *guitarist* but the *drums* are played by a *drummer,* not a **drummist*.

The innocent-looking question, 'How productive is this particular process?' turns out to be very troublesome (cf. Bauer, 2001). This is because, as we mentioned at the beginning, the term productivity suffers from an inherent ambiguity. On the one hand, a process is said to be productive if it is very general, that is, affects a vast number of forms and creates very many words. In this sense, the agentive morpheme *-er* (as in *worker*, *writer*, etc.) is very productive, since an overwhelming majority of verbs can be turned into nouns by this suffix. It certainly is more productive than the semantically related suffix *-ent* found in *president*, *student*, *correspondent*, *proponent*, etc. There are thousands of bases to which *-er* can be added as compared to the dozens which take *-ent*. If, on the other hand, we forget about the total number of words created using a given process and instead focus on the proportion of bases that are eligible to undergo a process compared to those which actually *do* undergo it, the results may be somewhat different. The chances of a particular affix appearing may crucially depend on characteristics of the base to which it is to be attached. Thus, while it is true that *-ness* (as in *goodness*) is more common than *-ity* (as in *gravity, banality*) in the English language as a whole, in the case of an adjective ending in *-ile* (e.g., *servile*, *docile*, *fertile*, *futile*, etc.) *-ity* is the preferred suffix (Aronoff, 1976). This is an instance of the phonological properties of the base influencing the likelihood of a morphological process taking place (see Section 4.2.1.1 below).

Exercise
Study the following:

[4.3]

appendicitis	bronchitis	vaginitis
tympanitis	hepatitis	meningitis
cephalitis	pneumonitis	tonsillitis
arthritis	dermatitis	neuritis
sclerotitis	pleuritis	bursitis

(i) What is the meaning of *-itis?*
(ii) Is the *-itis* suffix comparable to *-er, -ist* or *-id* in productivity?

The suffix *-itis* is borrowed from Greek, where it formed the feminine of adjectives. Already in Greek, it was used to form words referring to inflammatory diseases like *arthritis*. It is used in modern medical English to form names of diseases, especially inflammatory ones.

While it is true that the words formed by suffixing *-itis* are fewer than those formed by suffixing *-er,* nevertheless the suffix *-itis* attaches with an extremely high degree of regularity to most suitable bases. (And it can be

generalised beyond the semantic niche of inflammatory diseases. In jocular parlance it is extended to even psychological ailments like *skiveritis* and *Monday-morningitis*.)

Furthermore, as we noted above, in the process of refining our understanding of productivity we must consider the time dimension. Let us assume, to begin with, that a word-formation process is productive if it is in current use. Frozen or atrophied processes, like the suffixation of *-id* in *tepid, frigid*, etc., may be regarded, for practical purposes, as virtually unproductive in contemporary English. In contrast, the suffixation of the agentive *-er* suffix as in *worker,* which is attached freely and unfussily to most eligible forms, is said to be very productive. In between these two extremes there lies a vast grey zone in which there lurks a milliard of possible variations. This grey zone will be investigated in much of the rest of this chapter and in Chapter 5.

Let us begin by observing that productivity is affected by fashion. For a time, one method of forming words may be in vogue, but subsequently it may become less fashionable, or be abandoned completely. On first entering the language, an affix may affect only a fraction of bases to which it will eventually be attachable. For instance, the form *loadsa-* [ləʊdzə] (colloquial for *loads of*) was used in informal British English in the late 1980s in the word *loadsamoney,* a noun referring to the *nouveau riche* with conspicuously unrefined manners and tastes. In 1988, it began being used as a prefix in a few newly created words such as *loadsasermons, loadsaglasnost,* etc., that appeared in London newspapers (Spiegl, 1987). Even if this prefix does survive, and even if it manages to spread to many other words, it remains true that it has only affected a handful of bases to which it could be potentially attached.

Conversely, an affix that historically was used widely may atrophy or cease being applied to new forms altogether.

