Principles of English spelling formation

Thesis submitted in partial fulfilment
of a Doctoral degree in Linguistics

Trinity College Dublin
31 July, 2017

Des Ryan
11265114
Declaration

I declare that this thesis has not been submitted as an exercise for a degree at this or any other university and it is entirely my own work.

I agree to deposit this thesis in the University’s open access institutional repository or allow the library to do so on my behalf, subject to Irish Copyright Legislation and Trinity College Library conditions of use and acknowledgement.

______________________________
Abstract

Studies of English spelling have primarily focussed on correspondences between spelling and sound among core, standard spellings. Segmental-level correspondences have been examined in detail (Venezky 1970, Cummings 1988, Carney 1994, Rollings 2004), while recent work shows that English spellings also encode supra-segmental information (Evertz and Primus 2013, Evertz 2014). An outstanding problem is the degree to which morphemic spelling is applied across the system (c.f. Venezky 1970: 120; Carney 1994: 18). Berg et al. (2014) observe that the spelling of affixes is more stable than the spelling of bases, hence <profane> alternates with <profanity>, not *<profanety> or *<profaneity>. Yet none of these theories address in detail why certain spellings are chosen over others.

This thesis examines how English spellings are formed, particularly where compromises are made in the representation of both morphological and phonological information. The primary focus is on the spelling of inflections, but also derivations, names and non-standard spellings. The study also examines how we know if a spelling is a good phonographic match for a word. Five principles of spelling formation are proposed.

1. Any-Spelling principle: All words must have some spelling.
2. Distinctiveness principle (DSTNCT): Different words should have different spellings.
3. Identity Preservation principle (IdP): Related words should have related spellings. Hence complex words take their spelling from the spellings of their subcomponents.
4. Phonographic Matching principle (PhM): A word’s spelling should represent its phonological form.
5. Invariance principle (INVRNC): Established spellings cannot be changed.

The principles are manifested differently according to the category of word being spelt. Compound words adhere strongly to IdP. Regular inflections follow IdP by default, <jump, jumped, jumping>, but may involve PhM amendments, hence <dope, doped, doping> not <dope, *dopeed, *dopeing>. Affixed derivatives show similar patterns, <mode, modal>, regardless of stress-shifting, <origin, original>, while etymological influences complicate the picture, <possible, probable>. DSTNCT sometimes differentiates lexical homophones, <flour>, <flower>, but it is common among certain names, <Webb>, <Blu-Tac>,
Non-standard spellings violate INVRNC by definition, and the conditions for spelling variation arise in part due to phonological changes. Abbreviations frequently obey IdP, e.g. <a.k.a.>, but clippings may not, <telly>, <fridge>. Etymological spelling is deemed to be an example of IdP, either among unchanged base forms, <macchiato>, or adapted polymorphemic words, <philosophy>. This re-assessment helps to solve the long-standing problem of how to integrate etymological spelling into synchronic theories of spelling. A flexible unit of spelling, the complex pleremic unit, is identified and it accounts for IdP’s various manifestations. Native, monomorphemic base forms are not examined in detail as the priority is how new and complex words are spelt. This method is taken from word-formation studies (Marchand 1969, Bauer 1983, Plag 2003).

The initial model of spelling formation assumes that polymorphic words obey IdP by default, but that PhM amendments can be made, where necessary, if possible. The second iteration builds on Evertz’s (2014) graphematic hierarchy to show how unsatisfactory spellings can be identified by comparing spelling and phonology at all hierarchical levels, including phonemes, syllables, and feet. The third iteration shows how Optimality Theory can be used to explain how one imperfect spelling, e.g. <doped>, can be chosen over an even less perfect spelling, *<dopeed>. Two further applications arise from the model: the pronunciation of many polysyllabic words can be predicted from the spelling with greater accuracy; and a path is provided by which spelling pronunciation can be predicted from potential ambiguities in decoding.
Acknowledgements

It has been a privilege to spend a sizeable chunk of my life engaged in the research that I wanted to do, and this has been dependent on the faith of several people and the financial backing of others. Nevertheless, doctoral studies are notoriously isolating and repetitive, like cycling around the same island for several years, so every scrap of intellectual and personal help has been very welcome.

Breffni O’Rourke, my very pleasant supervisor, believed in me and my project from the beginning and has helped me for many years, through countless conversations and insightful, detailed comments on my work. Vivian Cook has been my long-term co-editor on a complementary project, *The Routledge Handbook of the English writing system*, and I have learnt so much from his vast experience. He too put trust in me from a very early stage. Thanks also to the thirty-plus unpaid contributors who have given so much of their time to that project, and the growth in our understanding of the field. Heinz Giegerich was a great help during my two stints working with him in Edinburgh, as were Patrick Honeybone, Warren Maguire and Meg Laing. Thanks to the many people who have organised conferences or let me give guest lectures; every talk has helped: Beatrice Primus, Martin Evertz, Frank Kirchhoff, Martin Neef, Anneke Neijt, Nanna Fuhrhop; Rhona Alcorn, George Walkden, David Denison. Communities of research wouldn’t exist without such efforts.

Thanks to the people in the linguistics community in Trinity, especially Jenny O’Reilly, and the people in the Long Room Hub, where I was fortunate enough to have a desk for over four years. Shout outs to Fergus Robson, Tim Murtagh, Manus O’Dwyer, Guilherme Braga, Feng Wei, Deirdre Dunlevy, Peter Sheekey, Robin Fuller, Frank Leahy, Mary Stefanazzi, Sarah Barry, Sarah Dunne, and Caitriona Curtis.

Thanks to people who have read and discussed my work in detail at various stages: Jeff Kallen and Heath Rose in Trinity; Jesper Kruse, Robin Fuller, and anonymous reviewers whose fresh criticisms taught me a lot; also to my two external examiners, Lynne Cahill and Mark Sebba, whose recommendations have forced me to dig deeper and get even more out of myself. Most of all, thanks to my sister, Maryanne, who has cast vigilant, doctoral eyes on many drafts of the text, at times when I needed it most.

I am very fortunate to have a large family, most of whom are nearby, and all of whom are regularly on the phone. Thanks to my mam for being such an inspiration and so supportive, to my dad for his constant good humour and work
Thanks to my five big sisters, Catherine, Barbara, Maryanne, Emer and Margaret, whose support grows and grows; to their respective husbands (Mick, Darren, Colm, Gerry and David), and their kids, some of whom are old enough now to take me Christmas shopping: Eoin, Emilie, Eleanor, Niall; Stephen, Laura; Juliette, Darragh, Oisin; Eli, Benjamin and Donnchadh. Thanks to the several people who I’ve lived with, especially Gabe Murphy, Brian Storan, Jack Olohan, Maria Costas, to all my friends in Dublin, especially Graham Kearns, Mob, Sean Flynn, Jasper O’Connor, Imogen Rabone, Mono, Michael McCormack, Nick Leonard, Jamie Blake-Knox, and David White; in Edinburgh: James Maxwell-Penny and Bill Thompson. Thanks to Seán Farrell, Elske Rahill and Phoenix, Fintan and Brocc for putting me up for the summer of 2013. RIP Michel. A lot more people have left Dublin in recent years but it’s good they are finally coming back. I have missed you all. Thanks to the uncountable number of people who have been interested in my work, especially since my early days writing a blog about spelling at a2dez.com, and to all the people who have even asked to read the thesis. That keeps me going.

Thanks to the many great baristas working in Dublin’s growing coffee scene, in the old 3FE, Vice, Cup, Angel, Kaph, Two Pups and, latterly, in Sligo’s Sweat Beat. Thanks to the growing army of podcasters who have kept me entertained for years: Second Captains, Radiolab, The Comedian’s Comedian, An Irishman Abroad, Hardcore History.

Huge thanks are due to the Irish Research Council, who gave me a full scholarship for three years, and to Dublin City Council who paid my fees in my first year, a grant that has sadly been abolished. Language teaching work has also been a big part of my life, so thanks to both my employers and to the endless flow of students who have taught me about the English language from a different perspective. Thanks also to all the medical people who have helped me in recent years. My knee surgeon Ray Moran, my physiotherapist and de facto counsellor, Andrew Watson. Recovering from a cruciate ligament injury was good mental preparation for doing a PhD. These things don’t fix themselves. A special thanks goes to the backroom teams in Trinity who do so much unseen work. To Kieran Lewis my occupational therapist, to Annemarie Naughton and Louise Hamilton at student counselling. To the library staff for all their help, to Helen Thornbury and other administrative staff. Thanks to Spiro Jamie for telling me about Irlen syndrome, to Alison Crowley for my diagnosis, and to Aislinn Ryan for helping me find my tint. My life is so much better with my coloured lenses. I hope we can all help more people. Light sensitivity is a subtle but profound problem which
affects everything from your mood to your relationships to your ability to read and write. I would never have finished this project without my Irlen glasses.

The greatest thanks goes to Sharifa, whose love, support and patience has been unending, and may our love last forever. Here’s to many fun times ahead, along with our newborn baby, Valentine Amon.

***
Conventions and abbreviations

*car*

- The word *car* is under discussion

<car>

- The spelling of the word is under discussion. Illegal or incorrect spellings are preceded by an asterisk, hence *<kar>*. Putative spellings have a question mark, ?<ka>.

/kɑː/, /kar/

- Phonemic forms. This particular word differs between British *Received Pronunciation* (transcribed here first) and *General American* (second). These accents are explained in more detail below.

//kar//

- A deeper ‘underlying form’. Only applicable in reference to certain theories (e.g. Chomsky and Halle 1968) or historical forms underlying different present-day accents.

[kɑː], [kɑr], [kær]

- Phonetic forms.

{car}

- The morpheme ‘car’, as occurs in *car, cars, car tax, carless*, etc.

<VC>

- A sequence of letters involving a vowel letter <a e i o u> and then a consonant letter: e.g. <at, et, ip, ok, uv>.

/VC/

- A phonological sequence involving a short vowel and then a consonant. Long vowels are marked with a colon. The words ‘vowel’ and ‘consonant’ only refer to sounds. Otherwise I talk about ‘vowel letters’ and ‘consonant letters’. On this important distinction, see Carney (1994: 9–11).
<VC₁C₂>

- A vowel letter followed by a consonant letter, followed by another (usually different) consonant letter, e.g. <arp, err, irk, onk, unk>.

#

- Word-final or syllable-final. <VC#> means spellings that end in a vowel letter then a consonant letter (it, fit, fix, quickfix etc.). /VC#/ similarly means words that end in a vowel then a single consonant (e.g it, pit, pick, pitch, tripswitch). It doesn’t matter here how many letters are used to represent the sounds.

§

- Following Carney (1994), this symbol indicates the subsystem to which a spelling belongs, be it §Greek, modern §French, §Italian, etc. The term §Latinate covers the interlocking influence of Latin and older forms of French. §Basic is a native or default system, as many borrowed words show no obvious subsystem markers (e.g. table). By contrast, chlorophyll is §Greek and chaise longue is §French. A catch-all group is known as §Exotic (e.g. Iraq) and a defective group is called §Romance, and seldom used here.

≡

- Correspondence between spelling and sound: e.g. <c> ≡ /k/, or <cat> ≡ /kæt

<table>
<thead>
<tr>
<th>σ</th>
<th>Syllable</th>
<th>g</th>
<th>Graphematic form</th>
</tr>
</thead>
<tbody>
<tr>
<td>On</td>
<td>Syllable onset</td>
<td>FG</td>
<td>Graphematic foot</td>
</tr>
<tr>
<td>Rh</td>
<td>Syllable rhyme</td>
<td>WG</td>
<td>Graphematic word</td>
</tr>
<tr>
<td>Nu</td>
<td>Syllable nucleus</td>
<td>()</td>
<td>Foot boundary</td>
</tr>
<tr>
<td>Co</td>
<td>Syllable coda</td>
<td>.</td>
<td>Syllable boundary</td>
</tr>
<tr>
<td>p</td>
<td>Phonological form</td>
<td>'</td>
<td>Primary stress</td>
</tr>
<tr>
<td>Fr</td>
<td>Phonological foot</td>
<td>'</td>
<td>Secondary stress</td>
</tr>
<tr>
<td>Wp</td>
<td>Phonological word</td>
<td>?</td>
<td>spelling candidate</td>
</tr>
</tbody>
</table>
OED (32395)

- This is a new convention used for referencing information taken from the *Oxford English Dictionary*, an invaluable source of information about words, including their spellings and their histories. Studies of spelling tend to be littered with mentions of the ‘immeasurable’ contribution of the OED. While this may be true to a certain extent, I have tried to fully acknowledge their input, where possible, and more importantly, provide readers with a follow-up reference. In the text I have introduced a convention of including the entry number from the *OED3* website, meaning the third and most recent form of the dictionary, available online. For example, the reference to the history of the spelling <chowder> is referenced as (OED 32395), and the reader can go to the following URL to find it: [http://www.oed.com/view/Entry/32395](http://www.oed.com/view/Entry/32395)

RP

- Received pronunciation. Commonly thought of as a British prestige pronunciation, the accent often serves as a model for others beyond Britain, and to language learners, so I have kept this term, following Wells (2000). I have thus avoided the newer term *Southern British Standard* (SBS) used by Carney (1994) and Kruse (2016), on the grounds that SBS is too region-specific. The RP vowel system is shown in Table 6.1.

GA


Where possible, I have chosen examples that don’t bring up discrepancies between the major reference accents. Where necessary, my transcription errs on the side of rhoticity, partly because English spelling encodes a rhotic accent but also because it is easier to construct a non-rhotic form from a rhotic transcription, than the reverse.

Finally, terminology is generally introduced in italics but discussions over choice of terminology use inverted commas.
# Contents

1. Introduction ............................................................................................................. 1

   1a. Background: how did English spelling become so complicated? .................. 2
      1a.1 Complications at the phonemic level ....................................................... 2
          1a.1.1 Phonemic spelling .............................................................................. 2
          1a.1.2 Compound spelling units ................................................................. 2
          1a.1.3 Multiple spellings per phoneme ....................................................... 2
          1a.1.4 Invariant spellings, changing phonologies ...................................... 3
      1a.2 Complications at the morphemic level ..................................................... 4
          1a.2.1 Compound words ............................................................................. 4
      1a.3 Creative spelling ....................................................................................... 5

   1b. Aims, methods, and structure of thesis ......................................................... 7
      1b.1 Methods ...................................................................................................... 7
      1b.2 Structure of thesis ..................................................................................... 7

   1c. Definitions ....................................................................................................... 11
      1c.1 Phonemes and phonology ....................................................................... 11
      1c.2 Morphemes ............................................................................................... 11
      1c.3 Spelling ..................................................................................................... 11
      1c.4 Spelling units and letters ......................................................................... 12
      1c.5 Writing and writing systems .................................................................... 12
      1c.6 Writing systems and scripts .................................................................... 13
      1c.7 The English writing system ...................................................................... 14
      1c.8 Orthography and graphematics ............................................................... 15
      1c.9 Spelling formation ..................................................................................... 16

Chapter 2 Literature Review ................................................................................. 17

   2a. English among the writing systems of the world ........................................ 18
      2a.1 Gelb: a three-way typology of writing systems ...................................... 18
      2a.2 Sound- and meaning-based systems ...................................................... 20
      2a.3 DeFrancis: all writing systems are sound-based ..................................... 22
      2a.4 Sproat’s two-dimensional classification .................................................. 24
      2a.5 Morphography, English spelling and complex pleremic units of writing . 25
          2a.6 Conclusion .............................................................................................. 28

   2b. Approaches to English spelling ................................................................. 30
      2b.1 Reading research ....................................................................................... 30
      2b.2 Historical studies of English spelling ...................................................... 30
      2b.3 Spelling reform ........................................................................................ 31
      2b.4 Major studies of English spelling: aims, methods and scope ............... 34
          2b.4.1 Venezky’s mappings from spelling to sound ...................................... 34
          2b.4.2 Carney’s bi-directional analysis ......................................................... 36
          2b.4.3 Cummings and Rollings: sound-to-spelling accounts ....................... 37
          2b.4.4 Graphematic approaches to English spelling ................................. 38
2b.4.5 Synchrony and diachrony ................................................................. 39

2c. Major research themes and findings ................................................... 41
2c.1 Keeping the spelling of the morpheme constant .............................. 41
  2c.1.1 Morphophonemic spelling ......................................................... 42
  2c.1.2 Lexical spelling ........................................................................... 42
  2c.1.3 Criticisms of the lexical spelling hypothesis ............................ 44
  2c.1.4 Further examples of morphemic spelling ................................. 44
  2c.1.5 The impossibility of a fully morphemic writing system .......... 45
  2c.2 Correspondences between spelling and sound ............................. 47
  2c.3 Etymology and the subsystems of English spelling .................... 51
  2c.4 Spelling patterns, markers and graphotactics .............................. 55
  2c.5 Conclusion: Outstanding issues in the study of English spelling formation ................................................................. 58

Chapter 3 A method for analysing spelling formation ...................... 60
  3a. Scope and focus of the study .......................................................... 61

  3b. A synchronic-diachronic approach to English spelling formation ...... 63
    3b.1 English word-formation and English spelling ......................... 64
    3b.1.1 New words and productivity ............................................... 64
    3b.1.2 Spelling formation .............................................................. 65
    3b.1.3 Inflection ............................................................................... 66
    3b.1.4 Other categories of English word-formation ...................... 67
    3b.1.5 Names .................................................................................. 68
    3b.1.6 Etymological spelling ......................................................... 68
    3b.2 Written forms not under discussion ...................................... 68
    3b.3 Other limits to the study ........................................................... 70

  3c. Data collection .................................................................................. 71

  3d. English spelling as a self-reorganizing system ............................ 74
    3d.1 The ever-changing writing system ........................................... 74
    3d.2 Berg and Aronoff’s method for studying spelling formation ...... 75

  3e. Principles of English spelling formation ...................................... 79
    3e.1. Precursors to the principles ................................................. 79
      3e.1.1 DeFrancis’s universals ....................................................... 79
      3e.1.2 Cummings’ demands ........................................................ 80
    3e.2 The principles ............................................................................ 81
      3e.2.1 Distinctiveness principle (DSTNCT) .................................. 82
      3e.2.2 The Identity Preservation principle (IdP) .......................... 83
      3e.2.3 Phonographic Matching principle (PhM) .......................... 83
      3e.2.4 Any-spelling principle ..................................................... 84
      3e.2.5 The Invariance principle (INVRNC) .................................. 86

  3f. Conclusion and hypothesis ............................................................ 88

Chapter 4 The interface of IdP and PhM ........................................... 90
  4a. Inflection in English ................................................................. 92
      4a.1 Regular suffixed inflection ................................................... 92
7. The Distinctiveness principle (DSTNCT) .........................................261

7a. DSTNCT among lexical spellings .................................................264
   7a.1.1 Semantic specialisation ..................................................264
   7a.1.2 The three-letter rule ....................................................264

7b. DSTNCT in the spelling of traditional names ................................267
   7b.1 Capital letters: systematic DSTCNT spelling ..........................267
   7b.2 Surnames ...........................................................................268
   7b.1.1 Constraints on the spelling of surnames ...............................269
   7b.1.2 <i> ➔ <y>, <y> ➔ <ie> ....................................................269
   7b.1.3 CLD-padding ...................................................................269
   7b.1.4 <e>-padding .....................................................................270
   7b.1.5 <VC1C1e> .........................................................................271
   7b.1.6 <VC1C2e> .........................................................................271
   7b.1.7 <Vye>, <Vve>, <VyCe> and <VwCe> .....................................272
   7b.1.8 <vvCe> ............................................................................272
   7b.1.9 What padding tells us about the nature of decoding ..........272

7c. Categorising non-standard spelling .............................................275
   7c.1 Dialect literature .....................................................................275
   7c.2 Other non-standard spelling strategies ....................................277
   7c.3 Semantic functions of non-standard spelling .........................278
   7c.3.1 Constructed homophony ...................................................278
   7c.4 A taxonomy of non-standard spelling ....................................280
   7c.5 Linguistic exaptation .............................................................282

7d. The spelling of American trade names .......................................284
   7d.1 Jacobsen ..............................................................................284
   7d.2 Praninskas ..........................................................................286

7e. The spelling of band names .........................................................289
1. Introduction

Research on English spelling has focused on either standard or non-standard spellings, and studies tend not to look at the commonalities underlying all kinds of English spelling. Two studies of standard spelling are of the utmost importance: Venezky (1970) tabulates the spelling-to-sound correspondences of ‘the 20,000 most common words’ of English, based on Thorndike’s (1941) wordlist, and he emphasises the fact that phonemes and morphemes share equal importance in spelling; Carney (1994) explores correspondences between letters and phonemes in both directions — like Venezky, he looks at how spellings are decoded but he also looks at how sounds are encoded. Cummings (1988) and Rollings (2004) study the encoding of each phoneme (of GA and RP, respectively), and their work has placed greater emphasis on the letter patterns themselves, a focus that has influenced very recent graphematic approaches to English spelling (e.g. Evertz and Primus 2013; Evertz 2014; Berg, Buchmann, Dybiec, & Fuhrhop 2014; Berg 2016), where major theoretical steps have been made.

All of the above studies share some very important assumptions. They view English spelling as having a large but static set of core spellings and other spellings are omitted, including rare words and wordforms, names, abbreviations, variant spellings and non-standard spellings. In the earlier studies, inflections were also omitted, including {ing} and {ed} forms (e.g. jumping, jumped), on the assumption that their spellings were predictable, although some recent studies have begun to address this problem (e.g. Berg 2016).

In this study, these outstanding groups are prioritised because they reveal a lot of information about how the writing system functions. Established spellings tend to preserve ancient patterns which may no longer be productive, and yet they cannot be changed. By contrast, the spelling of new words and names are subject to the formation principles of the present-day writing system. Furthermore, new and creative spellings are limited to patterns which extrapolate from core patterns — otherwise we would not be able to understand them. As a result, evidence from the periphery can teach us about the inner workings of the writing system, and this thesis brings together the ideas of several studies of peripheral spelling groups (e.g. Jacobsen 1966; Praninskas 1968; Preston 1982, 1985; Attridge 1988; Carney 1994: Ch. 6; Androutsopoulos 2000; Sebba 2007; Ryan 2010; Jaffe 2012; Honeybone and Watson 2013; Picone 2016).

Before looking at the aim and structure of this thesis, we shall begin by examining several reasons why English spelling has become so complicated over the centuries.
1a. Background: how did English spelling become so complicated?

1a.1 Complications at the phonemic level

1a.1.1 Phonemic spelling
One might imagine that spellings ought to represent the sound of words as unambiguously as possible, and this idea is best expressed by the phonemic principle, whereby there is a one-to-one set of mappings between letters and phonemes (see, for example, Pike 1947: 208). This occurs more frequently in English, notes Carney (1994: 32) than is often believed. Words like cat, sat and mat contain no obvious mismatches between their letters and their sounds: in each case, the three letters represent three phonemes (e.g. <c> ≡ /k/, <a> ≡ /æ/ and <t> ≡ /t/).

By the logic of the phonemic principle, the spelling of the word /tæk/ should therefore be *<tac>, because /k/ ought to always be spelt with <c>. However, monosyllabic words ending in /k/ generally end in <ck>, rather than <c>, for reasons that are not immediately obvious (see Section 4c.1.2). This simple example shows how English spelling quickly departs from any ideal of phonemic spelling, and a phoneme’s spelling may be determined by its position in a word.

1a.1.2 Compound spelling units
A spelling such as <chat> brings in a slightly different kind of complication because two letters, <c> and <h>, are used to represent just one phoneme /tʃ/. The same goes for the <th> and <sh> of that and shat (/ð/ and /ʃ/, respectively). In these words, the letter <h> is used not for its direct phonetic value of /h/ but as a diacritic, or marking device, which changes the representational value of <c>, <t> and <s>. This is typical of alphabetic systems, using two or more symbols together to represent one sound: French chat (‘cat’) has <ch> ≡ /ʃ/ while Irish chuaigh (‘went’) may have <ch> ≡ /x/.

1a.1.3 Multiple spellings per phoneme
Deviations from the phonemic principle can also result in different letters or clusters of letters being used for the same phoneme. For example, /k/ is spelt variously in almanac, anorak, rack, raccoon, Akkadian, arachnid, saccharin, khaki, Iraq, baroque, acquire and lax. At first, this can appear chaotic but deeper
analysis reveals a great deal of systematicity on various levels (see Carney 1994: 216-223).

**Relative position**

For example, <k> is used before <e>, <i> and <y>, as in *kettle, kitten* and *Kyle* and <c> before <a>, <o> and <u>, as in *cat, cot* and *cut*. This is because <c> $\equiv$ /s/ when followed by <e>, <i> and <y>, as in *cetacean, city* and *cygnet*. The reasons for this were once phonetically conditioned but now they are preserved (somewhat) consistently in the spelling. A spelling pattern that once represented a phonological distinction has mutated into being a design feature of the writing system, so a new(ish) word such as *cyborg* will be automatically read with <c> $\equiv$ /s/. Relative position of letters is thus a crucial factor in their reading.

**Etymology**

Much of this variation results from the different etymological inputs to English. *Baroque* takes its spelling from modern French (OED 15685) while *almanac* and *acquire* were influenced by earlier French and medieval Latin (OED 5564 / 1731); *arachnid* and *saccharin* come ultimately from Greek, again by way of French and Latin, whose writing systems also preserve the Greek etymology. The contrast between <ch> and <cch> derives ultimately from a distinction in Greek, even though there is no phonological difference in English (OED 169419). The <kh> of *khaki* and the <q> of *Iraq* were also chosen to indicate phonological distinctions in the donor languages, both of which are adapted as /k/ in English (OED 103170 / 99393). The same goes for the Anglicisation of foreign place names such as *Qatar* and *Kathmandu*, while the <k> of *anorak*, the <cc> of *raccoon* and the <kk> of *Akkadian* may have been attempts to achieve a similar effect (OED 8076 / 157186 / 4472). While the spelling of /k/ shows greater variation than any other phoneme (Carney 1994), the examples show how English spellings are complicated by the long-standing practice of retaining the spelling of borrowed words.

**1a.1.4 Invariant spellings, changing phonologies**

Once settled upon, English spellings are seldom altered to reflect changes in phonology, and this causes mismatches between spelling and sound that widen over time. Furthermore, the changes can vary from one accent to another, adding yet more layers of complexity. For some accents, the spelling of *whine* and *wine* reflects a distinct pair of consonants, /hw/ and /w/, but in most accents (nowadays) this distinction is either lost or recessive, and both words are pronounced identically with /w/ (Kruse 2016: 181-2).
In other cases, a sound change will (have) occur(red) in all accents but not in all phonological environments. For example, the word *watch* rhymes with *botch*, a change that has occurred in most words with historical /wæC/, where /C/ indicates a consonant, as in *wasp, waffle, swan, quality*. This sound change did not take place before the velar consonants, /k/, /ɡ/ and /ŋ/: in such words /æ/ remains, as in *wag, whack, swanky* and *quango*. Such phonemic splits can result in words sounding very different but keeping related spellings, hence *was* and *wag*. The reverse can also happen: when two phonemes merge, the writing system ends up with two spelling units for the same sound, as in *meet* and *meat*.

1a.2 Complications at the morphemic level

1a.2.1 Compound words
A very simple complication emerges in the spelling of compound words, which are formed by concatenating two existing words to make a new word. Generally, English compounds are spelt by agglomerating the two spellings, as in *catwalk* and *catnap*, but this practice leads to potential confusion in the word *cathouse*, because <th> is also a special unit of spelling (corresponding to the /θ/ of *think* and *thank*, or sometimes the /ð/ of *this* and *that*). Hence an ambiguity arises between bi-morphemic *cathouse* /ˈkæθaʊs/ and mono-morphemic *catholic* /ˈkæθəlk/.

Understanding how to read and spell thus requires knowledge of the internal structure of words, especially its morphemes (defined in Section 1c.2). The reader must recognise that <cat> and <house> are different morphemes in order to know how to pronounce the compound word. In some cases, the subunits have become opaque, so one might not immediately know that the placename *Chatham*, has two subcomponents <Chat> and <ham>, corresponding historically to /θ/₁, rather than /θ/₂. A further complication is that the /h/ is naturally dropped in connected speech, leaving just the /t/ of /ˈtʃætəm/. The pronunciation /ˈtʃæθəm/ is derived from the spelling.

1a.3 Other polymorphemic words
The fact that English spellings tend to represent both morphological and phonological information (Venezky 1970) adds a great deal of complexity to the writing system because a single spelling is often used to represent two or more pieces of linguistic information. In *cats* and *dogs*, the same spelling <s> is used to represent different sounds, the unvoiced /s/ of /kæts/ and the voiced /z/ of /dɒgz/. Here we see English prioritising the consistent spelling of morphemes over the
consistent spelling of phonemes, a practice that is fundamental to the workings of the writing system.

Another important design feature is that letters can be added or subtracted to clarify a potentially ambiguous pronunciation. In the word *catty*, a `<t>` is added to mark the `<a>` as a short vowel (/kæti/). This is in contrast to the historically long vowel sound in the names *Caty* and *Katy*, usually pronounced nowadays as /ketɪ/. Other changes only involve letter-sequencing: the `<y>` of *catty* changes to `<i>` in *cattily*, and *carry* changes to *carries*.

An associated feature of English spelling, therefore, is that it tends to maintain related spellings in related words, often at the expense of marking the pronunciation clearly. For example, `<a>` is present in both *democracy* and *democratic*, even though there are two very different vowel sounds in the penultimate syllable of each of these words, unstressed /ə/ and stressed /æ/, respectively. The word *hypocrisy* has a similar phonological ending to *democracy*, but evidence of the former’s `<i>`, as opposed to the latter’s `<a>`, can be garnered from the related word *hypocritical*, where `<i>` is a likely spelling of the stressed /ɪ/.

### 1a.3 Creative spelling

One advantage of having such a complicated writing system is that it becomes possible to use spelling for creative purposes. In a phonemic writing system, there is no opportunity for substitution. In English, however, there are many ways in which spelling units can be swapped in order to represent new meanings. One simple way of doing this is to make a direct substitution, something that happens in non-standard spelling (e.g. `<woz>` for `<was>`) and it is common in names. *Kit Kat* and *Kopikat* have spellings which differentiate them from the lexical counterparts. Creativity can also include symbols which don’t occur in normal spelling. The names *K9*, *Altern-8* and *C@ in the h@* are pronounced like *canine*, *alternate* (verb) and *cat in the hat*. Such symbols are a common feature of the abbreviated forms used in text messages (Crystal 2008, Shortis 2016) and increasingly in other domains of writing, especially computer-mediated communication (see Squires, 2016 for a review). This creativity has prompted a renewed interest in a topic — writing systems — which had become marginal in linguistics. A static and uncreative system is not very interesting, whereas a changing system with the potential for creating new meanings is interesting both for users and anyone curious about the system’s inner workings.

There are many other ways in which creativity occurs in English spelling, mostly in peripheral domains. Variant spellings can be used to represent regional pronunciations (e.g. `<wint>` for `<went>`), or to represent non-linguistic sounds
(<miaow> or <mrkrgnao> for a cat), or to blend words in ways that can only be seen but not heard (e.g. *funferal*), and so on.

***
1b. Aims, methods, and structure of thesis

The primary aim in this thesis is to develop a method for analysing how English spellings are formed, and we shall see that all spellings are united by the need to represent similar information, notably the sound and meaning of words, even though certain word groups tend to prioritise particular information. Additionally, some very important theoretical problems will be addressed, such as how we know if a spelling is a good match for a word, and how can spelling pronunciation be explained.

1b.1 Methods

The methodological approach mixes established methods for studying English spelling along with methods taken from word-formation studies. In that subfield, the focus is on the formation of new and unestablished words because they reiterate the system’s productive patterns, and this helps to explain how the overall system functions. The same rationale applies to this study. If we examine the spelling of new words, we can understand the writing system better.

At first, the plan was to study each category of word-formation in turn. In time, it became clear that the spelling of inflections provided the best insight into the spelling formation, as the competing influences of phonology and morphology can be clearly seen therein. These words are used to first develop a model of spelling formation and this is applied, in later chapters, to other categories of word formation. The analysis is not based on a single corpus of spellings and no effort is made to enumerate the frequency of patterns. It was felt that the most pressing need in the field is to understand basic spelling formation and it is hoped that the insights developed here can later be applied to a large corpus of English spellings in order to create a fuller description of the system.

1b.2 Structure of thesis

The analysis begins by looking at where English fits among the world’s writing systems (Section 2a). This section allows us to see that writing systems employ several strategies for representing words, and that there is, in principle, no reason why this cannot apply to subgroups of English. This section also identifies a very important unit of spelling, known here as the complex pleremic unit (CPU), which has seldom been identified in writing-systems typologies and has not been applied at all to English. Section 2b looks at why others have studied English spelling. Of most relevance are historical studies (e.g. Scragg 1974, Upward and Davidson
and the reform movement, which has ebbed and flowed for half a millennium. Studies of these two topics have heavily influenced modern theories of spelling, and thus they are of direct importance. Of less relevance is research on literacy acquisition, as it accepts the writing system as it is and asks how learners cope with its inherent difficulties.

Section 2c forms the main body of the chapter and it looks at linguistically-informed theories of English spelling, most notably the corpus-based studies conducted by Venezky (1970) and Carney (1994), as well as the work of Cummings (1988) and Rollings (2004). Here we shall see the efforts at describing the main set of issues — morphemic spelling, etymological spelling, letter-to-phoneme correspondences, and graphotactics, which looks at letter patterns. Graphematic theories of English spelling are introduced in Chapter 2, but explored in greater detail later in the study, where relevant (esp. Evertz 2014, Berg et al. 2014, Berg 2016, Berg and Aronoff 2017). A number of outstanding problems are identified and put forward as questions for examination throughout the study.

Chapter 3 looks at the methods used in this study, the data examined, and five principles of spelling formation are introduced. The Identity Preservation principle (IdP) shows how related words are connected within the writing system, while the Distinctiveness Principle accounts for how unrelated words are disconnected. These two principles refer directly to spelling-formation patterns. The Phonographic Matching principle recognises the importance of having spellings which represent phonological forms as unambiguously as possible, and this is a much more complicated task. The other two principles are existential in nature. The Invariance principle accounts for the stability of each word’s spelling and the Any-Spelling principle acknowledges the fact that words must have some spelling (although not all spellings represent words). At the end of Chapter 3, these principles are incorporated into a very rudimentary model of English spelling formation to be developed throughout the study. The remainder of the thesis is structured around four of the principles, while the more trivial, Any-Spelling principle is discussed briefly in Section 3e.2.4.

Chapter 4 tests the model of spelling formation by focussing on the interaction of IdP and PhM in the spelling of regular English inflections. These word forms allows us to examine the simplest sort of spelling formation because inflections are highly productive and not influenced by potentially confusing, phonological influences (e.g. stress-shifting).

Chapter 5 explores and develops some theoretical issues arising in the model of spelling formation, especially in relation to PhM. This is the central
chapter of the thesis. One issue which has never been sufficiently theorised before is how we know that a spelling is a good match for a word’s phonological form. This chapter explores supra-segmental English phonology (Giegerich 1992) and Evertz’s (2014) supra-segmental analysis of English spelling, what he calls the graphematic hierarchy. Two theoretical models are introduced to test the efficacy of a spelling. The first is a visual model for comparing the phonological and graphematic hierarchies, allowing us to identify mismatches between spelling and sound at any level of analysis up to the word (e.g. the phoneme, the syllable, sub-syllabic elements, and the foot). The second model is presented within Optimality Theory (OT) and its power is potentially greater than the visual model because it shows how one imperfect spelling can be chosen over an even less perfect spelling. Chapter 5 also makes steps towards explaining some outstanding theoretical problems in the field. We shall see how the spelling of polysyllabic words can be more easily decoded, by breaking them down into stressed and unstressed syllables. Finally, spelling pronunciation, a well-known but under-theorised phenomenon, will also be explained using the theoretical machinery developed herein.

Chapter 6 takes a different direction and looks at the relationships between invariant spellings and changing phonologies. The discussion looks at the effects of orthographic inertia on the correspondences between writing and speech. Instead of seeing all gaps between spelling and phonology as bad, it is argued that some phonological changes favour the reader (e.g. phonemic mergers), while some favour the writer (e.g. phonemic splits). Phonological rotations, by contrast, have no effect on the structural relations between letters and phonemes, a fact that helps to explain why English spellings have continued to function for so long across such a wide range of accents.

The complexity of the writing system can have some positive side effects, as we shall see in Chapter 7, because it paves the way for creative spelling. The fact that <ee> and <ea> nowadays represent the same sound /i:/ means that the two spelling units are interchangeable in domains of spelling which allow users to alter spellings (e.g. beetle and The Beatles). Such substitutions frequently occur in non-standard spellings and names, where there is a strong urge to create distinct orthographic identities. Chapter 7 also includes an updated taxonomy of non-standard spellings, building on those put forward by Androutsopoulos (2000), Sebba (2007) and Jaffe (2012). It is also argued that distinctive spelling is a fundamental design feature of the writing system, meaning that there is a strong desire for each lexical item, or name, to have its own distinct spelling, even though this situation is not always practical.
Finally, Chapter 8 shows how the model of spelling formation can be applied to a broader set of word-formation categories. There are two case studies of affixed derivation. These show the effects of etymological spelling on the system, and how it can be fully incorporated into a theory of English spelling, something that previous researchers have struggled to do. Other word groups are examined briefly (e.g. clippings, abbreviations, compound words, neo-classical compounds, blends), and the aim here is to show that these categories merit being studied under the same umbrella as the more central groups (e.g base forms, inflections, affixed derivations).

In short, we shall see that there is a great deal of systematicity across all kinds of English spelling, from long-established words to the most ephemeral of non-standard spellings.

***
1c. Definitions

1c.1 Phonemes and phonology
It will be assumed that the reader has a basic knowledge of segmental English phonology, and thus understand why /kæt/ is phonemic transcription for the word *cat*. It will also be assumed that readers will understand why the word *car* can be transcribed as /kɑː:/ in RP, but /kɑːr/ in GA. Giegerich (1992) provides an excellent introduction to English phonology, both segmental and supra-segmental, and he also explains the basic structural differences between RP, GA and Scottish Standard English. Wells (1982) remains the standard text on different English accents, and it is heavily referenced in the present work, especially in Chapter 6. In Section 5b, Giegerich’s model of supra-segmental phonology (e.g. syllables and feet) will be presented as background material for the spelling theory developed throughout the remainder of that chapter.

1c.2 Morphemes
A morpheme has traditionally been defined as the smallest meaning-bearing unit, and it can either be a whole word or a meaningful subcomponent of a word — *unhappy* has two morphemes, {un} and {happy} (see Plag 2003: 12). A morpheme is, however, a unit of language, rather than a unit of writing, and it is possible for its spelling to be manifested in different ways, according to the rules of the writing system. This occurs in <unhappiness>, where the <y> of *happy* has been replaced by <i>, with no effect on the spelling-to-sound relations.

One important distinction among morphemes is between free and bound morphemes. Free morphemes can exist alone as whole words (e.g. *happy, sad, mother*), whereas bound morphemes cannot: {un-}, {-ness}, {-less}, {-er}, {-s}, {-ly}. A rather different kind of bound morpheme (in English) occurs in ‘cranberry’ forms such as {cran-}, {matern-} and {poss-}, which cannot occur alone but only in conjunction with other morphemes, as in *cranberry, maternal, maternity* and *possible, possibility*. This is particularly common in §Latinate vocabulary.

1c.3 Spelling
A *spelling* is the invariant string of symbols used for representing a word in writing. Most English spellings are comprised of the twenty-six letters of the English alphabet, although some spellings contain capital letters or hyphens, as in *Turkey* and *multi-coloured*. In names, spellings might contain other symbols such as numerals or punctuation, as in *Altern-8, D*REAM, *dEUS, U2* and *R.E.M.*
While names show a greater range of symbols, they also tend to be invariant. Despite the tendency for each word to have one invariant spelling, there may be some variation, so that both <medieval> and <mediaeval> are variant spellings of the same word.

This study does not examine *mis-spellings*, which occur when a particular writer tries to spell a word but does so incorrectly. The definition of incorrect is dependent upon whose standard is being held up as correct. In the absence of a recognised authority for English, we may wish to defer to the makers of established historical dictionaries who gather variant spellings and make decisions about which variant to use as the headword. For the most part, the OED has been used as the primary reference in this work but that is not always enough. There is obviously some variation between British and American spellings (see Cummings 2016), but we shall see that many words continue to have variable spellings due to uncertainty over which information ought to be represented (esp. Chapter 4). We need to also be aware of the difference between a word’s spelling (i.e. the ‘correct’ spelling, or set of variants in use) and the actual *act of spelling a word*, be it correct, incorrect, or deliberately altered to create a linguistic or sociolinguistic effect, as often happens in non-standard spelling (see Section 1c.8 below).

### 1c.4 Spelling units and letters

This term is used to describe any functional *phonograms*, such as <t>, <c>, <h>, <th>, <ch> and <tch>, all of which consistently represent phonemes. The term *letter* is used where helpful. The sequence *<htc>* is not a spelling unit, even though it happens to be a name of a phone company. We shall see that the distinction between units of spelling for representing sound and units of spelling for representing meaning is not quite clearcut. I have avoided the contentious term ‘grapheme’ (see Kohrt 1986; Ryan 2011, 2015), which can either mean a single letter in all its abstract forms or a functional spelling unit within the writing system.

### 1c.5 Writing and writing systems

A common way to define *writing* has been to define it solely in relation to spoken natural language.
Writing is defined as a system of more or less permanent marks used to represent an utterance in such a way that it can be recovered more or less exactly without the intervention of the utterer.

Daniels (1996: 3)

The quote comes from Daniels’ review of the literature on writing systems, in his introduction to the voluminous collection *The world’s writing systems* (Daniels & Bright 1996). For Daniels, writing is intimately bound up with speech, so it does not include pictography and all of ‘those graphic expressions that do not reflect the sounds of the language’. This definition therefore excludes the precursors to writing (discussed by Gelb 1963) but also very modern innovations such as smileys or emojis which have become very common since Daniels provided this definition. Also excluded are notation systems used to represent music, dance and mathematics (although these are discussed at the end of Daniels and Bright’s collection). Sproat (1996) laments the absence of computer languages from the collection, as this too could be considered to be a form of writing. Harris (1996, 2000) emphasises the importance of graphic dimensions in writing (i.e. shape, size and position), while Ryan (2015) adds colour to this.

The definition of writing in use for present purposes will be closer to Daniels’ definition which focuses on the relations between abstract linguistic units (such as phonemes, syllables and morphemes) and the corresponding written forms (such as phonograms, syllabograms and morphograms). However, it is perhaps better to think of these relations as a *writing system*, a system of inscriptions used for representing language. This frees up the term *writing* to include the much broader set of written forms that interested Harris.

1c.6 Writing systems and scripts

The terms *writing system* and *script* are frequently confused or interchanged. By ‘writing system’, I mean the sets of relations or correspondences between units of writing (e.g. letters and words) and units of language (e.g. phonemes, syllables, morphemes and words). By ‘script’, I mean the set of symbols available for use in a particular writing system. For example, English and Spanish both use variants of the same script, the Roman alphabet, but they have different writing systems — Spanish represents sound in a more phonemically transparent way than English. By contrast, English and Russian use different scripts (the Roman and Cyrillic alphabets, respectively) but their writing systems share similarities in how they represent morphemic information (Sproat 2000: 78). It is possible for a language to be written with more than one script, so Serbo-Croatian can be written in the Cyrillic or the Roman alphabet, while Turkish switched from the Arabic script to
Roman, in the early-mid twentieth century. For an introduction to the world’s writing systems, see *The Blackwell encyclopedia of the world’s writing systems* (Coulmas 1996); for a more up-to-date review of the typologies of writing systems, see Joyce and Borgwaldt (2011).

