Chapter 1

1. The trouble with the truth is that we cannot recognise it; at some point we accept something as close enough to the truth for practical purposes, but this does not necessarily mean that it is the truth in any absolute sense. Newtonian physics works very well in day-to-day interactions with the world, but Einsteinian physics shows it to be, at best, an incomplete version of the truth and at worst, false. So truth is a lot to ask, and in Linguistics, there may be more than one truth, depending on how realist we are in our view of what the ‘truth’ entails.

None of this has anything to do with how many people believe something. At one point, heliocentricity was not believed at all. So if we now think that heliocentricity is true, it is not because more people believe in it; it is for other reasons. This has implications in academic argument: argument from authority (Person X says this, and they are an authority, therefore it must be true) is not necessarily a good argument. Ignoring previous scholarship is, of course, dangerous. We don’t want to reinvent the wheel all the time. But even if some authority knows a lot more than we do in general terms, it does not follow that such a person has found the ‘truth’.

2. Virtually any simple statement of a hypothesis will raise a huge number of questions. What do the terms mean (including things like ‘good’)? If we define ‘good’, how many people does it have to apply to for the predicate ‘good for people’ to be true? Can something be simultaneously ‘bad’ and ‘good’? What can we accept as evidence for or against the hypothesis? No media sound-bite can hope to include all such factors, and so any such brief summary inevitably leaves a number of questions unanswered, some of which could reverse your impression of how successful the hypothesis is.

Discussions about linguistic structure must be subject to the same problems. We cannot evaluate a claim such as ‘People talk in
sentences’ unless we can define a sentence, say what kind of people and under what circumstances and determine what is included in ‘talk’ (does it, for instance, include writing?). For example, a definition of a sentence as ‘A self-contained section of speech which contains at least one finite verb’ or as ‘A sequence of words stretching from after one intonational nucleus to the next’ will give rise to completely different results.

3. The answer to this question depends on how you think a definition works. If you think that a definition of ‘word’ must specify all the conditions under which a segment of speech is a word and exclude all others, then failure to provide a definition means that you cannot be sure that any conclusions on the structure of the word will hold. On the other hand, if you think that you can provide a definition of ‘word’ which tells you things that usually happen in words and things which would usually exclude something from being a word, even if all of these things are not present all of the time, you may end up with a fairly good idea of what the kernel of the notion of ‘word’ is like, even if there are marginal examples about which you cannot be sure. Under such circumstances, you may have a fairly good idea about the structure of words, even if there are instances where you may be insecure about a definitive answer. The notion of word is apparently so complex that it is not possible (in our current state of knowledge, and with our current understanding of ‘word’) to give a set of necessary and sufficient conditions to guarantee that something will be a word. But we certainly know enough about words to have some basic idea as to what is involved, and to have some idea about what is usual and what is not.

4. The two are not identical because – as the notation shows – the first is a statement about a lexeme, the second a statement about a word-form. Consider a situation where you are discussing the sentence in (a).

(a) Kim and Lee walk home every day.

Here it is clearly true that *walk* is an intransitive verb. But is it also true that *walk* is an intransitive verb? This may depend on where you believe that transitivity resides. Does the sentence in (b), for instance, contain the same lexeme *walk* as the sentence in (a)?

(b) Kim and Lee walk the dog every day.

In (b) *walk* is a transitive verb (it has a direct object *the dog*). But does this word-form belong to the same lexeme as the *walk* in (a)? If there is just one lexeme, then it is false that *walk* is an intransitive
verb, since it can be either and is inherently neither. If, on the other hand, you believe that there are two distinct lexemes WALK here (say, WALK₁ and WALK₂, with WALK₂ appearing in (b)), then it is true to say that WALK in (a) is an intransitive verb.

So, you probably want to know whether one of these or the other is ‘right’. This is one of those areas where the theory does not specify clearly which point of view we should take. My guess is that the overwhelming preference these days is to say that there is only one lexeme WALK, and that context coerces one reading or the other.