Exercise

Study the forms of the verb *take* in Early Modern English prior to 1800 and determine which inflectional endings are no longer productive:

[4.4]

Singular	Plural
I take	we take
thou takest	ou take
he, she taketh	they take

The second- and third-person singular present tense form of the verb were realised by the regular suffixes *-est* and *-eth* respectively, but these suffixes have dropped out of common use. They survive as relics in antiquated religious language and on stage when a pre-1800 play is performed.

4.1.2 Semi-Productivity

Some linguists, like Matthews (1974), recognise a special category they call **semi-productivity** to cover idiosyncratic affixes which inexplicably fail to attach to apparently eligible forms. (Note the seeming incoherence of such a term if we accept that productivity is gradient.) Furthermore, where such affixes are used, the meaning of the resulting word may be unpredictable.

Exercise
Study the following data and show that the suffix *-ant* is capricious in these respects:

(i) in the selection of bases to which it attaches,
(ii) in the meaning of words which result from suffixing it.

[4.5]

a.	communicant	defendant	applicant
	assailant	servant	supplicant
	entrant	contestant	participant
	claimant	accountant	assistant
	dependant	inhabitant	consultant
b.	*writ(e)ant	*buildant	*shoutant

The suffix *-ant* turns a verbal base into an agentive nominal. (It is similar in meaning to *-er*.) But it is very fussy. It accepts the bases in [4.5a] but not those in [4.5b]. The reasons for the particular restrictions on the bases to which *-ant* may be suffixed are, at least in part, historical. This suffix is descended from the Latin present participle ending *-antem*/*-entem*. Hence, it attaches to Latinate bases only. Germanic bases like *write, build* and *shout* are ineligible.

Even then, attachment to latinate bases is unpredictable. For no apparent reason many bases of Latin (or French) descent fail to combine with *-ant*:

[4.6]

destroy	*destroyant (Old French *destruire*; Modern French *détruire*)
adapt	*adaptant (Old French *adapter*)

Semantically *-ant* has unpredictable effects. The meaning of words created by suffixing *-ant* is inconsistent. For instance, a *defendant* has the narrow interpretation of a person sued in a law court, not just any one who defends oneself; an *accountant* is not merely anyone who renders an account or calculation, but a professional who makes up business accounts, and so on.

Unlike Matthews, we shall not give theoretical recognition to the concept of semi-productivity in this book because in practice it would be

difficult to know which word-formation processes are properly classed as semi-productive. There is not a neat three-way opposition, between productive, semi-productive and unproductive processes. Rather, there are different shades of grey, with some processes being relatively more productive than others.

4.1.3 Productivity and Creativity

The term productivity has sometimes been used to refer to **creativity**, i.e., the capacity of all human languages to use finite means to produce an infinite number of words and utterances (see Section 1.3). In the domain of morphology, creativity manifests itself in two distinct ways: rule-governed creativity and rule-bending creativity.

For the most part words are formed following general rules and principles internalised by speakers in the process of language acquisition which most of this book is devoted to exploring. For instance, if the suffix *-ly* is added to an adjective (e.g., *quick*), an adverb (*quickly*) is produced; if the prefix *post-* is attached to a noun base (as in *post-war*), an adjective with meaning 'after' is formed, and so on.

However, speakers have the ability to extend the stock of words idiomatically by producing words without meticulously following the standard rules of word-formation. This can be seen in the way in which certain compounds are constructed:

[4.7] a. stool pigeon (police informer)
b. redlegs (poor whites in Tobago)
deadline

No synchronic rules can be devised to account for the meaning of a semantically unpredictable compound like *stool pigeon*. But, in some cases, delving into history might show that some of these compounds originally had a literal meaning that was superseded by later metaphorical extensions. To take one example, during the American Civil War, a *deadline* was the line round the perimeter fence beyond which soldiers were not allowed to go. A soldier who wandered beyond that line risked being shot dead for desertion. (Thankfully, today, going beyond *a deadline* is unlikely to be fatal.) As for *redlegs,* it may be true that poor whites working in the hot sun as labourers on plantations in Tobago did literally have *legs* that were *red*; nevertheless, the compound *redlegs* is semantically opaque. It is very unlikely that anyone could work out the meaning of *redlegs* from the meaning of the words *red* and *leg*. Comparable examples in present-day English are not difficult to find. Consider words like *walkman* and *tallboy*. The former is not a kind of man but miniature personal stereo equipment and the latter is not a boy but a piece of furniture.