It should be noted that the present use of ‘writing system’ and ‘script’ is not the only way these terms have been used. They have often been used interchangeably in the literature or even with the opposite definitions, although the use of the term ‘writing system’ is the more variable. Coulmas (2013: 17–18) gives two definitions of ‘writing system’. One is similar to my own: “the specific rules according to which the units of the system are interpreted in a given language”. This definition is helpful when talking about a particular language’s writing system, hence the English writing system or the Japanese writing system. The other definition considers it to be an ‘abstract type of graphic system that functions at a particular [linguistic] level’ (*ibid*). This definition is useful when comparing writing systems (e.g. T. Joyce 2015), as it allows one to distinguish between systems that primarily represent morphemes (using *morphograms*) from those that mostly represent syllables (using *syllabograms*), from those that mostly represent phonemes (using *phonograms*). In such a definition, English, Spanish, Latin and Greek can all be grouped together as having a similar kind of writing system, since the basic units, letters, represent phonemes, while other information is expressed less directly.

1c.7 The English writing system

The *English writing system* is a very similar concept to *English spelling* (in general) but the writing system is very different from a single spelling or the actual act of spelling. As a result, the terms ‘English spelling’ and ‘the English writing system’ are often used interchangeably, with the latter having become the more accepted term.¹

A subtle distinction between ‘spelling’ and ‘writing system’ is provided by Scholfield:

[T]he fact that the spelling of English words has ‘not changed’ much in recent centuries disguises the fact that, as a consequence, the English writing system, in the sense of the letter-sound correspondences, has massively changed and become even more complex, while morphemic reference has strengthened.

*(Scholfield 2016: 160)*

In the present study, the term ‘English writing system’ includes the following:
- All correspondences between spelling and sound
- The spelling of all words and morphemes, both free and bound.
- Common spelling variants (e.g. <enquire / inquire>)
- Intelligible, non-standard spellings (see below)

1c.8 Orthography and graphematics

Neef (2015: 716) distinguishes between established, correct spellings and intelligible, recognisable spellings that are not listed in dictionaries. He considers conventional orthography to be synonymous with ‘correct spelling’, in accordance with the norms of the ‘whole writing community’. This is distinguished from systematic orthography, which includes spellings that adhere to the system’s correspondence patterns, but may not actually exist, or may be barred from use. Hence <rite> and <ryte> are possible spellings of the word rite, but only the former is conventionally accepted. Similarly, <rite> cannot be used for the words right or write, even though the systematic orthography would permit it.

The concept of non-standard spelling exists in the space between conventional and systematic orthography. Sebba (2007: 34) observes that deliberately altered spellings can only function as variant or non-standard spellings if the word being spelt is recognisable. Accordingly, the parameters of variation are very strict. A spelling such as <woz> may be a recognisable, and thus possible, variant spelling for the word was, at least in appropriate written domains. However, the spelling <waz> may not be. Non-standard spellings therefore require an orthographic precision which is absent when a word is merely spelt incorrectly.

Notwithstanding Neef’s distinction, I have refrained from using the term ‘orthography’ because of its restrictive historical associations with correctness. The term ‘standard spelling’ replaces ‘conventional orthography’, while ‘writing system’ includes standard spelling and intelligible, non-standard spellings. Nevertheless, I have retained the term ‘orthographic’ — for want of a better word — when an adjective is required to describe spelling, as in ‘orthographic environment’.

In Chapter 5, the term graphematic will be adopted, when discussing, and building upon, work done by Evertz (2014) and others working within graphematic approaches to writing systems. This term is useful for comparing equivalent units of speech and writing, such as phonological syllables and graphematic syllables, etc.
1c.9 Spelling formation

No distinction is usually made between spelling and spelling formation because the focus in the literature has been on describing existing spellings rather than trying to explain how new spellings are formed. To appreciate how different these are, we might compare the jeweller and the geologist. The jeweller might describe a rock in terms of its colour, shape and texture whereas the geologist might be interested in how a particular rock developed its physical features, and how various forces aligned in its formation. This study takes the geologist’s perspective, and tries to understand the interacting forces that have given rise to present-day spellings.

***

***
Chapter 2

Literature Review

The chapter begins by exploring the typology of writing systems (Section 2a). This allows us to see English within a broader context, as its writing system tries to represent the same linguistic information (sound and meaning) as every other fully functioning writing system. Recent typologists (e.g. DeFrancis 1989, Sproat 2000) have come to see the commonalities among writing systems, whereas earlier studies explored their basic differences (Gelb 1952/1963, Sampson 1985). The section concludes by arguing that the English writing system contains a unit of spelling, known henceforth as the complex pleremic unit (CPU), which has seldom been identified in the typologies of writing systems and never in studies focussed on English spelling. The unit is smaller than the written word and larger than the letter (although it can be either of these) and it indicates the connections between related words.

Section 2b provides an overview of the different approaches to English spelling, some of which are more relevant to this study than others. In increasing order of significance, they are: literacy acquisition, historical studies of English spelling (e.g. Scragg 1974, Upward and Davidson 2011), and the theories underpinning attempts at spelling reform. The review continues by discussing the aims, methods and scope of the major studies of English spelling (e.g. Venezky 1970, Cummings 1988, Carney 1994, Rollings 2004, Evertz 2014), where the ever-more detailed descriptions of the writing system have been informed by progress in linguistic theory and, to a lesser extent, progress in the computational technology available for studying large corpora of spellings.

The main part of this chapter (Sections 2c) looks at their findings. The goal is to identify some important gaps in the literature for discussion in the remainder of the thesis. In particular, we shall see that the long-established idea of morphemic spelling (Venezky 1970, C. Chomsky 1970, Carney 1994) has not been clearly defined. Moreover, there has been an insufficient theorisation of the interactions between morphemic spelling and the need to represent both phonological and etymological information.
2a. English among the writing systems of the world

Much of the literature on writing, both popular and professional, has been devoted to the question of whether writing systems like Chinese represent language in a fundamentally different way from obviously phonologically-based systems. Lost in this debate is the fact that the primary purpose of writing is not phonetic transcription, but the representation of words and morphemes; purely phonologically-based systems are arguably merely the most efficient way of achieving this goal.

(Sproat, 1996)

In other words, the representation of meaning is even more fundamental to writing systems than representing sound. Sproat’s words prompt some very important questions for the typology of the world’s writing systems, including:

- How do writing systems differ in their representation of sound and meaning?
- What are the different linguistic units that can be represented in writing?

There has been a handful of typological studies on the world’s writing systems, published intermittently over the past sixty-five years (e.g. Gelb 1952/1963; Haas 1976, Sampson 1985, DeFrancis 1989, Sproat 2000). This section examines how typologists have identified the various ways in which sound and meaning are represented across the world’s writing systems, and a new unit of writing, which has eluded most typologists, will be defined, enabling us to describe English spelling formation with greater simplicity.

2a.1 Gelb: a three-way typology of writing systems

The foundations for the linguistic study of writing systems were set by Ignace Gelb, often called ‘the father of grammatology’, who argued that the history of writing systems has been marked by ‘three great steps by which writing evolved from the primitive stages to a full alphabet’ (Gelb 1963: 203). The first step is from *pictography*, where ideas are represented by pictures, to *logography* where the pictures have been conventionalised as symbols representing whole words.
This first step occurred in early Chinese writing, where many characters have clear pictographic origins. Gelb does not consider Chinese to be a ‘pure’ logographic system because that would require every single word to have its own character. Instead, characters are recycled for their sound value and new characters are formed by combining two existing characters, one which represents an approximation of the word’s sound and the other which represents its semantic category (see DeFrancis 1989). For example, imagine the (complex) character for bee involved the character for the word be attached to a character representing ‘insect’. Gelb categorises Chinese as logo-syllabic. An example of the next ‘great step’ occurs within Japanese writing, where some of the Chinese-derived characters, known as kanji, developed into the native Japanese syllabaries, known as kana. The third great step occurs when syllabic systems evolve into alphabetic systems, as Gelb claims happened in the change from Phoenician syllabaries to the Greek alphabet. Gelb’s teleological view of history sees alphabetic systems as the zenith of writing and it seemed to him that the path from logographic to syllabic to alphabetic is the only path along which writing systems can develop.

Gelb’s problematic and Western-centric idea, known as alphabetism, could not be fully debunked until Daniels (1990) demonstrated that writing systems can develop from syllabaries into non-alphabetic systems (see also Share 2014). Daniels identifies two kinds, naming them abugida and abjad, after the respective names of the first few symbols in the Ethiopic and Arabic scripts (or signaries, as he calls them). Abugidas represent a consonant plus an unmarked default vowel, and other vowels are marked using diacritics. An abjad is a consonantal system, such as Hebrew or Classical Arabic, where vowels are not marked at all. Daniels (1990) argues that such Western Semitic writing systems made the great ‘breakthrough’ in writing, and were the first to isolate the phoneme. Phoenician writing was not a ‘syllabary’ as Gelb claims, rather an abjad where the marking of individual consonants means that phonemes — or segments — were identified explicitly in the writing. The later addition of vowel segments in Greek writing was merely a natural development of the consonantal/phonemic analysis, rather than a ‘great breakthrough’.

Despite the criticisms, Gelb’s taxonomy still resonates in two important and related ways. One is that writing systems are categorised according to the predominant level of linguistic representation, be it words (logographic), syllables (syllabic) or segments (alphabetic). Even more fundamental is that the taxonomy is based on linguistic categories and does not pay any attention to the graphic forms of the writing. The confusion of graphic form and linguistic function seems to have been pervasive at that time. In his review of Gelb’s study,
Paper (1955: 91) notes the long-standing tendency to categorise ancient writing systems according to their script (what Paper calls their *outer gestalt*), rather than by the linguistic units represented in the writing. For example, all writing systems using a script such as cuneiform were grouped together, no matter what level of linguistic representation predominates in the writing system.

![Figure 2.1](image)

**Figure 2.1**
Sampson’s (1985) taxonomy of writing systems

### 2a.2 Sound- and meaning-based systems

Sampson (1985) also groups writing systems according to their predominant level of linguistic representation (see Figure 2.1). Almost all writing systems are considered to be *glottographic*, meaning writing that represents speech. This includes both *logographic* and *phonographic* writing. Sampson considers Chinese to be logographic, with symbols representing whole words or morphemes. Phonographic systems include syllabic systems (such as Cretan Linear B), consonantal systems (such as the Western Semitic examples above: Hebrew, Arabic and Phoenician), and segmental, or phonemic, systems (such as Greek and English).²

Sampson contrasts glottographic writing against *semasiographic* writing, whereby meaning is represented directly, without recourse to spoken language (the glottographic v. semasiographic distinction also occurs in Gelb (1962) and Pulgram (1976). He provides one example (‘written’ in Yukaghir — note the question mark Figure 2.1) but he removes the example from the second edition of
the book, perhaps due to DeFrancis’s (1989) argument that it is not possible to represent a full language using symbols that only represent meaning. Yet Sampson (2015) maintains the category of semasiographic writing, and this time he provides the slightly more convincing example of Blissymbolics, a meaning-based system used to help children with severe communicative disorders. Sproat (2010) discusses Blissymbolics in detail and argues that it too is not a full writing system as it lacks the full range of nuance and expression of full writing systems (see also Sproat 2016). And yet Sproat notes (2000: 129) that the category of semasiographic writing is only suggested to show that it may be possible, in principle, to have a writing system based solely on meaning, even if there are no functional examples of it (or at least none that have survived).

It is worth pondering briefly whether a writing system based solely on sound is possible in principle, and what its implications would be. The obvious example is the International Phonetic Alphabet (IPA), which Neef (2016) defines as a set of tools for writing down particular languages, rather than as a writing system per se. According to the definitions in use here, IPA is a script, not a writing system. Every language requires its own adaptation of IPA, with particular conventions in use, and these adaptations function as quasi-writing systems, because they represent sound conventionally, but not meaning. However, it seems that once the conventions have been set, and transcriptions become relatively stable, then there will naturally develop a strong visual connection between the transcription and the meaning of the word. That is to say, the transcription would develop a new role as the visual representation of the word and would thus become the de facto spelling. An example of this occurs with the brand name fiˈziːk, which uses IPA symbols to re-spell the English word physique. In such an instance, the notation doubles up as the name’s visual representation. In other words, the IPA symbols used to represent the value of each sound have combined to form a visual whole which functions as a unit of meaning within the writing system. In this case <fiˈziːk> refers to something different from the spelling <physique>, despite the fact that they sound the same. To summarise, a writing system based on sound will gradually develop visual connections between the notation and the words being represented. In this way, meaning becomes inexorably attached to notation derived from sounds. Similarly, we can expect strings of letters in an alphabetic system (i.e. spellings) to become attached to words. This is why the spellings <sing> and <sign> represent different words, even though they use the same set of letters. A solution to this problem will be posed in Section 2a.5 below.
2a.3 DeFrancis: all writing systems are sound-based

The next important taxonomic work comes from the Sinologist John DeFrancis, who provides a genuinely groundbreaking approach. For him, all writing systems are formed primarily through phonographic units, including Chinese. This means that all full writing systems are fundamentally the same, hence the title of his (1989) book *Visible speech: the diverse oneness of writing systems*. Representing sound is seen as a universal feature of writing systems, even though systems vary hugely in the levels of representation and the amount of phonetic faithfulness. Similarly, representing meaning is also considered to be a universal, although again systems vary enormously. Despite this emphasis on sound and meaning, DeFrancis (1989: 58) persists with an arboreal classification (see Figure 2.2), so that the level of sound-representation is prioritised. Yet within each basic category there is another distinction for meaning: between systems that are ‘pure’ and those that also encode a lot of morphological information. Greek, Latin and Finnish are considered to be pure phonemic systems but Korean Hangeul is considered to be morphophonemic, as is English (following Venezky 1970, as we shall see in Section 2c). The Japanese kana syllabaries are pure because they contain just phonological information. Chinese characters are *morpho-syllabic* (an update on Gelb’s *logo-syllabic* category), because they usually contain two pieces of information, one semantic, one phonetic. For example, 橡 *xiàng*, meaning ‘oak’, is composed of two pieces: the semantic part is 木 *mù* ‘wood, tree’, and this character is typically used for characters denoting names of trees; the other character 象 *xiàng* ‘elephant’ is used purely for its phonetic value.\(^4\) The phonetic faithfulness of the clue varies enormously (this example is a good one, being a pure homophone) but the key part is that the system is primarily sound-based in the formation of new characters. DeFrancis (*ibid*: 99) claims that 97% of Chinese characters were formed in this way. This means that DeFrancis’s analysis is not based on taking a synchronic snapshot of Chinese writing at a particular time in history but asking how it has developed, how new characters are formed and added to the system. Here we see a marriage of synchronic and diachronic approaches to writing systems, one which will inform the methods used in this study.
DeFrancis (1989: 262–4) also posits a third universal of writing systems, called ‘the dead hand of the past’. This is the enduring influence of ancient written forms amidst ever-changing phonologies. His short discussion is focussed on Chinese (ibid: 264–8) where the gap between writing and present-day speech is particularly glaring, despite the large-scale reforms of the mid-twentieth century. While Chinese may be the contemporary language that exhibits the largest gap between writing and speech, a more extreme example occurs in the continued use of Latin writing long after the demise of the spoken language. Most languages and their writing systems have a gap that is ‘less striking but more widespread’ (ibid: 263). Swedish, by contrast, makes periodic changes to spelling in order to keep its gap to a minimum. The ‘dead hand of the past’ can help to explain many of the complexities in the English writing system and, in particular, how the gap grows between invariant spellings and changing phonologies. This will be the subject matter of Chapter 6.

For now, we need to remember three universals common to all writing systems. All writing systems represent sound and meaning, albeit in different proportions, and, over time, the gap grows between unchanged written forms and ever-changing speech.
Sproat’s two-dimensional classification

Sproat (2000: 135–38) builds on DeFrancis’s distinctions between ‘pure’ and ‘morpho-’ writing systems, introducing a planar taxonomy with two dimensions: ‘type of phonography’ on one axis and ‘amount of logography’ on the other (see Figure 2.3). The former is based on the assumption, shared by all typologists, that it is useful to categorise writing systems according to the phonological unit which is most commonly represented in the writing system, be it the word, syllable, phoneme or even the consonant. Sproat (2000: Ch. 3) formalises this as the ‘orthographically relevant level’.

The amount of logography is defined as follows:

[A] component of a writing system [has] a logographic function if it formally encodes a portion of non-phonological linguistic structure, whether it be a whole morpheme or merely some semantic portion of that morpheme.

(Sproat 2000: 134)

Sproat admits (2000: 137) that this definition is ‘largely impressionistic’, although he notes (2016: 31) that efforts have since begun to try and measure the degree of logography in a script (e.g. Penn and Choma (2006) try to estimate how logographic Chinese and English are). Sproat only uses two dimensions for ease.
of presentation but he points out that there could well be more. For example, English and Korean are morphophonemic, unlike Greek, so they could be separated along that dimension.

The term ‘logography’ remains controversial, as many writing systems (e.g. Chinese, Japanese) have written units that represent morphemes rather than words. Hence, Rogers (2005) uses the term ‘amount of morphography’ in his updated version of Sproat’s (2000) planar classification. T. Joyce (2011) also argues convincingly for the use of the term ‘morphography’ (over ‘logography’) given the prevalence of morphograms in Japanese writing (see also French, 1976: 105–6). In the next section, I shall argue that the difficulties in pinning down the meaning of ‘logography/morphography’ derive from the conflation in the literature of two very different types of logography/morphography. For simplicity, and in agreement with the more recent scholarship, I shall henceforth only use the term ‘morphography’ even where the scholar used the term ‘logography’ (e.g. Sproat, Sampson).

2a.5 Morphography, English spelling and complex pleremic units of writing

It remains unclear in the literature to what extent English can be considered morphographic. Sproat (2016: 39) concludes that it is ‘a bit segmental, but at the same time a bit [morphographic]’. Several authors have considered English to be somewhat morphographic. Sampson (1985) claims it is morphographic in its use of different spellings for homophonous words such as dear and deer. This argument echoes Bloomfield and Barnhardt (1961: 27) who consider nit and knit to be examples of ‘word writing’ (quoted by Sproat 2016: 36). However, this kind of morphography is very different from the use of individual symbols such as <&> and <$>, which resembles the use of morphograms in Chinese and Japanese writing.

One helpful distinction is provided by Haas (1976) who distinguishes between pleremic or ‘semantically informed’ units of writing, and cenemic or ‘empty’ units of writing. Japanese kanji are pleremic, because the symbols have inherent meaning (i.e. <王> means ‘king’) but the Japanese kana syllabaries are cenemic, because the symbols simply represent sound. Alphabetic letters are also cenemic because they carry no inherent meaning. Meaning in alphabetic writing only occurs through the ordered assembly of letters. Thus the letters <k>, <g>, <n> and <i> have no meaning but, taken together, they can form the meaning
bearing units \(<\text{king}\)>). However, there does not exist a term for directly comparing the English spelling \(<\text{king}\)> with the Japanese kanji \(<\text{王}\>>.

One solution is to have a three-way, rather than two-way, distinction. Instead of contrasting phonographic and morphographic, or cenemic and pleremic, it makes more sense to have two different categories of morphographic, one simple and one complex. The more basic morphograms used in Chinese and Japanese can be considered as simple pleremic units, while the letter-strings used in English (such as \textit{dear} and \textit{deer}) can be considered as complex pleremic units. This is because they recur with the same shape indicating the same meaning. Hence \textit{dearest} relates to \textit{dear}, while \textit{deerhunter} relates to \textit{deer}. In Haas’s (1976) two-way distinction, \(<\$\>) is a (simple) pleremic unit because the unit has a meaning attached to it but the meaningful letter-sequences \(<\text{dollar}\>) and \(<\text{s}\>) (indicating plurality) have no specific status. On their own, the letters, \(<\text{d}\>\), \(<\text{o}\>\), \(<\text{l}\>\), \(<\text{a}\>\), \(<\text{r}\>\) and even \(<\text{s}\>\) are cenemic because they have no meaning attached to them. But the sequences \(<\text{dollar}\>) and \(<\text{s}\>) take on meanings within the writing system, allowing them to recur in, for example, \textit{petrodollar} and \textit{petrodollars}. In the three-way distinction presented here, the indivisible letter sequence \(<\text{dollar}\>) is a complex pleremic unit (henceforth CPU), as is the positionally-determined use of \(<\text{s}\>\).

A similar three-way distinction is also observed by French (1976: 126), although he prefers the term \textit{complex cenemic} and uses that term to refer to spellings representing morphophonological alternations in writing systems such as Czech and English. In English, the words \textit{German} and \textit{Germanic} share a common spelling \(<\text{German}\rangle\), despite having different stress patterns and vowel qualities (French 1976: 124–5). In Czech, the unit \(<\text{pod}\>\), recurring in the words \textit{podzemni} and \textit{podpis}, show assimilative voicing and de-voicing, respectively, and this does not need to be represented in writing (as \(<\text{d}\>\) and \(<\text{t}\>\)) because it is predictable.

The term ‘complex pleremic’ is preferred in this study to ‘complex cenemic’ on the grounds that the empty symbols (letters) have coalesced to form something meaningful within the system. An example of a complex cenemic letter sequence, accordingly, might be a random string of letters used as a code. A password, for example, has an ordered sequence of letters whose meaning cannot recur within a system.

The use of CPUs is pervasive in the spelling of complex English words, as we shall see in great detail throughout this study. For example, the sequence \(<\text{symbol}\>\) recurs in \(<\text{symbols}\>\), \(<\text{symbolise}\>\), \(<\text{symbolic}\>\), \(<\text{symbology}\>\), etc. The last example, however, involves a degree of compromise between \(<\text{symbol}\>\) and \(<\text{(o)logy}\>\), a CPU that recurs in \textit{biology}, \textit{sociology}, \textit{anthropology}.
mineralogy, etc. Hence, we shall see that CPUs can be changed at the margins. A simple example occurs in happy – happiness, where the $<y>$ changes to $<i>$. The primary advantage of CPUs is that several pronunciations can be captured by a single string of letters representing one core meaning. Cummings (1988: 15–16) gives the example of the sequence $<\text{sign}>$, which has four different pronunciations, as in sign, signal, design and designate (respectively /sən/, /'sɪɡn/, /zaɪn/ and /zɪɡn/). The string $<\text{sign}>$ is a meaning-bearing unit of sorts, somewhat independent of its pronunciation but totally reliant on the order of the four cenemic units from which it is composed, $<s>$, $<i>$, $<g>$ and $<n>$. Hence $<\text{sign}>$ has a totally different meaning to $<\text{sing}>$. The spelling of affixes can also be considered to be CPUs, including the $<\text{-al}>$, $<\text{de-}>$ and $<\text{-ate}>$ of signal and designate.

Distinguishing between simple and complex pleremic units makes it easier to compare writing systems, so we can say that English $\langle\text{wood}\rangle$ is a complex pleremic unit (recurring in woody, woodland, Cherrywood, etc.) while the equivalent Chinese form $\langle\text{木}\rangle$ is a simple pleremic unit. The fact that the English spelling gives a strong indication of the word’s pronunciation — while the Chinese character gives none at all — is a completely separate issue. A failure to recognise this fact is what has led to the two definitions of ‘logography’. The English spelling $\langle\text{oak}\rangle$ is a complex pleremic unit, but so is the Chinese character $\langle\text{橡}\rangle$.

DeFrancis (1989: 254) also discusses the idea, without introducing terminology, when he writes:

One could also speak of a ‘double articulation’ within a specific writing system. In that case, the ‘first articulation’ would comprise the written units that represent morphemes — in Chinese, chiefly single characters with some of two characters (e.g. shānhú, ‘coral’); in English, strings of letters, such as the $<\text{teach}>$, $<\text{er}>$, and $<\text{s}>$ of teachers. The ‘second articulation’ would of course then be the phonetic elements in Chinese characters and the letters or groups of letters in English.

In other words, the English word teachers contains three CPUs, $<\text{teach}>$, $<\text{er}>$ and $<\text{s}>$. The concept of double articulation is not original to DeFrancis, who notes (1989: 254) that Sampson (1985) uses the idea, and he in turn borrows it from Martinet (1962: 22–30, 1964: 22–28), who uses the idea to describe speech: a spoken word has both a semantic meaning and also an internal sound structure. Sampson (1985; 2015: 24–5) applies this to writing systems, using it to justify the
distinction between morphographic and phonographic writing systems, on the grounds that each kind of writing system prioritises only one articulation. However, there is no reason to make them mutually exclusive. It makes sense to consider the double articulation of writing to be the fact the written units can be both morphographic and phonographic, representing meaning and sound at once. This is the key fact that has been largely unidentified in the literature (notwithstanding French, 1976) More specifically, we can consider all word-units to be morphographic (i.e. morphograms and spellings) while some of them are more phonographic than others. Another example should clarify the point. If we consider both <8> and <eight> to be equivalent, and both to be morphographic, then we can say that <8> is a simple pleremic unit and <eight> is a complex pleremic unit. Independently of this distinction, we can say that <8> has no phonography, while <eight> has a relatively large amount of phonography (although not as much as <ten>, which represents the pronunciation with less ambiguity). This way we can track the change in the mapping relations between the spelling <eight> and its earlier phonological forms (perhaps /ext/ > /eːt/ > /eɪt/), and the following odd question never arises: when did the spelling change from being phonographic to being morphographic. Instead all spellings are morphographic; however <eight> is more phonographic than <8>, but less phonographic than <ten>.

The concept of a CPU also allows us to ask the more fundamental questions of how many functional units there are in an entire writing system. The English writing system has twenty-six cenemic units (i.e. letters) and a handful of simple pleremic units (e.g. <&>, <$>, etc.), but there are thousands of CPUs, including the spellings of all the meaning-bearing units. Spellings such as <sing>, <sink> <sign>, <ing>, <ed>, <s>, <ity>, etc. all function as meaningful units within the writing system, even though their component parts (letters) have no inherent meaning. No attempt will be made here to estimate the number of CPUs in the English writing system. Instead, the focus will be on how these units are combined to make new spellings.

2a.6 Conclusion

The categories proposed for classifying writing systems have increased in subtlety since Gelb first proposed his typology, but there is one very important trend uniting the more recent theories: all writing systems try to represent both the sound and the meaning of words, albeit in different ways and in different proportions.
DeFrancis (1989) identifies a third universal of writing systems which he calls the ‘dead hand of the past’. This is the gap that grows over time between conservative spellings and ever-changing phonologies. In Chapter 6, we shall see that, in English, the dead hand of the past has resulted in a huge disparity between the phonological information encoded in spellings and the phonetic reality of how people around the world speak nowadays.

Finally, it will be very useful to recognise that the units of English spelling are not just cenemic letters and clusters-of-letters which represent sound. Instead there exist meaningful strings of letters that function as recurring CPUs within the entire writing system (e.g. <sign>, <symbol>, <s>, <al>, <ate>, <de>, etc.). The CPU is not an established unit in typologies of writing systems but something similar does occur in French’s (1976) three-way classification. Thus there are simple cenemic units (i.e. letters and clusters-of-letters, <a, b, bb, c, ch, ck, d, dd, e, ee, f, ff>, etc), there are simple pleremic units (i.e. <&, $, 木>) and there are complex pleremic units (i.e. <wood>, <橡>). This idea will prepare us for how English spellings are actually formed, but first we must look first at other reasons why scholars have been interested in English spelling.

***
2b. Approaches to English spelling

This section assesses the most relevant research approaches to English spelling. Some of the work on English spelling has been historical, and a good deal more of it has been motivated by the desire to reform the system’s perceived inconsistencies. While the reform movement may be less strong than it once was, the reformer’s need to understand the workings of the writing system has in fact precipitated much of the linguistically-informed research on English spelling over the past half-century. The aims, methods and scope of such work is presented here in detail as a precursor to Section 2c, where the major research themes and findings are discussed. Literacy acquisition studies are of less interest, as such research examines how users cope with the writing system, rather than to question the nature of the system itself. Finally, the growing body of research on non-standard English spelling will be discussed in Chapter 7, where it is of immediate concern.

2b.1 Reading research

All accepted models of reading aloud recognise that there are two routes by which a printed word can be recognised (Coltheart 2012). One route involves recognising a word as a whole unit, while the other route involves the computation of correspondences between letters and sounds. That is to say, meaning can be accessed directly by sight through the visual path or it can be accessed indirectly via the phonological path, by sounding out the letters (either aloud or in one’s head). This latter path will be referred to henceforth as decoding, and this is the more relevant path in this study. Experience allows readers to recognise more words by sight. The two reading paths have strong parallels in the structure of writing — letters correspond to decoding, while entire spellings correspond to the visual path. One ramification is that the existence of the visual path seems to allow for one spelling to be associated with one meaning, regardless of differences in pronunciation. The abovementioned pair of German – Germanic exemplifies this perfectly. The CPU <German> represents (roughly) the same meaning in both words and it is unrelated to the homophonous name <Jermyn>. This principle of associating one spelling with one meaning is deeply ingrained in the English writing system.

2b.2 Historical studies of English spelling

The history of English spelling has been documented in several books. The first study of the history of English spelling was Craigie’s (1927) treatise, which brings
together much of the information garnered in the making of the OED. The next major work comes from Vallins (1965), a work which attempts two jobs: charting the history of English spelling; and outlining the patterns of spelling in the present-day writing system. The second edition (first edition, 1954) was revised by Scragg, who went on to produce the first book-length study of the topic (Scragg 1974). Three more histories of English spelling have appeared this decade. Crystal’s (2012) and, to a lesser extent, Horobin’s (2013) books are aimed at a popular audience, while Upward and Davidson (2011) provide a thorough attempt to unravel the different etymological inputs to English: chapters are devoted to the impact of Greek, Latin and French spelling, at several stages in the language, as well as showing the residual presence of Old and Middle English spelling on the present-day system (this includes a great amount of influence from Old and Norman French). All histories of English spelling look in detail at the Early Modern period when the writing system made the transition from the instability of Middle English spelling to the relative stability of Modern English spelling. The work of Brengelman (1980) and Salmon (1992) are particularly informative on this topic, as is Scholfield’s (2016) review of the changes which have occurred in the writing system since then. Work on Middle English spelling has become surprisingly relevant in recent times as the lack of stability in medieval spelling is replicated somewhat in the non-standard spellings used in new kinds of written domains, such as instant messaging on phones and on the internet. The work of Lass and Laing (2008–13) on Middle English spelling provides many of the theoretical tools required for describing words with variable spellings.

2b.3 Spelling reform

The history of English spelling reform is long but the movement has seldom had much success. Scragg (1974: Ch. 6) charts many efforts to change the writing system, dating back as far as the late sixteenth century (beginning with Smith, 1568 and Hart, 1569) when change seemed both necessary and possible. Of course the concept of re-form was rather different at that point, as individual spellings were not themselves fixed. It was really a matter of forming the spelling along certain principles suggested by the scholars. However, as the spelling of individual words stabilised throughout the seventeenth century, the movement shifted from a desire to radically overhaul the entire system towards a more tempered attempt at making partial improvements and providing aids to learners. By the time of Johnson’s (1755) dictionary, which is highly conservative in its choice of spellings (Upward and Davidson 2011: 299–302), the movement hit a
low point not seen again until the present-day, in the sense that reform was not part of public debate. In between these ebbs, the nineteenth and twentieth centuries saw many diverse attempts to make changes, and this often sprang from the utilitarian need to make the writing system accessible, for the first time, to large swathes of the population, now that everyone needed to learn how to read and write.

Many efforts at reform have relied upon the basic assumption that spelling should reflect speech, although that often meant the speech of particularly prestigious societal groups. Once a particular accent was chosen (usually RP or GA, or their forerunners), it became possible to propose a single letter for each sound, with new symbols being added where necessary; other reformers made do with a quasi-phonemic system based on the existing twenty-six letters of the English alphabet, allowing for digraphs such as <th> and <ch>, etc. The very valid argument in favour of phonemic-based reform is that a simpler system would make literacy acquisition much easier for growing populations in need of a skill that was a pre-requisite for advancement in society. Nevertheless, the phonemic approach to reform has led to arguments that are more repetitive than insightful, notes Scragg (1974: 117), and there is little need here to enumerate the various failed efforts at reform. For more, see Scragg (1974: Ch. 6), Carney (1994: 473–489), Upward and Davidson (2011: Ch. 9), and Yule and Ishi (2016).

What is relevant to the present study is that much of what we now know about English spelling was precipitated by the reformer’s need to understand the principles that underlie the writing system (Scragg 1974: 114), especially the non-phonetic information that can make a writing system more practical than a purely sound-based system. Dating back as far as the work of Mulcaster (1582) and his protégé, Alexander Gil (1619), many reformers have been aware of the fact the English spelling often represents etymological information and it can include such subtleties as making homophones distinct so that words with different meanings usually have different spellings (e.g. way v. weigh) (see Upward and Davidson 2011: 296). Gil’s arguments are given modern expression by Bradley (1928), a lexicographer and editor of the OED, who recognises that the writing system often represents meaning at the expense of sound. Between Bradley’s essay and Venezky’s groundbreaking study (1970), only a small amount of research pursued this line of thinking, and Scragg’s (1974: 117) recommendation was that researchers ought to find ‘more information about the relative value of existing orthographies, be they traditional, phonetic, or compromise’.

Since the 1970s, a sizeable body of research on writing systems has accumulated and the results of this research have led to more conservative
approaches to the problem of reform in English. Ironically, some of the most important research on English spelling was indirectly prompted by the reform movement, and this is best seen in the work that has emerged from Manchester University, which has given us the work of Haas, Scragg, Carney.\(^6\)

The one case of successful — or at least partial — spelling reform in the history of English has resulted in a minor split between American and British spelling, so that a sizeable number of words have different spellings. These often occur in predictable environments (e.g. `<color, honor, labor>` v. `<colour, honour, labour>` and `<liter, meter, theater>` v. `<litre, metre, theatre>`). Coulmas (2016) observes that the reasons for the success of this campaign were more social than linguistic, and came about because America attained its independence from Britain much earlier than other English-speaking countries. The overall quantity of orthographic divergence between British and American spelling is negligible in the greater scheme of things, but there are enough differences for readers to identify which system is in use from reading a few pages of text. Hence, Coulmas argues, the reform was more about representing identity than representing the language in a significantly better way. This viewpoint is somewhat corroborated by Cummings’ (2016) study, which shows that the differences between the two systems are far from cut and dry. Examining the spellings recommended by two British and two American dictionaries, Cummings (2016) shows how much variation there is, not only between Britain and America, but within the two British dictionaries (The *Oxford English Dictionary* and the *Concise Oxford Dictionary*) and the two American dictionaries (*Webster’s Third International Unabridged* and the *American Heritage Dictionary*). He concludes that the two systems are mutually evolving, with one continually influencing the other.

The current reform movement is hardly in a healthy position (see Yule and Ishi 2016) and the existence of spell checkers means that incorrect spellings can (mostly) be solved automatically (Mitton 2010). It is also possible that the lack of push for reform may also stem from the fact that we are still (academically) in a phase of investigation (of the kind recommended by Scragg) and perhaps reform will once again be on the agenda when we understand more about the English writing system and its users’ needs. Scragg writes (1974: 111) that the ‘principal bugbear’ of the reform movement has always been the ‘multiplicity of approaches’. Perhaps with greater theoretical consensus, partial reform may be possible in the future.
2b.4 Major studies of English spelling: aims, methods and scope

The following section looks at the aims, methods and scope of the major studies of English spelling, providing some important context for the discussion (in Section 2c) of the descriptions of the writing system provided by linguists.

2b.4.1 Venezky’s mappings from spelling to sound

Major progress was made in the study of English spelling during the 1960s when researchers could make use of advances in both technology and linguistic theory. An important pioneer was Richard Venezky who began work in the early 1960s, eventually producing *The Structure of English Orthography* (Venezky 1970), a near-reprint of his doctoral thesis (Venezky 1965), and it remains a classic text in the field.7 Venezky tabulated the spelling-to-sound correspondences of 20,000 of ‘the most common words’ in English, using a CDC-1604 computer, a very innovative approach at the time, allowing him to cover such a large corpus. His description of the English writing system built on two strands of thought: one was established American structuralist theories of language (e.g. Bloomfield 1933 and Hockett 1958), and the other was the increasing awareness (Bradley 1928, Vachek 1949) that English spellings are ‘partly phonemic and partly morphemic’ (Hockett 1958: 542).8

Venezky’s awareness of the morpheme as a unit of organisation in spelling makes for a much more durable description of the writing system than the set of sound-to-spelling rules proposed by his contemporaries, Hanna et al. (1966), who also performed a computational study upon a large corpus of English spellings. Their unit of investigation was not the morpheme, but the syllable, using a definition that has been heavily criticised (see Carney 1994: 86–95 for an in-depth discussion of their work, and its limitations). Since Venezky’s time, all worthwhile research has taken the morpheme into account, although the syllable (under very different definitions) has only recently begun to receive attention as a unit of organisation in English spelling (e.g. Evertz 2014; this will be discussed throughout Chapter 5).

Venezky’s aim was ‘to construct a theoretical framework for deriving sound from spelling and to search for the most general patterns in the orthography and the most plausible linkages for fitting these relationships into the total language structure’ (Venezky 1967: 80).9 It is worth pausing briefly to analyse some of these terms. By *spelling*, he means the accepted (American) spellings for the words in his corpus, and by *orthography* he means what is nowadays called
the writing system, meaning the relations between spellings and language. Significantly, Venezky’s framework only works in one direction, from spelling to sound. He does not concern himself with how people transform sound into spelling or, at a deeper level, how they map linguistic knowledge into writing. In other words, he wanted to know how we read, not how we write.10

Venezky’s corpus was based on the Thorndike Frequency List (Thorndike 1941), the first modern word-frequency list, which was compiled without computers, and was used for pedagogic reasons.11 The list contains words that are relatively frequent and assumed to be important for literacy acquisition. Nevertheless, Venezky does not follow the corpus slavishly, stating:

> Many low-frequency words from that list, especially proper nouns, were omitted by the present writer, and a number of words not included in the original Thorndike list were included.

(Venezky 1970: 13n)

The words added were useful because they 'illustrate interesting or unusual patterns' (Venezky 1970: 13). Clearly, Venezky recognises the importance of including some strange or aberrant spellings in a theory of English spelling and that approach is expanded upon in this thesis (especially in Chapters 7 and 8). The functioning periphery tells us a lot about the essence of the core spellings. Venezky is not explicit about what his additions are and he only mentions briefly his omissions. These include 'contractions, hyphenated words, proper names and variant pronunciations of the same spellings' (loc. cit.).

One thing to note here is that Venezky does not explicitly partition his corpus into different kinds of spellings, yet one of the key distinctions in his conclusion (1970: 120) is that compound and derived words (i.e. a large proportion of the polymorphemic words formed within English) involve different spelling patterns to what he calls ‘base forms’ (i.e. monomorphemic words). Venezky has thus begun the process of subdividing the writing system according to patterns of spelling formation, and this is a line of enquiry that will be pursued in this study. Furthermore, Venezky’s later, more populist book begins with a chapter on creative spelling, (Venezky 1999: 42), although such spellings are categorised according to the ways in which they deviate from the standard spellings, rather than according to the patterns inherent in the creative spellings themselves. While he makes no great theoretical development in this book, the juxtaposition of standard and creative spelling suggests potential for the two kinds of spellings to be studied within a single framework.
2b.4.2 Carney’s bi-directional analysis

The next work to provide a large-scale statistical analysis of English spelling correspondences is Edward Carney's monumental *A Survey of English Spelling* (1994), a work so gigantic and thorough that it is hard to imagine how (or why) anyone might take on the problem again based on the same set of assumptions, namely that it is actually possible to find consistent mappings from letters to phonemes and back. Unlike all others before him, Carney attempts to study spelling in both directions: from both spelling to sound and from sound to spelling. He does not seek to explain the historical reasons why English spellings can be so bamboozling, instead he explores the present-day system synchronically, accepting it as it is, and looking to assess what information readers and writers can use to help with reading and spelling. We shall see that he still manages to introduce a useful way of accounting for the etymological information stored in so many English spellings.

Like Venezky, Carney wrote computer programs to analyse correspondences and these provide the data for his two core chapters: one on how spellings are *encoded* and the other on how printed words are *decoded*. The encoding chapter provides numerical data on how each RP phoneme is spelt and the decoding chapter provides information on the converse mappings. ‘Encoding’ and ‘decoding’ are roughly synonymous with spelling and reading, although this only includes one part of the reading process. Carney’s concern is with the correspondence path, which is the path used in reading new words, unfamiliar words, or pseudo-words, as a reader can decode the pronunciation from the spelling. This is primarily what Carney means when he talks about the ‘reader’. He is not normally concerned with the fluent reader who bypasses the auditory route in favour of visual word recognition. Similarly, when he talks about the ‘writer’, he is not concerned with the writer of fluid prose, but instead with the very act of spelling a word.

Carney also has a lengthy chapter dedicated to homographs and homophones, a chapter on names, and a chapter that looks at the standardization process and the more recent efforts at spelling reform. All of these chapters contribute to the book’s immensity and usefulness. Indeed, his book is really two studies: one large study of lexical items (Ch. 1–4) and one smaller study of proper nouns (Ch. 6). The proper nouns are extracted from the *Longman Pronouncing Dictionary* (1st ed., Wells 1990) and the *English Pronouncing Dictionary* (13th ed., Jones 1980). While Carney does not attempt to include names within his description of core English spelling, the juxtaposition again suggests the possibility of integrating both into a single description.
The lexical items, a colossal 26,048 uninflected forms, are systematically extracted from five diverse corpora of over 25 million words (Carney 1994: 104–6). Like Venezky, and indeed Hanna et al. (1966), Carney’s main corpus is also Thorndike’s word list (Thorndike and Lorge 1944), but he tries to counter-balance its literary bias by the careful choice of four other (smaller) corpora. Carney is a little more concerned than Venezky with the nature of the words included, whether they be base forms, borrowed words, compound words, derived or inflected forms. Carney's corpus is ultimately defined by exclusion and he weeds out many kinds of spellings that exist on the fringes of the system. These are worth looking at because they provide some justification for the categories studied in the latter parts of the present study. He lists eight different kinds of spellings omitted from his database, and provides examples (Carney 1994: 32):

1. Vocal gestures (a subset of onomatopoeia, including ah, aha, ouch, tra-la, o-oooh-aye)
2. Manifest archaisms (ycleped, forms in <eth> and <est>)
3. Certain abbreviations (i.e., viz.)
4. ‘Unassimilated’ foreign words (bunt, capita, contra, da, de, der, des)
5. Manipulated jargon and trade names (INXS, Kleenex, Trace-tac)
6. Variant spellings (<plow> and <plough>, etc.)
7. Homographs (invalid, insult, bass, row, etc.)
8. Hyphenated compounds (eye-lash, beer-barrel, etc.)

The very act of omitting spellings due to certain criteria suggests that there are different kinds of spelling patterns for different kinds of words, an assumption that will be central to this thesis. Each of these groups reveals details about the nature of (English) spelling that cannot easily be observed when the focus is on core spelling patterns. Many of these groups will be discussed in Chapters 7 and 8. In this chapter, however, the focus will be on what researchers have actually found, based on the spellings they actually did study.

2b.4.3 Cummings and Rollings: sound-to-spelling accounts

There have been several other studies of English spelling that make significant contributions to our current understanding of the system. Cummings’s (1988) work examines sound-to-spelling correspondences of American English spelling. His approach is ‘humanistic’ and a rejection of algorithmic, formal linguistics (ibid: xxvi). Instead of analysing spelling numerically, he seeks to explore how four conflicting ‘demands’ on English spellings — phonetic, semantic, etymological and systemic — function ‘complementarily’. These ‘demands’ are
very similar to the ‘principles’ discussed in this study, and their respective merits will be discussed in Chapter 3. Cummings’ work has been mostly overlooked since its publication but it contains some very important ideas that will be built upon in the present description.