<table>
<thead>
<tr>
<th>Orthographic words</th>
<th>My films are not comedies, but there’s comedy in them</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lexemes</td>
<td>1 2 3 4 5 6 7 8 9 10</td>
</tr>
<tr>
<td>Word-forms</td>
<td>1 2 3 4 5 6 7 8 9 10</td>
</tr>
<tr>
<td>Morphosyntactic words</td>
<td>1 2 3 4 5 6 7 8 9 10</td>
</tr>
<tr>
<td>Grammatical word</td>
<td>1 2 3 4 5 6 7 8 9 10</td>
</tr>
<tr>
<td>Lexical words</td>
<td>1 2 3 2 3</td>
</tr>
</tbody>
</table>

There are some places here where we might argue: is there’s one word-form or two? Is not only a grammatical word, or does it also have lexical content?

6. This is a tough question. In the sense that words surround us, in speech and writing, and that speakers of many languages are willing to speak a sentence ‘one word at a time’, we can argue that they are part of what we observe. However, the examples like alright/all right show that sometimes the word is determined by something other than what we feel. That ‘something’ is a theory of some kind, even if it is not an overt theory. Our spelling system provides us with a theory of what a word is, a theory that may not match our experience. Children learning to write sometimes do not know where the boundaries between words go, and we might argue that they have not yet learnt the dominant theory. But there may be alternative interpretations of this observation. One way to align these points is to suggest that the word is observed in spoken language, but is theorised in written language. Even that may be difficult to maintain. Items like coffee pot/coffeepot may be treated as one or two words because of the way they appear in the written language.
Chapter 2

1. Alternative answers are possible, but must be well justified.
   (a)  
   -ther was almost certainly a morpheme in Indo-European (and -ter would have to be an allomorph following a consonant), but in modern English this seem to create too many unique morphs for the size of the problem, so we would probably conclude that each of these is monomorphemic.

   (b) Each of these spellings covers two words, one with a prefix /ri:/ and one with a non-prefix usually pronounced in a different way. Recount ‘tell a story’ is different from re-count ‘count again’. Only the second clearly has a prefix. The <re> in the other cases has not really got a consistent meaning IN ENGLISH.

   (c) Here re- is almost certainly not a morpheme in English. What would the rest mean? What would the meaning of re- be?

   (d) Here re- is consistently a morpheme, with a consistent meaning ‘again’. Note refresh has an adjective in the base and a different pronunciation, so perhaps it is different from the others.

   (e) There are those who argue that th- is a morpheme in this set (meaning something like ‘deictic’), the final /ə/ in the being a default vowel. Comparison with other Germanic languages might support such an analysis, and certainly suggests that in Proto-Germanic there was a possibility of such an analysis. In modern English, though, these are not perceived as bipartite, there are too many unique morphs. I certainly do not believe these to be bimorphemic in current English. Do morphemes have to have some psychological reality? If these are not morphemes, they are just bits of etymology.

   (f) This is much harder. Each of these has a corresponding tb-form and a corresponding wb- form, where tb- is distal and wb- is interrogative. The system is small, but there is a good argument for a bimorphic analysis. Modern usage suggests that this is not the case, though; if it were, ‘from hence’ would be a tautology instead of normal usage. A monomorphemic analysis is more likely. Try to extend the system: that and what look as though they might fit, but we argued in (5) that if anything the tb- there was deictic not distal, and we can see from this and that the -at is the distal part of that.

   (g) The first black is clearly a colour word. Blackguard and blackmail may be black ‘illegitimate’ as in black market. Note the pronunciation of blackguard which suggests no morph. Blacksmith is arguable. The point is that a morpheme is supposed to share
form and meaning; these *black* elements don’t seem to. Shared form is no guarantee of shared morpheme structure.

(h) The answer is that gymkhana is irrelevant and for the others, you get *-iana* only if you get *-ian*. The form of the affix is awkward: *-a*, *-ana* and *-iana* might all be argued for. The meaning is ‘collection of materials associated with ~’.

(i) There is a prefix *dis-* in the first three examples, and not in the last, as there is no independent form *turb*. Distress typically provides difficulties. My answer would be that distress has a prefix *dis-* if and only if it means ‘remove the hair of’ (this gives an analysis *dis-* + *tress*). Since distress does not mean ‘shave the head of’, this analysis does not hold. It cannot mean *dis-* + *stress* with simplification of /ss/ at the boundary, because the meaning would be the opposite of what we find in distress. Distress must contain just one morpheme.