Our primary concern in this book will be synchronic rule-governed word-formation. One of the perplexing problems we will deal with is the fact that a word-formation process rarely applies consistently across the board to all the forms which, on the face of it, qualify for the application of a particular rule.

4.2 CONSTRAINTS ON PRODUCTIVITY

Although there is no limit to the number of words that can be produced in a language, not every conceivable word that could be formed is allowed. In this section we examine the factors that limit productivity.

4.2.1 Blocking

First, we outline in general terms some of the factors which frustrate the application of a word-formation process whose conditions of application appear to be met. The cover term **blocking** is used for these factors.

Blocking may be due to the prior existence of another word with the meaning that the putative word would have (Aronoff, 1976). Usually perfect synonyms are avoided. Thus, it may be because *thief* already exists that suffixing the otherwise very productive agentive suffix *-er* to the verb *steal* in order to form **stealer* is blocked. See Section 6.2.4.2 for further discussion.

Interestingly, where there exist two semantically similar morphemes, one of which is more productive than the other, the more productive morpheme is less susceptible to blocking than its less productive counterpart. This can be seen in the behaviour of the suffixes *-ity* and *-ness*. Aronoff (1976) has shown that the suffixation of *-ness* is more productive than the suffixation of *-ity*. He goes on to point out (cf. Aronoff, 1976) that where there is an existing noun derived from an adjective base ending in *-ous,* it is not possible to create a new noun by adding *-ity*. However, the existence of an established noun does not stop the derivation of a fresh noun using the more productive suffix *-ness*:

[4.8]

X + ous (Adjective)	Pre-existing Noun	Noun (-ity)	Noun (-ness)
acrimonious	acrimony	*acimoniosity	acrimoniousness
glorious	glory	*gloriosity	gloriousness
fallacious	fallacy	*fallacity	fallaciousness
spacious	space	*spaciosity	spaciousness
furious	fury	*furiosity	furiousness

The concept of blocking can be further refined by highlighting a number of factors that play a role in it. These factors may be phonological, morphological, or semantic.

4.2.1.1 Phonological factors

Blocking can be motivated by phonological considerations. Siegel (1971) and Halle (1973) have observed that verbs with an **inchoative** meaning, roughly interpretable as 'to begin to X', can usually be formed from adjectives by suffixing *-en* to an adjectival base provided it meets the following phonetic conditions:

(i) the base must be monosyllabic;
(ii) the base must end in an **obstruent** (i.e., stop, fricative or affricate), which may be optionally preceded by a **sonorant** (for example, a nasal consonant or an approximant like /l/ or /r/). These phonological constraints mean that the derived verbs in [4.9a] are allowed but those in [4.9b] are not:

[4.9] a.

black-en	/blæk-ən/	whit-en	/waɪt-ən/
damp-en	/dæmp-ən/	hard-en	/hɑːd-ən/
quiet-en	/kwaɪt-ən/	length-en	/leŋθ-ən/
tough-en	/tʌf-ən/	rough-en	/ɹʌf-ən/
soft-en	/sɒf-en/	fast-en	/fɑːs-ən/

b.

*dry-en	/draɪ-ən/	*dimm-en	/dɪm-ən/
*green-en	/gri:n-ən/	*blue-en	/bluːən/

Obviously, as Halle (1973) remarks, given the existence of numerous well-formed words like /laɪjən/ (*lion*) and /ʌnjən/ *onion,* the phonetic restriction on /-ənn/ following sonorants is not general. It is peculiar to inchoative verbs formed from adjectives.

Next, we shall focus on *-ly*. We have already seen that this derivational suffix is attached in a highly predictable manner to adjectives to form adverbs such as:

[4.10]

Adjective	Adverb	Adjective	Adverb
kind	kindly	elegant	elegantly
fierce	fiercely	serious	seriously

Exercise

Note, the adverbs in [4.11] are disallowed, or at best awkward, even though they might be listed in dictionaries. Suggest a phonological motivation for this restriction.