Rollings (2004) also looks at mappings from sound to spelling and his study maps out the relations between RP and (British) English spelling. Neither Rollings nor Cummings uses a corpus. Rather, they use a combination of observation mixed with information taken from dictionaries — which are themselves ready-made databases. It is very important to recognise that this has been a valid and productive approach to English spelling. Describing his own methods, Rollings writes:

Data were collected from a wide variety of sources as the research proceeded. As observations led to generalisations, further examples and counter-examples were looked for in order to test and, where applicable, support their validity. Although one of the aims was to distinguish regular from irregular spellings, statistics are not the raison d’etre of the study. The purpose is to discover and discuss orthographic patterns, rather than to achieve statistical precision. (Rollings 2004: 3)

Rollings tries to describe English spelling within the generative tradition of linguistics — which is as opposed to Cummings as it is possible to be — and he provides several kinds of rules for the spelling of each phoneme. The basic rules are simple (e.g. \( /i/ \rightarrow <i> \)) but the sub-rules can be overly convoluted. For example, the spelling of the words synonym, synchronise, syllable and symmetry are explained as ‘re-write’ rules, so that \( /i/ \) is first mapped onto \(<i>\) and then to \(<y>\) (Rollings 2004: 150). Clearly, it is simpler to recognise that this spelling pattern is almost entirely confined to words derived from Greek. The difficulty of accounting for the writing system’s historical nature is a recurring problem in synchronically-motivated studies, as we shall see in the next section.

2b.4.4 Graphematic approaches to English spelling
Recent years have seen a major growth in graphematic approaches to English spelling, a term that has become increasingly popular to describe the study of writing systems. The work most relevant to this study comes from Evertz (2014) and various papers by Kristian Berg (often in collaboration). Evertz (2014) tries to explain English (as well as German) spelling within two frameworks, one known as the ‘graphematic hierarchy’, which provides a graphematic (i.e. orthographic)
parallel to the established idea of the phonological hierarchy. This will be
discussed at length in Chapter 5. Evertz applies Optimality Theory (OT) to
English spelling, as we shall see in Chapter 5. His work on data taken from the
CELEX corpus (Baayen et al. 1995) where spellings are listed along with
phonological and morphological information, and this provides three important
kinds of information about any given word.

Two basic assumptions of the graphematic approach are that: a) it is
possible (and preferable) to study writing systems on their own merits with little
or no reference to phonology; and b) synchronic study of writing systems should
not take etymological information into account. Graphematic approaches have led
to some very interesting insights about the English writing system, as we shall see
throughout this study, but the assumptions lead to problems. In Chapter 4, we
shall see the difficulties faced if phonology is removed from the account of
English spelling. In Chapter 8, we shall see a solution to the problem of
etymological spelling, one which can be partially reconciled with graphematic
approaches. However, it is worth looking briefly at the relative importance of
synchronic and diachronic approaches to writing systems.

2b.4.5 Synchrony and diachrony
A common and long-standing approach to the English writing system is to
completely ignore its historical nature, even among linguists working today. The
following quote is illustrative:

[O]ur investigation is a synchronic one. Both synchronic and
diachronic data are potentially relevant for the exploration of writing
systems and their relation to other linguistic systems. It is, however,
important to keep them analytically distinct. A description of the
actual state of the system under investigation should rest solely on
synchronic data. After all, speakers/writers can only resort to
etymological information in a very limited way. Diachronic data, on
the other hand, may explain how the given system came to be the way
it is. From the perspective of today’s readers and writers, however,
this is of secondary nature.

(Berg, Buchmann, Dybiec, & Fuhrhop 2014: 285; my emphasis)
It is not clear why a ‘description of the actual state of the system’ should rely on
the information that ‘speakers/writers can resort to’. The two things are very
different and the facts examined by the researcher do not have to coincide with the
experiences of everyday users. After all, one does not expect a physicist to
describe the nature of gravity solely according to human experience of it. We must therefore separate our description of how the system functions from how users cope with its intricacies (or not), even though it would be ideal to be able to integrate both issues into the one model. Surely anyone interested in the system itself must pay attention to one of its central organising principles, that English spellings are frequently borrowed from other languages, with or without adaptation, regardless of complications caused by introducing new spelling-to-sound correspondences. If users fail to recognise that, and teachers fail to teach it, then that does not mean the information is not present in the spelling. Hence researchers cannot simply disregard it. In Section 2c.3, we shall see some of the efforts that have been made to integrate etymology into descriptions of the writing system, and in Chapter 3 I shall introduce a method for integrating both synchronic and diachronic approaches to the topic, borrowing heavily from methods used in the study of word-formation (e.g. Bauer 1983). Section 8a provides a solution to the problem of integrating etymological spelling.

With all of this background in mind, we can now explore the major research themes and findings of the past fifty years, and look to see which problems remain unanswered.
2c. Major research themes and findings

This section examines the last fifty years of research on English spelling, which has been based on modern linguistic theory. The aim of this section is to identify the most important problems or gaps, and to pose some questions for analysis in the remainder of the thesis.

2c.1 Keeping the spelling of the morpheme constant

The English writing system is not simply concerned with mapping phonemes onto letters. To a large extent it tries to offer the reader a constant spelling for a morpheme, in spite of the varying pronunciation of the morpheme in different contexts.

Carney (1994: 18)

The spelling of compounds and derived forms tends to be morphemic; the established graphemic form of the base is retained as much as possible, regardless of the phonemic alternations involved.

Venezky (1970: 120)

(emphasis added in both cases)

The most important discovery about English spelling, since the development of linguistics as a science, is that the writing system is largely consistent in its spelling of the morpheme. In this section, we will look at the diverse ways in which morphemic information is expressed within the writing system, first by looking how linguists have slowly gained a greater awareness of this fact and then by looking at some of the problems in their analyses. The idea that the English writing system represents meaning, rather than just sound, was given full, detailed, expression in the work of Venezky (1965, 1967, 1970) but the idea cropped up at various points in the decades preceding his work.

Bradley (1928) argues that the purpose of writing is to represent meaning, and that the use of sound-based symbols is just a vehicle for doing that. Taking many examples of homophonous pairs, he writes that:
It is because the expression of meaning is felt to be the real purpose of written language that these distinctions [e.g. between to sew and to sow] still survive, in spite of the disastrous effect that they have had on the phonetic intelligibility of written words.

Bradley (1928: 7)

Bradley argues that a spelling is not just a phonetic representation of the sound of a word but that a spelling functions as an ideograph of sorts, a unit which provides direct access to meaning for the experienced reader. In other words, a spelling is a CPU.

2c.1.1 Morphophonemic spelling

A more nuanced form of morphemic spelling has often been called morphophonemic spelling, where the same spelling is used to represent different allomorphs of the same morpheme. Vachek (1959: 15) notes the similarity of spelling in the words equal and equality, despite the different qualities of the vowel in the second syllable in each word, and he also notes similar phenomena in the Czech and Russian writing systems. Morphophonemic spelling is only possible because the morphophonological alternations are (often) systematic, so a sequence electric – electricity – electrician shows similar alternation patterns to elastic – elasticity and logic – logician. Hence the writing system rewards readers and writers with a greater knowledge of a) the language as a whole and, b) the internal workings of the writing system.

2c.1.2 Lexical spelling

[Noam] Chomsky and Halle (1968: 48) claim that 'English spelling, despite its often-cited inconsistencies, comes remarkably close to being an optimal orthographic system for English'. Their reasoning is that an optimal system ought to have one orthographic representation (i.e. spelling) for each lexical entry, so Carol Chomsky (1970), who builds on their work, introduces the term lexical spelling. The near-synonymy between ‘lexical’ and ‘morphemic’ reflects the debate among typologists between ‘logographic’ and ‘morphographic’ writing (see Section 2a.1.6 above). According to this theory, the lexical spelling of a word should ideally represent an abstract underlying form which Chomsky and Halle claim native speakers have internalised along with its morphophonological alternations. This means that phonetic or surface forms do not need to be indicated by the spelling whenever they are predictable. The related pair revise – revision will demonstrate everything of relevance. The abstract underlying form of the word is represented by the lexical spelling <revis> and its surface realization is
manifested in two different ways, depending on the phonological environment into which it falls. In *revise*, the vowel is long and the consonant is /z/: /rəˈvaɪz/. In *revision*, the stressed vowel is shortened and the consonant is palatalized to /ʒ/: hence /rəˈvɪʒən/. Such processes are 'automatic and predictable' for the native speaker so they don't need to be marked in the spelling (C. Chomsky 1970: 292). By contrast, an unpredictable alternation such as *run* – *ran* has to be spelt out explicitly.

C. Chomsky provides many examples of both consonant and vowel alternations (see Tables 2.1 and 2.2). Alternations also occur on stressed vowels, and this can interact with other phonological processes. For example, vowels are often shortened when followed by a bi-syllabic suffix, as in *profane* – *profanity*. Chomsky and Halle would argue that the letter <a> indicates an underlying vowel //æ// which surfaces as /eɪ/ or /æ/, depending on its environment.

<table>
<thead>
<tr>
<th>Spelling</th>
<th>Surface form</th>
<th>Underlying form</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;c&gt;</td>
<td>medic-ate  /ˈmedɪkət/ /k/          //k//</td>
<td></td>
</tr>
<tr>
<td></td>
<td>medic-ine  /ˈmedɪsɪn/ /s/          //s//</td>
<td></td>
</tr>
<tr>
<td>&lt;g&gt;</td>
<td>prodig-al  /ˈprɒdɪɡəl/ /g/          //g//</td>
<td></td>
</tr>
<tr>
<td></td>
<td>prodig-ious /prəˈdɪdʒəs/ /dʒ/        //dʒ//</td>
<td></td>
</tr>
<tr>
<td>&lt;t&gt;</td>
<td>resident   /ˈrezɪdənt/ /t/          //t//</td>
<td></td>
</tr>
<tr>
<td></td>
<td>resident-ial /reziˈdentʃəl/ /ʃ/      //ʃ//</td>
<td></td>
</tr>
<tr>
<td>&lt;s&gt;</td>
<td>revis-e    /rəˈvaɪz/ /z/            //z//</td>
<td></td>
</tr>
<tr>
<td></td>
<td>revis-ion  /rəˈvɪʒən/ /ʒ/            //ʒ//</td>
<td></td>
</tr>
<tr>
<td>&lt;s&gt;</td>
<td>sign       /ˈsaɪn/ /s/              //s//</td>
<td></td>
</tr>
<tr>
<td></td>
<td>de-sign    /dəˈzæm/ /z/             //z//</td>
<td></td>
</tr>
</tbody>
</table>

Table 2.1
Consonant alternations with invariant spelling (from C. Chomsky, 1970)
<table>
<thead>
<tr>
<th>Spelling</th>
<th>Surface form</th>
<th>Underlying form</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;a&gt; profan-e</td>
<td>/prəˈfeɪn/</td>
<td>/ə/ //a//</td>
</tr>
<tr>
<td>profan-ity</td>
<td>/prəˈfæntɪ/</td>
<td>/æ/</td>
</tr>
<tr>
<td>&lt;e&gt; seren-e</td>
<td>/soʊˈriːn/</td>
<td>/iː/ //e//</td>
</tr>
<tr>
<td>seren-ity</td>
<td>/soʊˈrentɪ/</td>
<td>/e/</td>
</tr>
<tr>
<td>&lt;i&gt; expedit-e</td>
<td>/ˈekspədɪt/</td>
<td>/ai/ //i//</td>
</tr>
<tr>
<td>expedit-ion</td>
<td>/ˌekspəˈdɪʃn/</td>
<td>/i/</td>
</tr>
<tr>
<td>&lt;o&gt; phon-e</td>
<td>/ˈfɔʊn/</td>
<td>/ɔː/ //a//</td>
</tr>
<tr>
<td>phon-ic</td>
<td>/ˈfɒnɪk/</td>
<td>/ɒ/</td>
</tr>
</tbody>
</table>

Table 2.2
Vowel alternations with invariant spelling, adapted from C. Chomsky (1970)

2c.1.3 Criticisms of the lexical spelling hypothesis
The most basic criticism of the idea of lexical spelling is that there are so many exceptions. Sampson (1985: 201) points to alternations such as *speak* – *speech*, *palace* – *palatial*, *joke* – *jocular*, *collide* – *collision* and he notes that *prodigal* and *prodigious*, cited by C. Chomsky (1970: 292) are not even related (etymologically or semantically). Sproat (2016: 35) says that 'statistics provide at best weak support for [Chomsky and Halle’s] hypothesis' (see also Sproat 2000: 78–85). The lexical spelling hypothesis is far from a complete model of English spelling but it is also not totally wrong. It provides a useful first approximation towards the workings of the writing system, reminding us that English spelling helps us to make visual connections between words with related meanings, even when their pronunciations show great surface variation. The fact that such variation is often systematic and predictable furthers the case for lexical, or morphemic, spelling.

2c.1.4 Further examples of morphemic spelling
Beyond the examples discussed thus far, Venezky (1967: 91) observes several cases where an awareness of the morpheme is crucial to our reading of the written word. A very simple example comes from polymorphemic words such as *uphill* and *topheavy*, where the <ph> corresponds to /ph/, rather than /f/, and this is because the <p> and <h> belong to the spelling of two recognisably different morphemes. This rule alone mops up a huge number of the apparent inconsistencies of the system. It applies to the <th> of *hothead* and *cathouse*, the <ng> of *ingenuous* and many, many more examples. A similar rule applies to long,
or geminate consonant sounds, which can occur at morpheme boundaries. Monomorphemic *missal* has one /s/ but polymorphemic *misspelt* may have /ss/.

Another group that can be explained with reference to the morpheme are pairs like *autumn* – *autumnal* and *bomb* – *bombard* where the <n> of *autumn* and the latter <b> of *bomb* have traditionally been called ‘silent letters’. Carney (1994: 40) calls these *inert letters*, because they have no phonological correspondence in some allomorphs (*bomb, autumn, sign*) but do in others (*bombard, autumnal, signal*). Importantly, the patterns in derived forms often differ from the patterns in inflected forms. The inert letter remains unpronounced in *bombed* and *bombing* as well as in *signed* and *signing*. This suggests that a description of English spelling must take word formation patterns into account.

Consistency in the spelling morphemes also extends to affixes. Venezky (1967: 91) notes the constant use of <s> for the spelling of three plural allomorphs, seen in *cats*, *boys* and *judges*, pronounced as /s/, /z/ and /ɪz/, respectively. The past-tense morpheme {<e>d} behaves in a somewhat similar way but it is ‘not entirely regular’ (ibid). This suggests that the idea of morphemic spelling may be working in conjunction with other forces.

2c.1.5 The impossibility of a fully morphemic writing system

It was highlighted above that English spelling is morphemic ‘to a large extent’ or ‘as much as possible’, but it has seldom, if ever, been asked *to what extent* English spelling is, or even can be, morphemic. The past-tense morpheme is a case in point. Carney (1994: 19) points out that the suffix triggers several different spelling patterns, depending on the phonological and orthographic environments. If the preceding vowel is long and the spelling ends in <e>, then only <d> is added, as in <hope – hope(d)> and <plane – plane(d)>.

This is in striking contrast to the three letters added to <hop – hop(пед)> and <plan – plan(ned)>. Consistent morphemic spelling would render *<hopeed>* instead of *hoped* and *<hoped>* for *hopped*, and such spellings would provide poor matches for the phonological form. Hence, compromises have to be found between the need to represent the morphemes as consistently as possible and the need to represent the phonological forms in a way that readers can easily decode. This is recognised in all the literature but the interaction between morphemic and phonemic spelling is not sufficiently theorised. The failure to do so is most clear in the following example. The final <e> of *profane, serene, revise, divine* is not present in *profanity, serenity, revision, divinity*, and yet both sets of words are supposed to display ‘morphemic spelling’. Clearly the spellings show a great deal of overlap in their representation of morphemic information but there is clearly some interaction...
between morphemic spelling and the need to represent the phonological forms, otherwise there would either be no need for final <e> in the base forms or no need to drop it in the complex forms. Therefore, the principle of morphemic spelling does not always operate in isolation and, under certain conditions, spellings seem to involve interaction between the twin needs of representing morphemic and phonological information. This raises some important questions that are among those to be investigated in this work:

- To what extent is English spelling morphemic?
- How does morphemic spelling interact with phonology?

Partial answers to these questions have been provided by some recent literature. Berg et al. (2014) observe that the spelling of English affixes is much less variable than the spelling of the stem. Hence compromises tend to occur in the spelling of the base, rather than in the affix (a situation that differs from German, they note). For example, in the spelling <profanity>, it is the spelling <profane> which is compromised, not the spelling of the affix <ity>. However, this observation only answers half of the problem. It explains why the complex word is not spelt *<profanety>, but it does not explain why *<profaneity> is disqualified. Let’s look briefly at what would happen if such spellings did exist, and English spellings fully obeyed the morphemic principle. We would have pairs such as:

- profane – *profaneity
- serene – *sereneity
- cone – *coneical
- revise – *reviseion

One obvious problem is that these spellings would lead to confusion between words with similar spellings but different pronunciations. Compare:

- *profaneity – spontaneity
- *sereneity – homogeneity
- *coneical – onomatopoeical
- *reviseion – pereoin

It is intuitively obvious that *<profaneity, sereneity>, etc. are not well-formed spellings, in the sense that they do not seem to provide an appropriate set of correspondences for the known phonological form. The phonological form is known (notwithstanding dialectal variation) because English words are formed
within speech. However, there does not yet exist a method for checking whether a spelling that adheres to the morphemic principle results in a well-formed English spelling, and providing such a method will be one of the central aims of this thesis. We are left here with more questions for later discussion:

- How do we know if an English spelling is well-formed?
- What is the actual process of spelling formation?
- How can spelling formation be modelled?

2c.2 Correspondences between spelling and sound

There is one obvious basic assumption which underlies the whole business of relating phonemes to letters: simply that it can be done.

(Carney 1994: 34)

English frequently departs from the phonemic principle of having one letter for each phoneme, and it even departs from the quasi-phonemic alphabetic principle whereby complex units such <ch>, <th> and <sh> represent just one sound. In this section we shall look at how previous researchers have discussed regularities of correspondence that exist at the phonemic level and, where relevant, other sub-syllabic units.

Undoubtedly the best analysis of the correspondences between spelling and sound comes from Carney (1994) who attempts to apply the 'usual grounds of linguistic analysis' to English spelling, namely 'simplicity, consistency, exhaustiveness and explicitness' (Carney 1994: 33). The theoretical section (p. 1–102), which precedes his lengthy analysis, is one of the most useful pieces ever written about English spelling, even though his 'missionary zeal' (p. xix) can seem too dismissive of most previous work on the topic. And yet it can be easy to forget how little linguistically informed research there was until then, something that can be seen by glancing at the very short bibliography in Scholfield’s (1994) review of the then-relevant literature. He makes the important observation that linguists up until that point had failed to provide the statistical data required for theorists to provide a more probabilistic model of spelling correspondences (Scholfield 1993: 71). Carney’s study provides the first effort at filling that gap.

Carney’s explanation of his methods of segmentation (Carney 1994: 32–48) provides a solid platform for the exhaustive description of his very large wordlist (the abovementioned 26,048 uninflected forms). His analysis is applied to all RP phonemes and ‘all’ spelling units (meaning the units that satisfy his
criteria). His primary task is to segment English spellings into units that map in both directions, from spelling to sound and back again. Doing this involves recognising many recurring concepts. The first of these is divergence from the phonemic principle (Haas 1970), which operates in both directions. For example, the spelling unit <th> is divergent because it can represent either /θ/ or /ð/, as in thin and this, while the reverse is true for the phoneme /ð/, which can be spelt as <f>, <ff>, <ph> or <gh>. Among these four, there are conditioned variants, such as <f> and <ff>, whose distribution is largely predictable from the environment (e.g. graft v. gaffer), and competing variants, such as <ph> (e.g. graph). There are many positionally-conditioned mappings, including <c> ≡ /s/ when followed by <e>, <i> or <y>. Similarly, <gh> ≡ /f/ only occurs at the end of syllables and even then it is not predictable, as it may represent another phoneme (e.g. tough v. though). Carney provides statistics on the text frequency and lexical frequency of correspondences. For example, he shows that, while there are not very many words with the strange correspondence of <gh> ≡ /f/, many of those that do exist (e.g. cough, laugh, enough, etc.) occur quite frequently, so the correspondence has a high text frequency. By contrast, <ff> ≡ /f/ occurs in a much larger number of rarer words (effete, suffragan, etc.), so the correspondence has a higher lexical frequency. There is no metric provided for how predictable each correspondence is. For example, <ff> ≡ /f/ is predictable, but <gh> ≡ /f/ is not (hence though, eight, night, etc.). Carney omits inflected forms, massively reducing the number of incidences of double consonant letters such as the <ff> in cheffing and cheffed.

It is not always obvious as to what counts as a unit of spelling and Carney’s choices are based on several overlapping criteria. The ‘simplicity criterion’ states that there should be as few different correspondences as possible. The spelling unit <gh> of nigh and night should be a single unit because it is consistent throughout the system, a similar analysis applies to <eigh> ≡ /ei/, seen in eight, weight and neighbour. Being consistent does not mean it is an efficient or ‘phonetically transparent’ unit. In a word such as campaign, there are further options where to chop up the spelling: the second vowel (/ei/) might be matched with either <ai> or <aig>. Albrow (1972: 37) chooses the latter but Carney reckons that allowing for too many one-off correspondences like <aig> ≡ /ei/ ultimately leads to inconsistency and complications in a system-wide analysis. He thus segments the spelling into <ai> and <g>, on the grounds that <ai> already occurs in pain and rain, and there is a similar kind of <g> in reign. This allows for greater economy of description. Hence Albrow has four units <ai>, <ei>, <aig> and <eig>, all representing /ei/, whereas Carney has just three: <ai> and <ei> both represent /ei/, while <g> ≡ /∅/, a kind of ‘empty’ letter.
Empty letters have zero phonological correspondence and thus they form another subcategory of silent letter, different from the inert letters we saw in Section 2.2.1 (as in bomb and sign). Other examples of empty letters include the <b> of debt and doubt. Empty letters must meet strict criteria (Carney 1994: 45) and they may still retain some function. For example, the empty <l> of salmon has a numerical function: just like the <mm> of gammon, the <lm> of salmon tells us that <a> represents a short /æ/, unlike the /eɪ/ of Damon. Another group of silent letters are auxiliary letters, 'extra letters which help to make up complex graphic units' (Carney 1994: 40). There are two kinds: one group includes the <k> of knit and the <w> of writ which are expendable yet clearly part of the unit: i.e. <kn> ~ <n> ≡ /n/ and <wr> ~ <r> ≡ /r/; these are called endocentric spelling units. The other group, including <th> and <sh> are considered to be exocentric because, when subdivided, they result in radically different phonemic correspondences: compare thin and shin with tin, sin and hint. Note that <gh> can be endocentric (ghost, ghastly) or exocentric (laugh, tough). Once again, position within the syllable is important here. Such a slew of categories may seem overwhelming at first but it does, ultimately, provide for a more economical and largely practical set of mappings from spelling to sound and back, albeit only at the segmental level.

Carney continues by stating that segmentation should also be discrete and exhaustive, so it has to be without remainder or overlaps, and it cannot cross morpheme boundaries. This is a reminder that it is possible to blend morpheme-level and phoneme-level analyses within the one model, and one must recognise morpheme boundaries before applying phoneme-level correspondence rules.

Kressler and Treiman (2003) provide a slightly different approach to complex units, and they point out that letter sequences such as <igh> and <eigh> only occur morpheme-finally or before <t>, as in nigh, night and weigh, weight. Similarly, /au/ ≡ <i> generally occurs before <ld> and <nd> (e.g. wild, mind). One might expect pint to be on this list but their study is confined to ‘first grade text vocabulary’ or ‘child vocabulary’ (ibid: 285). Kressler and Treiman’s argument suggests that the analysis of spelling-to-sound correspondences is not just confined to relative position within a word but can correlate strongly with the position within the phonological syllable. Harris (2000: 111–2) argues for a similar analysis of the French word oiseau, whose spelling, according to Saussure (1916, quoted by Harris), fails to represent any of the spoken word’s sounds. Yet Harris reminds us that the correspondences <oi> ≡ /wa/ and <eau> ≡ /o/ are consistently applied across the French writing system, as in the words oisif, oiseux, Oise, toise, and beau, eau, peau, etc. Hence the complex spelling units
function as simple correspondences. One might make the same argument about
<et> in the French-derived English words duvet, bouquet, etc.

Problems arise for Carney’s otherwise consistent analysis when it comes
to describing levels of representation above the phoneme (Carney 1994: 37). For
example, in the word picture, Albow (1972: 41) would argue that the <u> is an
important part of the correspondence for /tʃ/ and for the vowel correspondence.
Such overlaps break Carney’s principle of discreteness. Carney thus sets up
<e> ≡ /tʃ/, and leaves <ure> as being the vowel correspondence (in the non-rhotic,
RP accent he is describing). However, the occurrence of such consonant
palatalization is dependent on this very phonetic environment, as can be seen in
the words virtue, structure, virtual, voluptuous, perpetuate, question, suggestion,
etc., where <e> ≡ /tʃ/ occurs immediately after a stressed syllable and before the
same small set of vowels. Hence the spelling-to-sound relations can only be
explained satisfactorily with reference to higher phonological levels, such as the
syllable, the foot and word stress.

Carney’s attempt to provide mappings between phonemes and letters thus
results in a detailed but slightly incomplete description of English spelling. In
order to improve the thoroughness of the description, it is necessary to take into
account higher levels of correspondence than just the phoneme. There does exist a
small body of important work that studies the role of these levels in English
spelling, notably the work of Rollings (2004) who introduces an orthographic
syllable to explain consonant-letter doubling, while Evertz & Primus (2013) and
Evertz (2014) build on this, introducing a foot-level of analysis. This research will
be discussed in detail in Chapter 5, where there is space to present the theories of
supra-segmental phonology upon which their analyses are based.

Finally, it should be noted that Brooks (2015) presents the spelling
variants for each phoneme of what he calls ‘British English’ and, while his data is
presented in a slightly more digestible manner than Carney’s, there are no great
theoretical developments in his work. He disregards the syllable as a unit of
spelling (Brooks 2015: 458) and only occasionally mentions the morpheme as
being an important level of organisation in English spelling. Indeed the book’s
theoretical section is placed in a short appendix at the end of the book, rather than
preceding the analysis. Nevertheless, it is a useful reference work and it is part of
the growing trend towards providing numerical accounts of the occurrence of each
spelling unit.

***
2c.3 Etymology and the subsystems of English spelling

A large proportion of the English lexicon derives from other languages and, since the Renaissance, there has been a general tendency for adopting the spelling of borrowed words, especially words that derive from Greek, Latin, French and Italian. These words have affected all parts of the language, but the writing system has been particularly affected because the borrowed spellings frequently introduce new correspondences for words that have been adapted into English phonotactics. At the furthest end of the spectrum are examples such as \(<ps> \equiv /s/\) (word-initially), which occurs in Greek-derived words such as *psychology* and *psoriasis*. Upward and Davidson (2011) provide the most detailed account of the contribution of different languages to the English writing system. Nevertheless, synchronic analyses of English spelling have struggled to account for this mass influx of borrowed spellings, which is not surprising, given that synchronic and diachronic are widely understood to be mutually exclusive categories of analysis in linguistics.

The distribution of correspondences for \(<ch>\) is a perfect example of where reference to etymology can be very useful in explaining English spelling, although this explanation has frequently been avoided. For example, Venezky (1970: 66–67) simply lists examples of seemingly irregular correspondences. He makes no formal distinction between his short list of cases where \(<ch>\) has no correspondence (e.g. *fuchsia, yacht, drachm*), and the forty-three cases where \(<ch> \equiv /k/\) (e.g. *chameleon, chaos, character, eucharist, parochial*, etc.). He does not include in his description the fact that most of the latter group have Greek roots, nor that the thirty-seven cases (in his dataset) of \(<ch> \equiv /ʃ/\) are all of modern French origin (*cache, chandelier, parachute, sachet*, etc.). Clearly, Venezky was aware of the words’ origins but his descriptive concern was simply with letter sequencing. As a result, he considers *chlorine, Christmas and technology* to be regular on the grounds that \(<ch> \equiv /k/\) before \(<l>, \langle r\rangle\ and \langle n\rangle\). But the rest are irregular, within the limitation of his theory, despite the obvious subgroupings. Carney, by contrast, uses the subsystems as a safety check for spelling/decoding. If a reader/writer recognises that a word is \(\S\)Greek, then the appropriate correspondences should apply. The same goes for \(\S\)French, \(\S\)Italian, \(\S\)Latinate, and so on. This is why contestants in Spelling Bee competitions are allowed to ask for a word’s origin.

Albrow (1972) provides the first scholarly effort to defragment the entire writing system into sets of subsystems. His work is not 'maximally rigorous', by
his own admission (ibid: 3) but he hopes to be the first word on the matter and not the last. He proposes three subsystems of spelling correspondences that map from spelling units to phonemes and labels them neutrally as 1, 2 and 3. The idea is borrowed directly from the Firthian idea of polysystemic analysis whereby all levels of language have their own subsystems (ibid: 9). For example, English phonology has subsystems of consonants and vowels, the vowels in turn have systems of short- and long-vowels. Albrow’s System 1 contains the most fundamental of correspondences, such as the <f>, <a>, <c> and <t> of fact, but it also contains many supposedly ‘English’ sequences such as <ough>, despite the unpredictability in the mappings from spelling to sound and back. System 2 includes the commonly occurring spelling units that derive from Latin, French, and Greek, as well as the more marked native features (such as <igh> ≡ /au/, as in high). The <ph> and <c> of pharmacy would thus be among system 2, but the <a> and <t> in System 1. System 3 mops up all the remaining correspondences, including, for example, the <ch> of chauffeur and the <ps> of pseudo. The motivation behind Albrow’s taxonomy is not explained rigorously and he has no set criteria for inclusion in a particular subsystem. Etymology is important but not dominant. Frequency is probably a factor but it is not measured. No mention is made of the predictability of a correspondence, in either direction. For example, <ps> is unambiguously /s/ at the start of a word whereas <ch> could be /ʃ/ or /k/ (compare chowder and chaos) (see also Ryan 2011). Despite such problems, Albrow’s work had an important influence on Carney and, by extension, the present work.

Carney (1994: 96–101) takes a lexical approach to the issue, proposing a slew of etymologically-derived lexical subsystems based on the spelling of entire words, namely §Basic, §Romance, §Latinate, §Greek, §French, §Italian and §Exotic (‘§’ indicates ‘subsystem’, and some of these terms, notably ‘exotic’, appear in Craigie’s (1927) treatise). Not only do English morphemes keep their spelling constant, their correspondences are all part of the one subsystem. So while fill and fulfil are §Basic throughout, since they contain the default spellings for the respective phonemes, chlorophyll is §Greek throughout, as it has several Greek-derived correspondences. Carney’s reliance on the whole word rather than its subcomponents solves Albrow’s problem of what to do with problematic units such as <ch> which have several correspondences. The rest of the word often provides clues: §Greek words (or combining forms) often have <ph> and word-internal <y> so this instance of <ch> is more likely to be /k/ (as in chlorophyll and psychopharmacology). In the terms presented in Section 2a, the subsystems are built out of recurring CPUs, including combining forms such as <chloro>,


<phylll>, <psycho>, <pharmaco>, etc. These spellings represent bound morphemes which only appear in combination (Plag 2003: 93). It is only by preserving the CPU that the correspondences can maintain their relations. By contrast, the <ow> of §Basic chowder helps us identify <ch> ≡ /ʃ/. It is not a foolproof method but, as Carney suggests, spellings ought to be treated as ‘complex data items’. Writers have to make use of several kinds of knowledge in order to spell. They need phonological and lexical knowledge but subsystem awareness is a huge advantage. Knowledge of French, Latin and Greek is very helpful and the abandonment of the study of these languages undoubtedly has a major effect on people’s ability to spell. In the most advanced cases, readers of Greek might know that chlorophyll has <λλ> (in its latter Greek component <φύλλον>) whereas prophylactic only has one <λ> (προφύλακτικός) (OED 32067 / 152714). Similarly, writers can avail of their cultural knowledge: if one knows that /ʃi:k/ is a fashion term, then one might guess that its spelling is §French <chic> rather than §Basic *<sheek> (Carney 1994: 101). It should be noted that the §French subsystem involves words derived from modern French words, while many words deriving from earlier forms of French are included in the §Latinate subsystem (e.g. serenity) and others are part of the §Basic subsystem (e.g. table). A few French-derived words, such as measure and treasure, are categorised as §Romance, as their spelling has diverged considerably from Latin (Carney 1994: 101). However, Carney does not enumerate exactly which spelling units fit into which subsystem; instead he uses the labels as reference points when discussing particular spelling units in his core analysis (i.e. Ch. 3–4)

Membership of a subsystem is not entirely determined by etymology. Many words are §Basic by default: map and dot have no ‘observable markers’ so they can go in no other system (Carney 1994: 101). Implicit in Carney’s argument is that there is great overlap in the spelling of the various input languages so the subsystems may actually overlap more than they differ. Furthermore, apparent etymology can supersede actual etymology: Carney (loc. cit.) notes that totem – totemic is of Algonquian origin (probably Ojibwa, OED 203813), but it patterns like system – systemic so it falls into the §Greek subsystem.

Another implication of Carney’s argument is that some spellings are liable to be categorised in either of two subsystems, giving rise to some variable, and thus ‘incorrect’, pronunciation. This is likely to happen when it is not fully clear which subsystem a spelling belongs to. Words like serif and homage are not §French, but they may appear to be, so some speakers may pronounce them as /səˈrɪ:f/ and /ˈhɒmɪdʒ/, similar to sheriff and image, and in
accordance with §Basic correspondences. If this were true (my examples are anecdotal), then this suggests that readers can re-analyse spellings according to the correspondence patterns of different subsystems and thence construct a new pronunciation. This would be particularly likely to happen when people are unfamiliar with a word’s etymology. This idea will be discussed in detail in Chapter 5, when we look at spelling pronunciation, because if we can understand how readers can re-analyse words, then this will provide a huge amount of insight into how they analyse them in the first place.

Finally, borrowings from certain languages may be more compatible with English than others, and it would be instructive to find how Anglicised borrowings have interfaced with the (ever-changing) writing system over the history of the language. German and Spanish words might be adopted without change (nowadays) but Gaelic spellings are not so easily absorbed into the system, at least not outside Ireland and Scotland. Carney puts Gaelic and Welsh spellings into his §Exotic category, regardless of how speakers actually pronounce them (e.g. Taoiseach ≡ /ˈtəɪsəx/ for speakers of Irish and Eisteddfod ≡ /əˈsteɪdʊd/ for speakers of Welsh, but others may have a pronunciation that fits the phonotactics of English better). The concept of spellings being ‘exotic’ is consistent with Carney’s use of RP as his reference accent (although he calls it ‘Southern British Standard’, or SBS), but these Gaelic and Welsh correspondences are perfectly manageable for people with the appropriate linguistic background. This suggests that knowledge of peripheral spelling patterns may differ greatly among users of English worldwide, regardless of the depth of their core knowledge. It will not be possible to cover such vast territory in this thesis, except to acknowledge that the worldwide influences on the English lexicon have given rise to many of the §Exotic correspondences, including units such as <kh> or <dh> (e.g. khaki and dhal). Such spellings are explored in detail by Upward and Davison (2011: Ch. 8) who build on Carney’s ideas, which, despite their importance, are only sketched out over five pages (Carney 1994: 96–101).

Upward and Davison (2013) provide a book-length examination of the etymological inputs into English spelling and their focus is on correspondences rather than words or morphemes. This is a very valuable reference book but it provides very little theoretical development, which is to say that the subsystems are presented as a disconnected set of sub-regularities, rather than as different manifestations of some deeper principle of spelling. Their work does not provide a thorough investigation of which words and combining forms fit into which subsystem. A fuller study of English subsystems would need to combine Upward
and Davidson’s detailed analysis of individual correspondences with Carney’s observation that the subsystems operate across entire words and morphemes.

To summarise, it is well-established that English spellings contain a lot of etymological information and some of the more recent studies of the writing system have begun to include that information in their descriptions. However, there remains the problem of how to fully integrate these facts into the theory of English spelling. A simple solution to this problem will be posed in Section 8a, where it will be argued that etymological spelling and morphemic spelling are deeply connected and can be explained as being manifestations of the same principle of spelling formation.

2c.4 Spelling patterns, markers and graphotactics

It should be clear by now that English spellings represent some combination of sound and meaning, and this is because spoken language involves these two elements. One outstanding issue is how those elements are combined in the writing system.

The English writing system has a range of patterns for disambiguating potential confusions between spelling and sound. One example is the use of double consonant letters, which are used to mark vowels as short (hence, *gammon* versus *Damon*). Another, somewhat related device, is the use of *<e>* to mark the length and quality of vowels (*mate* v. *mat*). Morpheme-final *<e>* is also used to mark the quality of consonant letters (*tens* v. *tense*), and this job can sometimes be done by other letters (*ague, trafficking*), etc. No theoretical consensus has yet been reached about their status of such patterns within the system.

Venezky divides English spelling units into two categories: *relational units* and *markers*. Relational units are ones that have direct phonemic correspondences, including letters, digraphs and trigraphs such as *<c, h, t, ch, th, tch>* , etc. Markers are letters used 'to indicate the function of correspondences of other graphemes and to preserve graphotactical or morphological patterns' (Venezky 1970: 55). By ‘graphotactical patterns’, he means 'patterns which relate spelling to sound' (Venezky 1967: 75), and this essentially means all spelling-to-sound correspondences. Venezky groups markers according to their function.

Word-final *<e>* can indicate vowel length (*mate, mete, mite, mote, mute*) or consonant quality (*notice, manage, tense*, etc.). *<u>* can indicate *<g>≡/ɡ/ (guess, guide, ague, etc.*) and *<k>* marks *<c>≡/k/ (mimicking, panickey, etc.*) The *<e>* in *notice* and *manage* is perhaps the most illustrative example of Venezky’s concept
of markers. In both words, the <c> indicates the value of the preceding letter, so <c> ≡ /s/ and <g> ≡ /dʒ/ (respectively /nəʊtəs/ and /mæндʒər/). In noticing and managing, the <e> can be dropped because the suffixed <i> can perform the same marking functions. By contrast, the <e> cannot be dropped in noticeable and manageable because the <a> is not a potential marker, hence there is no *<noticable> or *<managable>. The resultant <ea> sequence is cumbersome but separable because it occurs at a morpheme boundary.

Carney argues that Venezky’s concept of markers does not stand up when it is applied across the entire writing system. For example, it is inconsistent in relation to <gu>. In guest, the <u> is a marker, indicating <g> ≡ /ɡ/ but in guard it performs no such function, and *<gard> would serve the same function. Carney prefers to set up <gu> ≡ /ɡ/ as a unit of spelling, regardless of what follows it, and this is corroborated by the existence of the <bu> ≡ /b/ in build and the <cu> ≡ /k/ in biscuit. This is because Carney’s theory requires mappings to go in both directions, whereas Venezky can have rules and sub-rules for the process of decoding spellings, without worrying about whether they work in the opposite direction. Carney’s description thus requires one-to-one mappings from spelling to sound so that they can be reversed without ambiguity. Venezky’s description, by contrast, can have many-to-one mappings because his description does not have to be reversed. Hence the redundancy of the <u> in guard is not a problem since readers only have to recognise the word pronounced /ɡɑːd/ or /ɡɑːrd/. He is not concerned with the fact that writers have to remember to include a <u> when spelling the word.

Carney does not entirely abolish the use of the term marker. He writes about subsystem markers (such as the conspicuous presence of <ph> in §Greek words) and further <e>-marking functions, such as the <e> of tense, which disambiguates the word from tens and also marks <s> ≡ /s/ not /z/ (Carney 1994: 129). Rollings (2004: 66–70) uses the term marker for an even wider range of spelling devices, effectively everything that deviates from the principle of one letter to one phoneme. Hence he includes double-consonant letters and its absence (hopped v. hoped); all digraphs and trigraphs indicating a single sound (bath, batch, back); and silent letters ‘in general’.

Another elusive term is graphotactics, and the related forms graphotactic rules and graphotactic constraints. In a short paper, Stubbs (1986) distinguishes between correspondence rules (letter to sound mappings), adaptation rules (e.g. the dropping of <l> in careful, spoonful, etc.) and graphotactic rules, which means rules of letter-patterning that have no bearing on the phonological correspondence (e.g. <silly, silliness>). Carney (1994: 66–7) argues that most of Stubbs’s
adaptation and graphotactic rules can be re-cast as correspondence rules, and any outstanding rules are included within Carney's category of graphotactic constraints. Most of them are defined negatively, describing letter sequences which cannot occur. For example, words cannot end in <i>, <j>, <q>, <u> or <v>, and this explains the final <e> of blue and love, the <y> of sly and the <w> of law, etc. Spellings such as gnu, Iraq and raj are thus considered exotic, since they break some simple graphotactic rules.

The concept of a graphotactic constraint, argues Carney (1994: 67), is somewhat dependent on the concept of graphotactics, which in turn is a concept derived from phonotactics. If we consider phonotactics to be the rules governing the possible sequences of phonemes in a language (Giegerich 1992: 151; c.f. Section 5b), then graphotactics would be concerned with the possible sequences of graphemes in a language. Unfortunately, the term 'grapheme' has two incompatible uses in the literature. One is the abstract form of a letter, meaning that <t>, <c> and <h> are all graphemes; Kohrt (1986) calls this the analogical view of the grapheme, because the concept is formed on analogy to the phoneme. Hence allographs are different graphetic forms of a particular letter, such as ‘t’, ‘t’, ‘t’, and ‘T’, etc. By contrast, the referential view of the grapheme includes all functional spelling units which relate to a phoneme, so <th>, <ch> and <tch> are all graphemes, as are all of the letters (see also Henderson 1985; Ryan 2011, 2015). Carney notes that the concept of a graphotactic constraint is thus dependent upon whether we are talking about permissible sequences of letters, or permissible sequences of spelling units used to represent phonemes. Venezky (1967: 75) uses the term graphotactics as one of the two most fundamental elements in orthography. To him, there are graphemes, or 'classes of letters', and there is graphotactics, the 'patterns which relate spelling to sound'. The term is basically synonymous with his subject matter and his book could equally be called The structure of English graphotactics. Yet he scarcely uses the term again. For Rollings, 'the borderline between graphotactic constraints and markers might be considered hazy' (Rollings 2004: 27).

One thing that unites all of the various definitions of marking devices and graphotactic constraints is that they are seen as a descriptive residue, since they describe what remains after the normal correspondences have been enumerated. This viewpoint is ultimately a result of having a description based on synchronic principles. If we try to take a snapshot of the writing system at a particular time in its history, as Venezky and Carney have done (i.e. looking at 20,000 words or more), then there will always be a large set of regular, predictable patterns and varying degrees of irregularities and sub-regularities, however these are
categorised (as ‘graphotactic constraints’ or ‘markers’ or otherwise). A purely
synchronic view makes it impossible to tell exactly which patterns are productive
and which are vestigial, or non-productive. It is only by examining the formation
of new spellings, and the spelling of polymorphemic words, that we can see which
patterns are functional and productive and which are not. In this study, the use of
double consonant letters and <e>-dropping/retention take centre stage because
they are viewed as the glue which binds spellings together in the formation of new
and polymorphemic spellings. Their function is to solve problems of intelligibility
between the application of morphemic spelling and the need to make spellings
readable.