(j) We can agree that there is a suffix *-al* which creates adjectives, but the question arises as to when it has an allomorph *-ial*. This is not predictable in the examples given (stress may be a factor, but not a decisive one, since suicidal is stressed on the last syllable of the base, just like dictatorial). Indeed, it is probably not predictable overall. This makes it a lexically conditioned allomorph, which some authorities do not like.

(k) If the *-m* and the *-t* in poem and poet are morphs, this implies there is a base *poe*, which has no independent existence. Moreover, there are no recurrences of these morphs with the same meanings. Therefore we cannot analyse these words morphologically. It looks as though there might be a prefix *d-* added to the word evil, but not only is the pronunciation wrong, we have no recurrence of an initial *d-* meaning ‘one who instigates or practices ~’. Again, we must conclude that these words are unanalysable.

2. Basque

<table>
<thead>
<tr>
<th>n</th>
<th>‘1SG’</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>‘absolutive’</td>
</tr>
<tr>
<td>z</td>
<td>‘2SG’</td>
</tr>
<tr>
<td>u</td>
<td>‘ergative’</td>
</tr>
<tr>
<td>z</td>
<td>‘plural’</td>
</tr>
<tr>
<td>r</td>
<td>‘singular’</td>
</tr>
<tr>
<td>d</td>
<td>‘3’</td>
</tr>
<tr>
<td>ka</td>
<td>‘bring’</td>
</tr>
<tr>
<td>ki</td>
<td>‘dative’</td>
</tr>
<tr>
<td>to</td>
<td>‘come’</td>
</tr>
</tbody>
</table>
You will probably find that you looked for minimal similarities and minimal differences in both the foreign language data and in the glosses. This means that, implicitly, you expect one bit of form consistently to carry the same meaning, and the meaning to be realised by the same form. In other words, you work on the assumption that the morphemes have both fixed form and meaning. That is precisely what the classical morpheme would lead you to expect.

3. It would be a very unusual language that used an affix for the singular and not for the plural. As far as I know, such languages do not occur. The reason comes down to markedness. In a language like English, the plural is marked by an affix, while the singular is not marked by an affix. In a language like Zulu, both singular and plural are marked by affixes. But where a property is not formally marked in this way, it is generally the case that it is the property with the widest distribution, it is widespread across languages and resistant to change across languages, and it is the cognitively dominant form that is not marked. This leads to a wider usage of the term ‘marked’, opposed to ‘natural’, where ‘marked’ means ‘unusual’, ‘difficult to process’, ‘relatively uncommon’, and so on. While you cannot absolutely predict what will be unmarked (it is odd that the third person singular of the present tense in English verbs is marked with an -s affix, for example, in forms like infers), the patterns are well-enough recognised that the lack of marking for a present as opposed to a past tense or a singular as opposed to a plural is simply regarded as one of the things that are repeatedly found across languages. A form contrasting with the lack of a form is thus normal, not problematical, as long as they are the expected way round.

This assumes that the words are really singular and plural, in the examples in the text. Some languages have different ways of marking number. In particular, there are languages which mark

---

<table>
<thead>
<tr>
<th>Swahili</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>ni</td>
<td>‘I’</td>
</tr>
<tr>
<td>wa</td>
<td>‘they’</td>
</tr>
<tr>
<td>li</td>
<td>‘past’</td>
</tr>
<tr>
<td>ta</td>
<td>‘future’</td>
</tr>
<tr>
<td>sa</td>
<td>‘present’</td>
</tr>
<tr>
<td>taka</td>
<td>‘want’</td>
</tr>
<tr>
<td>soma</td>
<td>‘read’</td>
</tr>
<tr>
<td>andika</td>
<td>‘write’</td>
</tr>
<tr>
<td>imba</td>
<td>‘sing’</td>
</tr>
</tbody>
</table>
singulatives and pluratives. In such languages, a noun is marked in the singulative if it denotes an object that you expect to find in large numbers (like ants, for instance), but marked in the plutorial if it denotes an object you expect to find just one at a time (like magpies or eagles, for example). This is clearly natural, too, but in a different way.

4. There is no problem for classical morpheme theory here, we simply have homophonous morphemes which are distinguished by their position.