[4.11]

Adjective	Adverb	Adjective	Adverb
silly	*sillily	friendly	*friendlily
miserly	*miserlily	sisterly	*sisterlily

What [4.11] shows is that the segmental phonology of the base can determine whether a form can undergo *-ly* suffixation. The *-ly* suffix tends to be avoided where an adjective ends in *-ly* (/-li/). Suffixing *-ly* would result in a dispreferred /-lili/ sequence in the derived adverb (cf. the discussion of morphological haplology in Section 3.6). But *-ly* is used freely where the adjective does not end in *-ly* (Aronoff, 1976).

Our final example of phonological constraints on word-formation is taken from French where the diminutive suffix *-et* (masculine)/*-ette* (feminine) is used freely to form diminutive nouns like these:

[4.12]	fille	[fij]	'girl'	fillette	[fijɛt]	'little girl'
	camion	[kamjɔ̃]	'truck'	camionette	[kamjɔnɛt]	'light truck'
	livre	[livʁ]	'book'	livret	[livʁɛ]	'booklet'

Exercise

Suggest a phonological explanation for the blocking of the suffixation of *-et*/*-ette* in the following:

[4.13]	contrefort	(masculine)	*contrefortet	'little buttress'
	bastide	(feminine)	*bastidette	'little country-house'
	ride	(feminine)	*ridette	'little wrinkle'
	cachet	(masculine)	*cachetet	'little stamp, seal'
	carotte	(feminine)	*carottette	'little carrot'

The suffixation of *-et*/*-ette* is frustrated if the last consonant in the base is an alveolar plosive, /t/ or /d/. This is reminiscent of the restriction in English on deriving adverbs from adjectives by suffixing *-ly* (cf. *friendlily* in [4.11]). Languages seem to have rules of 'euphony' which tend to bar certain jarring sound sequences in word-formation.

4.2.1.2 Morphological factors

The morphological properties of a base may prevent the application of morphological rules. Often **native morphemes** behave differently from foreign morphemes. Some affixes may only be added either to native bases or to bases of foreign origin. For example, as we saw above in [4.5], *-ant* (as in *defendant*) is suffixed to bases of French origin.

Similarly, the rule of **velar softening** which changes /k/ (usually spelled with the letter *c*) to [s] is essentially restricted to words of Latin and French origin:

[4.14] Velar Softening
/k/ → [s] before a suffix commencing with a nonlow front vowel (e.g., *i*)

The effects of velar softening can be seen in [4.15]:

[4.15]

[k]		[s]
cyni<u>c</u>, cyni<u>c</u>al	→	cyni<u>c</u>ism
criti<u>c</u>, criti<u>c</u>al	→	criti<u>c</u>ism, criti<u>c</u>ise
fanati<u>c</u>	→	fanati<u>c</u>ism
asceti<u>c</u>	→	asceti<u>c</u>ism
scepti<u>c</u>	→	scepti<u>c</u>ism

Velar softening only affects words with Romance roots. So, if a thinker called *Isaac* developed a new philosophy, we might call it *Isaacism* [aɪzəkɪzm̩]. But we could not call it * [aɪzəsɪzm̩], since *Isaac* is not a Romance root (cf. also Section 6.2.4).

Exercise

[4.16]

a.	boy-hood	brother-hood	man-hood	maiden-hood
	girl-hood	sister-hood	woman-hood	maid-hood
	child-hood	king-hood	priest-hood	knight-hood
b.	*judge-hood	*governor-hood	*colonel-hood	*minister-hood
	*director-hood	*author-hood	*prisoner-hood	*general-hood

(i) What is the meaning of *-hood* in [4.16]?
(ii) Show the relevance of the distinction between native and foreign bases in the selection of bases to which *-hood* is suffixed.

<u>Hint</u>: Consult a good etymological dictionary. This exercise requires some knowledge of the historical sources of English words.

Hopefully you have correctly observed that normally *-hood,* (which means 'rank, state, quality') co-occurs with native roots like those in [4.16a] and is disallowed after latinate roots like those in [4.16b].