2c.5 Conclusion: Outstanding issues in the study
of English spelling formation

The consensus from previous research is that polymorphemic English words tend
to keep the spelling of the morpheme constant, as much as possible, and that the
spelling of monomorphemic words, or base forms, has either been inherited from
within the system’s long history, or has been borrowed from other writing systems
(especially French, Latin and Greek). It is not fully clear how important the
phonological form of a word is in the formation of spellings. In some cases (e.g.
*hopped, hoped, profanity) there appears to be interaction between the need to
represent the morpheme consistently and the need to represent the phonological
form in a decodable way. Yet it is not exactly clear what conditions underlie these
interactions. One hypothesis is that the spelling of affixes is more stable than the
spelling of stems (Berg et al. 2014), but even this insight does not explain why
certain spellings are disqualified (e.g. *<profaneity> and *<serenity>). We thus
need a model of English spelling formation which can explain not just the
spellings in use, but also why certain spellings are not permissible. The model will
need to explain the relationships between spelling, morphology and phonology,
the nature of spelling formation itself, and how conflicts are resolved between the
need to represent both sound and meaning. And finally, we will need to include
the writing system’s wide-ranging use of ‘marking devices’ and ‘graphotactic
constraints’.

To conclude, it is worth repeating the research questions that will inform
the remainder of this thesis.

1. To what extent does English spelling represent morphemes
   consistently?
2. How does this interact with the need to represent the phonological form of a word?
3. How can etymological spelling be included in a theory of English spelling formation?
4. How do we know if an English spelling is well-formed?
5. What is the actual process of spelling formation?
6. How can spelling formation be modelled?

***

***
Chapter 3

A method for analysing spelling formation.

In this chapter, a hypothetical model of English spelling formation will be developed for examining how polymorphemic English words are spelt. This approach is rather different to previous studies of English spelling and thus we must first look at a number of important issues.

Sections 3a looks at the problem of delimiting the object of focus. Previous studies have examined the core lexicon en masse, whereas the focus in this study will be on certain revealing subgroups of words. This method is taken directly from word-formation studies where new words are the object of study, and predicting their formation patterns relies on understanding patterns among existing words. For example, if we know that sisterhood and brotherhood are existing words, then we might wonder whether unclehood and cousinhood are permissible words, should there be a context for their use.

Section 3b shows how word-formation theory can be applied to spelling-formation theory. In particular, we shall see that it is beneficial to employ a mix of synchronic and diachronic approaches to the topic, a new approach that follows Marchand’s (1969) effort to do the same for word-formation theory. We shall see that English words have been subcategorised by morphologists according to ways in which they were formed (e.g. compounds, derived forms, inflection, etc.) and these categories will provide a useful structure for examining how spellings are formed. Section 3c looks at how data was collected for analysis.

The latter parts of the chapter are rather different. Section 3d takes a look at one other attempt to explain how English spellings are formed, including Berg and Aronoff’s (2017) historical study of how certain affixes have come to be spelt. Like Cummings (1988), they argue that the English writing system is self-(re)organizing. That is to say, spelling patterns change over time due to variation in usage. In later chapters, this historical, numerical approach will be used as supplement to the theory developed here. The goal of Section 3e is to enumerate and explain five principles of spelling formation. These principles vary in their nature and they relate, in several ways, to DeFrancis’s universals of writing systems and Cummings’ (1988) ‘demands’ on spelling. The chapter concludes by hypothesising that English spelling formation can be explained by following a simple algorithm, one which will be developed in later chapters.
3a. Scope and focus of the study

Exhaustiveness is one of the fundamental desiderata of linguistic theory and Carney (1994: 32) notes the long-standing failure to apply such standards to spelling research. A description of a linguistic system, such as a writing system, ought to be an attempt, therefore, to describe the spelling of all existing words. We might also imagine that a more developed theory would predict future developments. And yet it remains necessary — and practical — to delimit the object of study, to be exhaustive in one particular approach.

Venezky and Carney choose to study the entire core of English spelling, by examining approximately 20,000 and 26,000 spellings (respectively), while excluding several kinds of spelling, notably inflections, certain abbreviations, onomatopoeia, etc. (c.f. Section 2b.3). By contrast, Rollings (2004) and Cummings (1988) use phonemic theory as a structuring model, investigating all the spelling units used for all the phonemes of one English accent (RP and GA, respectively). In all four studies, however, the central aim is to examine the correspondence patterns between spelling and sound, in one or both directions.

These approaches can tell us a lot about the core spelling patterns in the writing system, but they cannot explain peripheral spelling patterns, nor how the core and the periphery relate to one another. By defining a core stock of standard spellings, there arises an analytical segregation between standard and non-standard spellings, making the latter seem aberrant, rather than subject to different formation principles. Furthermore, it may be a lot harder to define what counts as an English spelling, and what counts as a foreign word (Upward and Davidson 2011). Each variety of English tends to borrow words from the local languages with which it has contact. Hence, spellings and correspondences that may seem (relatively) normal in one variety will seem most foreign elsewhere. Even neighbours such as Irish and Welsh English will have radically different vocabulary — and spelling patterns — borrowed from their respective indigenous languages.

The need for exhaustiveness will be met in a new way in this study. It will be assumed that it is useful to study all the different kinds of spellings and hence the different ways that spellings can be formed. There will be two contrasting strands to this investigation. One is to partition the lexicon according to the manner in which words were formed. For example, it will be assumed that the spelling of compound words and affixed derivatives are formed somewhat differently from one another, while inflections will be different again, and so on for various types of word formation. This assumption follows on from Venezky’s
conclusion (1970: 121) that the spelling of base forms is rather different to the spelling of compound and derived forms, with the former being broadly phonemic and the latter being broadly morphemic. In the next section we shall draw upon theories of English word-formation in order to partition the lexicon according to how words were formed. No attempt will be made to enumerate the incidence of any one group (e.g. compound words). Instead, the goal will be to see what patterns underlie their respective formation patterns.

The other strand of investigation will be to look at the spelling of everything that doesn’t fall inside the core lexicon. This includes names, non-standard spelling, certain foreign borrowings, and so on. These categories will be discussed in Chapters 7 and 8, in light of the findings of earlier chapters.

For reasons that will become clear in the next section, much of the focus will be on the spelling of English inflections (Ch. 4) and certain kinds of names (Ch. 7). It will not be possible to study all kinds of words in as much detail, but an effort will be made to show that each group is worthy of study in its own right, and different spelling patterns may apply to different kinds of words.

***
3b. A synchronic-diachronic approach to English spelling formation

In his brief history of the study of English word formation, Bauer writes:

The distinction between synchrony and diachrony drawn by Saussure, which has had a profound effect on linguistic studies since 1916, effectively precluded the study of word-formation, where synchrony and diachrony are most fruitfully considered together.

(Bauer 1983: 3)

It is poignant that Marchand’s landmark study, *The categories and types of present-day English word formation* (1960, revised 1969), had the subtitle: *a synchronic-diachronic approach*. In word formation, synchrony and diachrony are deeply intertwined. The reason is this: after a new word is coined it may stay within the speech community; it then enters people’s mental lexicons and sometimes even dictionaries, thus giving it legitimacy in people’s minds. Words are created by synchronic rules, but also capable of diachronic survival. Sentences, on the other hand, are uttered, perceived and then forgotten. Hence, syntax is suited to synchronic study, but word-formation requires a blend of synchrony and diachrony.

New spellings are also formed by applying synchronic rules of spelling formation and, if they survive, they are recorded in dictionaries, where they gain a more permanent status and legitimacy. The English writing system can therefore be viewed as an *accumulative system*, whereby spellings are formed and the accepted ones are ‘added to the pile’, one by one, over a long period of time. Each new spelling thus contributes minutely to the overall system. Moreover, the writing system can change in several ways, even when the spelling stays the same. One change, pointed out by Scholfield (2016), is that the correspondences between invariant spellings and changing phonologies become more complicated over time (the subject matter of Chapter 6). Another kind of change is that the amount of variation in the spelling of any given word or morpheme seems to diminish over time (see Berg and Aronoff 2017). A new word may have multiple competing spelling variants but older words tend to stabilise in their spelling. Finally, as new words get added to the language, new spellings get added to the writing system, one by one, with each new word potentially adding to the tangle of correspondences, especially if the new spelling has been borrowed from another writing system (e.g. Italian, French, Latin, etc.).
In previous studies, the writing system has usually been studied as a synchronic mass which does not take into account the fact that different formation rules may have applied at different periods in the history of the language and this is why the present-day system contains so many vestigial patterns. It is therefore problematic to apply the exact same analysis to an ancient spelling as a modern one, despite the fact that learners of the system have to acquire the information en masse (which is why the issue of diachrony is avoided by so many linguists studying writing systems). All we can really do is ask how new words are spelt nowadays and how complex words have been spelt in recent times. It is not even necessary to put a date on this (e.g. Johnson’s (1755) dictionary) because spelling alterations are incorporated into modern dictionaries, so subtle changes may be lost on the contemporary reader/writer.

The spelling of simple, monomorphemic words (or base forms) cannot be accounted for by this approach, as we shall see in the next section. The spellings of these words have been handed down to us, often with great irregularity, and they cannot be changed. However, new words do not have to perpetuate these trends nor do polymorphemic words which can be subtly altered to keep them in line with the spelling of related words.

### 3b.1 English word-formation and English spelling

#### 3b.1.1 New words and productivity

The study of English word-formation focuses on how new words are constructed; existing words, or base forms, are simply accepted as inputs to the system. For example, a word such as nation is not of interest per se, and the focus in word-formation studies is on how the base form can be recycled in related words such as national and nationhood. Topics of interest may include phonological changes caused by the addition of the affix. The suffix {hood} never affects the pronunciation of the base form but the suffix {al} sometimes does: compare ‘function – functional, with no change, against ‘hexagon – hexagonal (Plag 2003: 119).

Word-formation theorists may also be interested in the productivity of an affix. Given the existence of the words nation and nationality or ration and rationality, we may ask whether the word passion has a corresponding form *passionality*. Some affixes are productive while others are not. For example, the suffixes {th}, {ity} or {ness} all form abstract nouns in English. The first of these is a dead pattern, occurring only in ancient forms such as long – length, grow –
growth, slow – sloth, moon – month (OED 200044). The fact that the spelling of the base is not preserved in the all of the complex forms is an indicator that such spellings derive from a time when this was not such a priority in the writing system. The affix {ity} only occurs in borrowed §Latinate forms, as seen in the pairs serene – serenity, ‘possible – possi bility (note also the stress shift), or singly in maternity. The third form, {ness}, occurring in happiness, sadness, calmness, etc., is productive, occurring in new formations, however silly they may seem: yellowness, greenness, blue-ness, or indeed longness, should there be a need to distinguish a special kind of length, such as the ‘longness’ of a vowel. The suffixes {th} and {ity} do not occur in new formations, so we cannot have *yellowth or *yellowity.

The focus in word-formation studies is not just on which words happen to exist but which words are possible, and this can only be done by: a) looking at new formations and, b) extrapolating from existing patterns in the lexicon (i.e. the stock of existing words in the language).

3b.1.2 Spelling formation

The present study is similarly focused on how the spellings of new words are formed, and (almost) no attempt will be made to explain how the spelling of base forms emerged in the first place. For example, the irregularity of the spelling <sloth> is not of interest, and we do not have to account for the ‘missing’ <w>. Instead, the focus is on how the letters <sloth> get recycled in related words, such as slothful and slothfulness, or why bathe loses an <e> in the inflected form bathing. This approach is hugely at odds with studies of English spelling which focus solely on the relations between spelling and sound in a closed set of words (e.g. the ‘20,000 most common English words’). Instead, it will be assumed here that English spelling is an open set, just like the lexicon, and that new words require new spellings. This makes it possible to focus on subcategories of the lexicon which demonstrate revealing spelling patterns. For example, the inflections {ed} and {ing} are highly productive and can occur to any noun or verb, often giving rise to new orthographic challenges. Should {cliché} + {ed} be spelt <clichéd>, <cliched> or <clichéed>? The focus will often be on such extreme cases which test the boundaries of the writing system and yet demonstrate what information tends to be prioritised in the resulting spelling. There is seldom any need to pay attention to the number of words in a particular subgroup (e.g. how many compound words are there) nor to ask how often a particular correspondence occurs (how often is /f/ spelt <ph>, for example). Instead, the aim will be to look at how conflicts are resolved between opposing linguistic forces.
(e.g. the representation of sound and meaning). In order to do this, we need to first identify the different ways in which English words are formed and this will provide a framework for examining whether each group displays different spelling patterns.

3b.1.3 Inflection
Like other morphologists, Plag does not include inflection as a category of word formation because an inflection is not a new words per se but ‘a different form of the same word’, encoding grammatical information (Huddleston and Pullum 2005: 15). For example, nouns may have singular and plural forms (cat, cats). Verbs may be marked for tense (take, took, taken, taking) or for person (takes). Derivation, by contrast, results in entirely new words, even though they may be closely related (e.g. electric, electricity, electrician). Inflection and affixed derivation may be very different from a syntactic and semantic viewpoint (Plag 2003: 18–23), but both are subject to very similar patterns of spelling, because both involve suffixation.

Chapter 4, and much of Chapter 5, will focus on the spelling of inflected forms (such as hops, hopped and hopping), because these ‘words’ (as they will be referred to henceforth, for ease of reference) exhibit some very common and revealing spelling patterns. Regular, suffixed inflections are simply appended to the base phonological form and they have no effect on the stress pattern (loc. cit.). Furthermore, inflectional patterns are formed entirely within English and they are not affected by subsystemic spelling patterns, hence we see the spelling patterns in §Latinate nationalises, nationalised and nationalising, as we do in §Basic raises, raised, raising.

It is therefore feasible and relatively easy to study in isolation how sound and meaning are represented, in interconnected ways, in this one category of English spelling. In Chapter 4, we shall look at the highly productive inflections {ed}, {ing} and {s} (these suffixes will be defined when necessary). The method of analysis will be to assume that the complex forms are subject to morphemic spelling and then to analyse the situations in which amendments are made. For example, morphemic spelling occurs in <raises>, <jumps>, <jumped> and <jumping>, but not in <raised> and <raising>, where slight amendments are made. Otherwise, the words would be spelt *<raiseed> and *<raiseing>, resulting in ‘bad’ representations of the known phonological forms. The goal will be to uncover the basic principles by which these spellings are formed, mixing morphemic spelling with the need for a spelling to represent a word’s sound. It will then be possible, in later chapters, to see whether the principles which
underlie the spelling of inflections are applicable to the spelling of other, often more complicated categories of word formation.

3b.1.4 Other categories of English word-formation

In his introduction to *Word-formation in English*, Plag (2003) provides the following major categories for describing how new words are formed, and examples of each of these are presented below:

- **Compounds**
  - milkman, milkshake, milk thistle
  - upturn, uptown, uphill

- **Derivation with affixation**
  - milk.y
  - nation – nation.hood – nation.al – nation.al.ise
  - hexagon-hexagon.al, mode – mod.al, patella – patella.r

- **Conversion**
  - *table* (noun) and *table* (verb)
  - *house* (noun), *house* (verb)

- **Truncated forms (clippings)**
  - television – *telly*; telephone – *phone*
  - refrigerator – *fridge*; vegetable(s) – *veg*
  - microphone – *mic / mike*

- **Neo-classical compounds**
  - tele-vision, tele-phon-e, tele-path-y, phon-o-log-y,
  - ortho-graph-y, photo-graph-y

- **Blends**
  - *Oxbridge* (*Oxford + Cambridge*)
  - *Tanzania* (*Tanganyika + Zanzibar*)
  - *Brexit* (*Britain + Exit*)

- **Abbreviations and acronyms**
  - *FBI* (from ‘Federal Bureau of Investigation’)
  - *Laser* (from ‘Light Amplification by Stimulated Emission of Radiation’)

Each of Plag’s categories will be examined in turn in Chapter 8, where we shall see that there are several ways in which English spellings can be formed.
However, there are two other groups of words which will form the backbone of the present study: inflections and names.

### 3b.1.5 Names
Plag does not discuss names (in detail) since they are not usually subject to the same morphological processes as lexical items. However, names have spellings and, even though they may be subject to slightly different patterns, readers of the language must be able to understand those spellings, and writers must be able to reproduce them. Hence they will be of interest in this study, especially in Chapter 7, when we look at the spelling techniques used to make distinctive spellings which do not suggest a different pronunciation (e.g. *OutKast* instead of *outcast*). Such spellings will be examined in light of a fuller discussion of the different categories of non-standard spelling.

### 3b.1.6 Etymological spelling
It was noted above that base forms are seldom the object of interest. However, base forms borrowed from other languages are frequently spelt according to the spelling patterns of those languages, and this is something that has never been adequately theorised in studies of English spelling (see Section 2c.3). For example, the word *nation* is modelled on the spelling of the Latin word from which it derives. If the spelling were formed through the §Basic, or native, subsystem of spelling, it might have been spelt with <sh> rather than <t> (compare *fashion, cushion*).

Furthermore, borrowed words may be polymorphemic in the donor language and this can prevent us from making a clear distinction between monomorphemic and polymorphemic forms in English. For example, the words *apply, offer, suffer, suggest, serene*, etc., all contain fossilized Latin suffixes which cannot be separated within English (i.e. there is no *gest or *rene). And yet sometimes polymorphemic borrowed words may be re-analysed within English, so that the word *edit* is formed out of the Latin-derived word *editor* (OED 59546).

One of the tasks of this study will be to account for how such spellings enter the English writing system (Section 8a), and it will be important to distinguish this from how existing English spellings behave within the writing system.

### 3b.2 Written forms not under discussion
The present study is focussed on spelling in the strict sense of abstract written forms that represent abstract linguistic forms. Visual writing is dependent on the
letters being visibly distinct and recognisable, and this in turn requires a predictable path from the actual letter forms to the abstract letter forms. The path is not quite one-to-one, as one abstract linguistic form, represented by \(<a>\), can be represented by several different glyphs, including ‘a’, ‘a’ and ‘A’. (The significance of capital letters will be discussed in Section 7b). Thus no attention is paid to the graphic realisation of the written forms, and hence there will be no study of their significance. This is in keeping with almost all linguists interested in writing systems (since at least Gelb, 1962). While the relationships between glyphs and abstract letters can easily be taken for granted, there are situations where that relationship can be disrupted. For example, writing can be rendered through sound (Morse Code), touch (Braille) or using flags (Semaphore). These systems are reliant on the abstractness of the letters, not on the actual medium by which those symbols are relayed (see Harris 1996).

![Google Doodle celebrating New Year’s Day, 2007](image)

**Figure 3.1**
The Google Doodle celebrating New Year’s Day, 2007

Some scholars may be interested in the choice of typeface (e.g. Kress & van Leeuwen 2002; Spitzmüller 2012), while Ryan (2015) focuses on cases where glyphs have two or more functions, on account of their size, shape, position or colour. This often occurs in Google Doodles, such as the one shown in Figure 3.1. The ‘G’ and ‘e’ perform their usual alphabetic functions of representing \(<G>\) and \(<e>\), but the middle four glyphs each have three functions. They spell out the letters \(<oogl>\); they spell out the year 2007; and they also function as pictures of fireworks. While such examples may seem extreme, they seem (anecdotally) to be getting more popular, especially in professional graphic designs. There is no reason to assume this trend will not spread to normal, ‘amateur’ writing, and it already exists a lot in the linguistic landscape (see Kallen 2016). It should also be noted that medieval manuscripts, which were not constrained by the graphic limits
of print, often availed of the graphic dimensions, using colour, shape, size and position to great effect (e.g. Machan 2011).

3b.3 Other limits to the study

No effort is made to test whether the analytic methods developed henceforth are applicable to other writing systems. They may at first appear to be useful but the way in which language is represented in other writing systems may be controlled by very different circumstances. For example, English writing is mostly based on a closed set of symbols (twenty-six letters, plus an assortment of marginal extras) so solutions to representational problems must be found from within that set. Other Roman alphabetic writing systems tend to use diacritics to resolve certain problems, while Cyrillic alphabets have a much greater tendency to invent new letter forms (Daniels 2009: 38). Character invention is fundamental to the Chinese and Japanese writing systems (DeFrancis 1989; J. Joyce 2011). Such changes can have a radical effect on the writing system and, furthermore, each language presents its own linguistic problems.

No effort is made here to analyse English writing at early stages in its history, so it certainly cannot be assumed that methods of spelling formation in use today were exactly the same in the distant past (or will be in the future). What might be assumed is that the representation of sound used to be a greater priority than the representation of meaning, as exemplified above by long – length, slow – sloth, etc.18 Furthermore, the grammatical system was radically different at earlier stages in the language and this also may have had a huge impact on the needs of the writing system. The relative scarcity of inflection in Modern English grammar (presumably) makes it easier to represent morphemes invariantly in the writing system.

***
3c. Data collection

While a clearly-defined corpus of spellings is necessary for computer-based analyses of English spelling, not all theoretical innovations have come from studies based on a single corpus (e.g. Albrow 1972, Cummings 1988, Rollings 2004). There are some simple reasons why this is possible. All literate users of the language have a large database of spellings stored in their lexicon and these spellings count as valid data, much more so than syntactic structures, because spellings are fixed and agreed upon by users of the language. This makes informed introspection quite feasible, especially when the focus is taxonomical, looking at different spelling practices, rather than focusing on the statistical prevalence of particular correspondences or even the prevalence of particular word-formation categories. Such introspection and observation can then be backed up with searches in dictionaries (and other wordlists) because they function as ordered databases of spellings. Dictionaries are particularly useful if one is looking for related words whose spellings begin with the same letter sequences (e.g. theory, theorist, theoretical, etc.). It is thus very easy to look for words beginning with the same prefix but not so easy to find words ending in the same suffix. It is also difficult to find incidences of any non-initial letter sequence (e.g. which English words contain <eig>, in that order).

One way to conduct a targeted search for a suffix is to use a dictionary in reverse alphabetical order. Venezky (1970: 12) notes how he simply reversed the spellings in his database, while Plag (2003: 93) lists some reverse dictionaries designed for the purpose. In particular, Muthmann’s (1999) dictionary sorts words according to their morphological composition, so that all words ending in the suffix {ion} are listed together, but words such as lion, which happen to end in <ion>, are weeded out. Nowadays, targeted searches can be performed much more easily, as dictionaries are available online. Another option is to search on websites designed for the purpose, such as morewords.com. Here, one can easily search for any sequence of letters in a word, whether they are in consecutive order or not. One might search for all words ending in <ion> (using the search term <*ion>, where ‘*’ means ‘zero or more letters occurring up to that point in the spelling), or one can also look for co-occurrences of particular sequences. This makes it possible to check, for example, the co-occurrence of, say, <ch>, <ph> and word-medial <y>, as one might find in §Greek words. This is how the examples of chlorophyll and psychopharmaceutical were found (by using the respective search terms <*ch*ph*y*> and <*y*ch*ph*>. Other combinations were also attempted, (such as <*ph*y*ch*>, etc.). Morewords.com also stores inflectional forms
because, unlike paper dictionaries or even Carney and Venezky’s early databases, present-day computers have no trouble storing the extra (supposedly predictable) spellings which inflected forms add to the corpus.

The database used on Morewords.com often throws up results that are either unfamiliar or not real English words, so it is often necessary to check its results against the OED or other dictionaries (I have frequently used thefreedictionary.com as it hosts several dictionaries and encyclopedias on one website). For example, I discovered the obscure word brach while searching on morewords.com for words ending in <ach>. I could then look up the existence, meaning, pronunciation and etymology of brach in the OED (22390). This kind of word may not be among the most common words of English, but it still contributes to overall patterns in the writing system. This is important because when one encounters a word as rare as this, one must make reading decisions based on one’s knowledge of the writing system. Should the word be rhymed with batch, back or bash? Or perhaps it might be pronounced with /ʃ/, a pronunciation often attached to the name of the German composer Johann Sebastian Bach.

The OED3, the internet version of the Oxford English Dictionary, hosted at OED.com, is especially useful for etymology, the history of words, the history of their spelling and present-day variant spellings. The OED can be useful for pronunciations (both RP and GA) but there are specialist dictionaries dedicated to this purpose, including the Longman Pronunciation Dictionary (three editions, all edited by Wells, 1990, 2000, 2008) and the English Pronouncing Dictionary (fifteen editions, edited by Jones, Gimson, and now a whole team). For etymology alone, the website etymonline.com is very useful.

The vast majority of searches conducted have been based around letter sequences, but sometimes rhymezone.com was used for the purpose of finding words that rhyme or have other similar patterns. Carney’s detailed analysis is the best resource when looking for particular correspondences (e.g. pronunciations of <ough> or spellings of /ʒ/, etc). Another resource for searching phonological and spelling patterns can be found at the MRC Psycholinguistic Database, although it is based on the 1963 version of Jones’s English pronouncing dictionary and is thus tied strongly to one accent. Such resources are useful for doing combined searches for the co-occurrence of particular letter sequences alongside particular prosodic patterns.

Finally, in Chapter 7, a corpus of spellings is used for the study of band names and details are provided at that point. Data from other studies are also examined in that chapter.
The data-collection methods discussed above have proved less fruitful in the study of spellings which have not received much treatment in the literature and which tend not to be listed in mainstream dictionaries. This includes abbreviations, clippings, various kinds of non-standard spellings and compound words whose spellings include a space. Examples discussed have been taken from studies of these words or from personal observation. The one advantage of this ‘magpie approach’ is that it allows one to continually look for new and unusual patterns, rather than being confined to a corpus that was collected early in the research before ideas were well formulated. A new pattern will often reveal further patterns and this in turn opens up new portals of research into spellings that may have otherwise not seemed of interest.
3d. English spelling as a self-reorganizing system

Before looking at the approach to English spelling formation which will be developed throughout this study, it is worth looking at a rather different approach to how English spellings are formed. Cummings (1988) considers the system to be self-reorganizing, while Berg and Aronoff (2017) call it self-organising.

3d.1 The ever-changing writing system

Over time the English writing system seems to gain more words than it loses and this is (presumably) because written records preserve words that might otherwise fall out of use. Dictionaries also play a role in the preservation of ancient and obscure words, and historical dictionaries (such as the OED) take on the added job of preserving variant spellings, including those that have fallen out of use. Over time, the number of words with variant spellings has diminished, and older spellings may be brought into line with present-day patterns. Cummings (1988: 6) gives the example of millionaire which used to be spelt <millionnaire>, as it is in French, whence it came (OED 118560). Significantly, the new spelling can be re-analysed as an amalgam of two existing spellings: the word <million> and the bound morpheme <aire> which recurs in billionaire and commissionaire, etc. (OED 236398). In this way, Cummings considers the writing system to be self-reorganizing, with each word providing its own tiny contribution to the overall system. The assumption underlying his observation is that writers actually perform these kinds of linguistic analyses in any given act of spelling.
A more familiar, and more drastic, example of self-reorganization has seen the spelling <gaol> being almost entirely replaced by <jail>. The OED (100650) gives <jail> as the primary spelling and <gaol> as an acceptable variant, but Google Ngrams shows the massive disparity in usage that has emerged between the two spellings over the past two centuries (see Figure 3.2). The effect on the writing system as a whole is that the frequency of the correspondence <g(a)> ≡ /dʒ/ is further diminished, although it survives in *margarine* (at least in the more common pronunciation, OED 114020). Similarly, the incidence of <ao> ≡ /eɪ/ is reduced to nil, according to Cummings (1988: 257), although one might include *aorist* as having this correspondence. Once again, the accumulative effects of individual spelling choices result in old correspondences being slowly eroded from the writing system. The availability of tools such as Google Ngram Viewer has begun to allow linguists to chart and measure the progress by which one spelling variant prevails over another.

### 3d.2 Berg and Aronoff’s method for studying spelling formation

It has long been assumed that the stability of modern spelling can be put down to the influence of printers, dictionary makers, and (some rather obscure) early grammarians. This view, expressed most fully by Scragg (1974), is rejected by Berg and Aronoff, who argue (2017: 37–38) that the (relative) stability of modern English spelling has emerged slowly out of an earlier state of near-chaotic variation. That is to say, each word used to have several possible spellings but, slowly, through use and consensus, a single spelling emerges as *the* spelling of the word. As a result, spelling variation is almost always obsolescent because of the tendency towards a uniform spelling of any given word or morpheme. While this idea was recognised by Cummings (1988), Berg and Aronoff (2017) are the first to study numerically how one spelling variant prevails over others. Their approach is very new and thus there is a lot of detail about a small set of cases, but the method could, in principle, be applied to explain any how English spellings are formed. I have thus included a brief discussion of their findings as an alternative to my own method for studying how English spellings are formed.

Instead of focussing on the spelling of base forms, Berg and Aronoff focus on the spelling of a few affixes, including {ic}, {ous} and {al}. They show how the spellings <ic>, <ous> and <al> have come to prevail over variant forms such as <ick, ique>, <ouse, us, ows> and <all, alle, el>, etc. The result is that one spelling has become consistently associated with each affix while the other
spellings were either ‘washed out of the system’ or re-distributed for other purposes.

Figure 3.3
The incidence of <ic> (in white), <ique> (in grey) and <ick> (dotted), in the spelling of 39 ‘word types’. Taken from Berg and Aronoff (2017: 50).

Figure 3.4.
Relative amount of tokens in the Helsinki corpus with the suffix spellings <ick>, <ik/-yk>, <icke/-ycke>, <ike/-yke>, <ique>, and <ic>.
Basis: Words ending in <ic> today.
Taken from Berg and Aronoff (2017: 49)

Figure 3.3 shows the erstwhile spelling of 39 words that are nowadays spelt with <ic>. The spelling <ic> (in white) was once a minority variant, compared to <ick> (dotted) and even <ique> (in grey). Between 1650 and 1850,
<ic> came to dominate (for these words). Before 1650, there was a lot more variation, as seen in Figure 3.4, although several variants were already obsolescent, including <ik, yk> (in black), <ike, yke> (dark grey), <icke, ycke> (lines).

Nowadays, almost all occurrences of word-final <ic> are associated with adjectives, including allergic, demonic, etc. (ibid: 47). Only 3% of <ic#> spellings are not adjectives: this includes the nouns attic, critic, republic, logic; and a few nouns which have been converted to verbs, including fabric, panic, traffic (ibid). The earlier spellings variants of {ic} have not all vanished entirely and many have been retained for different purposes. <ique> is nowadays associated with §French words, including boutique and mystique (pronounced with /i:k/ and end-stress). The <ick> spelling survives in polysyllabic nouns such as gimmick, derrick and maverick, and of course in monosyllabic words of any grammatical class — stick, flick, quick, thick, etc. The presence of <k> in gimmick, etc. indicates that these words do not contain the {ic} morpheme and hence they are not adjectives. The <ick#> spelling also contrasts against <ik#> spellings, which represents another quasi-morpheme, seen in sputnik, apparatchik and beatnik. Berg and Aronoff (2017: 47) consider there to be some other non-<ic> spellings used to represent unstressed final /ik/ in nouns, including the <ock> of buttock, the <ack> of barrack and the <uch> of §Greek eunuch. However, their analysis requires a merger of all incidences of final unstressed /ok/ and /ik/, so the argument is less convincing for these cases. Regardless, the spelling of counter-elements does not need to be consistent. The only requirement is that nouns ending in /ik/ are not spelt <ic#>, allowing that spelling to be reserved for adjectives.

The authors conclude that the tendency towards systematicity is a result of ‘competitiveness, the struggle for existence’. The wording may seem lofty but the idea is simple. A writing system which associates one spelling with each morpheme must also, by necessity, require different spellings to be used in words which sound like they contain that morpheme but actually don’t. It is the partial-word equivalent of having two different spellings for homophones. Hence the English writing system uses invariant strings of letters to connect related words and other strings to letters to distinguish unrelated words.

The process by which <ic> has become the spelling of the adjective-forming suffix {ic} is similar to the way that <jail> has prevailed over <gaol>. The success of one spelling is not the work of an individual, nor is it (entirely) the work of a dictionary which may recommend one variant over another. Instead, fixed spellings emerge out of competition between competing variants and, over time, one variant always seems to win. The survival of some variation even in the
present day is a reminder that English spellings are formed by consent over time. There is no reason why Berg and Aronoff’s method could not be applied to a much wider set of spellings, if not the entire writing system. However, such brute-force statistical methods of analysis are not the only way to analyse how English spellings are formed.

***
3e. Principles of English spelling formation

The theoretical approach to English spelling formation taken in this study will be very different from the one taken in the last section. Instead of focussing on the emergence of one spelling variant over another, it will be assumed first that the spelling of each word has stabilised and is now invariant. It will then be hypothesised that the spelling of polymorphemic words can be reverse-engineered into their constituent parts. For example, <jumping> can be broken into <jump> and <ing>, while <raising> can be broken into <rais> and <ing>. This process allows us to analyse what information is contained in any given spelling, what information is prioritised under different conditions, and why other variant spellings are not in use. Reverse-engineering provides an economical, rather than a numerical, approach to the problem of explaining English spelling formation because it requires an analysis which identifies commonly occurring patterns, rather than relying on masses of data for every word. Furthermore, one can extrapolate from the patterns to predict the spelling of new and future words, on the assumption that the same principles will apply in the future.

Nevertheless, Berg and Aronoff’s (2017) method will be applicable in cases where there remains spelling variation, as with <jail> and <gaol>, or <zinced> and <zincked>. Evidence will be drawn (in Chapter 4) from Google Ngram Viewer to see how the variant proportions have changed over time. Such diachronic data will allow us see tendencies in the prioritisation of one type of information over another and this will shed light on the deeper patterns that underlie English spelling formation. If, for example, <zincked> had become more common than <zinced>, then we might see this as evidence that indicating sound unambiguously had become more important than keeping the spelling of the morphemes unchanged (i.e. <zinc> + <ed> → <zinced>, without amendment). Such competition between competing forces suggests that English spellings are subject to some deeper principles, five of which will be enumerated below.

3e.1. Precursors to the principles

3e.1.1 DeFrancis’s universals

In Section 2a, we saw that all writing systems represent both the sound and meaning of words, and these universals apply in their own way to English. We also saw a third universal of writing systems, which DeFrancis calls ‘the dead
hand of the past’, whereby a gap grows over time between the ever-changing spoken language and the linguistic information included in fixed spellings. Each of these will be incorporated into the five principles.

3e.1.2 Cummings’ demands

Cummings (1988: 12–24) argues that English spellings are subject to competing ‘demands’. In keeping with other writing systems, the core demands are the phonetic and semantic demands, whereby sounds and ‘units of semantic content’ are both ‘spelled consistently from word to word’. This is nothing new: for sound, it means that a phoneme such as /b/ is usually spelt <b> and sometimes <bb>; for meaning, it means that a spelling such as debt is consistently spelt <debt> in related words such as <indebted>, <debtor>, <debt collector>, etc. The third of Cummings’ demands is the etymological demand. Not only has this given rise to the subsystems of spelling, identified by Carney (1994: 96–101), the inconsistency by which etymology is applied in the writing system shows the role of human intervention — this is why a dubious spelling such as <debt> has gained prominence (Cummings 1988: 19, OED 47935).

Cummings (1988: 21–24) also considers there to be an important ‘visual bias’ in English spelling, which reflects the increased standardization of modern writing. The visual bias is not a ‘demand’ because it does not contribute to the formation of the spelling. Instead, the present-day visual bias is contrasted against an earlier or pre-Modern ‘auditory bias’ whereby spellings attempted to better reflect the sound of words. He argues that the emergence of silent reading contributed to the growth of this visual bias in reading over an erstwhile auditory bias. Cummings gives two examples as evidence that the written form of a word has taken priority over the spoken form (in English at least). Nowadays, the spelling of one’s surname cannot easily be changed even though one can ask for it to be pronounced differently. This is in stark contrast to the time of Shakespeare when the spelling could be altered at will, varying between <Shakspeare>, <Shakspere>, <Shaksper>, etc. Secondly, the existence of spelling pronunciation is further evidence of a shift towards a visual bias. The use of a spelling as evidence for a particular pronunciation is predicated on the assumption that the written form has greater authority than the spoken form. At first glance, the growing visual bias may appear to be Cummings’ version of ‘the dead hand of the past’ because invariant spellings become associated with the meaning of words, no matter how much the pronunciation changes. However, the visual bias is also present in morphemic spelling, since the same visual unit is used across different related words. Hence invariance and morphemic spelling need to be disentangled,
and we need two concepts to account for them. Furthermore, morphemic spelling is a very large part of Cummings’ semantic demand.

3e.2 The principles
One innovation of this study is to re-organise the ‘demands’ or ‘universals’ into principles which are applicable to the way in which spellings are actually formed. They are, as follows:

1. Any-Spelling Principle
   All words must have some spelling.

2. Distinctiveness principle (DSTNCT)
   Different words should have different spellings.

3. Identity Preservation Principle (IdP)
   Related words should have related spellings.

4. Phonographic Matching principle (PhM)
   A word’s spelling should represent its sound.

5. Invariance principle
   Once a spelling has been settled upon it cannot be changed.

The exact nature of the principles varies. At first sight, one might think that the first and last principles do not refer to linguistic information. However, the Any-Spelling principle shows that it is actually possible to for writing to represent words, while the Invariance Principle implies that the sound and meaning of words tend to be sufficiently stable not to warrant change — despite calls for spelling reform.

The other three principles refer to the manner in which linguistic information is contained within the spellings themselves. PhM is rather like Cummings’ phonetic demand because both involve the matching of symbols to sounds. IdP, however, is a little different from Cummings’ semantic demand. The principle refers to the relations of spellings to one another: the preservation of the same CPU in <medic>, <medic.al>, <medic.ine>, <medic.inal> and <para.medic> demonstrates the strong semantic connections between these words. Yet IdP is really the logical conclusion of DSTNCT which is itself structuralist in nature. Meaning can be understood through difference: <A> only means something
because it is not <B>, and neither of them is <C> or <D>, etc. Similarly, at the word-level, we can see, from the spelling, that <aye> is not <bye> which is not <dye>, etc. Both IdP and DSTNCT are semantic in nature — one preserves relations of meaning, the other separates them. Since Principles 2, 3 and 4 are the most important in the present study, they will be discussed first.

3.2.1 Distinctiveness principle (DSTNCT)

In most cases, DSTNCT is so fundamental that it is hardly worth stating — different words should have different spellings. While this may be obvious when words sound completely different, as in cat and dog, it is more problematic when two words sound identical, as in flour and flower, where the spelling allows us to distinguish the two.

An awareness of distinctive spelling appears in the literature in several shapes but it is always seen as marginal, rather than as a fundamental principle of spelling, and it is never viewed as having a reciprocal relationship with morphemic spelling. Carney (1994: 131–4) gathers many cases of homophonous words with different spellings. Scholfield (2016) notes the long-standing tendency for attaching variant spellings to homophones with different meanings. Quite a bit of attention has been paid to the ‘three-letter rule’, whereby lexical words must have at least three letters, unlike function words (Jespersen [1909] 1928: §4.96; Carney 1994: 131–4; Venezky 1999: 86 Cook 2004: 56–7; Evertz 2014: 161–6; Berg 2016: 456), but distinctions between in and inn or by and bye are an epiphenomenon of this rule, rather than its primary purpose. The tendency for affixes to have distinct spellings has also become clearer in recent studies. Carney’s (1994: 417–441) section on homophonic affixes is particularly revealing and similar ideas are developed by Berg and Aronoff (2017) who recognise the importance of not indicating a morpheme as much as indicating it.

When one includes non-standard spelling and the spelling of names, then it becomes clear that attaching different spellings to words with different meanings is an important design principle of English spelling formation. The focus in this study, primarily in Chapter 7, will be on the orthographic techniques used to distinguish homophonous pairs, either when there exist two lexical words which have the same pronunciation (in and inn) or when a name derives its form from a lexical word (dog and Snoop Dogg). Hence such words can be disambiguated through spelling, despite a lack of auditory distinctiveness.
3e.2.2 The Identity Preservation principle (IdP)
If different words ought to have different spellings then it follows that related words might well have related spellings. However, this may not be practical from the viewpoint of phonology because related words may not sound identical, as in the pair *magic – magician*. Hence the idea of maintaining orthographic connections between words makes more sense on semantic and morphological grounds. We have seen, throughout Chapter 2, that the spelling of polymorphemic words tends to preserve the spelling of the constituent morphemes and this idea will be generalised henceforth into the broader *Identity Preservation principle* (IdP) which requires the spelling of new words to take their form from the spelling of the constituent parts — thus <magician> = <magic> + <ian>.

We shall see that IdP takes several forms and it applies to several different kinds of words, not just the compound and derived forms identified by Venezky (1970: 120). This includes inflections (Chapter 4), etymological spelling, abbreviations, clippings and blend words (Chapter 8). Potentially, IdP can occur in the spelling of any sort of word formation as long as there is a need to make a visual connection between the spelling of an existing word and the spelling of the new word or form.

3e.2.3 Phonographic Matching principle (PhM)
Modern descriptions of the English writing system have emphasized the importance of how meaning is represented in English spellings but, in doing so, they have failed to establish how we know if a word’s spelling adequately represents its sound. In other words, there is no metric for measuring the phonetic or phonological adequacy of any given spelling. The question arose, in Chapter 2, of how we know that *<profaneity>* is not a good spelling of the word normally spelt *<profanity>*. Using terms introduced here, we can ask: how do we know whether a spelling which adheres to IdP (i.e. *<profane + ity>* ) matches its sound or not? Chapter 4 examines the interlocking roles of IdP and PhM while Chapter 5 makes steps towards the goal of measuring whether a potential spelling is a good fit for the phonological form of a word.

At its core, PhM involves pairing sounds with spelling units, and the writing system’s enormous range of spelling-to-sound correspondences has been enumerated in previous studies (most fully by Carney, 1994 and Brooks, 2015). There was a time when words could be spelt according to how they sounded. A word such as *bird* is spelt <b-i-r-d> because those letters attempted to capture the word’s phonological form, at an earlier period in the language (in this case, the sixteenth century, OED 19327). In Modern English spelling, new words are
seldom formed by mapping from a phonological form to the spelling. Examples include words like *kangaroo* (OED 102472) which was taken from languages without a writing system, or a name such as *The Pogues* which is an Anglicisation of the Irish Gaelic word *póg.*

If the spellings of existing morphemes are invariant and phonological changes are not marked in spelling (see Chapter 6), then how can PhM be satisfied? The answer lies in the use of *phonographic amendments*, which are systematic solutions to problems arising within the formation of spellings. For example, the <pp> of *bop – bopping* is used to mark the vowel quality, while the absence of <e> in *dope – doping* has a similar function. The writing system contains several such amendment patterns and often these are the only tools available for indicating a word’s pronunciation. Previous studies have seen such ‘marking devices’ and ‘graphotactic rules’ as marginal to the writing system. In this study, phonographic amendments will be seen here as the glue which holds the entire system together. In spelling formation, they are used to disambiguate the correspondences, allowing for compromises to be made between the need to represent both sound and meaning. Some solutions, such as consonant-letter doubling are applied broadly across the system, while others are *ad hoc* spelling solutions, applicable only to one phoneme. However, all are used to maintain readable relations between spelling and phonology.

The focus on phonographic amendments begs another question which needs to be addressed. Can we draw a line between orthographic patterns that clearly attempt to represent sound and ones which have no phonological or morphological motivation. Quite clearly, some patterns are solely orthographic: changing <y> to <i> in *happy – happiness* has no phonological function. Therefore, we can reverse the question and ask whether there are spelling patterns which are conditioned solely by phonology and which are not predictable from the letter patterns. If not, then the phonographic matching principle is just an orthographic principle, whereby the spelling of complex words can be predicted solely from the spelling of the input forms. If some spellings are phonologically conditioned, however, then we need to have a way of accounting for both phonographic- and orthographic-matching patterns.

### 3e.2.4 Any-spelling principle

It was noted at the start of Chapter 2 that the ‘primary purpose of writing’, according to Sproat (1996), is to represent words and morphemes. The representation of a word’s sound is therefore a secondary concern, even though it is particularly helpful for spellings to include phonological information. It is
therefore possible, at least in principle, for written forms to represent meaning (primarily) without representing sound and, of course, several of the world’s writing systems are based around this principle (see Section 2a). In English it is rare for written forms to not represent a word’s sound and, indeed, it seems odd to call such a form a ‘spelling’. The handful of logographs in use fall into this category, including <@>, <$>, <&>, etc., as do the numerals <0, 1, 2, 3…>. Such symbols are totally arbitrary from a phonological point of view.