5. If we assume a beads-on-a-string model of classical morphemes, then we must have independent items which occur in a given position on the string. When two elements within one of those units swap positions, the model cannot cope. So if we assume that we have a process swapping the positions of items, classical morpheme theory cannot deal with this. This is the position taken by Anderson.

To have a version of this that is consistent with the classical morpheme we would have to list each of these stems as separate morphemes. This is tantamount to saying that there is no relationship between them: if we simply list the forms of the stem, those forms could be as distinct as go and went, for example. So this solution would lose a great deal. We could try to say that there is a process which derives one stem from another, but we have already seen that processes are incompatible with the classical morpheme, although there are versions of morpheme theory which allow this (see further in Chapter 5).

Stonham’s argument, however, is different in nature. If we have an underlying form which consists of just consonants and the vowels are inserted by rules, then we can have the vowels inserted in different positions in the two aspects. Superficially, it looks as though one of those processes would have to be a process of infixation, which is also a problem for the classical morpheme, but Stonham would argue that the vowel is inserted after the first consonant by default (to make a syllable), and at the end when an extra vowel position has been suffixed to the stem. In this case, there is no metathesis, just differential vowel insertion. The question is whether all instances of apparent metathesis can be explained in similar ways. Stonham implies that they can, Anderson thinks not. Even if Stonham is right, you might want to argue that the forms that face the speaker/listener are metathetical and that speakers are likely to interpret the data in this light even if a smarter solution that does not involve metathesis is available. This then becomes like the problem of subtraction in French.
Chapter 3

1. The main point of this exercise is to make you think about the differences between morpheme-based morphology and word-based morphology. In a morpheme-based morphology, you will seek morphemes such as {e}, {ais}, {er}, {ai} (or possibly, {ai} and {s} to account for the sequence -ais). In a word-based morphology you will look for the stems (donn, parl) and rules to provide the forms required elsewhere. You will not expect, necessarily, to find that -er- means anything, for instance. That is quite helpful, since it is not entirely clear what -er- and -ais (whether as a single morpheme or a sequence of two) would mean in a morpheme-based morphology. A morphemic treatment would allow you to recognise these as units of form without any single meaning, but recurring in different contexts. That might be a useful compromise solution.

2. In the Latin, as we have already seen with different data, you will need to say that, for example, -ās means ‘accusative plural’ without dividing this up into two separate bits. The implication is then that Turkish -lerde will equally not be split up into two meaningful elements, but will simply be taken to mean ‘locative plural’. But there is a difference: in Latin there is no justification of a split in the -ās element, in Turkish -lerde can be split. In a word-based morphology, this difference will be captured in the simplicity of the rules for generating the phonology required to mark plurality. In the Turkish (at least in the bits of Turkish illustrated) there will only be one rule needed to produce relevant forms of the meaning ‘plural’, while in Latin, the rules for marking plurality will be inextricably entwined with the rules for marking case. Morpheme-based morphology will have to deal with this differently.

   How things change when you add more material is something you can consider yourself. In Turkish the rules for number and case become more phonologically complicated as bases with more different phonological patterns are added, but the independence of case and number persists. In Latin, the marking of case and number is also involved in the difference between declensions. It is not immediately clear which of these might be the more economical system, if economy is a desideratum.

3. Presumably view will be inserted directly into the sentence from the lexicon, while criticism will have to be inserted in a way which indicates that it is the nominalisation of criticise. But while the meanings of inflectional properties are specified in the syntax, there is no difference between the syntax of (a) and (b), only a difference
in the morphology. So where does the information that *criticism* is a nominalisation (or has the meaning ‘nominalisation’) come from? One possible solution is that it arises from the fact that the appropriate element is dominated by an N node in the syntactic tree. This implies that the verbal meaning of *criticise* also comes from the tree rather than from the lexicon. But it is not clear that this would allow you to distinguish between *king* and *kingdom* in a tree, where the syntax may not need to know about the different kind of noun, but the morphology does. There is probably a perfectly good solution to all of this somewhere, but it is not necessarily an obvious one.