Clearly, the distinction between native and borrowed morphemes is very important. However, we should be careful not to press this too far. The roots in [4.17] below are borrowed from French, yet they can take the suffix *-hood*. With the passage of time, foreign morphemes can be fully assimilated and nativised so that they behave in the same way as indigenous morphemes:

[4.17] parenthood statehood nationhood

Finally, the selection of affixes that co-occur with a particular base may depend on that base being a member of a particular **paradigm**, that is, a purely morphological subclass. Morphemes belonging to different paradigms take different affixes. This is very often the case in inflectional morphology. Consider the French examples below. The regular verbs belong to one of these morphological classes: (i) *-er* verbs [4.18a]; (ii) *-ir* verbs [4.18b]; and (iii) *-re* verbs [4.18c]:

[4.18]

a. -er verbs

donn-er	'to give' (inf.)	demand-er	'to ask' (inf.)
je donn-e	'I give'	je demand-e	'I ask for'
tu donn-es	'you (sg.) give'	tu demand-es	'you(sg.)ask for'
nous donn-ons	'we give'	nous demand-ons	'we ask for'

b. -ir verbs

fin-ir	'to finish'(inf.)	gém-ir	'to moan'(inf.)
je fin-is	'I finish'	je gém-is	'I moan'
tu fin-is	'you (sg.) finish'	tu gém-is	'you (sg.) moan'
nous fin-issons	'we finish'	nous gém-issons	'we moan'

c. -re verbs

vend-re	'to sell' (inf.)	romp-re	'to break' (inf.)
je vend-s	'I sell'	je romp-s	'I break'
tu vend-s	'you (sg.) sell	tu romp-s	'you (sg.) break'
nous vend-ons	'we sell'	nous romp-ons	'we break'

As you have seen, depending on which paradigm a verb belongs to, it co-occurs with different allomorphs of inflectional suffix morphemes. This discussion will be resumed in Section 4.3 and in Chapter 11. The existence of paradigms is very important for understanding the nature of allomorphy.

4.2.1.3 Semantic factors

Semantic considerations, too, may impinge on the application of word-formation processes. This is seen in the way the otherwise general process of forming compounds from adjective plus past participle (V-ed), which is shown in [4.19] is blocked in [4.20]:

[4.19]

short-sleeved (shirt)	one-armed (bandit)
short-sighted (man)	three-legged (stool)
green-roofed (house)	red-nosed (reindeer)
blue-eyed (boy	red-haired (woman)

[4.20]

*two-carred (family)	(for 'a family with two cars')
*big-Alsatianed (woman)	(for 'a woman with a big Alsatian')

Compound adjectives derived from the past participle (V-ed) form of the verb are most likely to be permitted where the root to which *-ed* is added is **inalienably possessed** (i.e., inherently possessed) by the head noun that it modifies. The compound words in [4.19] are permissible because someone's eyes are an integral part of their body. Similarly, the legs of a stool, the sleeves of a shirt and the roof of a building are an obligatorily possessed part of some piece of furniture, garment or building. But it certainly is not the case that an Alsatian dog or a car must necessarily be possessed by someone.

Exercise

The use of the italicised words in the dialogue below is odd. What would be the natural and preferred word choices that one would probably use instead of *unill, unsad, unpessimistic* and *undirty?* Why?

[4.21] SURGEON: How are you today, Leslie?
PATIENT: I am feeling much better. It's just wonderful to be so *unill* again.
SURGEON: Oh, I'm so *unsad* to see you making such good progress. I am very *unpessimistic* about your chances of making a full recovery. The main thing now is to make sure we keep the wound *undirty* to avoid infection.

This example illustrates how semantics may restrict the application of morphological rules. If there are two adjectives with opposite meanings, one of which has a more positive meaning than the other, normally the negative prefix *un-* attaches to the positive adjective (see [4.22a]). If *un-* is attached to the negative member of the pair, as in [4.22b], the resulting word is usually ill-formed.