It will be useful to consider there to be a trivial Any Spelling principle of English spelling formation. No matter how good or bad a spelling may be, it is the spelling of that word. This is why intractable spellings such as <yacht>, <debt> or <thorough> can continue to function within the English writing system. Abbreviations may be even less phonologically transparent, but they too may become the spelling, as in <AKA>, short for also known as. No chapter will be devoted to this principle, as the focus here will be on the more normal relations between spelling and sound. It is therefore worth looking briefly at some of the issues raised by assuming that all words should have a spelling.

Firstly, it is possible for a word, or at least a name, to be composed of symbols with no phonological value. The American rock band !!! has no obvious pronunciation when analysed within the English writing system. Significantly, the band has, over time, developed a spoken form, which is often spelt <chkchkchkk>. An even more extreme example occurs in the symbol adopted for some time by the late musician, Prince. Not wanting to use his normal moniker, he adopted the symbol seen in Figure 3.5 below because he had to have some presence in writing, even though he did not want his name to be pronounceable (Till 2010: 63–65). Yet since that symbol could not be printed within normal text, using a keyboard, a need developed for a phonographic spelling of some kind to be used. One that emerged was love symbol, a description of the symbol itself, another was <Tafkap>, an acronym for The artist formerly known as Prince. In both cases, there was a need for some written form to exist, and for it to be attached to a pronounceable form. This is a reminder that there is a strong urge for written forms and spoken forms to match one another, even where the spelling provides no guidelines.

One complication to the Any-spelling principle is that it is not bi-directional. All words must have some spelling but not all spellings represent words. Sometimes, spelling is used to represent non-verbal or non-linguistic sounds, as occurs in the phenomenon of onomatopoeia, which itself can come in several forms.
We saw in Section 2b.3 that Carney omitted ‘vocal gestures’ from his corpus, although he did not omit examples such as *boo* which were ‘syntactically integrated’ (Carney 1994: 105–6), as in the forms *boos, booed, booing*. Attridge (1988, 2009) discusses several more extreme types of onomatopoeic spellings in the novel *Ulysses*, where the author, James Joyce (1984 [1922]), plays with our understanding of the writing system’s conventions in order to create a more vivid auditory image of the world described in the book. One particular example mixes human speech with the static sound of a poor recording: *Kraahraark! Hellohellohello amawfullyglad kraark awfullygladaseeagain*. Another represents the sound of animals (a cat mews three times: *mkgnao, mrkgnao* and *mrkrgnnao*). Such examples call into question the emerging distinction in the field between writing systems, which represent language (e.g. T. Joyce 2015, Neef 2015: 708), and writing in general, which includes notation systems such as mathematics, music, dance and computer languages. Spellings which represent non-linguistic sounds are still constrained by the conventions of the writing system in use, as Attridge (1988) convincingly argues, and yet they do not seem to fall within the definition of writing systems put forth by Neef and T. Joyce, and certainly not into the earlier definition of writing proposed by Daniels (1996: 3), whereby writing represents only spoken language.

### 3e.2.5 The Invariance principle (INVRNC)

It has long been known that English spellings are invariant, in the sense that it is not allowed, or at least accepted, for one to change the spelling of a word without good reason. Hence we cannot have the spelling *<invaryance>*. The actual process by which one variant spelling comes to dominate over others was not, until the work of Berg and Aronoff (2017) understood in any real detail, and it will be assumed here that their hypothesis is applicable to all English spellings, not just the small number of suffixed words that they examine. Of course, the way in which an invariant spelling is preserved is rather different as it relies on the
work of publishers, house styles, dictionary makers and social norms, the very ideas discussed by Scragg (1974) in his account of the standardisation process.

Assuming that English spellings are invariant, regardless of how that has come to pass, the outstanding question, which motivates Chapter 6 of the present study, is this:

- If English spellings are invariant but phonological systems change over time and across varieties of English, then how does that affect the correspondence relations between spelling and sound?

One advantage of invariant, or fixed, spellings is that they allow a stronger visual path to develop between writing and meaning, and there is less need for sound to function as an intermediary, as was required in earlier forms of the written language. This means that writing can represent language directly (T. Joyce 2015). IdP therefore presupposes Invariance. Related words can’t have related spellings if those spellings are variable.

There is one other very important assumption made in this study which needs to be clarified here. Invariant spelling is not bound by the medium through which the writing is represented. Morse Code, Braille and even Semaphore all use different media to represent the same abstract units.

***
3f. Conclusion and hypothesis

In this chapter we have looked at some of the research methods that will inform the analysis of English spelling to follow. The aim has been to develop an analytic method that can be applied, in Chapters 4 and 5, to the spelling of inflections and, in later chapters, to other categories of word (e.g. compounds, affixed derivation, inflection, neo-classical compounds, blends, acronyms, names, etc). Existing spellings will be reverse-engineered, to see what their internal components are, how they were formed, and what compromises have been made in order to create spellings which simultaneously represent the sound and the meaning of words, preferably uniquely and invariantly. Since written forms are neither sounds nor semantic concepts, the information can only be relayed through existing strings of letters, and we need some way of referring to the different ways in which these strings can be deployed. At this point, we can state the following hypothesis:

- Polymorphemic English spellings are formed by concatenating the spelling of the input morphemes. This ‘draft spelling’ is then tested against the known phonological form and amendments are made, if necessary.

This can be put into a simple algorithm:

1) **Identity Preservation (IdP)**
Concatenate the spellings of the two input morphemes to create a draft spelling:

- `<jump> + <ing> → ?<jumping>`
- `<bop> + <ing> → ?<boping>`
- `<dope> + <ing> → ?<dopeing>`

2) **Phonographic Matching (PhM)**
Compare the draft spelling against the known phonological form. If the spelling-to-sound correspondences do not provide a good phonographic match, then amend if possible.

- `?<jumping> → <jumping>` (no change)
- `?<boping> → <bopping> (<p> is added>)`
- `?<dopeing> → <doping> (<e> is dropped)`
3) Compare the amended spelling against the known phonological form.
4) Choose the better spelling.

There are many questions that remain unanswered. How do we know which amendment options are available? Many of these will be identified in Chapter 4. How do we check if the draft or amended spelling best fits the phonological form? A model for this will be proposed in Chapter 5. There are other important steps which cannot be added to the model yet, and there is no mechanism yet for creating DSTNCT spellings. Nevertheless, it should be clear that representing the sound of a polymorphemic word is secondary to representing the spelling of the input morphemes (IdP), and yet PhM amendments are often used, where possible, in order to help indicate the pronunciation of the complex word more clearly.
Chapter 4
The interface of IdP and PhM

The question that informs this chapter is one that has not been satisfactorily answered in the literature: to what extent are English spellings morphemic? A slight variation on this question, prompted by the observations of Berg et al. (2014) is this: how much more stable are the spellings of affixes compared to the spelling of base forms? In the terms just introduced, we can also ask, to what extent is IdP applied in the spelling of complex words, and what role, if any, does PhM play?

In many cases, IdP is applied and the phonological form of the word is indicated without any further ambiguity, as in <jump+s>, <jump+ing>. However, the question becomes interesting in cases where there is a tension between the application of IdP and the need to represent the known phonological form in a way that can be read without ambiguity. The spelling <jumped> appears to contain two syllables, spelt <jum> and <ped>, making a clear mismatch with the monosyllabic phonological form /jʌmp/. This suggests that representing phonology unambiguously is a lower priority than representing the morphemes clearly. And yet IdP is not applied to the spelling of all inflections. For example, the spellings <hoped> and <hoping> do not follow IdP: if they did, they would be *<hopeed> and *<hopeing>. The latter spellings do not provide a good match for the phonological form, and the compromise spellings <hoped> and <hoping> are used. This chapter explores how such tensions are resolved and it concludes with a slightly fuller model of English spelling formation. This model can explain better the complex interactions between the need to represent morphological information (using IdP) and the need to represent phonological information (using PhM). Outstanding problems will be addressed in Chapter 5, where the trickier concept of PhM is examined in isolation.

The focus in this chapter will be on the spelling of regular suffixed inflections because they provide exceptionally good data for analysing spelling patterns (c.f. Section 3b.1.4). They are highly productive and they do not affect the phonological form of the base. For example, in the sequence elect, elects, elected, electing, there are no changes to the pronunciation of the base, and this holds for all regular inflections. Compare the affixed derivatives, eˈlectric, elecˈtricity, elecˈtrician, where there is a stress-shift between the base form electric and its two derivatives. In this way, inflections allow us to analyse the spelling without complicating phonological factors. Furthermore, inflectional
patterns are generally independent of a word’s origin, and this gives us a chance to see how native spelling patterns interface with spelling patterns borrowed from other writing systems. (There are some Latin and Greek exceptions to this, such as stadium, stadia and criterion, criteria).

The inflections studied will be {ing} forms, {ed} forms, and three kinds of {s} forms. These shorthand terms are explained in Section 4a, where we look at the limited amount of inflection in the English language. Section 4b looks at the literature on consonant-letter doubling (CLD), a very important and highly-productive spelling pattern which some scholars have considered to be phonologically conditioned and others have thought it to be predictable from the spelling of the base. In the main body of analysis (Sections 4c and 4d), it will be shown that IdP and PhM interface with one another in several ways in the spelling of inflections, and several phonographic amendments will be identified. However, there are many problematic spellings for which there are no solutions. The analysis will attempt to examine the spelling patterns arising in all possible base forms, by checking words ending in all possible letter patterns and all phonological shapes. The chapter concludes by producing a fuller model of spelling formation based on the facts.

***
4a. Inflection in English

4a.1 Regular suffixed inflection

English inflection is rather limited, and suffixation has always been the main inflectional device in English (Lass 1999: 137). Some suffixed inflections are syllabic, others involve a consonantal appendix, while some suffixes show allomorphic variation between the two.

The phonological form of the gerund-participle, known here as the {ing} form, is always an unstressed syllable /ɪŋ/, as in jump, jumping; walk, walking and need, needing.

Regular past participles, the primary subgroup of {ed} forms, vary between syllabic and consonantal forms. After /t/ or /d/, the allomorph is syllabic /ɪd/, as in waited and needed; otherwise it is a dental stop which assimilates for voicing to the end of the base form. After unvoiced consonants, past participles are manifested as /t/, as heard in jumped, walked, laughed. In all other environments, they become /d/, as in raged, bombed, roared, prayed. English actually has another suffix which has the same spelling and pronunciation as the past participle (OED 59471) and the two have, understandably, been confused for a very long time. This suffix is formed directly from nouns, as in moneyed, cultured, jaundiced, dog-eared, etc. Such forms may be very low in frequency (e.g. umbrellaed, dialogued), and hence they should provide excellent data for testing the principles of spelling formation. This is based on the assumption that rare forms are more likely to follow the writing system’s normal patterns because they might not be easily decoded otherwise. There is no need to distinguish the two suffixes, since they are indistinguishable from a phonological and an orthographic point of view, and they will be referred to here as the {ed} form.

Similar allomorphemic variation occurs in the three kinds of {s} forms: syllabic /iz/ after sibilant consonants; /s/ after unvoiced consonants; and /z/ in all other environments. This applies equally to the third person singular {3PS s} as in kisses, jumps, needs; plural nouns {Plural s} as in buses, cups, tables; and the genitive form {Genitive s} the bus’s wheel, the cup’s handle and the table’s leg.

There are two suffixes for grading adjectives, as in bright, brighter, brightest. Neither of these will be studied here, partly for reasons of space, but also because it was assumed that they would not reveal any further details about the writing system. A fuller study may well find otherwise.
4a.2 Other inflections

There exist some irregular suffixes, such as the {en} suffix occurring in a few irregular past participles, including forgiven and forgotten. Furthermore, there are three other kinds of irregular, non-suffixed, shape changes (Lass 1999: 137), and these can also occur in combination:

1. Word-internal changes
   a. sing, sang, sung
   b. foot, feet

2. Suppletion
   a. go / went
   b. bad / worse

3. Zero-inflection
   a. hit, beat (present, past, past participle verb forms)
   b. fish, sheep (singular, plural nouns).

4. Combinations
   a. wrote, writ(t).en
   b. knife, knive.s
   c. leave, lef.t
   d. good, bett.er
   e. go, wen.t

The spellings of these various inflections remind us that where phonological changes are not predictable by rule, the spelling must indicate them explicitly, as Chomsky and Halle note (1968: 49). In other words, irregular inflections adhere to PhM over IdP. Moreover, no effort is made to distinguish the inflection left from the adjective left, suggesting that DSTNCT is not prioritised. It is worth noting that many verbs ending in /t/ undergo an irregular zero inflection, as in bet, set, put, hit, etc., and these are not spelt *<betted>, *<setted>, etc., whereas regular forms follow the usual patterns, as in whetted, vetted, fretted, etc.

Plurals taken from Latin and Greek (e.g. appendix – appendices, criterion – criteria, etc.) frequently follow patterns that are borrowed directly from Latin and Greek, and these might be thought of as grammatical subsystems comparable to the spelling subsystems. As noted in the introduction, borrowed words are generally absorbed into the English inflectional patterns, so that one can drink two espressos, rather than *espressi, as in Italian. A full list of English inflections, both regular and irregular, can be found in Palmer, Huddleston & Pullum (2002: 1569–1570; 1580–1590):
4b. Consonant-letter doubling

Along with the dropping of <e>, consonant-letter doubling (henceforth CLD) is the most important pattern in the spelling of English inflections. Its primary purpose, across the entire writing system, is to mark the length and quality of the preceding vowel. In {ed} forms, for example, vowel length is indicated by the presence or absence of CLD, as in mat(t)ed – mated, pen(n)ed – convened, pin(n)ed – pined, hop(p)ed – hoped and rebut(t)ed – reputed. Marking consonant quality is a secondary concern as single and double consonants usually correspond to the same consonant. As a result, double consonants are frequently absent where we might expect them: compare gammon and lemon. Problems can arise for <g> and <c> which often have different values when occurring before <e, i, y>. Thus, words which don’t have <gg> and <ck> (the de facto double for <c>) often run the risk of being decoded wrong.

There remains a lack of consensus in the literature about how CLD is conditioned. It was long assumed to be conditioned phonologically, as one of its primary functions is to mark the length/quality of the preceding vowel letter, but recent graphematic approaches contend that the patterns can be explained solely in relation to the letter sequences, (e.g. Berg et al. 2014; Berg 2016), henceforth the ‘letter-sequencing hypothesis’. Before looking at these analyses, an important distinction needs to be made between two ways in which double letters can occur.

4b.1 CLD and historical doubles

It will be important for present purposes to distinguish between two kinds of double consonant: CLD is a productive pattern in the spelling of polymorphemic words; while historical doubles, as they shall be referred to here, occur in monomorphemic words. Some historical doubles have emerged within English (e.g. apple and hammer) or were borrowed from other writing systems (e.g. §Latinate apply or §Italian cappuccino). Both kinds of double letters must be distinguished in turn from the phonological phenomenon of long or ‘geminate’ consonants. In modern English, geminate consonants only occur at word- or morpheme boundaries, either in polymorphemic words such as unnecessary or in compounds such as bit torrent. Hence it is useful to distinguish between the short /t/ in bitter and the long /t:/ in bit torrent. Many historical doubles (in spelling) derive ultimately from geminate consonants (in speech), either within earlier forms of English (see Britton 2011) or in the foreign languages from which the words were borrowed. This is why we have §Greek cirrhosis, §Italian arpeggio.
and §Latinate apply, effect, support, etc. (see Upward and Davidson, 2011 for a full analysis of the various inputs into English spelling).

4b.2 Carney’s account of historical doubles
Carney (1994: 112–129) provides a detailed account of double consonant letters, but he does not explicitly distinguish between CLD and historical doubles. Since inflected forms are omitted from his corpus, he does not focus much on the intricacies of conditioned CLD; instead, his complex list of rules reflects the extremely varied distribution of historical doubles. These in turn reflect the range of phonological, morphological and etymological inputs to English spelling, as well as its irregular letter-patterning. For example, certain consonant letters cannot be doubled (e.g. <x> and <q>), and some are only doubled in new or slang forms (compare the <vv> of skivvy and navvy with the <v> of even and seven). Not all phonemes trigger doubling — some are represented by a kind of quasi-doubling in monomorphemic bases: words ending in /k/ often have <ck> as in lack, lick, lock, luck; words ending in /tʃ/ have <tch> as in batch, latch, etc; and words ending in /dʒ/ have <dge>, as in badge, edge, ridge (ibid: 123). Doubling can’t occur after two vowel letters, even when the vowel is short (e.g. trouble v. rubble) or after repeated consonant letters (e.g. <fully> rather than *<fullly>). Doubling seldom occurs in the ante-penultimate syllable of a word (e.g. gratitude, criminal). Some letters occur doubly at the end of a morpheme (staff, smell, boss, buzz), but others don’t, except in order to create a three-letter word (e.g. ebb, odd, err, etc.). Carney does not go so far as to say that doubling is confined to §Basic words and absent from §Latinate words, but he does say that it is a ‘tempting hypothesis’ (ibid: 126). His analysis is of great importance to the study of English spellings but it tells us a lot more about the spelling of base forms than of polymorphemic forms, and it is thus not of great help here.

4b.3 Accounts of CLD
Explaining CLD is less complicated than accounting for historical doubles, and yet the descriptions do vary, primarily because the two were never treated separately in the literature until Berg’s (2016) analysis.

A typical explanation is that CLD is phonologically motivated, and it occurs in /VC#/ words, as in beg – begging, rob – robbing, etc. A long vowel, conversely, is marked by the absence of CLD (e.g. Carney 1994: 19). More recent approaches have changed the perspective. A letter-sequencing account is provided by Rollings (2004: 32–34), who argues that CLD can be explained with reference to an ‘orthographic syllable’, and he describes this with the examples of plan –
planning and plane – planing. Readers can divide <plan.ning> and <pla.ning> into two syllabic parts, separated here by a dot. The presence of <n> in <plan> marks a short vowel, while its absence in <pla> indicates a long vowel.

Berg (2016: 454) notes that the pattern only works in one direction. The occurrence of CLD regularly predicts a short phonological vowel but short vowels do not always indicate double consonants: compare limit and lemon to bulimic and demon. Carney’s solution to this problem is to formulate the consonant-doubling rule in reverse; instead of saying the double consonants occur after short vowels, he states that double consonants tend not to occur after long vowels (Carney 1994: 114–115). However, the distinction between CLD and historical doubles (or polymorphemic and monomorphemic words) also helps to resolve this problem: <limit> and <lemon> are §Latinate spellings (OED 108477 / 107204) and that explains their ostensible irregularity. Berg (2016: 463) notes that CLD ‘is highly correlated with the graphemic shape of the word ending’. Words ending in <ic, it, id, ish, ace, ous, al, age, ule> frequently have a single consonant letter where we might expect two. Chomsky and Halle (1968: 181n), in their discussion of phonology, observe this sub-pattern in words ending in <id> (acid, rapid), <it> (credit, visit) and <ish> (radish, abolish). Regardless, such a problem would not occur in polymorphemic words formed within English, hence we have slimming and stemming, contrasted against limit and lemon.

Berg’s solution (following Cummings 1988: 161–176) is to formulate a letter-doubling rule in terms of letter-sequencing patterns. He states that doubling only occurs for the following consonant letters.

<b, d, f, g, k, l, m, n, p, r, s, t, v, z>

It should be noted that <cc> is also possible, as in specced, but such examples presumably don’t occur in Berg’s corpus (Baayen et al. 1995). Furthermore, doubling only occurs when a single consonant letter follows a single vowel letter. Hence, we can list the following examples of CLD occurring in {ing} with <VC#> base.

bobbing, speccing, bedding, cheffing, bugging, trekking, gelling,
bumming, banning, bopping, barring, bussing, betting, revving, fezzing

Berg (2016: 457) makes the important point that CLD does not occur if the orthographic conditions are not met, even if the phonological conditions are. Hence we have come – coming and look – looking, rather than *<comming> and *<lookking>. Likewise, CLD is not an option for digraphs such as <ch>, <ph>,

96
Finally, there is no need for CLD in words such as staff, smell, boss or buzz, where the historical double performs the function of CLD in the complex forms: staffing, smelling, bossing, buzzing.

It is thus clear that CLD is not conditioned solely by phonological conditions. Berg et al. (2014: 299) write:

This leads to a simple analysis: written English morphology operates on the graphemic form of the base in the first place and to a lesser degree on the phonological form of the complex word. Berg (2016: 457–461) goes further again and argues that CLD can be explained entirely by morphological and orthographic conditions, except in a few cases. For him, CLD is best viewed as a special case of morphemic (or ‘morphological’) spelling, with the extra consonant letter simply being conditioned by the orthographic environment. However, there are phonologically-motivated exceptions, as he acknowledges himself. For example, there is usually no CLD after unstressed syllables, so we have ‘edit – ‘editing, different from e’mit – e’mitting. There is also no phonological need for <xx> because <x> usually maps onto two phonemes. Hence the {ing} form of fix is just fixing, rather than *<fixxing>.

Throughout Section 4c, I shall argue that the formation of {ing} and {ed} forms in English spelling is reliant on all three conditioning factors, namely morphological, phonological and orthographic. It is neither possible nor useful to entirely remove phonological information from a theory of how polymorphemic English spellings are formed and, furthermore, it is seldom possible to separate phonological and orthographic factors, since they are so deeply intertwined.

***
4c. The spelling of \{ing\} and \{ed\} forms

The spelling patterns shown in \{ing\} and \{ed\} forms are almost identical and they are treated together in the following in-depth analysis. While the vast majority of these spellings can be explained solely by referring to the spelling of the base form, as Berg (2016) shows, we shall see that the details can get a lot more complicated. There are two situations where problems most commonly arise: when the mappings from spelling to sound (in the base form) are non-canonical; and when the spellings themselves are variable (either in the base form or the complex forms). In order to demonstrate this variability, it will be necessary to draw on two things: examples taken from some very rare word forms and evidence taken from Google Ngram Viewer. The more illustrative evidence is presented in figures, otherwise a link is added in an endnote. It may be a surprise that there remains so much variation in the spelling of English inflections although this makes more sense if we accept the idea that the writing system is self-reorganizing (Cummings, 1988: 6; Berg and Aronoff 2017). If spellings have to represent several types of linguistic information at once, then it follows that the system will change as the language changes. Diachronic evidence of such variation often shows the respective change in the proportion of each variant spelling.

In the interests of exhaustiveness, I have tried to examine every pair of orthographic and phonological word-endings. Hence we need to look at words ending in all consonant sounds and all vowel sounds, as well as words ending in all consonant letters, vowel letters and \(<e>\). Unsurprisingly, this section is rather complicated, and its structure is laid out below, with some clarifying examples.

1. Base forms ending in a consonant sound I: /C#/ ≡ <C#>
   1./VC#/ ≡ <VC#> base forms with final stress
      \{sit, emit, zigzag\}
   2.Other /C#/ ≡ <C#> base forms with final stress
      \{boss; jump, gird, sign, bomb, zinc, sync; itch; aim, weed, dread\}
   3./VC#/ ≡ <VC#> base forms with penultimate stress
      \{edit, author, focus, cancel\}
   4.Other phonological correspondences for <VC#> spellings
      \{ok, parquet, motif, mic\}
   5.Summary and Table of Examples
2. Base forms ending in a consonant sound II: /C#/ = <Ce#>
   1. <VCe#> base forms
      : rage, engage, mete, delete
   2. <VV(C)Ce#> base forms
      : peace, cease, bruise
   3. <VCCe#> base forms
      : ache, table, singe, tense

4. Summary

3. Base forms ending in a vowel sound
   1. /V#/ = <V#> base forms
      : subpoena, acne, taxi, go, snafu, spy
   2. /Ve#/ = <Ve#> base forms
      : spae, free, die, canoe, rue, dye
   3. Table of examples

4. Summary

4c.1 Base forms ending in a consonant sound I:
   /C#/ = <C#>

4c.1.1 <VC#> = /VC#/ base forms with final stress
We saw above that CLD is highly productive when {ing} is added to base forms with final stress and a <VC#> spelling. Identical patterns apply to {ed} forms. Hence we have:

bobbed, speced, bedded, cheffed, buged, trekked, gelled, bummed, banned, bopped, barred, bussed, vetted, revved, fezzed.

There is no indication from the spelling that {ed} is pronounced as three different allomorphs: syllabically in bedded and vetted; as /t/ after an unvoiced consonant, as in cheffed, trekked and bopped; and as /d/ elsewhere. Clearly, the spellings are morphophonological, rather than conditioned by the phonological form of the final consonant. Hence we do not have *<cheft>, *<trekt> and *<bopt>. Furthermore, CLD is productive among letters which seldom occur as historical doubles, hence speced, trekked, revved, fezzed. There is no rule against such formations. It should be noted that spellings such as <cheffed>, <trekked> and <bopped> are not ideal spellings from the viewpoint of PhM. However, they are the best available spellings within the confines of the English writing system. The only other options are <chefed>, <treked> and <boped>. Only in Chapter 5 will we see exactly why <cheffed> is a better, or less
bad, phonographic match than <chefed>, etc. The reader is asked to accept this for the time being.

One set of counterexamples involve <Vx#> words such as taxed, taxing, flexed, flexing, fixed, fixing, etc. where no CLD occurs. It is undeniable that this pattern is phonologically motivated. Since <x> only occurs after short vowels, there is no need for <xx> to mark the quality or length of the preceding vowel letter.

Polysyllabic words with final stress show similar spelling patterns. Iambic base forms, which have one unstressed syllable followed by a stressed one, are identical. Hence we have reˈfer, reˈferred, reˈferring, comˈpel, comˈpelled, comˈpelling and eˈmit, eˈmitted, eˈmitting. The word input varies in its past participle, between input and the less common inputted (OED 96483), and it seems safe to assume this kind of spelling variation reflects the presence or absence of /ɪd/. In other words, the spelling chosen reflects the desired phonological form. A slightly different issue arises for compound words where the final syllable has secondary stress. Berg (2016: 460) argues that compounds undergo CLD if the final syllable of the base is a monosyllabic root, hence the CLD in sandbag, sandbagged, sandbagging or kidnap, kidnapped, kidnapping. This tendency may extend to compounds whose components are opaque; hence we have humbugged, humbugging and zigzagged, zigzagging.

4c.1.2 Other /C#/ == <C#> base forms with final stress
CLD does not occur for spellings other than <VC#>, even in the few cases where it might be useful to indicate the phonological form more clearly. Nevertheless, there is a very small number of words where there is a strong argument for including another kind of PhM amendment in the spelling of the complex form.

CLD is obviously not needed where the base form is <VC₁C₁#>, as in <boss-ing>, <buzz-ing>, <smell-ing> and <staff-ing>, although it is worth noting how both <VC₁#> and <VC₁C₁#> end up with streamlined {ed} and {ing} forms. For example, fill, filled, filling and fulfil, fulfilled, fulfilling, etc. This suggests that CLD is a regularising force on the writing system. Inconsistencies among base forms (e.g. fill v. fulfil) are weeded out in the polymorphemic forms.

When the <VC₁C₂> spelling corresponds to a short vowel and consonant cluster, there are usually no complications, as in jump, jumped, jumping and gird, girded, girding. Carney (1994: 113–4) notes that there is a more general rule which states that English spellings avoid <C₁C₂C₂> sequences, except at morpheme boundaries, as in granddaughter, withhold, etc. While this would
explain the absence of *<jumpping> and *<girdding>, there is also no PhM-need for such spellings.

Other <VC₁C₂#> ≡ /VC#/ words result from the use of a digraph, as in wash, lack, sing, froth. All of these spellings obey IdP on addition of {ing} and {ed}, even where a potentially new PhM ambiguity can arise. For example, frothed and frothing, with a short vowel and an unvoiced consonant, contrast against bathed and bathing with a long vowel and a voiced consonant. The problem in these cases is not so much with the spelling <frothed, frothing>, but rather with the dropping of <e> in <bathed, bathing> (c.f. Section 4c.2.3). It is notable that the writing system does not have solutions to prevent mis-readings in these cases.

<VCC#> words such as bomb and sign retain their irregularities in their inflected forms, so the inert letters remain unpronounced in bombed, bombing and signed, signing, unlike the derived forms bombard, signal, etc.

Words spelt <VC₁C₂C₃#>, <vvC#>, <vvCC#> or <vvCCC#> all show the default IdP pattern. The first case includes itch, itched, itching and birth, birthed, birthing. Some <vvC> spellings are unproblematic, as in aim, aimed, aiming and weed, weeded, weeding, because the <vv> digraph represents a long vowel. But book and dread cannot undergo CLD, despite having a short phonological vowel, so we end up with booked, booking and dreaded, dreading. The latter of these might suggest a pronunciation like pleading and leading, rather than heading and spreading. <vvCC#> and <vvCCC#> base forms show the same pattern, regardless of vowel length, as in reach, reached, reaching; touch, touched, touching, and search, searched, searching.

The one set of counterexamples involve <quVC#> words such as quit and quiz, which render quitted(ted), quitting, quizzed and quizzing. (The OED lists <quitted> as a variant spelling, and it has made a comeback this century, OED 156786). Again, this is phonologically motivated. One could of course introduce an exceptional explanation according to the letter-sequencing patterns, considering <u> to be a consonant letter here, and thus the CLD patterns of quit and quiz would be considered identical to win and swim. Doing this, however, requires us to explain why there is no CLD in suit, suited, suitng or ruin, ruined, ruining. Clearly, the only consistent explanation is that the presence or absence of CLD is phonologically motivated. In order to fulfil the letter-sequencing hypothesis at the same time, it would be necessary to classify <u> as consonant letter when it follows <q>, making <quit> comparable to <wit>.

There are some other <VC₁C₂#> base forms where a PhM is possible but not necessarily availed of. Potential problems arise in <VCc#> ≡ /VCk#/ words,
where the \{ing\} and \{ed\} forms introduce the possibility of \(\text{ci}\) and \(\text{ce}\) indicating /k/. Usually, the solution to this lies in the spelling of the base verb which is almost always \(<k>\). Hence we have rank, ranked, ranking; bulk, bulked, bulking; twerk, twerked, twerking; and even walk, walked, walking, where the \(<l>\) is a dummy letter. The need to avoid \(<cing#\) and \(<ced#\) spellings may also explain why monosyllabic /Vk#/ words are spelt \(<Vck#\), as in lick, licked, licking, etc. Nevertheless, there do exist a few problematic \(<VCCc#\) words, including arc, talc and zinc, all of which are nouns. The OED (10246 / 197193 / 232864) suggests that their \{ing\} and \{ed\} forms generally adhere to IdP at the expense of PhM, resulting in \(<arced, arcing>, <talced, talcing> and <zinced, zincing>\). The OED (197193) does list \(<talcoked>\) as a variant spelling but there are no instances of this in Ngrams, nor of \(^*<\text{talcoking}>\), \(^*<\text{arcked}>\) and \(^*<\text{arcking}>\). The spellings \(<arke\text{d, arking}>\) and \(<talked, talking>\) are blocked by the existence of the unrelated words ark, arked, arking and talk, talked, talking. However, zinc shows the most variation, resulting in the less common variant spellings \(<zincked, zincking>\) and \(<zinked, zinki\text{ng}>\), all four of which obey PhM but not IdP. The \{ed\} forms can be seen in Figure 4.1. The IdP spelling \(<zinced>\) is the most common, but \(<zinked>\) and \(<zincked>\) also both occur, suggesting that there is some demand for the consonant /k/ to be marked unambiguously. In this way, we see how different forces are manifested in the variant spellings of rare and unestablished word forms. Some writers feel the need to preserve IdP, others feel the need to represent the sound more clearly.

**Figure 4.1**
Between 1958 and 2008, \(<zinced>\) (in blue) has been the most common spelling but \(<zinked\> (in red), and \(<zincked\> (in green), also occur.
A similar case comes from the clipped (or truncated) form of *synchronise*, pronounced /sɪŋk/, but spelt either <sync> or <synch>. Once again, the spellings of the inflections show a slightly different pattern to the spelling of the base clipping. Figure 4.2 shows that <sync> has become a far more common spelling of the base than <synch>, at least in the years between 1970 and 2008; and yet there is no dominant spelling in either of the pairs <syncing, synching> and <synced, synched>. Notably, none of these six spellings breaks IdP, in the sense that the spellings of the clipped forms are taken directly from the spellings of the full form (this will be called *IdP-reduction*, as opposed to *IdP-extension*; c.f. Section 5d.4). The presence of so many variants which obey IdP allows us to compare the options for obeying PhM. It turns out that there are good reasons for the usage patterns: <sync> is unambiguous in its representation of /sɪŋk/ whereas <synch> is ambiguous between /sɪŋk/ and /sɪntʃ/ (compare pinch), so it makes sense for <sync> to predominate in the spelling of the base. The spellings of the complex forms are all ambiguous, however, and the question comes down to which spellings are worse: <synced, synching> could be read like minced, mincing, with \(<c> \equiv /s/\); while <synched, synching> might again be read like pinched, pinching.
Figure 4.2
In the base form (top), <sync> (in red) predominates over <synch> (in blue), but in polymorphemic forms, <synced, syncing> are about equal to <synched, synching>.

104
The fact that <h> is common in the complex forms but not the base clipping suggests that <synched, synching> are better spellings than <synced, syncing>. Even though the <ched, ching> spellings are ambiguous, the <ced, cing> spellings are much more misleading. This is because <y> marks the value of the §Greek correspondence <ch> ≡ /k/ (Carney 1994: 302; c.f. Section 2c.3), whereas <ce> and <ci> correspond to /s/ in over 90% of incidences (ibid: 304).26

Amidst all of these details, the point here is that where there are two valid spellings of the base form, one of these may be preferred in the complex form, and it would make sense for the better (or less bad) phonographic match to be chosen (i.e. <synched, synching>). Despite this, there may also be a strong tendency for the more common base form to survive in the complex form, simply because it follows the usual letter-sequencing patterns (i.e. <synced, syncing>).

**4c.1.3 <VC#> ≡ /VC/ base forms with penultimate stress**

Base forms with penultimate stress have much more variable spellings of their {ing} and {ed} forms than base forms with ultimate stress.

In general, CLD is absent in trochaic base forms, giving us ‘edit, edited, ‘editing; ‘rabbit, ‘rabbited, ‘rabbiting; ‘author, ‘authoring, ‘authored, etc. Berg (2016: 460–1) explains this on semantic grounds, observing that the functional need for CLD in polysyllabic words is less than in monosyllabic words because there are so few pairs of spellings distinguished solely by the presence or absence of <e>. In his corpus, he finds 121 monosyllabic pairs such as bid, bide and mop, mope, etc., but only 12 polysyllabic cases, most of which are semantically related, as in human, humane and unit, unite. The need for homography avoidance is thus much higher among monosyllabic base forms, hence the CLD of bidding and mopping, distinguishing them from biding and moping.

Nevertheless, trochees ending in <l>, <s>, <c> and <g> are more variable, and may be subject to CLD. For example, British spelling generally has ‘cancelled, ‘cancelling> and ‘equalled, ‘equalling> (see Figure 4.3). American spelling is usually associated with a single <l> but Figure 4.4 shows that <cancelled> has always coexisted with <canceled>. For some reason, <(un)paralleled> never undergoes CLD, even in British spelling, perhaps as a way of avoiding letter-sequences that are too similar (i.e. *(un)paralleled).
British spelling has always shown a clear preference for `<canceled>` (in blue) over `<canceled>` (in red).

American spelling has seen the rise of `<canceled>` between 1800 and 2000, but `<canceled>` (in blue) has always been steadily present.

Trochaic bases ending in `<s>` also show variable incidences of CLD. The spelling `<focused, focusing>` predominates over `<focussed, focussing>`, in both British and American spelling. Figure 4.5 shows how, in the period between 1850 and 1950, the spellings without CLD became strongly dominant in American spelling but the CLD spellings were more prominent in British spelling until about 1930. Figure 4.6 shows how the CLD spellings have since been ousted by the simpler spellings in British spelling.
Figure 4.5

In American spelling (top) <focused> (in blue) and <focusing> (in red) have dominated since the twentieth century. In British spelling (below), <focussed> (in green) and <focussing> (in orange) were dominant until about 1930.

Figure 4.6

Since 1950, CLD has seldom been used in British spelling.
It would appear that the tension lies between the need, on the one hand, to follow the normal patterns for trochaic bases, thus marking the stress indirectly and, on the other hand, the need to indicate the value of /s/ clearly (compare the /z/ of *a’bused, a’busing*). It seems that the need to mark stress outranks the need to mark consonant quality, at least in this case.

Polysyllabic *<c#>* words require a special case of CLD, with *<k>* being added irrespective of the stress patterns. Examples include *panicked, panicking, bivouacked, bivouacking* and *trafficked, trafficking*. The spellings *<icing>* or *
*<acing>* would suggest the correspondence *<c> ≡ /s/, as in noticing and menacing*. Others have treated this as a very special case (Berg 2016: 45; Venezky 1999: 83; Carney 1994: 223). Another way of looking at the problem is to ask whether the marking of consonant quality is a higher priority than the marking of vowel quality. In *<editing>* or even *<rabbiting>* there remains the possibility of reading *<i> ≡ /ai/, as in biting or unre’quiting*, but there is no danger of the consonant letter being misread. It may be the case that the English writing system is concerned with marking consonant quality more explicitly than vowel quality. We must therefore ask whether there are other consonant letters that require marking in complex forms.

The obvious candidate is *<g>*. Its patterns often resemble those of *<c>*. There are very few polysyllabic monomorphemic base forms ending in *<Vg#> ≡ /Vɡ#/*, but one example is the American spelling *<catalog>* which came to dominate over *<catalogue>* in the latter half of the twentieth century, as can be seen in Figure 4.7. In the *{ed}* form, the *<ue>* spelling has remained slightly more common, and this suggests some important tensions between IdP and PhM. If *<cataloged>* satisfied PhM adequately then we would expect its relative incidence to reflect that of *<catalog>*. Instead, *<catalogued>* is more common, and this suggests that the *<u>* functions as a consonant marker, just as it does in *guest, guilt* and *rogue*. Of course, the *<u>* of *<catalogued>* is not a system-wide phonographic solution akin to CLD. It is only possible because there happens to be a minority spelling of the base and this variation is put to functional use in the spelling of the polymorphemic form. In this, it is much like the *<h>* of *sync(h), synched, synching*. Nevertheless, this marking function of *<u>* is not applied to the spelling of the *{ing}* form, as can be seen in Figure 4.7, where *<cataloguing>* remains less common than *<cataloging>*.
In American spelling, "catalog" (in blue, top) has come to dominate "catalogue" (in red). The "u" is frequently retained in the {ed} form (bottom) but not the {ing} form (middle).
Interim Summary of Section 4c.1.3

In general, trochaic bases obey IdP. Berg (2016) provides a semantic argument for this, noting that CLD has a lower functional load in polymorphemic forms. However, this explanation does not explain why CLD occurs after iambic base forms but not trochaic. The reasons must therefore be phonological and, furthermore, the spelling variation may reflect a tension between different phonological needs. It would appear that marking stress (explicitly or implicitly) is the most important phonographic priority. Just as CLD marks a preceding syllable as stressed, the absence of CLD marks the preceding syllable as unstressed, suggesting that the stress is earlier in the word. This difference is best exemplified by <emitted> and <edited>, respectively.

Nevertheless, the lingering presence of variation among trochaic base forms ending in <l>, <s>, <c> and <g> suggests that any (putative) hierarchy of spelling constraints has not fully stabilised among the words under discussion. It would appear that there is a tension between the need to represent the stress pattern and the need to represent consonant quality unambiguously. Adding extra letters to the spelling of the complex form (be it CLD or otherwise) clarifies the consonant value (hence <trafficked>, American <catalogued>) but this may lead readers to assume that the stress falls on the syllable before the suffix. Not adding the letters may lead to confusing consonantal correspondences. Marking vowels seems to be much less important and this is why spellings such as editing and authoring persist despite the possibility of a misreading. The survival of <ll> forms is a residualism (mostly in British spelling) which defies any other explanation within the present framework.

4c.1.4 Other phonological correspondences for <VC#> spellings

Berg (2016) does not look at the rare cases where <VC#> spellings correspond to phonological sequences other than /VC#/\). Such cases allow us to test whether CLD can always be predicted from the spelling of the base, or whether the phonological form can play a role. Examples are plentiful among §French words. There are many bases with <VC#> ≡ /V:/, including début and croquet (which have initial stress) and parquet (which has initial stress in British English but final stress in American English, OED 138121). In other cases, <VC#> ≡ /V:C/, as in motif, aperitif, again with final stress. Similar correspondences occur in the word OK, and clippings such as mic, short for microphone.

The presence of initial stress in début can account for the lack of CLD in débuted, débuting (OED 47596), regardless of the dummy <t>. Having <tt> would unnecessarily suggest a /t/ pronunciation, as in rebutted, rebutting. In
British English, the spellings <parqueted, parqueting> can also be explained this way, as the base form has initial stress, /ˈpɑːkeɪ/. In American English, this word is pronounced /parˈkeɪ/, so the spellings <parqueting> might be considered counter-examples to Berg’s hypothesis that CLD always occurs after stressed <VC> spellings, regardless of the phonological form. Clearly the non-CLD spelling is motivated by the phonological form, with the <t> as a dummy letter. However, there used to exist a pronunciation with /t/, as listed by the New English Dictionary of 1904 (i.e. the first edition of the OED), and this would motivate the minority spellings <parquetted, parquetting>. This earlier pronunciation would explain the greater proportion of <tt> spellings in the nineteenth and early twentieth centuries, as can be seen in Figure 4.8. Whichever pronunciation occurs, it seems that phonology plays a role in the spelling. If parquet has initial stress, then there should be no CLD, in keeping with other trochaic bases (e.g. edited). If it has final stress and the /t/ is pronounced, then we would expect CLD (like in omitted). However, if it has final stress but no /t/, then there should be no CLD, as <tt> would indicate /t/.

Figure 4.8
The spelling <parquetted> (in blue) has fizzled out over the last 60 years compared to <parqueted> (in red)

In the words, motif and aperitif. CLD would obscure the <VC> \( \equiv /V:C/ \) correspondence in the inflected forms. Ngram Viewer records a few instances of <motified> but none of *<motiffed> and no {ing} forms. Despite the scarcity of occurrence, we can conclude that Berg’s letter-sequencing hypothesis does not predict the form that actually occurs.

Other evidence comes from clippings and abbreviations, and, once again, the evidence is lacking for CLD motivated solely by the spelling of the base. The clipped form of microphone is predominantly spelt <mic> (nowadays) but it can
also be <mike> (OED 118031). The latter variant is by far the more common among {ed} and {ing} forms, rendering <miked, miking>, rather than <miced, micing>, as can be seen in Figure 4.9. As with <cataloged, catalogued> and <synced, synched> above, the existence of two variant bases allows for the ‘better’ spelling to be chosen in the polymorphemic form, meaning the one which best obeys PhM.

Figure 4.9
In the top picture, the spellings <mike> (blue) and <mic> (red) are both possible clippings of microphone but the {ed} form is only spelt <miked>.

A similar kind of variation occurs for the word OK. It can also be spelt <okay>, a spelling which obeys PhM more clearly, and the two have risen steadily in their usage in recent decades. However, the spelling <okayed> obeys PhM much more transparently than <oked>, so it predominates in the spelling of the {ed} form. The relative use of <okaying> and <oking> are skewed by the existence of the word oke. The spellings *<okked, okking> are not possibilities,
despite the letter-sequencing hypothesis. Again, the minority spelling of the base is preferred in the spelling of the polymorphemic forms, and one can only conclude this is due to its greater adherence to PhM.