4. There is no ‘right answer’ here, only a matter of deciding what counts as relevant data and what is irrelevant (and why), and then trying to weight the various factors that apply. Things to consider include:

- The form used in the perfect is almost always the same as the form used as an adjective, but there are exceptions. This is true for irregular and for regular cases.
- The adjectival form sometimes is phonologically different but orthographically identical to the form used in the perfect.
- The *-ed* form can be converted into a noun.
- There are apparent cases where inflectional affixes may occur outside (further from the root than) the past participle marker, but these are all suspicious for one reason or another.
- There are apparent cases where derivational markers occur outside the past participle marking, but these examples are also dubious.
- The *-ed* marker is not used exclusively on verbs.
- Variation in the forms used in the perfect is not always precisely parallel to variation in forms used adjectivally.

5. This question is very similar to the last, though the relevant factors are not all identical. Beard (1982) includes an argument that derivational affixes are easily borrowed, while inflectional ones are not. This argument has not been used in this book, and instances such as the one in hand are critical for determining the truth of the presupposition. If the English plural is inflectional, the assumption cannot be entirely true.

Chapter 4

1. The question is to make you consider the relevant factors rather than to get at a particular answer. Having said that, I would expect an
answer that the base is the head in all these words (whether prefixed or suffixed); but other answers can probably be justified.

2. If we take seriously the hyponymy criterion, then the right-hand element is the head, because the whole is not a hyponym of the left-hand element but identical to it. We might argue, though, that the left-hand element denotes the whole (and is the head) and the right-hand element merely ensures that we are in the right semantic area: a *pine tree* is contrasted with *pine wood*, though both are properly called *pine*. In many contexts, we could delete either element and keep the overall meaning, and in all cases, there are nouns in both elements, and it is hard to determine which one the noun-quality of the whole derives from. A rather different kind of argument is that compounds in English are generally right-headed, so that if these are compounds (which might be something worth discussing), then they should be right-headed, too. Note, though, the problem of *vegetable marrow*, which appears to have a reverse ordering, possibly because of the parallel with *bone marrow*.

3. You need at least to consider the extent to which the order of the affixes is determined by the word-class (part of speech) of the base to which each affix is attached. Note the circular nature of the attachment (B attaches to A, C attaches to B and A attaches to C), although this circularity is not free recursion, since any sequence of six such affixes is almost certainly of marginal acceptability. There are other parallel cases in English, but not many.

4. We never get sequences of *-ation + -ic*, *-ation + -ous*, or *-ation + -ar* (even though we do find words with the letters <ion> before *-ic*), while we do find *-al* attached to other kinds of base. This seems to imply that *-ation* chooses *-al*.

**Chapter 5**

1. The existence of *enmity* but nothing that resembles *enbity* suggests that the /nm/ in the middle of *enmesh* is possible because /nm/ clusters are possible inside English words, but /nb/ clusters are not.

2. Probably not. It looks as though there is a problem because a suffix is truncated in one case but not in another, but if you are creating the word *designee* from the beginning, you might simply not add any *-ate* sequence when there is final *-ee* suffix. This would not necessarily affect *-ise* sequences. *Dedicatee* and *rotatable* are explicable on other grounds (though the explanations may not hold everywhere). In the case of *dedicatee*, if we removed the *-ate*, we would get *dedicee*, and the spelling system would require that to be pronounced /dedɪsiː/.
Keeping the \(-ate\) avoids the problem. In \textit{rotatable} the \(-ate\) is not a suffix. You might like to explore the extent to which these explanations are valid or helpful.

3. The morphological conditioning for the rule suggests that it is not purely phonological. Without such evidence, the two might seem to be equivalent statements.

4. The point about etymology is that these extenders are not English but are borrowed, and can be argued not to be part of English phonology. I would expect you to find that they are not, in general terms, predictable.

5. I find that my pronunciation does not always match that given in pronunciation dictionaries for relevant forms. This has implications for the automaticity of the processes involved, and may suggest that there is change affecting the system, with some of these changes being gradually lost. There are various suggestions in the literature as to what the conditioning factor for cluster simplification might be in such cases, but they do not seem to be fully generalisable.

6. There may well be a reason for distinguishing between various kinds of morphophonological process, for classification purposes if no other. To what extent we find different kinds of processes in different morphological environments is a matter to be discovered by looking at the evidence to hand. The \textit{casts} kind of reduction may well be different from others illustrated because it is not confined to particular morphological contexts. It would work as well in \textit{Cast six stones, sixths, boosts} (noun or verb), \textit{bestseller}, and so on. It may also differ from at least some of the others in being phonetic rather than phonological: a lengthened \(/s/-\)sound is a possible outcome, but this is not phonemically contrastive.