[4.22]	a.	unwell	b.	*unill
		unloved		*unhated
		unhappy		*unsad
		unwise		*unfoolish
		unclean		*unfilthy, *undirty
		unoptimistic		*unpessimistic

As shown, if there are words representing the two poles of the same semantic dimension, we tend to prefer treating the positive end as **unmarked** (i.e., as normal or basic). We are happier to derive the **marked** (i.e., 'unusual'), less favourable meaning by prefixing the negative prefix to a positive base than doing the reverse. That is why *a happy* person is not said to be *unsad* (Zimmer, 1964). To make the dialogue in [4.21] normal, the marked words, which are italicised, must be replaced by their unmarked counterparts in [4.22a].

4.2.1.4 Aesthetic factors and the adoption of words

In some cases word-formation is inhibited by vague aesthetic factors. There are many examples of words that are in principle well-formed whose adoption has nevertheless been resisted.

In the 1970s, the word *stagflation* was coined to refer to the combination of economic *stagnation* and a high level of *inflation* that afflicted the world economy at that time. So far, this word seems to have failed to get a firm foothold in the language. Aesthetic considerations may have had something to do with it. Some commentators consider it 'ugly'.

Other 'ugly' words which raise hackles include *talkathon, swimathon, knitathon,* etc. These are made up by analogy to *marathon*. Erudite purists are outraged not only by the sight of a combination of a Greek pseudo-suffix with native Anglo-Saxon roots, but also by the misanalysis of *-athon* as a 'suffix' meaning 'undertaking a strenuous prolonged activity (specified in the part of the word that precedes *-athon*) for the benefit of a good cause'. In Greek *-athon* was not a morpheme. But the average speaker of English who is unaware of such niceties will probably contentedly coin more *-athon* words, regardless. This is English, not Greek, after all.

4.3 DOES PRODUCTIVITY SEPARATE INFLECTION FROM DERIVATION?

Productivity is often taken as a criterion for distinguishing inflection from derivation. Derivational processes are by and large much more unpredictable than inflectional ones. While inflectional processes usually affect most of the eligible forms in a regular manner, derivational rules tend to be capricious. They tend to have as input a class whose membership is subject to various arbitrary exclusions and to affect it spasmodically as we have already seen (see [4.5] and [4.6]). Unlike derivation, inflection is normally productive in these senses:

(i) It is general. The addition of particular inflectional affixes is not subject to various arbitrary restrictions. Stems that belong to a given class normally receive all the affixes that belong to that class (cf. Section 11.2.1.2).
(ii) The words resulting from the addition of inflectional affixes have regular and predictable meanings.

Another way to put it is to say that typically inflectional morphology displays **lexemic paradigms** but derivational morphology does not. Paradigms are regular and predictable sets of word-forms belonging to the same set of lexemes. They share morphological characteristics (e.g., prefixes, suffixes or infixes). Such words belong to a particular word-class or subclass.

The selection of a specific word-form is determined by the syntax. See [4.18] above where three **inflectional paradigms** of French verbs were shown.

English verb forms also belong to inflectional paradigms:

[4.23] walk ~ walks ~ walked ~ walking
love ~ loves ~ loved ~ loving

Usually inflectional morphology exemplifies automatic productivity. Most English verbs have these forms. So, if we encounter a new verb like the made-up verb *pockle* (meaning perhaps 'to go (away in a huff)'), which was introduced in Chapter 2, we can predict that it will have the forms *to pockle, pockles, pockled* and *pockling* with the standard meanings.

By contrast, as a rule, paradigms cannot be set up for derivational morphology. If we try to produce a paradigm with the derivational suffixes *-ate*, *-ant*, and *-ation*, we are soon thwarted. Our putative paradigm in [4.24] is riddled with gaps:

[4.24]

Verb (X)	Noun (one who does X)	Noun (act of X)
communicate	communicant	communication
donate	—	donation
navigate	—	navigation
rotate	—	rotation
militate	militant	—
applicate	applicant	application
(obsolete)	accountant	—
—	—	atation (art of swimming)

It is tempting to assume that wherever paradigms can be recognised, as in [4.23], one is dealing with inflection, but where no regular paradigms exist, as in [4.24], one is dealing with derivation.

However, although using the presence of paradigms involving a given affix as a rule of thumb for distinguishing between inflectional and derivational morphemes usually works, it will not unfailingly lead to safe results. This is because, on the one hand, there are exceptions to inflectional processes and, on the other, there are derivational processes which appear to fall into very general paradigm-like patterns. We will consider these two problems in turn.