**Figure 4.10**
In the top picture, the spellings <okay> (blue) and <ok> (red) occur with similar frequencies, but <okayed> dominates <oked>.
Table 4.1
The {ing} and {ed} forms of base forms ending in \(<C#>\)

<table>
<thead>
<tr>
<th>/VC#/ (\equiv) &lt;VC#&gt;</th>
<th>/VC/ (\equiv) &lt;VC(_1C(_2)&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>bob &gt; bobbed &gt; bobbing</td>
<td>staff &gt; staffed &gt; staffing</td>
</tr>
<tr>
<td>spec &gt; specced &gt; speccing</td>
<td>smell &gt; smelled &gt; smelling</td>
</tr>
<tr>
<td>bed &gt; bedded &gt; bedding</td>
<td>boss &gt; bossed &gt; bossing</td>
</tr>
<tr>
<td>chef &gt; cheffed &gt; cheffing</td>
<td>buzz &gt; buzzed &gt; buzzing</td>
</tr>
<tr>
<td>bug &gt; bugged &gt; bugging</td>
<td>/VC/ (\equiv) &lt;VC(_1C(_2)&gt;</td>
</tr>
<tr>
<td>trek &gt; trekked &gt; trekking</td>
<td>lack &gt; lacked &gt; lacking</td>
</tr>
<tr>
<td>gel &gt; gelled &gt; gelling</td>
<td>wash &gt; washed &gt; washing</td>
</tr>
<tr>
<td>gum &gt; gummy &gt; gummimg</td>
<td>sing &gt; - &gt; singing</td>
</tr>
<tr>
<td>ban &gt; banned &gt; banning</td>
<td>froth &gt; frothed &gt; frothing</td>
</tr>
<tr>
<td>hop &gt; bopped &gt; hopping</td>
<td>bomb &gt; bombed &gt; bombing</td>
</tr>
<tr>
<td>bar &gt; barred &gt; barring</td>
<td>sign &gt; signed &gt; signing</td>
</tr>
<tr>
<td>bus &gt; bussed &gt; bussing</td>
<td>/VCC/ (\equiv) &lt;VC(_1C(_2)&gt;</td>
</tr>
<tr>
<td>vet &gt; vetted &gt; vetoing</td>
<td>gird &gt; girded &gt; girding</td>
</tr>
<tr>
<td>rev &gt; revved &gt; revving</td>
<td>rank &gt; ranked &gt; ranking</td>
</tr>
<tr>
<td>fix &gt; fixed &gt; fixing</td>
<td>zinc &gt; zinc(_k)ed &gt; zinc(_k)ing</td>
</tr>
<tr>
<td>fez &gt; fezzed &gt; fezzing</td>
<td>sync(_h) &gt; sync(_h)ed &gt; sync(_h)ing</td>
</tr>
<tr>
<td>refer &gt; referred &gt; referring</td>
<td>weed &gt; weeded &gt; weeding</td>
</tr>
<tr>
<td>emit &gt; emitted &gt; emitting</td>
<td>suit &gt; suited &gt;suiting</td>
</tr>
<tr>
<td>compel &gt; compelled &gt; compelling</td>
<td>/VC/ (\equiv) &lt;VVC&gt;</td>
</tr>
<tr>
<td>sandbag &gt; sandbagged &gt; sandbagging</td>
<td>book &gt; booked &gt; booking</td>
</tr>
<tr>
<td>kidnap &gt; kidnapped &gt; kidnapping</td>
<td>quiz &gt; quizzed &gt; quizzing</td>
</tr>
<tr>
<td>hobnob &gt; hobnobbed &gt; hobnobbing</td>
<td>/V(:)CC(C)/ (\equiv) (&lt;V)\ VCC(C)&gt;</td>
</tr>
<tr>
<td>edit &gt; edited &gt; editing</td>
<td>itch &gt; itched &gt; itching</td>
</tr>
<tr>
<td>author &gt; authored &gt; authoring</td>
<td>/{other/ (\equiv) &lt;VC(#&gt;</td>
</tr>
<tr>
<td>cancel &gt; cancel(_l)ed &gt; cancel(_l)ing</td>
<td>mic/_mike &gt; miked &gt; miking</td>
</tr>
<tr>
<td>focus &gt; focus(_s)ed &gt; focus(_s)ing</td>
<td>motif &gt; motified &gt; motifying</td>
</tr>
<tr>
<td>traffic &gt; trafficked &gt; trafficking</td>
<td>parquet &gt; parqueted &gt; parqueting</td>
</tr>
<tr>
<td>catalog &gt; catalog(_u)ed &gt; catalog(_u)ing</td>
<td>ok(_{_}) &gt; okayed &gt; okaying</td>
</tr>
</tbody>
</table>

Base forms iambs

<table>
<thead>
<tr>
<th>/V:C/ (\equiv) &lt;VVC&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>refer &gt; referring</td>
</tr>
<tr>
<td>emit &gt; emitting</td>
</tr>
<tr>
<td>compel &gt; compelling</td>
</tr>
<tr>
<td>sandbag &gt; sandbagging</td>
</tr>
<tr>
<td>kidnap &gt; kidnapping</td>
</tr>
<tr>
<td>hobnob &gt; hobnobbing</td>
</tr>
<tr>
<td>edit &gt; editing</td>
</tr>
<tr>
<td>author &gt; authoring</td>
</tr>
<tr>
<td>cancel &gt; cancel(_l)ing</td>
</tr>
<tr>
<td>focus &gt; focus(_s)ing</td>
</tr>
<tr>
<td>traffic &gt; trafficking</td>
</tr>
<tr>
<td>catalog &gt; catalog(_u)ing</td>
</tr>
</tbody>
</table>

Key: Pink = CLD; white = IdP; beige = other PhM amendment.
4.1.5 Summary

Berg (2016: 473) does not deny that the phonological explanations of CLD are plausible, but instead he argues that they are not necessary, and the spelling of polymorphemic forms can be predicted solely from the morphological and orthographic information. The evidence above suggests a different picture. CLD only occurs when both the orthographic and the phonological conditions are met, which is to say:

\( '/_\text{VC}/ \equiv <=_\text{VC}#> \)

A full set of examples can be seen in Table 4.1, and incidences of CLD are marked in pink. In all other cases – with some very notable exceptions, marked in beige – the spelling simply follows IdP, and the spelling of the base is concatenated to <ed> or <ing>, as required. Hence we have:

<aim, aimed, aiming>
<dread, dreaded, dreading>
<jump, jumped, jumping>
<froth, frothed, frothing>
<search, searched, searching>
<edit, edited, editing>

There are several cases where the phonological conditions are met but not the orthographic ones. Berg (2016) notes some of these: <dread, dreaded, dreading> and <look, looked, looking>, etc. He does not note the case of <quiz, quizzed, quizzing> which meets only the phonological conditions. An exception can be made by arguing that the <u> represents a consonant but that is itself an appeal to phonological conditions, providing further evidence against a theory based exclusively on letter patterns, and evidence in favour of the ‘dual-condition hypothesis’, as the present analysis might be called. Berg also notes some of the situations where the orthographic conditions are met but not the phonological, including <fix, fixed, fixing>, but he does not note them all, possibly because they don’t occur in his relatively small corpus (of 52,447 items, Baayen et al. 1995).

None of the following base forms undergo CLD: ok, mic, motif, par’quet, and there may be many more, something that could only be discovered through a fuller analysis of a very large corpus, spanning several centuries. The first two of the four examples are particularly interesting because the {ed} and {ing} forms are formed from the variant spelling of the base, hence <okayed, okaying> and
<miked, miking> (shown in beige in Table 4.1). This provides evidence that the better phonographic match will be chosen, where there exists a choice.

However, it may not always be clear which is the best phonographic match, as there may be conflicting phonological demands. This is especially the case for trochaic bases where the need to mark stress (implicitly by not having CLD) may conflict with the need to mark consonant quality directly. The latter need is prioritised <traffic, trafficked, trafficking> and <catalog, catalogued, cataloguing> (both shown in beige in Table 4.1). The incidence of <ll> in <cancelled, cancelling>, etc. (marked in bold in the table) is not explicable on phonological grounds. Curiously, it is more common in American spelling than received wisdom suggests. Again, a fuller study is required to assess the relative weighting of these needs. A diachronic study might also show how certain spelling patterns may develop prominence over time. For example, the spelling <zinced> is preferred to <zincked> and <focussed> has lost ground over <focused>, suggesting two things that may be closely related. One is that stress-marking may have become more important than consonant marking, and the other is that IdP may have a stronger influence over time than PhM. This may be because the ‘visual bias’ in writing (Cummings 1988: 21) accustoms readers to spellings that are less phonologically transparent, so the spelling does not need to encode the sound as much as the meaning. However, the present study does not provide enough data to warrant such a grand conclusion.

At the end of the chapter, we shall see how this information can be included in an updated model of spelling formation. But first, we must look at the other two cases: <Ce#> ≡ /C#/ base forms; and then base forms ending in a vowel sound.

4c.2 Base forms ending in a consonant sound II: /C#/ ≡ <Ce#>

There are many words that end in a consonant sound but are spelt with final <e>, and the letter’s presence makes the spelling of the {ing} and {ed} forms a little more complicated. Unlike /C#/ words with a <C#> spelling, this time there does exist a small amount of variation between the spelling of {ing} and {ed} forms, as a handful of words retain <e> before <ing>.

4c.2.1<VCe#> base forms

In base forms with final stress, <VCe#> words generally have a long vowel and a /V:C/ shape, as in rage, engage, mete, delete, line, divine, tone, atone, puke, dispute. In their {ing} forms, the final <e> is dropped and <i> takes on the length-
marking function, as in raging, debating, meting, deleting, lining, divining, toning, atoning, puking, disputing. This is part of the broader tendency for the writing system to encode /V:CV/ ≡ /VCV⟩, as in notary, note, noting, proton, rotund, etc. {ed} forms simply add <d>, irrespective of the suffix’s pronunciation: as /d/ (raged, engaged, lined, divined, toned, atoned); /t/ (puked); or /id/ (meted, deleted, disputed). Once again, these are not ideal spellings from the viewpoint of PhM, but they may simply be the best available. The only other options are to keep <e>, rendering *<divineed>, *<atoneed> and *<pukeed>, etc., or to use digraphs and trigraphs, rendering *<divighned>, *<atoand>, <pyookt>, etc. These unused alternatives will be discussed in Chapter 5.

There are some exceptions where the <e> is not dropped. <ageing> is a British spelling which competes with <aging> (see Figure 4.11). This <e> is never retained in the longer words raging, waging, engaging, etc., nor in <icing>, which has no variant *<iceing>. The retention of <e> in <ageing> cannot be attributed solely to the need to mark consonant quality, since <g> ≡ /dʒ/ when followed by both <e> and <i>, and therefore the <eing> sequence does not provide a very efficient phonographic match. The spelling might suggest that there is some demand for base forms to always have three letters in order to be recognised. However, <ageing> is not the only example of <geing#> spellings. Words such as singeing, bingeing, etc. (discussed in Section 4c.2.3) may have an analogical effect, and this would explain why *<iceing> doesn’t occur in either British or American spelling, since there are no <-ceing#> spellings. The total preference of <aging> in American spelling suggests that PhM may be a higher priority than it is in British spelling.

There exist a few common <VCe#> spellings with a short vowel: come, give, have, live, love. The {ing} forms all drop <e>, rendering having, coming, giving, living, loving, but it is notable that three of them have irregular past forms, came, gave, had; the other two follow regular inflectional and orthographic patterns: lived, loved. The other relevant word here is axe (with <VCe⟩ ≡ /VCC/) and it is regularised in its inflectional spellings to axing and axed. The spelling <axeing> is very rare, according to Ngrams, and it is not listed in the OED (14105).

In trochees, <e> always seems to be dropped from <VCe#> spellings of the unstressed syllable, as in noticed, noticing, managed, managing, as well as the <VCCle> spellings of scribbled, scribbling, and bottled, bottling.
In British spelling (top), both <ageing> (in blue) and <aging> (in red) occur. In American spelling, <ageing> scarcely occurs.

4c.2.2 <VVC(C)e#> base forms
Carney (1994: 129–30) argues that <e> is often used in /V:C/ base forms to mark /s/, /z/ and /ð/, as in peace, cease, crease, bruise, browse, seethe, breathe. All of these words follow the usual pattern in the {ed} and {ing} forms, giving peaced, peacing; ceased, ceasing; bruised, bruising; seethed, seething. The same pattern extends to §French /ʒ/, as in rouge, rouged, rouging and §French /ʃ/, as in douche, douched, douching. In the polymorphemic forms, the function of <e> to mark §French <ch> ≡ /ʃ/ is lost, and the writing system has no solution for distinguishing such spellings from touch, touched, touching, where <ch> ≡ /tʃ/.

In light of Berg and Aronoff’s (2017) findings, we can re-assess this particular ‘<e>-marking function’, as Carney (1994: 129) calls it. The purpose of the <e> may not be to mark the phonological value of the preceding letter; instead it marks the grammatical category. Just as the <ick> of gimmick indicates the
word is not an adjective (c.f. Section 3d.2), the <e> of <browse> and <peace> indicates the words are not inflected forms, such as <brows> and <peas>.

4c.2.3 <VCCe#> base forms
Carney (1994: 130) notes that <e> can be used after consonant clusters to mark <s> ≡ /s/, as in tense and else (compare the /z/ of tens, lens and bells). Again, this can now be analysed as a DSTNCT spelling. Regardless, this <e>-marking function is lost in the polymorphemic forms tensed and tending, which have nothing to mark their consonant values as being distinct from lensed, lensing, where <s> ≡ /z/. In surge and bridge, <e> marks <(d)g> ≡ /dʒ/, giving surged, surging and bridged, bridging. <e> can also mark vowel length after digraphs (ache, bathe) and again this function is obscured in the polymorphemic forms: compare aching and stomaching. Sometimes <e> is a length marker after a cluster, as in table, which gives tabled, tabling.

Words ending <nge#> ≡ /ndʒ/ are highly variable and provide further evidence that the writing system is not fully stable. Mark Liberman (Language Log 7/6/15) writes that binge and singe almost always become <bingeing> and <singeing> but cringe always becomes <cringing>. His readers comment that this is perhaps because bing and sing are already words, so <binging> and <singing> are blocked. *cring is not a word, so <cringing> is available. Bob Ladd (in a comment) warns that lunge usually changes to lunging, despite lung being a word, although not a verb. The retention of <e> shows a desire for consonant correspondences to be indicated clearly, even in the face of problematic letter sequences (<eing>), but this seems to be counterbalanced somewhat by the need for orthographic consistency across the entire system. This is also a rare example of where DSTNCT can play a role. Retaining the <e> allows for a clear distinction to be made between singing and singeing, swinging and swingeing, and perhaps also binging versus bingeing.

4c.2.4 Summary
/C#/ ≡ <Ce#> base forms show very regular patterning in their {ed} and {ing} forms. In the former case, <d> is added. In the latter, <e> is dropped and <ing> is added. The only exceptions are the few cases where <e> is not dropped, as in ageing, singeing, swingeing and bingeing. It is not clear why <ageing> survives (only in British spelling) because the <e> has no phonological function. The other cases have two functions. After <ng>, <e> is required to mark the consonant /dʒ/, as opposed to /ŋ/, hence singeing versus singing. This PHM function is supported
by the need to mark the words as DSTNCT, and these dual needs may outweigh the single need to adhere to the usual letter-sequencing patterns.

4c.3 Base forms ending in a vowel sound

4c.3.1 /V#/ $\equiv$ <V#> base forms

Base forms ending in a vowel sound and <a>, <i>, <o> or <u> have relatively straightforward patterning for {ed} and {ing} forms. Words ending <Ce> $\equiv$ /CV/ will be dealt with here (e.g. adobe), as will <Cy#> and <ey#>. Words ending in <Ve#> will be dealt with in Section 4c.3.2.

While verbs ending in <a> are very rare, some nouns have {ed} forms, including umbrellaed and subpoenaed. Burchfield and Fowler (1996: 237) recommend using the archaic spelling <'d>, as in <umbrella’d> and <subpoea’d> but Figure 4.12 shows that this spelling is much less common (<subpoea’d> is unattested). This shows the powerful force of IdP and especially the importance of maintaining the spelling of the suffix in the face of the unusual letter sequence <…aed>.

![Figure 4.12](image)

<u>umbrellaed</u> (in blue) has long been dominant over <u>umbrella’d</u> (in red).

The patterns for <o#> and <u#> words are similar, so we get doing; going; booed, booing; vetoed, vetoing; videoed, videoing; plateaued, plateauing; and snafued, snafuing. <Cy#> spellings are changed to <Cied>, as in fry, fried, frying and marry, married, marrying, etc. and the spelling <skied> derives from both ski and sky, a reminder that DSTNCT is not a powerful force in the spelling of inflections. The change from <y> to <i> does not occur to <CVy#> base forms, as
in monkey, monkeyed, monkeying or array, arrayed, arraying. The word taxi usually changes to taxied and taxing, but there is a minority variant <taxying>.\(^{29}\) <CVw#> spellings are also regular, as in follow, followed, following. We might include <Vh#> spellings here, as in poohed and hurrahed, and <Vgh#> spellings such as sigh, sighed, sighing; weigh, weighed, weighing; and plough, ploughed, ploughing, which all follow the default IdP spelling. Alternatively, we could group these words under ‘base forms ending in consonant letter’ (Section 4c.1) and make conclusions accordingly. The <VC#> base form hurrah is the only one of these that might undergo CLD, but of course it does not, and we do not have *<hurrahed, hurahhing>. However, this is not evidence against Berg’s (2016) letter-sequencing hypothesis, as <h> is not listed among letters undergoing CLD.

A sizeable number of words end in <Ce> ≡ /CV/, including be, acne, adobe and cliché, which may or may not retain its §French accent (Evertz 2014: 81 counts ninety-seven such words in the CELEX database, or 1.1% of all words ending in <e>). Where there exists an {ing} form, <ing> is added without amendment, as in being. Obviously *<bing> is not a viable spelling from the viewpoints of IdP, PhM or even DSTNCT. The spelling of the {ed} forms are more problematic. The spellings <acned> and <adobed> provide strong evidence for any letter-sequencing hypothesis, although it doesn’t matter if these are analysed as <acne+d, adobe+d> or <acn(e)+ed, adob(e)+ed>. There is an argument that *<acnee+d> and *<adobeed> provide better phonographic matches but the <eed#> spellings suggest final stress, as occurs in agreed and guaranteed. Again, we see two PhM tensions between representing stress and representing phonemes unambiguously, in this case vowels.

The spelling <clichéd> provides a counter-argument to Berg et al.’s (2014) argument that the spelling of the base is compromised more than the spelling of the suffix. In this case, the spelling of the base form <clichê> is retained and {ed} is expressed with just <d>. This example is enough evidence to show that it makes no sense to distinguish between <acne+d> and <acn(e)+ed> or <adobe+d> and <adob(e)+ed>. It makes more sense to say that these English spellings are formed as a result of compromises between the need to express the spelling of the input forms (IdP) and the need to represent the phonological form (PhM), where possible. The reduction of <ee> to <e> is a compromise solution which attempts to represent both principles of spelling formation. Similarly, retention of the accent in <clichéd> solves all of the relevant problems at once. The spelling of the input forms is expressed adequately and the pronunciation is clarified with the retention of the accent. This perhaps explains why the accented form has gained in popularity in both the base and {ed} forms, as can be seen in Figure 4.13.\(^{30}\)
Accent-free <cliche> and <cliched> (in blue) are more common than <cliqué> and <clichéd> (in red), but the latter pair have gained ground in recent years.

4c.3.2 /Ve#/ ≡ <Ve#> base forms
The {ed} forms of <Ve#> words simply involve adding <d>, but things get more problematic for the spelling of {ing} forms. The rare case of <ae#> is straightforward, as in spae, spaeing, and so are <ee#> spellings, as in free, freeing and agree, agreeing, etc. Monosyllabic <Cie#> words usually change to <Cy>, as in dying, lying, vying, although the polysyllabic words birdie, stymie and sortie become birdieing, stymieing and sortieing. When the noun die (singular of dice) is converted to a verb, it is spelt dieing (Cummings 1988: 157). Carney (1994: 131) writes: ‘the use of these nouns as verbs is evidently so novel that the orthography draws attention to it by ignoring the usual graphotactic change and here, too, preserving the shape of the morpheme.’ If true, this would suggest that word-recognition, manifested through the use of IdP, is of supreme importance. However, there may be more reasons than simple IdP. <dieing> is clearly a case
of DSTNCT. Furthermore, *birdie and *sortie have initial stress, so the change may not be needed here. The same can be said of *stymie, but having *<y> twice in quick succession might also not be desirable, as in *(stymying) (c.f. (un)paralleled, above). There is also the question of why *<Cieing> spellings need to be changed at all. One answer is systemic inertia — the pattern is maintained simply because it seems to have always been there. Another possibility is that there is a visual need for sequences such as <iei> to be avoided, because they make it hard to recognise the syllable boundary. Fuhrhop, Buchmann and Berg (2011: 285–6) argue that syllable boundaries are frequently marked by letters with vertical lines (with an ascender, as in <b d f h k l> or descender, as in <g j p q y>). Vowel letters, which have neither ascenders or descenders, do not mark syllable boundaries well. Hence *<vying> is easier to split into its component morphemes than *<vieing>.

The very small number of *<Cye#> words do not follow the usual patterns. *<dye, dyeing> presumably has to keep the <e>, or else it would be confused with *<die, dying>. While this may appear to be an example of DSTNCT, it may also be the three-letter rule in operation (c.f. *ageing). Another three-letter example is *<eyeing> which is seldom spelt *<eying> (see Figure 4.14), not just because it might be confused with another word (such as *ey or *eie), but because the base form is an irregular spelling, and might be unrecognisable if altered. This contrasts against *<owe, owing> or even *<age, aging>, where the letters <ow> and <ag> provide enough information for the *<owing> and *<aging> to obey PhM. The spelling *<eying> cannot do this because *<ey> is not a good match for /aɪ/.31

Words ending in *<Coe#> keep their <e> in {ing} forms, rendering *hoeing, toeing, tiptoeing, shoeing, canoeing. The first two might be explained as examples of the three-letter rule, although *toing is also blocked by toing-and-froing, so it may also be taken as a DSTNCT spelling. Tiptoe is a compound, so we can expect its patterns to follow toe. Shoe and canoe have a non-canonical correspondence, and the /u:/ would be further obscured if the <e> were dropped, as in *<shoing> and *<canoing>. Again, PhM amendments are not helpful if they overly obscure the base form. Furthermore, the <oing> spelling would result in a collision with the regular correspondence <oi> ≡ /ɔɪ/, as in the onomatopoeic word *boing (OED 21027). The difference between *<shoeing> and *<going> is that the base form has an <e> which can be kept; *<go> does not contain that option so we cannot just add an excrescent <e>. DSTNCT spellings in inflected forms only seem to be possible when the option emerges within the spelling of the base. They cannot be fashioned from nothing by adding letters that were not there in the first place. This would explain why no DSCTNCT spellings occur in inflected forms such as

123
moped, evening and hops which all have the same spelling as a monomorphemic noun.

Figure 4.14
<eyeing> (in blue) versus <eying> (in red), between 1800 and 2008.

Words ending in <ue#> rarely keep their <e> in their {ing} forms, although its retention seems to have been more common in the past. The longer words always drop the <e>, as in intriguing and critiquing, and the same goes for monosyllabic sue which always renders suing, as shown in Figure 4.15. Ditto for ensuing. The spelling <ruing> is significantly more common than <rueing>, and nowadays the same goes for <gluing> over <glueing>. Both of these are shown in Figure 4.16. The correspondence <uing> ≡ /u:ɪŋ/ is evidently better than either <oing> ≡ /əʊɪŋ/ or <oing> ≡ /u:ɪŋ/. Hence readers can more easily recover sue, rue and glue from suing, ruing, gluing, than they can toe, hoe, shoe and canoe from *<toing, hoing, shoing, canoing>. The conclusion here is that default IdP spellings, with strange <eing> sequences, are preferable to amended spellings which obscure both the sound and spelling of the base form.
Figure 4.15
The spelling *sueing* (in blue) scarcely occurs, relative to *suing* (in red).

Figure 4.16
The spellings *ruing* and *gluing* (in blue) dominate *rueing* and *glueing* (in red)
### 4c.3.3 Summary

<table>
<thead>
<tr>
<th>Spelling</th>
<th>Base</th>
<th>{-ed}</th>
<th>{-ing}</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>umbrella</td>
<td>umbrellaed</td>
<td>umbrellaing</td>
</tr>
<tr>
<td>e</td>
<td>be</td>
<td>—</td>
<td>being</td>
</tr>
<tr>
<td>adobe</td>
<td>adobe</td>
<td>adobed</td>
<td>adobing</td>
</tr>
<tr>
<td>cliché</td>
<td>cliché</td>
<td>clichéd</td>
<td>cliching</td>
</tr>
<tr>
<td>i</td>
<td>ski</td>
<td>skied</td>
<td>skiing</td>
</tr>
<tr>
<td>o</td>
<td>go</td>
<td>—</td>
<td>going</td>
</tr>
<tr>
<td>do</td>
<td>do</td>
<td>—</td>
<td>doing</td>
</tr>
<tr>
<td>veto</td>
<td>veto</td>
<td>vetoed</td>
<td>vetoing</td>
</tr>
<tr>
<td>radios</td>
<td>radios</td>
<td>radioed</td>
<td>radioing</td>
</tr>
<tr>
<td>u</td>
<td>snafus</td>
<td>snafued</td>
<td>snafuing</td>
</tr>
<tr>
<td>y</td>
<td>sky</td>
<td>skied</td>
<td>skiing</td>
</tr>
<tr>
<td>plays</td>
<td>plays</td>
<td>played</td>
<td>playing</td>
</tr>
<tr>
<td>w</td>
<td>snows</td>
<td>snowed</td>
<td>snowing</td>
</tr>
<tr>
<td>ae</td>
<td>space</td>
<td>spae</td>
<td>spacing</td>
</tr>
<tr>
<td>ee</td>
<td>free</td>
<td>freed</td>
<td>freeing</td>
</tr>
<tr>
<td>ie</td>
<td>die (1)</td>
<td>died</td>
<td>dying</td>
</tr>
<tr>
<td></td>
<td>die (2)</td>
<td>died</td>
<td>dicing</td>
</tr>
<tr>
<td>oe</td>
<td>toe</td>
<td>toed</td>
<td>toeing</td>
</tr>
<tr>
<td>shoe</td>
<td>shoe</td>
<td>shoed</td>
<td>shoeing</td>
</tr>
<tr>
<td>canoe</td>
<td>canoe</td>
<td>canoed</td>
<td>canoeing</td>
</tr>
<tr>
<td>ue</td>
<td>sue</td>
<td>sued</td>
<td>sung</td>
</tr>
<tr>
<td>rue</td>
<td>rue</td>
<td>rued</td>
<td>ru(e)ing</td>
</tr>
<tr>
<td>intrigue</td>
<td>intrigued</td>
<td>intrigued</td>
<td>intriguing</td>
</tr>
<tr>
<td>we</td>
<td>owe</td>
<td>owed</td>
<td>owing</td>
</tr>
<tr>
<td>ye</td>
<td>eye</td>
<td>eyed</td>
<td>eyeing</td>
</tr>
<tr>
<td>dye</td>
<td>dye</td>
<td>dyed</td>
<td>dyeing</td>
</tr>
</tbody>
</table>

Table 4.2: The {-ed} and {-ing} forms of base forms ending in <V#>

#### Key:
- **Grey** = <e> dropping, **yellow** = <ie>-to-<y>

The default IdP spelling is used in the {-ed} and {-ing} forms for all /V#/ base forms ending in <a, i, o, u, Vv, Vw>, as in subpoenaed, taxied, taxiing, vetoed, vetoing, snafued, snafuing, monkeyed, monkeying, followed, following. /Cy#/ base forms require <y> to change to <i> when {-ed} is appended (e.g. marry, married). /Ve#/ ≡ /Ce#/ base forms follow the same patterns as /Ce#/ ≡ /Ce#/. We might expect *<adobeed> but, even
though this seems to mark the vowel /iː/ unambiguously, the spelling suggests the wrong stress pattern.

\(<Ve#>\) words are frequently troublesome, especially in the spelling of {ing} forms. The retention of \(<e>\) in hoeing, (tip)toeing, shoeing, canoeing, and eyeing suggests that PhM amendments are not made if they obscure the base form too much, and the default IdP spelling has to be retained in spite of the awkward sequence of vowel letters. DSTNCT can also play a role, hence the homophonous triplet of dying, dyeing and dieing.

Finally, the spelling \(<\text{clich\'{e}>}\) provides evidence that the spelling of {ed} forms is a result of compromises between PhM and IdP. The retention of the accent is a neat way of indicating the phonological form and it also shows that the spelling of the suffix is not retained at all costs. Instead a very viable compromise has been found between the need to represent the sound of the word, and the need to spell both the base form and the suffix.
4d. Final {s} inflections

The spelling patterns for {s} forms are quite different from {ed} and {ing} forms. This shorter section begins with the third-person singular {3PS s}. Its spelling patterns are almost identical to {Plural s} but {Genitive s} is quite different and it provides some different insights into the workings of the writing system.

4d.1 3\textsuperscript{rd} Person singular: {3PS s}

Phonologically, {3PS s} shows very similar allomorphic variation to {ed}: there is one unstressed syllabic suffix and two kinds of consonant appendix, due to voicing assimilation. The syllabic suffix occurs after sibilants, and a consonant appendix occurs in all other environments. However, {3PS s} and {ed} are treated very differently in the writing system, as Carney (1994: 19) points out. He argues that {ed} triggers CLD and {3PS s} does not, although it turns out that it does under very specific conditions. Furthermore, the spelling of the inflection is very frequently conditioned by the phonology, although the morphemic and the letter patterns all play a role too. In order to reflect the increased phonological conditioning in the spelling, the analysis will be ordered around the spelling of the suffix, rather than the spelling of the base forms.

4d.1.1 <es> spellings

When {3PS s} is added to base forms ending in a sibilant consonant, the suffix is manifested as /ɪz/ and these forms are spelt with <es>. Examples include *kiss\textsuperscript{s}, buzz\textsuperscript{s}, fix\textsuperscript{s}, wash\textsuperscript{s}, watch\textsuperscript{s}, reach\textsuperscript{s}, etc. Hence these spellings obey PhM over IdP. The absence of <e> in these cases would result in some very odd spellings: *<kiss, fixes, buzzes, washes, watches, reaches>, and yet these spellings are suggested by the application of IdP. These spellings are also predicted by Chomsky and Halle’s ‘fundamental principle of orthography’ which states that ‘phonetic information is not indicated where it is predictable by general rule’ (1968: 48). Since */kss/, */zs/, */ʃs/ and */tʃs/ are all disqualified phonological sequences (see Giegerich 1992: Ch. 6), there should be no need to mark the epenthetic vowel in writing. But the <e> is included, and this seems to satisfy the need to represent all phonological syllables in some minimal way. This suggests yet another level of complexity to the concept of PhM — not only do phonemes need to be marked, where possible, but so too do syllables. This would help to explain the importance of stress-marking, directly or indirectly. A somewhat comparable situation occurs in the spelling of the words bottle and metre where the <e> gives extra
orthographic weight to the spelling of the unstressed syllable, <tle> and <tre>, respectively. This is missing in prism and the German trade name Lidl. Compare bosom and fiddle or needle. Nevertheless the converse is not true: English spelling often has orthographic syllables where there is no phonological counterpart, so hopes looks like a word with two phonological syllables even though it only has one (compare the monomorphemic word herpes).

It is possible for the <es> spelling to co-occur with CLD, once the appropriate conditions are met, namely <VC#> base forms which also end in a sibilant consonant. Only <s> and <z> can fulfil this and it occurs in busses, gasses and quizzes. The spelling of other sibilants cannot undergo CLD, e.g. <ch>, <sh>, <dg(e)>, etc.

One graphotactic curiosity is that <es> is added after <Co#> base forms, as in vetoes, echoes, goes, does. Some words don’t take this <e>: radios and videos have just <s>, perhaps to avoid the unwanted sequences <ioes> and <eoes>, but perhaps also because they are derived from nouns. Verbs ending in <oo> do not have this <e> in their {3PS s} form, as in boo, boos.

The <es> spelling also occurs in hurries and marries which undergo the same <Cy> to <Ci> transformation as hurried and married. Once again, <CVy> words do not undergo the graphotactic change, as in enjoys, betrays, etc.

4d.1.2 <s> spellings

In all other cases, <s> alone is added, whether the pronunciation is /s/, /z/, or even /ç/. The latter case includes: binges, singes, lunges, cringes, manages, notices, tenses, etc., where the base form ends in a sibilant. This <e> might be viewed as an inert letter of sorts, with the <e> being silent in one allomorph, but pronounced in another, rather like the <n> of damn, damnation. As with {ed} forms, the function of <e> to mark consonantal values is lost in the polymorphemic forms, as in touches, douches or tenses, lense, etc. (c.f. Section 4c.2.2).

Other base forms ending in <e> simply have <s> added at the end, as in hopes, dies, dyes, eyes, rues, sues, frees, teethe and the same occurs after all non-sibilant consonants, regardless of voicing: hops, trots, locks, froths, robs, nads, begs, and even stomachs. This last example provides further evidence that the spelling <es> is phonologically conditioned, and not explicable from the orthographic patterns alone. Compare the <es> of coaches, beaches, etc.

Summary

The patterns in the spelling of {3PS s} forms show yet again, that English spelling is formed as a combination of morphological, phonological and graphotactic
factors, and none of these can be completely ruled out. The occurrence of <es> is conditioned either phonologically (after sibilant consonants, as in washes, fixes, sandwiches, etc.) or orthographically (after <o> and <y>, as in vetoes, hurries). The occurrence of CLD plus <es> would appear to be phonologically conditioned, as it only occurs after /s/ and /z/. The occurrence of <s> in all other cases is morphophonemic, as it can express three phonological forms: /s/ (as in hops, hopes); /z/ (as in robs, robes); or /ɪz/ (as in binges, douches). In these cases, IdP is obeyed and PhM seems to be adequately satisfied, but again, this satisfaction still defies measurement.

<table>
<thead>
<tr>
<th>Base form ending</th>
<th>Allomorph</th>
<th>Suffix spelling</th>
<th>Spelling</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>sibilant</td>
<td>/ɪz/</td>
<td>&lt;es&gt; kisses, bleaches fixes, washes</td>
</tr>
<tr>
<td>b</td>
<td>sibilant</td>
<td>/ɪz/</td>
<td>&lt;s&gt; lounges, bruises creases</td>
</tr>
<tr>
<td>c</td>
<td>vowel+sibilant</td>
<td>/ɪz/</td>
<td>CLD+&lt;es&gt; gasses quizzes</td>
</tr>
<tr>
<td>d</td>
<td>unvoiced consonant</td>
<td>/s/</td>
<td>&lt;s&gt; hops, hopes stomachs</td>
</tr>
<tr>
<td>e</td>
<td>others</td>
<td>/z/</td>
<td>&lt;s&gt; kills, welds prays, radios</td>
</tr>
<tr>
<td>f</td>
<td>&lt;o&gt; / &lt;y&gt;</td>
<td>/z/</td>
<td>&lt;es&gt; vetoes, hurries</td>
</tr>
</tbody>
</table>

Table 4.3
Spellings of the {3PS s} morpheme

4d.2 {Plural s}
The tendency for English to derive verbs from nouns and vice-versa means that there is huge number of spellings which can represent both {3PS s} and {Plural s}: e.g. bobs, beds, backs, books, etc. The variant spelling patterns are mostly identical so we have two kisses and no worries, etc. However, there are a few differences in how the writing system treats the two morphemes. Proper nouns ending in <y> are not changed to <ie>, so we have the Chomskys and the band, Dead Kennedys. Another difference is that CLD does not usually occur after <s>, as in two buses and toxic gases. Thus we see that grammatical category can sometimes affect spelling.

In abbreviations and numbers, an apostrophe is often added, as in CD's and the 1980's, although Burchfield and Fowler (1996: 61) recommend its absence nowadays. Yet they do suggest the apostrophe be kept where confusion
might arise, as in dot your i’s and cross your t’s (see also Huddleston and Pullum, 2002: 1586). This example is effectively a case of DSTNCT whereby <i’s> is kept distinct from <is>, but it is also part of a tendency to recycle minority variants where they may be useful (cf. <sync(h), synched, synching> and <catalog(ue), catalogued>).

The use of <es> after <o#> forms may be less common in {Plural s} than {3PS s}, although it is not clear exactly what the distributions of <oes> and <os> endings are: the <oes> ending seems to occur more in older words and it may indeed be recessive. Carney (1994: 19) claims the <oe> spelling is a marker of the long vowel but final <o#> ≡ /əʊ#/ is a predictable correspondence (e.g. avocado, cello, hello). Many words ending in <o> are §Italian and §Exotic nouns and their plural often has just <s>, as in armadillos and pistachios, etc. The <oes#> spelling may overlap with the long-standing tendency for §Basic English spellings to avoid words ending in the vowel letters <o>, <u> or <i>. In present-day English, this constraint can be maintained by the existence of <w> and <y> as variants of <u> and <i> (i.e. law and say, rather than *<lau> and *<sai>), but <o> does not have such a counterpart, merely the possibility of adding <w> (as in follow, yellow, etc). Obviously, this is not an option in armadillo, etc. due to the preference for maintaining the spelling of borrowed words. Adding <e> to the plural of these words may be less of an ‘offence’ than adding <w> to the base form, but there is little need for it too. It is quite possible that the <oes> spelling will be slowly ousted by <os>, in keeping with Berg and Aronoff’s (2017) explanation of how one variant tends to win out eventually. However, this tendency may be counterbalanced somewhat by the fact that there are several monomorphemic words ending in <os>, including kudos.

4d.3 Genitive {s}

{Genitive s} is a different kind of grammatical form to {3PS s} and {Plural s} because it can be attached to phrases, as in The King of Spain’s beard, but it is well worth comparing its use in the writing system. The spelling <’s>, provides a more consistent solution to the familiar allomorphic problem of representing /s/, /z/ and /ız/.
In other words, <‘s> has a wider phonological range than the use of <s> for {3PS s} and {Plural s}, which means that it is applicable to more phonological sequences. The <‘s> spelling applies to all sibilants, irrespective of a preceding <e>, and to <Co#> and <Cy#> words.

the bus’s wheel
the bruise’s size
the witch’s brew
the college’s year
the massage’s effect
Johnny Fox’s pub
the euro’s collapse
The Spy’s Choirbook

Nevertheless, the <‘s> spelling cannot cover every eventuality, and it is only partially compatible with the <s> used for {Plural s}. For example, the boys’ changing room contains <‘s>, a strange mix of the two. This spelling does at least reflect the single phonological monosyllable of /bɔɪz/. The issue is particularly acute in names. Burchfield and Fowler (1996) recommend <‘s> for monosyllabic and bisyllabic words, as well as longer words with penultimate stress: Des’s dissertation, Dickens’s oeuvre, St. James’s Gate and Elias’s birthday. Names ending in unstressed /ɪz/ are customarily spelt with just <‘>, as in Jesus’ mother and Claudius’ wife. The question is why. Is it to tell the reader to avoid cacophonous sequences like /ˈdʒɪzɪz/ and /ˈkl ə:dɪzɪz/? If so, then the problem shifts from phonology to syntax: pronouncing the possessive forms as /ˈdʒɪzəs/ and /ˈkl ə:dɪəz/ does not suggest possession and we end up with a weird juxtaposition of two nouns, ‘jesus mother’ and ‘claudius wife’. Neither solution seems ideal but the options are revealing from the viewpoint of spelling-to-sound relations. They suggest that writers have a certain amount of power in suggesting a pronunciation. Spelling them <Jesus’s> and <Claudius’s> actively suggests /ˈdʒɪzɪz/ and /ˈkl ə:dɪəz/. Spelling them <Jesus’> and <Claudius’> suggests two possibilities: a) don’t pronounce the final /ɪz/, or b) interpret <‘> as /ɪz/. The problem with (b) is that an entire syllable is represented with just <‘>, a spelling that is either super-efficient or not a very good phonographic match, depending on which way it is interpreted. The problem has arisen in the present study. The
spelling <DeFrancis’s> has been used because the base form ends in /Vs#/, while the /VCs/ ending has prompted the spelling <Cummings’>.

There are two important conclusions to be taken from this short subsection. One is that spelling can potentially have an effect on how readers pronounce a word. The other is that spelling units have a phonological range: the <’s> spelling can indicate /s/, /z/ or /ɪz/ after any base form, and the pronunciation can be predicted from the phonological context. By contrast, <s#> has a smaller range. It too can spell /s/ and /z/ (as in hops and hobs) but it cannot spell /ɪz/ unless the base form ends in <e> (compare binges and buses). Accordingly, <es> must be added to certain base forms, and this explains the difference between the witch’s brew and two witches, etc.

4e. Conclusion & model of English spelling formation

This chapter was built around the need to answer the question: to what extent is English spelling morphemic? In the terms put forth in Chapter 3, this was translated into the question of how much IdP is applied across the English writing system, and how is this tempered by PhM. And since it would be impossible to study the entire system at once, the focus here has been on the spelling of some revealing English inflections, with the aim of developing a theory of spelling formation which can be applied to the spelling of other word-formation categories.

Inflections formed by adding the {ed}, {ing} and {s} suffixes provide ideal data for examining spelling formation. They are all highly productive and they occur for a huge range of base forms, including all regular verbs and an indeterminate number of nouns (and noun phrases).

Furthermore, these suffixed inflections have no effect on the basic phonological shape of the word and hence we can examine the formation of the complex spellings without the complicating influence of phonological change (e.g. stress-shifting, vowel reduction, consonant alternations, etc.). And finally, borrowed words are subjected to the same inflectional patterns as native words and this gives us the opportunity to examine some unusual patterns of Anglicization in the spelling (exceptions occur in many Latin and Greek plurals).

IdP alone accounts for the spelling of a large proportion of {ed}, {ing} and {s} forms, although no effort has been made here to quantify the exact number. Instead, the focus in this chapter has primarily been on cases where there is a mismatch between the draft spelling suggested by IdP and the known
phonological form. The aim has been to unravel how those mismatches are resolved. Thus we see that there is an array of spelling patterns that cannot be explained by simple IdP. In some cases, the spelling of a complex word is conditioned by the phonological form of the base (kisses, washes, fixes). In other cases, changes require both phonological and orthographic conditions to be met (e.g. almost all incidences of CLD). It has been argued, therefore, that it is necessary to have a model of English spelling formation which takes phonological information into account, as well as morphological and orthographic information. This argument opposes the view of Berg (2016) who contends that CLD can be predicted from the spelling of the base form, without reference to the phonology. Berg et al. (2014) express a similar view. Letter patterns can explain the vast majority of cases, but not all, and the shortcomings of their theory are clearest when the correspondences between spelling and sound are non-canonical, such as <VC> ≡ /V:C/ or /V:/:. Such correspondences can occur in borrowed words (motif, motified; parˈquet, parˈqueted), clippings (mic, miked) and abbreviations (ok, okayed). One might argue that such examples are too peripheral to count as valid data. Yet it is their very peripherality which makes them expose the most basic patterns in the writing system.

Since orthographic precedents don’t always exist for the spelling of these extreme cases, solutions must be found from outside the established correspondence patterns, and this is where phonology can be a useful auxiliary in forming the spelling. If parquet has /t/ and final stress, then CLD makes good sense, just as it does in omitting. Otherwise, CLD would be misleading and might suggest an inappropriate pronunciation.

This line of reasoning has a very important implication, which is that phonology has a strong role to play in English spelling formation. This is not inconsistent with Venezky’s (1970) assertion that spellings are morphemic ‘as much as possible’, or ‘to a large extent’ (Carney 1994). Instead, the argument presented here is that morphemic spelling — or IdP — forms the basic structure of complex spelling but it interfaces with phonology a lot more than was previously recognised in the literature. It seems that phonological comparisons, or ‘phonographic matching’ as it has been termed here, fills the gap between Venezky’s morphemic hypothesis and the slightly more complicated reality. It is only by testing a possible spelling against the known phonological form that we can test if that spelling is a good one. If it is not a good spelling, it may be possible to amend the spelling in order to make it better fit the phonological form. That is the function of CLD, <e>-dropping, and the selection of other PhM
amendments discussed throughout this chapter (e.g. adding <c>, adding <k>, keeping <h> or <u>, etc.).