7. Some of the patterns are rare (perhaps only \textit{sawyer} is like \textit{lawyer}), others have many examples. For instance, the alternation between \(/ai/\) and \(/i/\) is common, and in many instances occurs in the same places as the alternations between \(/i:/\) and \(/e/\) discussed in the text.

**Chapter 6**

1. Clearly, what you discover will depend on the examples you find. There may be some traces of classical morphemes (words ending in \(-on\) tend to look technical, for instance), but they are likely to be rare. Phonaesthemes may or may not be used, but inappropriate phonaesthemes will be avoided. Rhyme, assonance, chiming, and so on may be used, but need not be. There will often be some sort of motivation, even if it is not transparent to all users.
2. As far as I know there are no rules, which is odd.
3. The extent to which the e- in email represents a morpheme is likely to be controversial. Its form (/iː/) suggests it is a letter, not a part of the word electronic, but it is not clear whether that matters. The initial <i> in iPad, iPhone, and so on seems to have taken on some meaning (precisely what may not be clear), and that is even more marginal. At some level we descend into matters of definition. Whether these words are distinct from acronyms is also a matter of opinion. At some point there is supposed to be a difference between an acronym and a clipped compound like modem (modulator-demodulator) or a blend like tigon (tiger-lion). Whether that boundary is fixed in terms of the number of letters or not is an open question. The difficulty in drawing such boundaries may suggest that these are not really distinct formation types, but are all types of abbreviation, but there are objections even to this position.

Chapter 7

1. When the foreign words were first introduced, they were introduced as entire words with no possible breakdown into morphs. Only once a sufficient number had been learnt was it possible to start using a suffix like -ation productively. A native suffix like -ful could be added to suitable nouns, and so could be added to -ation nouns, if appropriate. However, there were no exemplars of -ation being added to -ful, so no model for new words with such forms.

2. Given that there are two competing patterns, the swim, swam, swum one and the swing, swung, swung one, we need to know why the latter is more likely to be generalised than the former. One thing we might hypothesise about is frequency. Perhaps there are more verbs like swing and, in modern usage, they occur more frequently in the past tense than the swim-verbs. In principle, you can check this easily enough in a large corpus, though care is needed. You might like to try for yourself. I suspect that a different factor is more important. Most verbs in English (all the regular ones, for example, and see Anderwald 2009 on the irregular ones) have a past tense form which is identical to the past participle form. Swim-verbs are an exception to this wide generalisation, swing-verbs fit it. This means that there is a pattern of expectation that there is only one form in the non-stem, non-present space, and moving swim-verbs to the swing-verbs' space fits with that expectation.

3. As words get used more often, we do not need to analyse them to understand them. They become listed as wholes in our memory, and
we recognise them as wholes. If most of the words in a particular pattern are listed in this way, we may no longer see the link between the base and the derivative — especially if the derivative is more frequent than the base (Hay 2003). At that point, the words lose their overt motivation, and become as arbitrary in form as any monomorphemic word.

4. Here we have an extended example of the process discussed in relation to the last question. The example of *-th* shows that the lack of motivation in words with a particular affix does not occur suddenly, but through some process similar to lexical diffusion. Eventually motivation vanishes for all the relevant words, and a particular pattern of derivation becomes lost.

5. You were invited to speculate, so we must speculate. The forms *demonstrable* and *navigable* are loans from French (originally from Latin), and that form is one learnt with the words which illustrate it. The normal pattern in English, though, is to add suffixes without truncation. So we have two competing patterns: an inherited one with no *-ate* (when viewed from the English point of view) and a productive one where we expect to retain the *-ate*. Which is stronger in any given case may depend on a number of factors (including, for instance, how phonologically or semantically similar the new form is to one of the models we have already learnt), and can vary from person to person. *Educatable* is a later formation than *educable*, which is a loan. To the extent that *educable* has any English influence on it, it is influenced by parallel loans from French and Latin; *educatable* shows all the signs of being a genuinely English formation. I would predict that newer formations would retain the *-ate*, but this might not apply if the coiner is aware of French or Latin precursors, or if a number of loan words happen to provide good models for a new formation.