First, while most inflectional processes tend to fall neatly into paradigms, there are often some forms that escape. For instance, although the vast majority of nouns are count nouns, which means that they have both singular and plural forms and meanings (e.g., *book* ~ *books*, *dog* ~ *dogs*, *ass* ~ *asses*, etc.), there is a minority which are not marked for number. They belong to different subclasses of **noncount (or mass) nouns** referring to entities that are not individually counted:

(i) Some noncount nouns have a plural form but lack a plural meaning:

[4.25]
*alm	~ alms
*outskirt	~ outskirts
*oat	~ oats
*linguistic	~ linguistics
*new	~ news

(ii) Other noncount nouns lack a plural form altogether:

[4.26]
milk	~ *milks
health	~ *healths
equipment	~ *equipments
courage	~ *courages

Conversely, sometimes a derivational process such as the suffixation of *-ly* to adjectives to turn them into adverbs (e.g., *kindly*, *quietly*, etc.) is almost exceptionless (but see [4.11]). Similarly, derivation of nouns referring to containers (e.g., *cupful*, *basketful*, *canful*, *bagful*, etc.) by suffixing *-ful* to nouns is very regular and predictable. Hence, regularity cannot be used as a foolproof litmus test that distinguishes inflection from derivation.

4.4 THE NATURE OF THE LEXICON

We have seen in the foregoing sections that the lexicon is open-ended. Not all the words of a language can, or need be, listed in the lexicon. The question that arises is: on what basis, then, are words (and morphemes) selected for inclusion? And for the words selected, what kinds of information must the lexicon include? In brief, what is the nature and function of the lexicon?

4.4.1 Potential Words

As we have already observed, knowing a language involves, among other things, knowing the rules of word-formation. Speakers are able not only to identify the meaningful units which words contain, but also to create new words, and to understand the meanings of unfamiliar words that they have not encountered previously. If we came across 'words' like *grestifier* and *disperidate* we would recognise them as potential English words. But how do we know this? At least part of the answer lies in the nature of our mental lexicon.

The lexicon has a set of **phonotactic constraints** which function as a filter allowing entry only to phonologically well-formed words. Before any putative word can enter the lexicon, it must have a combination of sounds that is permissible in the language. So, 'words' like **ltarpment* and **mpandy*

are immediately rejected because the consonant combinations /lt/ and /mp/ are disallowed at the beginning of a word in English.

Non-nativised foreign words with sound sequences that are not permitted in English may be allowed in, as a special case. Presumably such words are kept in a special sector reserved for them. We need to make some allowance for, say, foreign place names like *Tblisi* or foreign personal names like *Zgusta* which begin with the consonant clusters [tbl] and [zg] which are unorthodox in English. Foreign words entering the language may be allowed to by-pass the phonotactic filter. In that case they keep their foreign pronunciation, virtually unchanged (e.g., *tsunami* for those speakers who pronounce the *t*, [tsunami]. But, more often than not, they tend to be significantly modified so that they fit in the general phonotactic patterns of the language. So, a schwa may be inserted between the first two consonants in *Zgusta* and *Tblisi* to produce [zəgʊstə] and [təblisi] respectively.

4.4.2 Knowledge of Language and the Role of the Lexicon

The lexicon is a mechanism for capturing broad regularities involving words in a language. For instance, using phonotactic constraints English speakers can distinguish very generally between, on the one hand, possible words (which may not be instantiated in sentences and utterances, e.g., *grestifier*) and, on the other hand, impossible words (e.g., **ltarpment*).

In the foregoing we have proposed that some of the general properties of morphemes and lexical items should be shown in the lexicon. However, traditionally the lexicon was not regarded as a place where regularities were captured. Rather, it was viewed as the repository of exceptions, in the form of a list. We read in Bloomfield (1933: 274) that 'the lexicon is really an appendix to the grammar, a list of basic irregularities.'