Problems remain, however. PhM is a much more complex concept than IdP because it involves matching entire spellings, or would-be spellings, against known phonological forms and comparing them at every level of analysis, including phonemes, syllables and feet. In this chapter, it has been assumed that the reader can do this from experience, meaning that they can tell that <jumping> is a good spelling of /jʌmpɪŋ/, and that <dopeing> is a bad spelling of /dəʊpɪŋ/. In the next chapter, a system will be developed for testing whether a spelling is a good phonographic match for the known phonological form (a concept which has not been discussed in any detail here). Only then can we tell whether the IdP draft or the amended PhM spelling ought to be chosen.

It is now possible to propose a fuller model of English spelling formation in order to explain the complex spellings discussed hitherto. The model, presented below as the final part of this chapter, takes the shape of a rudimentary algorithm, based around the principles of spelling formation.

***
4e.1 Updated model of English spelling formation

The following algorithm is an effort to summarise the how {ed}, {ing} and {s} inflections are spelt. The examples summarise what has been discussed in this chapter, and may not be exhaustive.

1. IdP draft

Form a draft spelling by applying IdP, adding the spelling of the base form to the default spelling of the affix, <s>, <ed> or <ing>.

<rank.s, rank.ed, rank.ing>
<jump.s, jump.ed, jump.ing>
<bomb.s, bomb.ed, bomb.ing>
<hop.s, hop.ed, hop.ing>
<zigzag.s, zigzag.ed, zigzag.ing>
***
<hope.s, hope.ed, hope.ing>
<canoe.s, canoe.ed, canoe.ing>
<singe.s, singe.ed, singe.ing>
<die.s, die.ed, die.ing>
<eye.s, eye.ed, eye.ing>
<cliché.s, cliché.ed, cliché.ing>
***
<edit.s, edit.ed, edit.ing>
<emit.s, emit.ed, emit.ing>
<traffic.s, traffic.ed, traffic.ing>
***
<wash.s, wash.ed, wash.ing>
<sandwichs. sandwich.ed, sandwich.ing>
<quiz.s, quiz.ed, quiz.ing>
<echo.s, echo.ed, echo.ing>
***
<marry.s, marry.ed, marry.ing>
2. PhM Testing

Compare the draft spelling to the known phonological form. If there are mismatches, then amend the draft, if necessary and possible.

{3PS s} and {Plural s} amendments:

i. Vowel insertion after sonorant consonants
   - washes, sandwiches
   - n.b. stomachs

ii. CLD + <e> insertion
   - quizzes, bus(es)

iii. Adding <es> after <o#>
   - echoes

iv. Changing <y> to <i> and adding <es>
   - marries
   - huskies

- *The Chomskys* (for proper nouns)

{ing} and {ed} amendments:

i. Removing <e>
   - hoped, hoping
   - canoed, singed,

ii. CLD
   - hopped, hopping,
   - zigzagged, zigzagging, etc.

iii. Other consonantal disambiguators
   - trafficked, trafficking,

iv. Graphotactic changes
   - <Cy> to <Ci> married,
   - <Cie> to <Cy> dying

v. Employment of the minority spelling of the base form
   - sync(h)ed, sync(h)ing,
   - okayed, okaying,
   - miked, miking,
   - catalog(u)ed
3. **Choosing**

Choose either the amended spelling or the draft spelling. The draft is chosen by default in the following cases:

i. If there are no possible or valid amendments:
   
   bombing, signing, stomaching,
   looking, dreading, hoeing, canoeing.
   hopes

ii. If the amendments suggest the wrong stress pattern:
    
    editing, hammering, etc.
    except <c#>: trafficking, mimicking

iii. If the amendment overly obscures the base form:
    
    eyeing, canoeing, ageing

iv. If the amended spelling is an existing inflected form:
    
    singeing, dieing

***

***
Chapter 5

The Phonographic matching principle

The primary aim of this chapter is to provide a method for testing whether a spelling is a good match for a known phonological form, and the method aims to formalise our intuitions about what counts as a good spelling. Armed with this tool, we can decide upon the suitability of a draft or amended spelling, and potentially between any two or more spellings. We know intuitively that `<dopeing>` is not a good spelling of `/dəʊpɪŋ/` and that `<boping>` is not a good spelling of `/bɒpɪŋ/` (i.e. `{bop+ing}`, but in this chapter, we shall see that it is possible to decide how one spelling can make for a better, if not always ideal, phonographic match between spelling and sound. This process will involve identifying problems at any phonological level up as far as the word, including the phoneme, the syllable and the foot.

Much of this chapter is spent explaining existing theoretical machinery used for showing the relations between spelling and sound. Section 5a begins by looking at Carney’s (1994) simple set of bi-directional mappings, a linear model which can be visualised as straight-line mappings between spelling units and phonemes (which themselves are imagined as discrete units in a line, or ‘beads on a string’, as the traditional analogy explains it). Venezky’s (1965/1967/1970) explanation of the decoding process is also linear but it implies elements of suprasegmental phonology which must be provided by the reader, including stress assignment, palatalization and vowel reduction. Recent theories have shown that English spellings contain information relating to all levels of analysis, including the syllable (Rollings 2004) and foot (Evertz 2014). In order to understand this theory, we must first look at non-linear, or suprasegmental, descriptions of phonology (e.g the foot, the syllable and its sub-units). In Section 5b, Giegerich’s (1992) model of English phonology will be presented as the necessary background phonological theory for later discussion. This phonological model is chosen partly because Rollings builds upon it, while Evertz builds on Rollings’ model of spelling, albeit using a slightly different phonological model. Furthermore, Giegerich’s model is particularly useful for comparing non-linear phonology to non-linear spelling, as we shall see. In Section 5c, we shall look in detail at Evertz’s groundbreaking work, which shows that the foot is a relevant level in the organisation of English spelling, and therefore it is possible to envisage a graphematic hierarchy in parallel with the established phonological hierarchy, allowing us to compare writing and speech at all relevant levels.
Evertz’s theory is modelled using Optimality Theory (OT) and I shall build on his work to show how OT can be applied to explain formally how we know if a spelling is a good phonographic match for the known phonological form, or not. To be more specific, OT can show why an imperfect spelling will be chosen over an even more imperfect spelling. It is obvious that <doped> is not a perfect phonographic match for /dəʊpt/, but it will be argued that it is the best of a bad lot. The OT model introduced here is used simply to highlight how the theory can be applied to a model of English spelling formation, and to solve one major outstanding problem. In order to temporarily solve this problem, a visual model will be presented here to show how spellings can be compared to phonology across their respective hierarchies. This model is useful for showing how and where problems arise in spelling formation, and why certain solutions (e.g. CLD, <e>-dropping) function so well. Another system of linear notation will be introduced here, and it will be used to show how readers can predict pronunciations from the spelling.

The theory developed in this chapter is based on the assumption, discussed throughout Chapter 4, that spelling formation is somewhat dependent on phonology and the two systems cannot be entirely divorced. One very important corollary is that phonology may, on occasion, be dependent on spelling. We shall see in Section 5d that ambiguities in the decoding of a spelling can lead to two or more pronunciations being used for the one word. Thus, it will be shown that variant pronunciations can be predicted from a word’s spelling. Evidence will be presented from several pronunciations of the word *omega*, all of which follow predictable paths from spelling to sound. The purpose of this section is partly to provide a better theory of spelling pronunciation, a phenomenon which has remained marginal in linguistic theory, but also to show that spelling and pronunciation are more deeply connected than was previously accepted, and both are important in understanding how spellings are formed.

***
5a. Linear models of English spelling

5a.1.1 Carney’s bi-directional letter-phoneme model

Carney’s (1994) model of English spelling is by far the best of the linear, phonemic descriptions of English spelling and it involves bi-directional mappings between spelling and sound at the phonemic level. A spelling such as <fill> has a very simple set of mappings, shown in Figure 5.1. The only complication is that /l/ is represented by <ll>, rather than just <l>. We have seen that this is common for monosyllables ending in /l/ (compare still, thrill, kill, Bill, etc), and it is thus predictable within the §Basic subsystem of spelling. Carney’s analysis of the bi-syllabic word filler is identical, with <er> ≡ /ə/. No effort is made to explain the fact that <er> represents an unstressed syllable, and there is no reference to any higher-level structures such as the syllable or foot. There is not even any mention of the fact that the absence of /r/ (in the British accent he analyses) can be explained by its position within the syllable. Instead, <er> simply maps onto /ə/ and vice-versa. Such a phonemic analysis also confines one to a very limited set of accents, as it cannot account for variation such as whether /r/ is pronounced or not (c.f. Section 6e.3).

\[
\begin{array}{cc}
< \text{fill} > & < \text{fill er} > \\
\uparrow \uparrow \uparrow & \uparrow \uparrow \uparrow \uparrow \uparrow \\
/ \text{fill} / & / \text{fill er} /
\end{array}
\]

Figure 5.1
Carney’s bi-directional mappings for the words fill and filler.

Simplicity is central to Carney’s work, but in order to maintain consistency across his entire description, he has to set up a sizeable number of descriptive categories, including dummy letters, inert letters, auxiliary letters, and so on. This can be seen in the word sign, where <g> is considered to be a dummy letter, with no phonemic mapping, and <i> ≡ /a/. While this letter alone is not the default correspondence for /a/ in this position, the correspondence is relatively common word- and syllable-finally (compare final and hi-fi). Such an analysis makes for a simpler system-wide description, he argues, than considering <ig> as a functional unit (c.f. Section 2c.1). The other two letters and phonemes have their
default correspondences, $<s> \equiv /s/ \text{ and } <n> \equiv /n/$.

\[
\begin{array}{c}
< s i g n > \\
\uparrow \uparrow \uparrow \uparrow \\
/ s a i \emptyset n /
\end{array}
\]

Figure 5.2
Carney’s bi-directional mappings for the word sign

The simplicity of Carney’s model is complicated by the presence of the etymologically derived subsystems of spelling (see Carney 1994: 96–101). He argues that in a §Greek word such as chlorophyll, the individual spelling units reinforce one another’s phonemic correspondence.

<table>
<thead>
<tr>
<th>Actual Greek spellings</th>
<th>χλωρός</th>
<th>φύλλον</th>
</tr>
</thead>
<tbody>
<tr>
<td>§Greek input spelling</td>
<td>chloro</td>
<td>+</td>
</tr>
</tbody>
</table>

\[
\begin{array}{c}
< \text{ch l o r o ph y l l} > \\
\uparrow \uparrow \uparrow \uparrow \uparrow \uparrow \uparrow \\
/ k l ɔː r ɔ f i l /
\end{array}
\]

Figure 5.3
Carney’s bi-directional mappings for $<\text{chlorophyll}> \Leftrightarrow /k\text{lɔːrəfil}/$.

The presence of $<\text{ph}>$ and word-internal $<y>$ suggests that $<\text{ch}>$ is more likely to be §Greek $/k/$, not §Basic $/tʃ/$ or §French $/ʃ/$. In effect, there are two interconnected levels of analysis: one involves the simple correspondences between phonemes and letters, the other involves the internal connections within the contributing CPUs (my term, not his), in this case $<\text{ch-l-o-r-o}>$ and $<\text{ph-y-ll}>$. This can be seen in Figure 5.3 (c.f. OED 32067).

5a.1.2 Venezky’s model of decoding
Venezky’s description of English spelling is really a model for decoding sounds from spellings, and it implicitly includes levels of analysis higher than the
He presents the example of *social*, which includes several kinds of mappings from spelling to sound (Venezky 1965: 53–4; 1967: 95–6; 1970: 46n). He posits two distinct phases to the decoding process. The first set of mappings involves the application of graphotactic rules, meaning rules mapping the letters to an intermediary morphophonemic level. The second set of mappings involves the application of linguistic rules, either phonological or grammatical. These are all shown in Figure 5.4.

All letters are assigned a morphophonemic value on the first level analysis, so `<s>` → `//s//`, `<o>` → `//o//`, `<c>` → `//s//`, `<i>` → `//i//`, `<a>` → `//æ//` and `<l>` → `//l//`. The mapping `<c>` → `//s//` is positionally-conditioned, since `<c>` is followed by `<i>`. Venezky’s use of double slant brackets `//` indicates the morphophonemic level, which is more abstract than the phonemic level. The arrow used here indicates that the correspondence is in one direction only, from spelling to sound. After the spelling `<social>` has been converted into the morphophonemic form `//sosɪæl//`, the linguistic rules can be applied. Stress is assigned to the first syllable and this has a knock-on effect on the two succeeding unstressed syllables, which collapse into one, *viz.* `//stæl//` reduces to `//sjəl//` and the resulting `//sj//` palatalizes to `/ʃ/`, leaving a phonemic form which might nowadays be transcribed (in GA) as `/ˈsʊʃəl/` or (in RP) as `/ˈsəʃəl/`. In Figure 5.3, I have collapsed the phonological rules into one level. In later phonological theories, most notably Chomsky and Halle (1968), the rules were ordered and one phonological change could feed another. If the rules were ordered differently, then the output might be significantly altered.

The problem with the second set of rules is that they have to be supplied by readers who must rely on their knowledge of both the language and the writing system. In the latter parts of this chapter, we shall see how stress, palatalization and vowel reduction can all be predicted from the letter-patterns alone, and thus allows us to a) have a more reliable model of decoding, and b) be able to test the adequacy of a spelling for a known phonological form.
<table>
<thead>
<tr>
<th>Spelling</th>
<th>Morphophonemic level</th>
<th>Phonemic output</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; … &gt;</td>
<td>// … //</td>
<td>/… /</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>graphotactic rule application</th>
<th>linguistic rule application</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>s</strong> → <strong>s</strong></td>
<td><strong>s</strong></td>
</tr>
<tr>
<td><em>stress</em></td>
<td></td>
</tr>
<tr>
<td><em>assignment</em></td>
<td></td>
</tr>
<tr>
<td><strong>o</strong> → <strong>o</strong></td>
<td><strong>əʊ</strong></td>
</tr>
<tr>
<td><em>diphthongisation</em></td>
<td></td>
</tr>
<tr>
<td><strong>c</strong> → <strong>s</strong></td>
<td><strong>ʃ</strong></td>
</tr>
<tr>
<td><em>palatalization</em></td>
<td></td>
</tr>
<tr>
<td><strong>i</strong> → <strong>ɪ</strong></td>
<td></td>
</tr>
<tr>
<td><strong>æ</strong> → <strong>ə</strong></td>
<td></td>
</tr>
<tr>
<td><strong>l</strong> → <strong>l</strong></td>
<td></td>
</tr>
</tbody>
</table>

**Figure 5.4**

Venezky’s mappings from spelling to sound via an intermediary ‘morphophonemic’ level. Example 1: <social>.

Venezky (1967: 95) goes on to explain how the polymorphemic word <signing> may be decoded (Figure 5.5). The spelling is first broken into two ‘graphemic allomorphs’, <sign> and <ing>, in a reversal of the word-formation process. These yield //sign// and //ŋɡ// on the first morphophonemic level and these two allomorphs recombine as //signɡ// for the application of phonotactic and phonological rules. The //gn// cluster levels to /n/ (compare gnat, gnome, etc) and the final //ŋɡ// cluster levels to //ŋ//. Venezky also claims that compensatory
lengthening accounts for the change of //i// to //ai// (even though this analysis only applies to <ig> spellings, such as align and paradigm, but not to phlegm and syntagm). The concept of a ‘graphemic allomorph’ is loosely defined but it is essentially the same as the concept of the CPU — it can be the spelling of a whole word or morpheme, or a slightly reduced version of that. There is only one graphemic allomorph of {sign} but {divine} would have two, spelt <divine> and <divin>, as in divinity. The main point here is that morphemes have to be identified before decoding can take place.

5a.1.3 Summary
Carney’s and Venezky’s models are both linear because they involve mappings from letters (or clusters of letters) to phonemes and, in Carney’s case, back again. Going both directions requires greater system-wide consistency, hence Carney’s greater attention to detail. However, Venezky’s model is much closer to a supra-segmental, non-linear analysis because it is reliant on stress, and thus the syllable, the level at which stress is expressed. However, the stress pattern must be supplied by the reader as the description does not provides rules for where stress occurs, and the same goes for the application of other linguistic rules (e.g. palatalization, consonant-cluster reduction, etc.). A fuller model of the decoding process should aim to predict more of the pronunciation, reducing the amount of information to be supplied by the reader. Ideally, then, the pronunciation would be fully predictable from the model. It will be shown throughout the remainder of this chapter that supra-segmental analyses can make steps towards that ideal. Before looking at attempts at doing this, we must look in some detail at supra-segmental analyses of phonology.
\[
\begin{array}{ccc}
\text{Spelling} & \text{Morphophonemic level} & \text{Phoneme} \\
< \ldots > & // \ldots // & /\ldots/ \\
\hline
\text{graphotactic rule application} & \text{linguistic rule application} \\
\hline
s & \rightarrow & s & \rightarrow & s \\
i & \rightarrow & i & \rightarrow & ai \\
g & \rightarrow & gn & \rightarrow & n \\
n & \rightarrow & \eta g & \rightarrow & \eta \\
+ & \rightarrow & \eta g & \rightarrow & \eta \\
i & \rightarrow & i & \rightarrow & i \\
n & \rightarrow & gn & \rightarrow & n \\
g & \rightarrow & gn & \rightarrow & n \\
\hline
\end{array}
\]

Figure 5.5

Venezky’s mappings from spelling to sound via an intermediary ‘morphophonemic’ level. Example 2: <signing>.
5b. English supra-segmental phonology

With regard to English phonology, we have mostly been concerned thus far with the phoneme, an understanding of which has been assumed, and the syllable, a relatively intuitive concept which has been left undefined. In this section, we shall explore the English syllable in some detail, looking at its subcomponents (onsets, nuclei, codas, and rhymes), and examining what constitutes a well-formed English monosyllable. We shall then move on to polysyllabic words and the foot, a less intuitive phonological level which will allow us to understand how syllables alternate between stressed and unstressed syllables. Feet are very important in English poetry and this is due to their importance as an organisational unit within words and across strings of words. However, we shall see that the phonological foot is a slightly different unit to the poetic foot, although the same terms are frequently applied to both. Once armed with this knowledge, we will be able to see how these supra-segmental levels are represented in English spelling.

5b.1 The phonological syllable

The analysis that follows is taken from Giegerich’s (1992) introductory textbook on phonology. There are several reasons for choosing this model. One is that Giegerich’s model is a relatively simple and phoneme-based model of phonology and the one I find best for comparing phonology and spelling in both directions. This may be because, as Linell (1982: 129–142) argues, ‘the phonologist’s view of phonetic structure is influenced by the perspective formed by alphabetic writing’; but it is also because ‘this is true of the layman’s thinking about speech to an even greater extent’. The functioning of English spelling seems to be somewhat dependent on the phonological categories that normal, literate people seem to have, so it makes sense to work with a phonological theory that does not stray too far from that. Furthermore, Rollings’ (2004) non-linear model of CLD builds on Giegerich’s model of English syllable structure, and Rollings’ work influences Evertz’s (2014) model, which in turn forms the basis of the analysis presented in this study.

5b.1.1 Well-formed monosyllables

We need to first understand why some English syllables are possible and others are not. This relies on the concept of sonority, which is the ‘relative loudness’ of a sound (Giegerich 1992: 132). Consonants are less sonorous than vowels. Among the consonants, stops [p b t d…] are less sonorous than fricatives [f v s z…], which are less sonorous than sonorants, [l r m n ŋ]), hence their name. Every
syllable must have a **sonority peak** and the segments that come before the peak must have increasing levels of sonority while segments coming after the peak must have decreasing levels of sonority. This is the **sonority sequencing principle**. The sonority sequence of [klæmp] is shown in Figure 5.6. The peak is the vowel [æ], and either side of it are sonorant consonants [l] and [m], respectively. On the margins are stop consonants [k] and [p], the least sonorous of consonants. Square brackets are used here since we are talking about sound segments, not phonemes.

![Figure 5.6](image)

The sonority sequencing of the syllable [klæmp]

The sequence /klæmp/ is, of course, an English word and ordered subsets of that sequence can also produce well-formed syllables, as long as the sonority sequencing principle is met. So we have /kæp/ and /læm/ (cap and lamb, respectively) or /kæm/ and /klæm/ (camp and clam) but we can’t have */lkæpm/ or */lkæ/, etc. Note that it is possible in English for a sonorant consonant to be the peak of an unstressed syllable, as in bottle, butter, bottom and button. Phonologically, these are often analysed as /əl/, /ər/, /əm/ and /ən/, as in Venezky’s /ˈsəʊʃəl/ above.34
The word /klæmp/ can then be subdivided into its onset and rhyme, /kl/ and /æmp/. The onset is everything that comes before the peak and it is optional. The word /æmp/ amp has no onset but it has the same rhyme as clamp, lamp and ramp. Rhymes are often more complicated than the nucleus. A stressed syllable must have at least two elements in the rhyme and this allows for a ‘branching nucleus’. This can be either a lax vowel plus a consonant, as in the words am and Sam, or it can be a tense vowel, such as the diphthong /aɪ/ (I, eye, bye, sigh) or the long monophthong /ɜ:/ (as in me, see, sea, etc.). Other consonantal phonemes are part of the coda, as in the /p/ of amp, ramp, clamp, etc. In Giegerich’s model, the /m/ is analysed as being part of both the nucleus and coda. Stressed syllables with the following vowels must always be closed: /æ, e, i, ɪ, ɒ, ʊ, ʌ/, as in hat, bet, bit, bot, put, but. All other vowels can occur in open syllables, as in bay, bee, bye, low, loo, law, bough, boy, bra and, in non-rhotic accents such as RP, bare, beer, fur (in Chapter 6 we shall look at the English vowel system in more detail).

Giegerich outlines the maximal English monosyllable, which includes major restrictions at its margins. Onsets can occupy up to three positions, here represented by the position holders X₃₋₅, although X₃ must be /s/. The coda can also have up to three appended segments X₄₋₃, although these must be non-sonorant coronals (e.g. /s, z, t, d, ð, ɹ/), segments that tend to feature in English inflections (paints, minds, bounced, lounged, length(s), width(s), sixth(s), etc.). The sections marked X₁₋₂ represent the obligatory branching nucleus and X₂ can be analysed as part of both the nucleus and coda.
The last two kinds of syllable forms in Figure 5.8 will be the most important when we come to look at Rollings’ analysis. /æm/ is an open syllable, with its tense vowel occupying two X-positions, whereas /æm/ is a closed syllable, with its lax vowel occupying just one X-position, X₁, and its consonant /m/ occupying X₂. The syllable template presented here does not apply to reduced syllables, and that includes the reduced forms of function words such as to, a, the, or, of, etc. (e.g. going to the shops to get a pint or two of milk). Such syllables,
having just one X-position in the rhyme, are called light (disregarding the optional /v/ in of). Again, the onset is optional: compare the and a. Syllable weight is an important factor in the positioning of stress in polysyllabic words.

5b.1.2 Polysyllabic words and ambisyllabic consonants
Well-formed polysyllabic words must be composed of well-formed syllables, although light syllables are interspersed among them. A very important consequence is that the possible sequences of sound segments, known as phonotactics, apply at the syllable level, not the word-level. This means that a discussion of polysyllabic words requires only slight amendments to what we already know about monosyllables. In particular, polysyllabic words are subject to the syllable boundary rule:

The syllable-boundary rule:

Within words, syllable boundaries are placed in such a way that onsets are maximal (in accordance with the phonotactic constraints of the language).

Giegerich (1992: 170)

Figure 5.9
Tree structure for the bi-syllabic word /meitrən/ ‘matron’
A word such as matron has a consonant cluster between its two syllable peaks, and the consonants fall into the onset of the latter syllable, as long as the cluster is a well-formed onset cluster. Since /tr/ is a well-formed syllable onset, the word matron syllabifies as /me.tron/, with the baseline dot marking the syllable boundary. In atlas, however, the /tl/ sequence is not a permissible syllable onset so it is split between the coda of the first syllable and the onset of the latter — hence /ˈat.las/ not */ˈa.tlas/.

The /t/ of /ˈat.las/ closes the first syllable, making it well-formed. A problem arises in words such as metric, petrol, apple, epic where there is a conflict between the need for onset maximisation — me.tric, pe.trol, a.pple, e.pic — and the phonotactic disallowance of stressed light syllables (/me/, /pe/, /æ/, /e/). Giegerich argues that the solution is to model the onset consonant as being ambisyllabic, so that the /p/ of apple and epic and the /t/ of petrol and metric are in both syllables at once (Figure 5.10). Ambisyllabic consonants have the dual function of closing the first syllable while also being part of the onset of the second syllable.

A consonant is ambisyllabic if it is (part of) a permissible onset (cluster) and it immediately follows a stressed lax vowel.

Giegerich (1992: 172)

The unstressed second syllable, despite being heavy, has a non-branching nucleus, consisting of a reduced vowel, followed by a coda consonant.
5b.2 The phonological foot

Ambisyllabic consonants can only occur across two syllables that are in the same phonological foot, a higher-level unit of organisation that plays a very important role in the timing of connected English speech. Every foot has exactly one stressed syllable, and this may, or may not, be followed by unstressed syllables. The words petrol, metric, epic and apple all have a stressed syllable followed by an unstressed syllable. This very common kind of foot is traditionally called a trochee, and other examples include bit.ter, but.ter, bet.ter, Bet.ty, yel.low, hap.py, hop.ping, quiz.zes, e.dit, Red.dit, rab.bit. These words all have a lax vowel in the stressed syllable, followed by an ambisyllabic consonant. Note how the unstressed second syllables can be either heavy or light. Reddit and hopping have branching nuclei, /ɪt/ and /ŋ/, respectively, so these syllables are heavy; happy and Betty both have a non-branching nucleus, /i/, so they are light. Where the stressed syllable has a tense vowel, it is open and the onset is confined to the second
syllable, as in *ma.tron, e.vil, ri.sing, ho.ping, u.ser, noi.some, au.dit, lou.sy.* Again, the unstressed syllable can have one or two X positions.

A dactyl has a stressed syllable followed by two unstressed syllables, a pattern that is much less common among monomorphemic words. Examples include *animal, criminal, militant, calendar, lavender,* etc. This kind of foot is very common among inflected words, such as *opening, edited, yellower* where a suffix is added with no effect on the stress pattern of the base form. Dactyls also occur with stress-neutral derivational suffixes such as \{y\} and \{ness\}, as in *flowery* and *happiness*. It is possible for there to be three unstressed syllables in a foot, as in *militancy* and *humourlessness*, but such words are rare.

All stressed monosyllabic words comprise one phonological foot, by default, as in *way* and *lone*. Many English words have a stressed syllable after an unstressed syllable, as in *away, alone, although, behold, relax, insist, emit,* etc, so the foot begins on the second syllable. In poetry, these are known as iambic feet, but they are not phonological feet per se because the unstressed syllable becomes connected to the previous foot. For example, the phrase *Go away!* has two feet, (Go a) and (way), and the first foot sounds like the placename *Goa*. Round brackets will be used henceforth to indicate a foot. Unstressed function words also get attached to the previous foot, and Giegerich (1992: 25) presents the following sentence with four feet:

\[(This\ is\ the)(house\ that)(Jack)(built)\]

These are, respectively, a dactyl, a trochee and two stressed monosyllables. The focus henceforth will only be on individual words, as that is the level at which English spelling operates.

Many English words and names contain more than one foot: *(cal.cu)(la.tor)* has two trochees, while *(cal.cu)(la.te)* has a trochee followed by a stressed monosyllable (although opinions vary on the status of the final syllable, as we shall see in Section 5d.2). In both words, the first foot has greater prominence than the second and this is indicated here with \<\>, the symbol usually used for marking stressed syllables.\(^{37}\) The word *(kanga)(roo)* also has a trochee followed by a stressed monosyllable, but this time the second foot has greater prominence. A large number of combinations are possible. One common pattern is for an unstressed syllable to be followed by a trochee, as in *Oc(to.ber), ve(ran.da)* and *(ro.ma)*. As with iambic words, the first syllable becomes part of the previous foot, as in *(ear.ly Oc)(to.ber).* Another pattern is for words to have two stressed syllables in succession, known in poetry as a spondee. This is very
common among compound words, including ‘(work)(shop), ‘(dead)(lock) and ‘(air)(craft), all of which have greater prominence on the first foot/syllable. Spondees can occur in monomorphemic words, including ‘(rab)(bi), (ho)’(tel) and (bam)’(boo). Notice that rabbi has greater prominence on the first foot, while hotel and bamboo have greater prominence on the second. This can shift, however, so that the main stress in ‘(ho)(tel)(ma.na.ger) moves to the first syllable of hotel. Stress shifting can also occur when certain affixes are added to words, and this is very common in the §Latinate vocabulary, as in (a.tom), a(to.mic) and ‘(for.mal) and for(ma.li.ty). The different kinds of phonological feet are exemplified in Table 5.1, along with some common combinations. This is an illustrative list, not an exhaustive one. An ‘X’ marks a syllable with primary word stress and an ‘o’ marks a reduced or unstressed syllable. Where there are two feet within one word, an ‘x’ indicates a stressed syllable in a foot with the lesser prominence. The exact status of such syllables is a matter of debate (see Section 5d.2) but the (temporary) notation is enough to highlight the relevant points. In the second part of the table, the distinction between monomorphemic and polymorphemic words can break down as many of these words are functionally monomorphemic in English (e.g. reality, conflict) but contain fossilized affixes as they were borrowed from other languages.
<table>
<thead>
<tr>
<th>Stress pattern</th>
<th>Monomorphemic examples</th>
<th>Polymorphemic examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>X</td>
<td>sit, close, lilt</td>
<td>sits, lipts</td>
</tr>
<tr>
<td>‘monosyllable’</td>
<td></td>
<td>closed</td>
</tr>
<tr>
<td>Xo</td>
<td>bitter, yellow, open,</td>
<td>hopping, opened</td>
</tr>
<tr>
<td>‘trochee’</td>
<td>wander, edit, conflict (n)</td>
<td>closing, closer</td>
</tr>
<tr>
<td>Xoo</td>
<td>lavender</td>
<td>editing, opening, wanderer,</td>
</tr>
<tr>
<td>‘dactyl’</td>
<td>calendar</td>
<td>martyrdom,</td>
</tr>
<tr>
<td>Xooo</td>
<td></td>
<td>manageable, militancy,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>humourlessness</td>
</tr>
</tbody>
</table>

Some combinations

<table>
<thead>
<tr>
<th>Xx</th>
<th>rabbi</th>
<th>workshop, aircraft</th>
</tr>
</thead>
<tbody>
<tr>
<td>‘sondce’</td>
<td>alloy</td>
<td>sandbag,</td>
</tr>
<tr>
<td>Xx</td>
<td>bamboo, hotel, omit,</td>
<td>unreal</td>
</tr>
<tr>
<td>oX</td>
<td>away, alone, allow</td>
<td></td>
</tr>
<tr>
<td>‘iamb’</td>
<td>insist, emit, conflict (v)</td>
<td></td>
</tr>
<tr>
<td>oXo</td>
<td>October, veranda, aroma, emotion</td>
<td>emitted</td>
</tr>
<tr>
<td></td>
<td>emitted</td>
<td>insisting, allowance</td>
</tr>
<tr>
<td>oXoo</td>
<td>reality</td>
<td>medicinal, emotional</td>
</tr>
<tr>
<td></td>
<td>allowances</td>
<td></td>
</tr>
<tr>
<td>Xox</td>
<td>lemonade, calculate</td>
<td>lioness</td>
</tr>
<tr>
<td>xoX</td>
<td>kangaroo</td>
<td>referee</td>
</tr>
<tr>
<td>Xoxo</td>
<td>Coca-cola</td>
<td>calculator</td>
</tr>
</tbody>
</table>

Table 5.1
Examples of the different kinds of phonological feet in English words
5c. Non-linear models of English spelling

5c.1 Rollings’ orthographic syllable

Trochees are very often spelt with double consonant letters, either within monomorphemic words (bitter, yellow, rabbit, etc.) or at the morpheme juncture in polymorphemic words (hopping, quizzes). Rollings (2004: 32–34) observes this pattern of English spelling and generalises it to all double consonant letters in English, whether they be historical doubles or examples of CLD in polymorphemic words. He argues that phonological ambisyllabicity is expressed through double consonant letters. He takes the examples of planning and planing and argues first for a parallel between phonology and orthography in the concept of open and closed syllables. In the case of /ˈpleɪnɪŋ/, the syllable /ˈpleɪ/ is phonologically open and so it is represented by the ‘orthographically open’ spelling <pla>. Hence /ˈpleɪ.nɪŋ/ ≡ <pla.ning>, unambiguously. In /ˈplænɪŋ/, there is a stressed lax vowel /æ/, followed by a permissible onset consonant /n/, so there must be ambisyllabicity. This is why the first syllable is spelt <plan>, with <n> closing the syllable, and the second syllable is spelt <ning>, with this <n> representing the onset segment. If we map Rollings’ analysis onto Giegerich’s visual model, we can see (in Figure 5.1) the bi-directional relationship between phonological and orthographic ambisyllabicity. The /n/ is both the coda of the first syllable and the onset of the second syllable, and this is represented by <nn> in the spelling. All other mappings from phonemes to letters are simple, assuming we take <ng> as a single unit.
Tree structure showing the relations between /plænn/ with ambisyllabic /n/ and the spelling <planning> with <nn>.

Figure 5.11
Figure 5.12
Tree structure showing the relations between \( /\text{ple}m\text{ŋ}/ \), with its open first syllable, and the spelling \(<\text{pla}>\) of \(<\text{planing}>\);

5c.1.1 Further descriptive benefits of the orthographic syllable
The idea of open orthographic syllables can be extended to explain more \(<\text{VCV}>\equiv /\text{V}:\text{CV}/ \) sequences, as in \(<\text{pla}n\text{ar}>, <\text{ve}n\text{us}>, <\text{wri}t\text{er}>, <\text{mo}t\text{or}>\), \(<\text{fu}t\text{ure}\.Orthographic syllabification can also explain the marking of vowel length before consonant clusters. In the word \textit{pantry}, for example, the cluster /\text{ntr}/ is not a permissible syllable onset, so the sequence must be split over two syllables, hence \(<\text{pan.}\text{try}\\equiv /\text{pæn.}\text{tri}/. The hypothetical word /\text{pen.}\text{tri}/ is phonotactically possible but the long vowel would have to be spelt with a digraph. This very sequence occurs in the placename \(<\text{Ain.}\text{tree}>\) and in the word \(<\text{re.}\text{main.}\text{der}>. The spelling of open syllables before a consonant cluster can be ambiguous however, a fact missed by Rollings. Take the word \textit{pastry}: since /\text{str}/ is a permissible onset, it seems that \(<\text{pa.}\text{stry}>\) is a good spelling of /\text{pei.}\text{stri}/, with the
long vowel being implied by the open orthographic syllable. However, in the word *gastric*, the vowel is lax, so the /s/ must be amabisyllabic, suggesting under-specification in the writing system. Applying the rules from <planning>, we might expect */gasstric*, but this cumbersome spelling doesn’t occur. One tempting hypothesis is that Rollings’ analysis only applies to §Basic spelling patterns, and not always §Latinate ones. For example, words ending in the Latin-derived {ic} suffix frequently have short vowels in their penultimate syllable, despite the <VCic> spelling. Examples include: *dynamic, atomic, tonic, mimic, arithmetic* (adj.), etc. Some of these words have long vowels, including *ludic*, while *systemic* varies between tense /i:/ and lax /e/, at least in British English (OED 196680).

Rollings’ introduction of an orthographic syllable is a very useful tool for explaining some of the more complex patterns of English spelling, such as CLD and non-CLD, which don’t lend themselves well to linear analyses. It would also appear that English spelling usually indicates phonological ambisyllabicity among §Basic spellings, if not always in §Latinate ones.

### 5c.2 Evertz’s graphematic hierarchy

Evertz (2014) builds on Rollings’ distinction between orthographically open and closed syllables and provides a very full analysis of what he calls graphematic syllables and graphematic feet (henceforth g-syllables and g-feet, contrasted against p-syllables and p-feet). Before discussing these in detail, it is worth having a look at where these levels fit within the larger graphematic hierarchy, shown in Figure 5.13 below (Evertz 2014). This hierarchy is defined in parallel with the phonological hierarchy and it is useful because it allows us to compare writing and speech at all levels of analysis. The hierarchy has been used to describe German spelling (Primus 2010) and English (Evertz & Primus 2013). Before the hierarchy existed, it was only possible to compare individual letters and clusters-of-letters against phonemes, while syllables and feet could not be discussed in any great detail.
5c.2.1 The graphematic hierarchy

The most familiar level in the hierarchy includes the ‘segments’ which are just letters. These are different from the graphemes which are defined as structural positions, effectively slots into which one or more letters are mapped (Evertz 2014: 44). Graphemes correspond to the structural positions marked ‘C’ or ‘V’ in Figure 5.13, and they are the graphematic equivalent of X-positions in Giegerich’s phonological analysis. Hence <sh> has two letters but it represents just one grapheme. In the graphematic approach, it is the structural positions which dominate the letters (*ibid*). The word shouter, shown in the figure, has two syllables, and the latter syllable is easier to understand. The onset dominates <t>, the non-branching nucleus dominates <e> and the coda dominates <r>. It is immaterial whether the word is pronounced with or without /r/ (e.g. GA v. RP) as the analysis is built solely around the relations among the letters. Graphemes frequently map to phonemes, but they don’t have to.

The first g-syllable is more complicated: the onset, which represents one structural position, dominates the complex grapheme <sh>. Since this is a stressed, or ‘strong’, g-syllable, it must have a branching nucleus (an issue discussed shortly) and therefore there are two structural positions to be occupied:
the <o> represents the syllable peak and the <u> is the other element of the nucleus. This analysis perhaps makes more sense if we take the example of rounder, where the nucleus again dominates two positions, spelt <o> and <u>, but this time the optional coda is present, spelt <n>.

Establishing the g-syllable boundary cannot be done using the syllable-boundary rule above, as it only applies to phonological consonant clusters. Instead, syllable boundaries are established in accordance with the ‘features’ of the letters themselves, effectively their graphic shape (Evertz 2014: 57–60). For example, letters which represent stop consonants <p b t d k g> tend to occur at syllable boundaries and in writing this is marked with a long ascender or descender, rising above or dropping below the x-height. By contrast, vowel letters <a e i o u> go neither above nor below the x-height (except for the dot on the <i>), and these represent the sonority peak of the syllable. Fricatives and sonorants, including <n>, are in between (Fuhrhop et al. 2011). This is called the ‘letter sequencing principle’, in parallel to the sonority sequencing principle, and it allows for the somewhat reliable predictability of graphematic syllable boundaries (for exceptions, see Evertz 2014: 58). These patterns provide graphic evidence of the importance of the syllable as a unit in written English (and other alphabetic languages). In this study, no reference will be made to letter features, as they add a layer of extra complexity to the discussion without providing a great deal of descriptive benefit. It will therefore be assumed that readers can deduce g-syllable boundaries in English spellings by applying the phonological syllable boundary rule and mapping to the spelling. In other words, the reader can ‘tell’ that the <shou.ter> and <roun.der> syllabify thus.

The graphematic word (g-word) is a relatively incontrovertible unit which involves a continuous sequence of letters bordered by spaces (Evertz 2014: 23). Issues only arise here in relation to hyphenation, spacing and compounding, which is beyond the present study. And, of course, graphematic words can be bordered by punctuation marks too.

**5c.2.2 Graphematic syllables**

There are several kinds of g-syllables, but all of them must have at least one vowel letter to represent the syllable peak. Other elements are optional (Evertz 2014: 61). G-syllables can either be full or reduced (ibid: 62), and there are two types of full g-syllables, strong and weak. **Strong full g-syllables** all have a branching nucleus, which means they must correspond to either: a) a lax or ‘unary’ p-vowel plus a single p-consonant or, b) a tense or ‘binary’ p-vowel (i.e. a diphthong or a long monophthong). P-vowels in
group (a) are usually spelt \(<\text{VC#}\rangle\) as in \textit{pat, pet, pit, pot, put}, and these are all closed syllables.\(^{38}\) P-vowels in (b) can be spelt in two ways. In monosyllables, or in the final syllable of a word, they are spelt with a digraph (or ‘graphematic diphthong’), as in \textit{say, paid, paw, tau, see, seat, seed, rein, feud, pie, loo, soy, soil, suit, sue, as well as away, guffaw, referee, curfew, bonhomie, kangaroo, alloy, issue}. Consonant letters in the rhymes of these words correspond to the coda of the syllable, not the nucleus (\textit{<seat, rein, feu, a.wa.i, de.void, etc.). In polysyllabic words, a single vowel letter can represent a binary p-vowel, as in \textit{la.ter, A.pril, me.tre, ve.to, wri.ter, di.va, vo.ter, ro.ta, u.ser and nu.clear}. As above, these are called open \textit{g}-syllables. It should be remembered that polysyllabic words can have closed \textit{g}-syllables too, as in \textit{lat.ter, as.trid, men.tal, ven.ted, win.ter, din.gy, vor.tex, ros.ter, un.der, num.ber}. When a \textit{<V>} letter is followed by a permissible onset cluster, then the syllable is open, as in \textit{<ta.ble>} and \textit{<pa.stry>}, but when it is not, then the syllable is closed by one of the consonant letters, as in \textit{<hum.ble>} and \textit{<spin.ter>} (Evertz 2014: 115). It seems that the cluster \textit{<st>} can be parsed in either of two ways. The word \textit{pasty} has two pronunciations, one syllabified as \textit{<pas.ty>}, the other as \textit{<pa.sty>}. Reduced \textit{g}-syllables must fulfill the following conditions. They must have a non-branching nucleus; their peak must be \textit{<e>}; they must be situated in a prosodically weak position in a foot; and they must be light (Evertz 2014: 79). Examples include the \textit{<ter>} of \textit{la.ter} and \textit{lat.ter}, as well as the \textit{<ken>} of \textit{darken}. A \textit{minimal g-syllable} contains just \textit{<(C)e>}, as in \textit{<la.te>} and \textit{<ho.pe>}, and these can only occur after a stressed syllable. Minimal syllables have no phonological correspondence but they have some useful functions in a) marking the preceding \textit{g}-syllable as open, and b) providing \textit{graphematic weight} to spelling, a concept we shall return to. It should be noted that spellings such as \textit{<la.ment>} and \textit{<con.fess>} do not contain reduced syllables, as they have \textit{<CC>} after the syllable peak \textit{<e>}. Syllables such as these are \textit{graphematically heavy} and have a tendency to attract the stress; hence they are less likely to appear in prosodically weak positions.\(^{39}\)

\textit{Weak full g-syllables} are not dealt with as thoroughly by Evertz. Like reduced syllables, weak full \textit{g}-syllables have a non-branching nucleus, which means that they have a single vowel letter representing a unary p-vowel in a weak prosodic position. These are often unstressed and correspond to /a/ or /u/. Examples include the \textit{<da>} of \textit{co.da}, the \textit{<di>} of \textit{di.vi.ne}, and both \textit{<a>} and \textit{<ma>} in \textit{a.ro.ma}. Membership of this group is largely by default: it includes \textit{g}-syllables which are neither strong nor reduced, which amounts to \textit{<(C)a>, <(C)i>, <(C)o> and <(C)u>} syllables. Problems arise with this syllabic classification in graphematic dactyls, as we shall see.
5c.2.3 Graphematic feet: canonical trochees

The word *shouter* is a ‘canonical graphematic trochee’ and it can be broken into two g-syllables. \(<\text{shou}\>\) is a strong full g-syllable (and thus stressed) while \(<\text{ter}\>\) is reduced g-syllable (and thus unstressed). All canonical trochees have \(<\text{e}\>\) in their second syllable.