This view is colourfully caricatured by Di Sciullo and Williams (1987) as being one where the lexicon is conceived of as a prison which 'contains only the lawless, and the only thing that its inmates have in common is lawlessness'. And the lawless are a disparate bunch including words (for example, *work*), morphemes (for example, *-ed*) and idioms (for example, '*eat one's words*').

Nonetheless, it is the case that the lexicon in a generative grammar must list various kinds of information about words (and morphemes and idioms) which have to be memorised. For example, speakers of English who know the word *aardvark* need to memorise at least this information:

(i) Meaning: it refers to a Southern-African insectivorous quadruped mammal.
(ii) Phonological properties: its pronunciation /ɑ:dvɑ:k/.
(iii) Grammatical properties: for example, it is a count noun (you can have one aardvark, two aardvarks).

Admittedly, what needs to be listed in speakers' mental lexicons may vary. While for most people *aardvark* needs to be memorised, some erudite speakers know that this word is a compound borrowed from Afrikaans and is composed of *aarde* 'earth' and *vark* 'pig'.

Today most generative linguists reject the view that the lexicon is merely a list of irregularities. If there is a need for lists in a grammar – and there clearly is, since we need to list basic morphemes then the lists belong to the lexicon. However, this does not mean that the lexicon consists just of lists. There are many extensive and far-reaching lexical regularities resulting from the operation of general principles.

Normally, the relationship between the meaning and form of a morpheme or word is completely arbitrary and idiosyncratic (notable exceptions being cases of onomatopoeia, e.g., *cuckoo* and *miaou*), but many other properties are not. There are numerous pervasive regularities in the phonological and syntactic behaviour of words.

In the next part of the book we are going to explore the organisation of the lexicon, concentrating on the representation of word-formation regularities that relate to the phonology. In the last part of the book we will come back to the lexicon and consider regularities that relate to the syntactic and semantic properties of words.

FURTHER READING

Bauer, Laurie. (2001). *Morphological Productivity*. In the series *Cambridge Studies in Linguistics*, vol. 95 (Cambridge: Cambridge University Press).

EXERCISES

1. (a) Make a list of ten words containing the suffix *ic,* as in *magnetic* and *allergic*.
   (b) What word-class do the bases to which *-ic* is added belong?
   (c) What is the word-class of the resulting word?
   (d) Is *-ic* an inflectional or derivational suffix? What is your evidence?
   (e) Does *-ic* attach freely to all eligible bases? If not give two examples of bases which are ostensibly suitable for *-ic* suffixation but which fail to take this suffix.

2. Study the examples below and, with help of a good dictionary, answer the questions that follow:

monoism	unique
monologue	unilateral
monolingual	unitary

monolith	unify
monogamy	unipolar
monogenesis	unicellular
monochrome	unidirectional

(a) What is the meaning of the morphemes *mono-* and *uni-?*
(b) What bases can these morphemes be attached to normally? List as many relevant factors that play a role in the selection of these prefixes as you can think of.
(c) What is the word-class of *mono-* and *uni* according to the data above? Why?
(d) What is the word-class of the words formed by attaching *mono-* or *uni-*?

3. This is an open-ended question.

(a) In the data below list three examples of (i) conversion, (ii) suffixation, (iii) compounding.
(b) Comment on problem cases which do not seem to fit neatly into any one of these categories.

Note: When you consider conversion, take into account the pronunciation of vowels and consonants as well as where stress falls in a word.

advice (noun)	advise (verb)
plan (noun, verb)	concrete (noun, adjective)
reject (noun, verb)	table (noun, verb)
greyhound (noun)	milkman (noun)
blood (noun)	bleed (verb)
song (noun)	sing (verb)
teapot (noun)	breakfast (noun, verb)
bio-science (noun)	telegraph (noun)
hardship	hardwood

4. Compare the following pairs:

author ~ authoress	lion ~ lioness	actor ~ actress
manager ~ manageress	mayor ~ mayoress	waiter ~ waitress
editor ~ editress	poet ~ poetess	emperor ~ empress

(a) What is the word-class of the bases that the *-ess* suffix attaches to?
(b) What class do the resulting words belong to?
(c) What is the meaning of this suffix?
(d) Do you find these words equally acceptable? If you do not, explain why.