The graphematic foot can be used to explain CLD, non-CLD and mute final \(<\text{e}\>\). For example, the words *latter* and *later* are both trochees and they syllabify as \(<\text{lat.ter}\>\) and \(<\text{la.ter}\>\). This use of round brackets, with extra spaces, is my own innovation and has been borrowed from phonological theory, where feet are indicated with round brackets. By the definition of the foot, the first syllable is stressed, and later syllables are unstressed. Both words \(<\text{lat.ter}\>\) and \(<\text{la.ter}\>\) have the same reduced second syllable, phonologically and graphematically, but the first syllable is the important one here. The syllable \(<\text{lat}\>\) is closed by the \(<\text{t}\>\) and this marks a unary p-vowel followed by an ambisyllabic consonant. In *later*, however, the syllable \(<\text{la}\>\) is open so the vowel marks a binary p-vowel and the consonant /t/ is simply the onset of the second syllable.

The same analysis can be applied to the other vowel letters, as in *better, meter; bitter, biter; hopper, hoper; rudder, ruder*. All of this is just a more technical way of saying what Rollings (2004) says about CLD and non-CLD, with the example of *planning* and *planing*, but Evertz (2014: 111–124) provides further analysis of the different ways that ambisyllabicity is regularly represented. This can also be expressed through normal letter doubling (e.g. \(<\text{bb}\>\), \(<\text{dd}\>\), \(<\text{ff}\>\), \(<\text{gg}\>\), etc.) or with \(<\text{ck}\>\) (*bucket*), \(<\text{tch}\>\) (*kitchen*), \(<\text{dg}\>\) (*budget*) and \(<\text{ng}\>\) (*singer*) (see also Rollings 2004: 127–8). Like others, Evertz does not distinguish explicitly between such historical doubles and CLD in polymorphemic forms. The spellings \(<\text{tch}\>\) and \(<\text{dg}\>\) are not productive and don’t occur across morpheme boundaries.

An important innovation of Evertz’s approach is that a word such as *late* is considered to have two g-syllables \(<\text{la.te}\>\), despite the fact that it only has one p-syllable (compare Figures 5.14 against Figure 5.15).
Figure 5.14
<la.te>: one graphematic foot, two graphematic syllables
Taken from Evertz (2014)

Figure 5.15
/leit/: one phonological foot, one phonological syllable
Taken from Evertz (2014)
The <te> syllable is a minimal g-syllable and thus mute. The foot-level analysis allows us to view the structural similarities between <la.te> and <la.ter>, with the syllable <la> indicating a long vowel in both words. The same analysis can be applied to other vowel letters, as in mete, bite, hope, rude, and the final or ‘mute’ <e> does not have to be explained as a split digraph <a..e>, <e..e>, etc. (c.f. Carney 1994: passim). Henceforth, I shall underline (parts of) spellings where there are two g-syllables for just one p-syllable, as in <la.te> and <un.du.la.te>.

5c.2.4 Graphematic feet: non-canonical trochees
Trochees which have a vowel letter other than <e> in the second syllable are considered by Evertz to be non-canonical. This includes affix, cello, ferry, summit, willow, which all have historical doubles, and cuckoo, jockey, itchy, clingy, edgy, pidgin, ethic, echo and pushy, which all represent ambisyllabicity with regularity, at least within the limited confines of English spelling patterns (Evertz 2014: 127). However, almost one in four non-canonical trochees have only one consonant letter in the expected position, including cabin, epic, relish and widow, so Evertz makes the understandable assumption that this is far too large a proportion for the spelling of such trochees to be considered predictable. By contrast, only one in seventeen canonical trochees is spelt unpredictably with respect to ambisyllabicity, including camel, desert, honest, linen, modest (Evertz 2014: 123).

The spelling of non-canonical trochees may be more predictable if we take into account some new facts put forth by Berg (2016: 468–471), that there is a sizeable number of word endings which are almost never preceded by a double consonant. Examples include the <it>, <id> and <ish> of limit, tepid and relish. Of course, these sub-regularities can often be explained with reference to etymology, and this brings us back to Carney’s observation (1994: 126) that double consonant letters feature a lot more in §Basic rather than §Latinate vocabulary. This issue is very important and merits a fuller study, especially since etymological information has been neglected in the graphematic approach (Berg takes a letter-sequencing approach to the issue, and does not note the origins of these words).

A further reason why one might disregard Evertz’s distinction between two kinds of trochee is the fact that a large number of canonical g-trochees are the result of stress-neutral suffixes spelt with <e>, as in faster, fastest, fasten and fasted. Furthermore, the distinction makes for an unwanted split between {ed} and {ing} forms which, as we saw in Chapter 4, display remarkably similar patterning
in their spelling. Hence *shouted* is a canonical graphematic trochee but *shouting* is non-canonical. For present purposes, no distinction will be made between canonical and non-canonical trochaic feet, so I shall not be concerned henceforth with which vowel letter is in the reduced syllable. Thus there will be no need for the category ‘weak full g-syllable’ just full g-syllables and reduced g-syllables. However, the idea of a ‘mute’ minimal g-syllable <Ce> is very useful and will be retained as a separate category.

5c.2.5 Graphematic dactyls

A spelling such as <wan.de.rer> is a graphematic dactyl as it has one strong full (i.e. stressed) syllable <wan>, followed by two reduced syllables <de> and <rer>. The unit is identified by Fuhrhop and Peters (2013, cited by Evertz 2014: 107) but Evertz does not include them in his analysis on the grounds that they can be explained as a combination of a graphematic trochee <wan.de> plus an extrametrical syllable <rer>. He argues that there are good morphological grounds for this analysis, as the reader can see that the final syllable contains an extra morpheme. It is indeed true that phonological dactyls are common among polymorphemic words, as in *wanderer, yellower, edited, opening, martyrdom, happiness*, etc, occurring when a stress-neutral suffix is added to the trochaic bases. However, English has many monomorphemic dactyls, especially among §Latinate nouns, and there is a very strong argument for including dactylic feet in a description of English spelling. Examples include ‘vertebra, ’capital, ’cinema, ’optimum, ’enema, ’animal, A’merica. Syllable weight is the explanation behind the unusual stress pattern of these words. While English words tend to be stressed on the penultimate syllable, making trochees so common, the dactylic words just listed all have a light penultimate syllable, spelt <CV>, which cannot take the stress. Hence the stress is pushed back to the antepenultimate syllable (Giegerich 1992: 186–8). It is not necessary for this syllable to be light, as can be seen from the fact that the <op> of *optimum* and the <ver> of *vertebra* both have branching nuclei, making them strong full syllables in Evertz’s analysis. However, the concept of a weak full g-syllable breaks down here as there is stress on the <ca> of <ca.pi.tal>, the <ci> of <ci.ne.ma>, the (first) <e> of <e.ne.ma> and the <a> of <a.ni.mal>. Phonologically, these light syllables become heavy through ambisyllabicity (Giegerich 1992: 188), but this is not shown in the spelling. Carney (1994: 123) notes the general lack of consonant doubling in words with antepenultimate stress, as in *celery, character, prodigal, bigamy*, etc. The functional need for double consonants in dactyls is also much lower than in trochees because of a very old phonological process called tri-syllabic laxing,
whereby stressed ante-penultimate syllables have a lax vowel, hence alternations such as ‘cone – ‘conical, se’rene – se’renity, etc. (see Lahiri & Fikkert (1999) for an account of tri-syllabic shortening from Old English up to the present day). There is no need for the spellings *<connical> and *<serennity> because the lax vowel is predictable from context.

In this study, the graphematic dactyl will be considered a valid and important kind of foot in English spelling, with one strong syllable and two weak ones. The shape of the dactylic foot is usually <(C)V(C).CV(C)>, which means that the ante-penultimate, stressed vowel may have a branching graphematic nucleus, marked in the spelling (<(op.ti.mum)>, <(ver.te.bra)>) or it can have an ambisyllabic consonant as part of the nucleus, unmarked in the spelling (as in <(ci.ne.ma)>, <(e.ne.ma) > and <(a.ni.mal)>). If these spellings followed the §Basic pattern they would be *<cinnema>, *<ennema> and *<annimal>. Giegerich (1992: 188) also provides examples of dactyls where the penultimate syllable is not light, again marked in bold: <(ca.len.dar)>, <(cy.lin.der)>, <(bad.min.ton)>, but he notes that these are exceptions. Compare the penultimate stress of <ve(ran.da)>, <e(nig.ma)>. Such details will have to be accounted for in a fuller description of the foot in English spelling than can be provided here

5c.2.6 Summary
The evidence provided by Evertz shows that the foot is undoubtedly a useful unit for analysing some English spellings. The foot plays a major role in CLD, non-CLD and the use of mute <e>, although it may well be the case that the foot is a more important unit in §Basic rather than §Latinate vocabulary, and this would account for the similarities between English and German, since §Basic vocabulary is substantially Germanic and trochaic. Phonological dactyls may not be encoded so clearly in English spelling, perhaps because they often derive from French and Latin.

For the remainder of this study, we need to recall the following. English feet can be monosyllables, trochees or dactyls, as exemplified by the words <lat>, <(lat.ter) > and <(la.te.ral)>. Syllable boundaries are marked with a dot, feet are marked with round brackets, and the first syllable is always stressed, while subsequent syllables are unstressed. G-syllables can be either open or closed, corresponding to a binary or unary vowel. Compare <(A.pril) > and <(wri.ter) > to <(as.trid) > and <(win.ter) >. In dactyls, this distinction does not apply and, as we shall see, this can cause ambiguity in the decoding. In monosyllables, closed syllables end in a consonant, as in <pat>, <pet>, <pit>, etc, but open syllables
may be spelt with two vowel letters, as in <pay>, <pee>, <pie>, <pow>. This also applies to polysyllabic words with a stressed final syllable, as in <a.way> and <fri.ca.see>. There is a mismatch between the phonological monosyllable late and its trochaic graphematic counterpart <(la.te)>, indicated here by underlining. The <te> is called a minimal g-syllable. Two other g-syllables will be recognised in the remainder of this study. Full g-syllables occur in stressed position, and must have at least two segments in the rhyme (e.g. <see>, <sea>, <set>, <say>), although there may be more, as in <sent>, <seat>, <cert>, <Spain>, <eight>, and these are almost always stressed. Reduced g-syllables occur in unstressed positions, and have either one or two segments in the rhyme, as in <(ben.der)>, <(ben.dy)>, <(pan.try)>, <(ban.dit)>. Evertz’s distinction between strong full syllables <CeC> and weak full syllables <CaC>, <CiC>, <CoC> and CoC> will not be retained in this study. While it is convenient for distinguishing canonical and non-canonical trochees, that distinction will also be abandoned here, but it is worth noting that the distinction may correlate somewhat with the difference between §Basic and §Latinate subsystems. Spelling like <insipid> and <lemon> do not have <pp> or <mm> because they were borrowed (c.f. Section 8a).

We shall see in the remainder of the chapter that the foot can be a very useful unit in understanding the mechanics of English spelling, in the formation of polymorphemic forms and in the process of decoding new and/or unfamiliar words. In particular, a hierarchical structure allows us to make comparisons between the spelling and phonology at all levels of analysis, making it possible to identify any kind of mismatch between spelling and sound.
5d. Modelling the spelling formation of \{ing\}, \{ed\} and \{s\} forms

We now have enough theoretical machinery to start answering the question posed in Chapter 4: how do we know if a draft spelling is a good match for the known phonological form? One answer is that an ideal spelling will have a one-to-one mapping between spelling and phonology at all levels of representation, including onsets, nuclei, codas, rhymes, syllables, feet and the word, and that this can be checked by comparing the phonological hierarchy (henceforth P-tree) against the graphematic hierarchy (henceforth G-tree). Initially, a visual model will be used to compare these two and to identify mismatches in cases where the spelling is not ideal. However, we will need more powerful machinery to explain why one non-ideal spelling is chosen over another.

5d.1 \{ing\} forms

In the first instance, we will look at how the draft spelling \(<\text{jum}ping>\) compares against the known phonological form /ˈjʌmpɪŋ/. This will be an easy way to introduce G-trees and their relationship with P-trees. Afterwards, we will look at cases where amendments are required, and the G-trees reveal why the draft spelling doesn’t work.
Constructing a P-tree for /ˈjʌmpɪŋ/ is easy, given what we saw in Section 5b. The word has one foot and two syllables, as can be seen in Figure 5.16. It is a simple trochee with the first syllable being stressed and the second unstressed. The stressed syllable is a closed, /CVC/ syllable, with the onset being represented by /dʒ/, and the rhyme has /ʌm/. This can be analysed as a branching nucleus or single nucleus /ʌ/ and coda /m/. In order for the P-tree to best match the G-tree, I have represented the nucleus as branching, with the coda marked in brackets in the diagram. The unstressed syllable has /p/ as the onset, a nucleus of /ɪ/ and a coda of /ŋ/.

![Diagram of FG tree for /ˈjʌmpɪŋ/]

Constructing the G-tree is almost as simple. The first syllable is <CVC>: the onset is <j>, the branching nucleus is <um>, although this splits into a peak of <u> and a coda of <m>. The second syllable is also <CVC>, as long as we consider <ng> to be one unit. Hence the onset is <p>, the nucleus is <i> and the coda is <ng>. Within Evertz’s model, <jum> is a strong full g-syllable and <ping> is a weak full g-syllable, but there is no actual evidence of this in the spelling, apart from their relative position in the word. The imaginary word *<pingjum> is perfectly possible and we might expect it to also be a trochee or spondee (compare dingbat). However, the morphological structure of jumping provides evidence beyond the position of the syllables: {ing} forms always end in unstressed /ɪŋ/.
Hence we can assume that the first syllable is a full g-syllable and the second is reduced. The G-tree for < (jum.ping) > is presented in Figure 5.17.

The next task is to map from the spelling to the sound in both directions and this is done by flipping the G-tree and putting it underneath the P-tree. We can see that almost every mapping works perfectly in both directions. At the syllable level, <jum> ≡ /dʒʌm/ and <ping> ≡ /pɪŋ/. At the segmental level, <j> ≡ /dʒ/, <u> ≡ /ʌ/, <m> ≡ /m/, <p> ≡ /p/, <i> ≡ /i/, and <ng> ≡ /ŋ/. Figure 5.18 shows how all the relevant levels match their phonological equivalents in the P-tree.

All of the graphemes provide good matches for the phonemes but the reverse is not necessarily so unambiguous. Sometimes initial <j> is decoded as /j/, as in the name Jung and, much more likely, <u> can be /ʊ/ as in butcher. Hence it may be possible to decode < (jum.ping) > as /ˈjʌmpɪŋ/ or /ˈdʒʌmpɪŋ/. Such details cannot be accounted for in the present model. The mapping <j> ≡ /j/ is dependent on the subsystem to which the word belongs, although we recall that there may be further evidence from within the word’s spelling, while the semi-predictable distribution of /ɔ/ and /ʌ/ is discussed in Section 6d.2.

The usefulness of G-trees becomes clearer when we look at draft spellings which do not provide a good phonographic match for the known phonological form. When IdP is applied to the words dope and bop, the draft spelling of the {ing} forms are ^2<dopeing> and ^2<boping>, respectively. As stated in the introduction, experienced literate speakers know intuitively that these are not ‘good’ spellings, but now we can check this by comparing the known phonology against the G-trees that are suggested by the draft spellings. In other words, we now have the tools to construct a hypothetical G-tree out of the draft spelling and then compare it against the known P-tree, with the aim of locating where the problems actually lie.

Let’s begin with ^2<boping>, the draft spelling of /ˈbɒŋ/ . From a phonemic point of view, all of the letters are the default spelling units for the respective phonemes: /b/ ≡ <b>, /ɔ/ ≡ <o>, /p/ ≡ <p>, /i/ ≡ <i>, /ŋ/ ≡ <ng>. The problem arises with syllabification, because we know that /p/ is ambisyllabic (it occurs after a stressed lax vowel /ɔ/). If we apply the syllable boundary principle, and assign stress to the first syllable, then ^2<(bo.ing)> syllabifies thus. This leaves the first syllable open, incorrectly indicating a long, or binary p-vowel. If <p> closes the vowel, giving < (bop.ing) > , then we break the syllable-boundary rule. One interpretation is that this suggests two feet, < (bop) ’ing > , with the primary stress possibly on the latter foot, a situation that arises when we need to emphasise the suffix, in order to make a contrast. For example, someone might say “I was doing the interview-ing, not being the interviewee”. Hyphens are a
useful spelling device for showing this stress pattern, but clearly it does not represent the trochee we are looking for here. Figure 5.19 shows that the problem begins at the syllable boundary.

The sections marked in red show the knock-on effect of a problem occurring at the segmental level. The coda is missing, which affects the rhyme, which affects the syllable, which affects the whole word. Thus the job of PhM amendment is a very important balancing act. It is not permissible to make changes to the complex pleremic units <bop> or <ing>, so there are only two potential candidates for the spelling of /bɒpɪŋ/: ?<boping> and ?<bopping>, with CLD. We can therefore put forward a second draft spelling ?<bopping> for comparison against the known phonological form. This time, everything matches up neatly, as can be seen in Figure 5.20. The extra <p> marks the ambisyllabic consonant /p/. This in turn closes the first g-syllable, confirms the value of <o> (as /ɒ/ rather than /əʊ/) and makes a stressed, full g-syllable <bop>. This leaves the second g-syllable as <ping>, just like the reduced g-syllable in jumping. Hence the spelling syllabifies as <bop.ping>. We can summarise the formation of <bopping> as follows:

- Add the spelling of the input morphemes: <bop> + <ing>
- Construct a G-tree for the draft spelling: ?<boping>
- Compare it against the P-tree of /bɒpɪŋ/, which we know.
- Identify any problems.
- Amend if possible.
- Construct a G-tree for the new draft spelling: ?<bopping>
- Compare it against the P-tree of /bɒpɪŋ/, and accept the new draft spelling, if it is preferable.
- Otherwise, use the draft spelling by default
Figure 5.18

G ↔ P trees for the word jumping.
Figure 5.19

G ↔ P trees for the draft spelling "< (bo.ping) > ≡ /bɒŋ/"
Figure 5.20

G ↔ P trees showing the bi-directional mapping < (bop.ping) > ≡ '/bɒpɪŋ/'
Figure 5.21

G ↔ P trees showing the mappings between

/ˈdɔupəŋ/ and the draft spelling ?<do.peing>.

We can apply a similar analysis to the word /ˈdɔupəŋ/. The draft spelling, formed by applying IdP, is ?<dopeing>. The syllable-boundary rule requires it to be syllabified as <do.peing>. This breakdown makes for a good first syllable:
<do> represents the stressed open syllable /ðəʊ/. The problems are thus concentrated in the second syllable ²<peing>. Clearly, the <ei> sequence is not an ideal spelling of /ɪ/. Furthermore, <peing> is too heavy for an unstressed g-syllable, as it has a digraph <ei>. This suggests a branching /VV/ nucleus, making a /VVC/ rhyme. The problems are shown in red in Figure 5.21. This spelling could attract stress and make it appear to be iambic, as in <do(peing)>, or a spondee <ˈ(do)(peing)>. The spondee structure occurs in protein, whose foot structure may be analysed as <ˈ(pro)(tein)>. The word detail (noun) has variable stress, being either a spondee <ˈ(de)(tail) > or an iamb < de(tail) > (see OED 51168). Again, the problems are marked in red.

The only possible amendment to ²<do(peing) > is to drop <e>, and this gives us a second draft spelling ²<doping> to be compared against the known phonological form /ˈdəʊpɪŋ/. Once again, <ping> is a valid reduced syllable and we know already that <do> is a good first syllable. Hence there exist no mismatches at any level and <ˈ(do.ping) > is a good fit for /ˈdəʊpɪŋ/, as shown in Figure 5.22. Therefore the amended spelling <doping> is retained as the word’s spelling. We can summarise its formation as follows:

- Add the spelling of the input morphemes: <dope> + <ing>
- Construct a G-tree for the draft spelling: ?<dopeing>
- Compare it against the P-tree of /ˈdəʊpɪŋ/, which we know.
- Identify any problems.
- Amend if possible.
- Construct a G-tree for the new draft spelling: ?<doping>
- Compare it against the P-tree of /ˈdəʊpɪŋ/, and accept the new draft spelling.
Figure 5.22

G ↔ P trees showing the mappings between /ˈdɔpɪŋ/ and the draft spelling ?< (do.peing) >
Polymorphemic dactyls

The absence of CLD after trochaic bases such as *edit*, *rabbit* and *open*, is explicable from the structure of graphematic dactyls, which are spelt *< (C)V(C).CV.CV(C) >*. Regular dactyls are not spelt *< (e.or.ting) >*,

Figure 5.23

G ↔ P trees showing the bi-directional mapping *< (do.ping) >* = */ðəʊpɪŋ/
*< (rab.bit.ting) > or *< (o.pen.ning) > because the heavy penultimate syllables suggest penultimate stress. These spellings would then be analysed as: *< e(dit.ting) >, *< rab(bit.ting) > and *< o(pen.ning) >, giving them the same structure as *< e(mit.ting) >, with an unparsed first syllable, followed by a trochee. It is only when consonant quality is at stake that amendments might be made in dactyls. This always seems to occur for /k/, as in mimicking, trafficking but variably elsewhere, as in focus(s)ing and catalog(u)ing. It is also worth re-assessing whether vowel quality is actually ambiguous in editing, rabbiting, opening. In general, dactyls don’t have binary p-vowels in the penultimate p-syllable, as can be seen from the quote from Wells in Section 5e.1.3.

5d.2 {3PS s} and {Plural s} forms

Simple {s} forms such as cats and dogs are easily explained within long-standing theories of English spelling. The morphophoneme //s// surfaces as a context-dependent phoneme /s/ or /z/. Voicing assimilation is one of those ‘automatic and predictable’ phonological processes that English spelling can afford to ignore (C. Chomsky 1970: 292). This example is a reminder that there are fewer difficulties in explaining the mapping of phonological monosyllables onto graphematic monosyllables, and our focus is on polysyllabic forms and mismatches in the syllable-level correspondences.

When {3PS s} and {Plural s} cause an unstressed phonological syllable to be added, there are two spelling possibilities. The simpler case to explain is when the inert final <e> of bridge, tense, douche, etc. becomes activated in the {s} forms, as in < (brid.ges) >, < (ten.ses) > and < (dou.ches) >. Thus the mute <Ce#> syllable becomes <Ces> and this maps onto a /Cɪz/ syllable.

In the case of kisses, washes, fixes, an <e> is added, and this seems to be because the peak of a g-syllable must be spelt with a vowel letter (Evertz 2014: 61). Otherwise, the phonological syllable would have to be decoded from spellings such as *<kiss.s>, *<fix.s>, *<watch.s>, etc. The resulting (correct) spellings can be analysed as trochees: < (kis.ses) >, < (wa.shes) >, < (fi.xes) >; or dactyls < (fi.nishes) >, < (ma.na.ges) >, etc. I have indicated the syllable boundary in washes, finishes and fixes with a vertical line below, to indicate that the units <sh> and <x> are relevant to both syllables, and thus equivalent to CLD. The metrical notation introduced in this chapter breaks down in washes and finishes. The syllable boundary in < (was.hes) > is not acceptable because <sh> cannot be broken into its component parts, so that neither <s> nor <h> can indicate /ʃ/ (see Evertz (2014: 213); also Section 2c.2 on Carney’s distinction between endocentric and exocentric spelling units). However < (wa.shes) > and < (fi.nishes) > suggest
the open syllables <wa#> and <ni#>, respectively. This is different form < (bud.ges) > and < (whin.ges) > as <g> ⇔ /dʒ/.

5d.3 {ed} forms

Base forms ending in /t/ or /d/ cause no trouble, so < (bat.ted) > and < (pad.ded) > both have appropriate correspondences at every level. The same goes for base forms with open syllables: < (ra.ted) > and < (wa.ded) >; digraphs, < (wai.ted) > and < (nee.ded) >; trochaic bases < (rab.bit.ed) > and < (vi.si.ted) >, and combinations of feet: < '(cal.cu)(la.ted) > and < '(bar.ri)(ca.ded) >, etc.

Problems arise when there are two g-syllables but just one p-syllable, even after PhM amendments have been made to the draft spelling. This is the basic problem with the correspondences < (do.ped) > ⇔ /dəʊpt/ and <(bopped)> for /bɒpt/. We need a more rigorous way of explaining why these spellings are chosen over all others, despite their mapping mismatches. The draft spelling of /dəʊpt/ is ?<dopeed> and this is a bad spelling for two reasons. The first reason, pointed out by Carney (1994: 129), is that <eed> might well be read like the <eed> of need or agreed. The second, related reason is that the second syllable bears too much graphematic weight (Evertz and Primus 2013; Evertz 2014: 132–189; Berg 2016: 458). The spelling ?<dopeed> might therefore be analysed as the iamb *< do.(peed) > (compare the verb proceed /ˈprəʊsi.d/), or the spondee *< '(do).(peed) > (compare the noun proceed /ˈprəʊsi.d/). By contrast, ?<doped>, the amended spelling (of /dəʊpit/) does at least have the advantage of suggesting a trochaic reading < (do.ped) >, even though novice readers might map this onto /ˈdəʊpid/ or /ˈdəʊped/ (compare stupid and Lopez). Significantly, the phonological jump between /ˈdəʊpid/ and /ˈdəʊpt/ is not that great, since phonological changes so often take place within an individual foot, especially to unstressed p-syllables in trochees and dactyls. Compare how camera has changed from being a dactyl /ˈkæmərə/ to a trochee /ˈkæmərə/, at least for many speakers (OED 26688). Figure 5.24 shows how cumbersome the mappings are between /dəʊpt/ and <doped>, even though this is the spelling in use.
A very similar analysis can be applied to *bopped*. Its draft spelling /ˈbəʊpɪd/, a trochee with the wrong vowel sound, or /ˈbəʊˌped/, a spondee like the noun *moped*. The amended spelling /ˈbɒpɪd/ (compare *solid*), which is also not perfect.
The problem with the visual model presented above is that it can only compare a good spelling against a bad spelling. It cannot show clearly why one bad spelling is better than an even worse spelling. For that job, we need Optimality Theory, a formal model designed to help choose between the better of two evils. It will thus be helpful in subdividing PhM so that we can make analyses based on much more specific criteria.

5d.4 Applying OT to the analysis of spelling formation

Optimality Theory (OT) is a constraint-based linguistic theory that will be applied here to explain why one potential spelling candidate is chosen over another. In the visual model presented above, we saw why one good candidate spelling, such as ¿<doping> was chosen over ¿<dopeing>, but OT can be used to explain why a less than perfect spelling such as ¿<doped> will be chosen over an even worse one, such as ¿<dopeed>. In this way, we can apply rigorous linguistic criteria to explain English spellings which, up to now, could only be explained by appealing to the reader’s intuitive understanding of the writing system.

5d.4.1 Optimality theory in phonology

OT was first employed in phonological theory (Prince & Smolensky 1993, McCarthy & Prince 1993) and it grew out of rule-based phonological theory. Thus it retains some aspects of that theoretical approach. For example, OT phonology assumes that there exists an underlying form and the surface, or output, form is rendered from this. Unlike rule-based phonology, however, OT evaluates the output in a single move.

There are several important components to the evaluation. First, there needs to be an input form. In an OT analysis of phonological words, the input will be taken from the lexicon, which contains the lexical representations, or abstract underlying forms, of morphemes. The generator, or GEN, generates output candidates for an input. In theory, there can be an infinite number of output candidates for any input. In practice, the experienced researcher can whittle these down to a few plausible candidates. The choice of constraints usually eliminates all but the most likely candidates. Constraints fall into two basic categories, faithfulness constraints, which require the input and output to be identical, and markedness constraints, which require the output to be well-formed, phonologically or otherwise. Frequently, constraints are in conflict with one another so they must be ranked in order to explain the output forms. For example, take the underlying form /bildin/, ‘building’. Faithfulness constraints would
require the output to be [bildin], but markedness constraints might require all syllables to have a CV form, removing consonant clusters, and rendering something like [bi:di:]. In most varieties of English, the relevant faithfulness constraint will dominate this particular markedness constraint, so the output candidate [bildin] will be chosen above [bi:di:]. OT assumes that all constraints occur in all languages, but the rankings differ, resulting in multifarious outputs. The evaluator, or eval, is the set of ranked constraints which evaluates the output candidates. Most importantly, each constraint is unaffected by the lower-ranked constraints and these cannot have a cumulative effect upon one another.

Several examples of OT evaluations are presented below, but first, it may be worth noting that the game of poker is based on similar evaluations. If someone has a royal flush, they always win. If no one has a royal flush, we check to see if anyone has a straight flush, then four of a kind, a full house, and so on through the ranking. This is why a pair of twos will always trump a single ace and king. However, if two players have a pair of twos, then a player who also has an ace will beat a player with a king.

5d.4.2 Applying OT to spelling formation

OT has been applied to several branches of linguistics and there is a growing body of work which applies OT to writing systems. Much of the work has been on the German writing system (Geilfus-Wolfgang 2007; Primus 2004; Wiese 2004), a little on Korean (Song & Wiese 2010) and there have been two recent works on English (Baroni 2013; Evertz 2014). Baroni’s (2013) paper argues that OT can be used to explain eye dialect, a type of non-standard spelling (which is discussed in Section 7c.1). Using OT constraints to describe mappings from phonemes to graphemes, he formalises the reasons why, for example, <luv> can replace <love>, at least in certain non-standard spelling domains where the spelling can be based entirely on the pronunciation. Evertz’s (2014) full-length study shows how OT can be applied at all levels of the graphematic hierarchy, an approach that is of great use in the present study. His theory allows us to compare graphemes against phonemes, g-syllables against p-syllables and g-feet against p-feet.

In order to introduce how OT can be applied to English spelling formation, we might expect an example from a previous study, but Baroni’s analysis is only helpful in cases where a word’s spelling is based solely on its pronunciation, while Evertz’s complex analysis requires a large amount of background (and this is presented in Section 5d.4.3 below). We shall therefore look at the example of <jumping>, a spelling which has already been accounted for satisfactorily in the
algorithm presented in Chapter 4. This will allow us to examine the components of an OT evaluation without the risk of further problems arising.

Firstly, we need an input form, some output candidates generated from it and constraints which can be used to evaluate the optimal candidate. The input could be conceptualised morphologically or phonologically. This depends on what we assume a spelling is trying to represent. If phonological, then we might consider /jʌmpɪŋ/ to be the input form and then posit various candidates for its spelling, such as <jumping>, <jumpping>, <jumpyng>, and <djumping>. If morphological, then we can consider {jump+ing} to be the input form, and again <jumping> is a likely candidate, based on the spelling of the input morphemes. It is not fully clear, from this example, which linguistic forms are to be represented, so both forms will be included in the input box on the top left corner of the table (e.g. Table 5.2). For ease of reference, the table captions only show morphological notation.

Secondly, we can treat IdP and PhM as OT constraints for now, defined as follows:

**IdP**: The spelling of a polymorphemic word is formed by concatenating the spelling of the input morphemes.

**PhM**: A spelling must match the phonological form at every level of analysis in the graphematic and phonological hierarchies.

It should be noted that IdP is a very specific constraint: \(<\text{BAC}> + \text{<DEF>} \rightarrow \text{<BACDEF>}.\) By contrast, PhM requires the alignment of several graphematic and phonological levels. For now, the reader is asked to rely on his or her own knowledge of the writing system in order to determine whether a candidate spelling obeys PhM or not. We shall return to this theoretical problem at the end of the current subsection. In order to mimic the order of the algorithm developed in Chapter 4, it will be assumed that IdP is the higher-ranked constraint. This means that any candidate spelling which does not obey IdP is automatically disqualified.

In Table 5.2 Candidate a., <jumping>, is chosen because it does not violate IdP, the highest-ranked constraint. It doesn’t matter if other candidates obey PhM they are disqualified at the first step of the evaluation. The successful candidate is indicated with the symbol ‘☞’. An asterisk marks a violation of a constraint, while an exclamation mark indicates that it is a fatal violation.
Let’s look now at \{jump+ed\}, pronounced /jʌmpt/ in most accents. Two likely candidates are <jumped> and <jumpt>, which respectively match the morphological and phonological forms. We can add <jumpd> and <djumpt> as extra test cases.

According to Table 5.3, the spelling <jumped> is chosen because it is the only candidate that satisfies IdP. The rest are all eliminated because they violate the highest-ranked constraint. It doesn’t matter that <jumped> is not a perfect phonographic match. It’s the best candidate available and is thus chosen.
This analysis becomes problematic when the spelling in use results from PhM amendment, such as the CLD of <bopped>, <bopping> and the <e>-dropping of in <doped>, <doping>. All of these spellings seem to adhere to PhM, but not IdP. Let’s take {dope+ed}, for example. The IdP draft spelling of this word is ?<dopeed>, formed from concatenating <dope> and <ed>. ?<dopeed> obeys IdP but not PhM (although this assumption remains based on our intuitive understanding of the writing system) and, according to the analysis in Table 5.4, ?<dopeed> is a better candidate than ?<doped> which incurs the fatal violation of not adhering to IdP, the highest-ranked constraint. However, we know that <dopeed> is not the spelling in use, so the output must derive from something other than the morphological input {dope+ed}. And yet the spelling ?<doapt> is excluded because it does not adhere to IdP. This means that the phonological form /dəʊpt/ cannot be the sole input. Therefore, the spelling in use, <doped>, must result from some combination of inputs, representing the morphological and phonological forms within a compromised spelling.

From the viewpoint of OT, there are two reasons why this analysis does not work. Either the constraints have been ranked in the wrong order or they may not have been defined properly. If we change the order of the constraints, as shown in Table 5.5, then ?<doapt> becomes the best candidate for {dope+ed} because it doesn’t violate PhM. A similar problem arises for {bop+ed}, shown in Table 5.6, where ?<bopt> fares better than ?<boped> and ?<bopped>. Since the first two of these violate PhM, they are eliminated and <bopt> is the chosen candidate. However, we know that this is not the normal spelling of the word, so it seems that the re-ordering of the constraints has not helped to explain why <doped> and <bopped> are the spellings in use.

<table>
<thead>
<tr>
<th>{dope+ed}</th>
<th>PhM</th>
<th>IdP</th>
</tr>
</thead>
<tbody>
<tr>
<td>/dəʊpt/</td>
<td></td>
<td></td>
</tr>
<tr>
<td>a. &lt;dopeed&gt;</td>
<td>*!</td>
<td></td>
</tr>
<tr>
<td>b. &lt;doped&gt;</td>
<td>*!</td>
<td></td>
</tr>
<tr>
<td>c. ⇨&lt;doapt&gt;</td>
<td></td>
<td>*</td>
</tr>
</tbody>
</table>

Table 5.5
Second OT analysis for the spelling of {dope+ed}
The next solution is to redefine the constraints but keep their original order. There is a strong argument for considering CLD to be a special case of IdP (c.f. Berg 2016 for the same argument with different terminology). After all, CLD involves adding an identical consonant to the base form, and thus identity preservation. Words ending in <p> have another <p> added; words ending in <d> have another <d> added; and so on. Hence CLD might be redefined as IdP-extension. Similarly, <e>-dropping might be considered a case of IdP-reduction. The base CPU is shortened. Its internal composition is not disrupted, and the bulk of the spelling is maintained.

**Complex-IdP:** The spelling of a polymorphemic word is formed by concatenating the spelling of the input morphemes, but letters may be reduplicated (e.g. CLD) or dropped (e.g. <e>-dropping).

This redefinition of the constraint is not simply an expedient way of fitting the theory to the data. In fact, IdP-reduction and IdP-extension occur in several other kinds of English spellings. For example, IdP-reduction occurs in the formation of clippings (Section 8g), neo-classical compounds (Section 8d) and blends (Section 8e). The word *mic* is reduced from *microphone*, while the letter-sequence <psy(ch)(o)> is retained in *psychology*, *psychedelic*, *psy-trance*, <psybient> and *PSYPRESS*. The clearest case of IdP-extension occurs in what Androutsopoulos (2000) calls prosodic respellings (discussed briefly in Section 7c). In such cases, the word *hey* might be extended to *heeeeeeey*, in order to convey some prosodic information, while *no way* might become *no waaaaaay*. Clearly the iterations of <e> and <a> are motivated by the spelling of the base CPU, not by the need to obey PhM. This is why *hey* and *way* do not become either *<haaaaay> or *<weeeeey>*.

We can now solve the problems arising above regarding the spelling of \{dope+ed\} and \{bop+ed\}. In the first case, shown in Table 5.7,
?<doped> can both be considered as spellings which adhere to Complex-IdP.

Table 5.7
Third OT analysis for the spelling of {dope+ed}

<table>
<thead>
<tr>
<th></th>
<th>Complex IdP</th>
<th>PhM</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. &lt;dopeed&gt;</td>
<td>*</td>
<td>?</td>
</tr>
<tr>
<td>b. (&lt;b&gt;) &lt;doped&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. &lt;doapt&gt;</td>
<td>*!</td>
<td></td>
</tr>
</tbody>
</table>

Table 5.8
Second OT analysis for the spelling of {bop+ed}

<table>
<thead>
<tr>
<th></th>
<th>Complex IdP</th>
<th>PhM</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. &lt;boped&gt;</td>
<td>*</td>
<td>?</td>
</tr>
<tr>
<td>b. (&lt;b&gt;) &lt;bopped&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. &lt;bopt&gt;</td>
<td>*!</td>
<td></td>
</tr>
</tbody>
</table>

In the tableaux above, both ?<doapt> and ?<bopt> are eliminated because neither obeys Complex IdP. However ?<dopeed> and ?<doped> do obey it, as do ?<boped> and ?<bopeed>. By redefining IdP, we reduce the list of reasonable output candidates to just two (although other examples of Complex-IdP are theoretically possible, according to the above definition, including <dopppeeeed> and <dopd>). This brings us back to the issue of whether the input form is morphemic {jump+ed} or phonological /jʌmpɪŋ/. It would appear that in the spelling of inflections, the morphemic information provides the primary input. This is ultimately why <bopt> and <doapt> etc. must be excluded as possibilities. In non-standard respellings of existing words, such as the eye dialect studied by Baroni, it would appear that the input form is phonological. And yet the existence of PhM amendments, under certain circumstances, means that the output form must be an effort to represent both kinds of linguistic information.

Now that we have excluded spellings derived by mapping solely from phonology to spelling, the next challenge is to show why the compromise forms ?<doped> and ?<bopped> are chosen over the simple IdP forms?<dopeed> and?<boped>. Clearly, none of these are perfect phonographic matches (compare the bisyllabic words moped and dogged), so we need to find a better way of assessing how well a spelling obeys PhM.
This task is in fact rather complicated because PhM is not a single constraint, but rather a bundle of constraints. In order to fully obey PhM, a spelling must match the phonological form at every level of analysis. A full-length study would be required in order to do justice to the enormous number of grapheme-to-phoneme mappings in existence, at all levels of the hierarchy, as well as the number of OT constraints required to explain them. However, it will be assumed that readers have a strong grasp of sub-syllabic mappings between spelling and sound, of the kind enumerated by Venezky (1970), Cummings (1988), Carney (1994), Rollings (2013) and Brown (2015), but that there remains a need to identify syllabic and foot-level mappings. We therefore need a more powerful theoretical model which can provide more concrete evidence for the degree to which a spelling is a good phonographic match for its phonological form at each level of analysis.

The focus in the next section will be on how phonological feet can be predicted from the spelling, building on Evertz’s (2014) analysis of the graphematic syllable and graphematic foot. We can then determine which syllables are stressed and which are unstressed. Hence, if we know that a spelling such as <vended> maps onto a phonological trochee, then we know that the first <e> maps onto a stressed vowel and the second maps onto a reduced vowel. By contrast, if we know that <ee> tends to occur in stressed syllables, then we can identify the fact that <veneer> and <agreed> both have final stress.

5.d.4.3 Evertz’s OT model of graphematic feet
There are two distinct parts to Evertz’s (2014) OT analysis of English spelling. The first part involves determining the optimal g-foot structure for any given polysyllabic English spelling and the second part involves mapping from the spelling to phonology (decoding), as well as the converse (coding). The major advantage of his method is that the assignment of g-foot structure provides very strong clues to p-foot structure, and thus it tells us whether any given g-syllable ought to be decoded as a stressed or reduced syllable. In this way, the tricky task of decoding polysyllabic words can be reduced to the simpler task of decoding individual syllables in turn. That final step has been the primary concern of most studies of English spelling, where the focus has been on grapheme-to-phoneme mappings and, to a lesser extent, on mappings between syllabic onsets and rhymes. Understanding foot structure therefore allows us to account for the innumerable ways in which reduced syllables are represented in English spelling.

A bi-product of Evertz’s analysis is that it will allow us to decide between spelling candidates such as ²<doped> and ²<dopeed>, or ²<bopped> and
Determining graphematic foot structure from the spelling

Evertz (2014: 169–180) provides a method for determining the most likely graphematic-foot structure to be decoded from a spelling, and this requires the introduction of a further seven constraints which explain graphematic well-formedness. We shall look first at the five higher-ranked constraints, and these are exemplified in Table 5.9. The other two, lower-ranked, constraints are then discussed, and all seven constraints feature in the evaluation presented in Table 5.10.

The highest-ranked constraint is called NonHead(RGS), which is defined as: ‘reduced graphematic syllables cannot be heads of graphematic feet’. In other words, weak graphematic syllables such as the <ter> of <lat.ter> and <la.ter> cannot be stressed, and similarly the <ped> of <do.ped>, <bo.ped> and <bop.ped> cannot be stressed. Thus, these spellings cannot be analysed as <do(ped)>, <bo(ped)> or <bop(ped)>.

The noun moped is an exception, as are refer and hotel. The next four constraints are equally ranked. Trochee states that ‘graphematic feet are head-initial’. Minimality states that ‘feet are mono- or bisyllabic’, and there are further details attached to this, which are not relevant to the present set of examples. Since we are dealing with bi-syllabic words, this constraint has no bearing on the outcome. The Weight-to-Head-principle (WHP) states that ‘heavy syllables are heads of feet’, meaning that they are stressed syllables. This constraint is decisive because it requires the <peed> of <dopeed> to be a stressed syllable, since there are three letters in the syllable rhyme, <eed>. Thus we would expect <dopeed> to be read like <a(greed)> or the verb <pro(ceed)>. The fourth constraint is called Align-gFoot-Right which means that every graphematic word should end with a foot. Again, this is not decisive because we are dealing with bi-syllabic words, and thus it is possible for a word to begin with, and end, in a trochee.

Table 5.9 shows the evaluation of the possible p-foot structures of the spelling <dopeed>, based on the five highest-ranked constraints identified by Evertz. Assuming that syllables can be either stressed or not, this provides four obvious candidates (i.e. stressed-stressed, stressed-unstressed, unstressed-stressed, unstressed-unstressed; or spondee, trochee, iamb, impossible). This also assumes...
that the syllable-boundary rule applies (see Section 5b.1.1 above). The latter four constraints are equally ranked, and thus presented with broken lines.

<table>
<thead>
<tr>
<th>&lt;dopeed&gt;</th>
<th>NonHead (RGS)</th>
<th>Trochee</th>
<th>Minimality</th>
<th>WHP</th>
<th>Align gfoot right</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. (do)(peed)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. (do.peed)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>*!</td>
</tr>
<tr>
<td>c. do(peed)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>d. do.peed</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>*!</td>
</tr>
</tbody>
</table>

Table 5.9
Partial OT-analysis for the foot-structure of <dopeed>

The first constraint NonHead(RGS) doesn’t apply to any of the candidates, as the final syllable is not <Ce(C)>. Of the other constraints, the only one that is broken is the Weight-to-Head Principle: Candidate b., ²< (do.peed) >, contains a g-trochee with a graphematically heavy, unstressed, second g-syllable, and this is a fatal violation. The word < (certain) > also violates this constraint, because we see <VV> in an unstressed syllable. Compare < per(tain) >. Candidate d., ²< do.peed >, also breaks WHP, with the only difference being that both syllables are unstressed, a situation that doesn’t occur in the citation form of English words. The other candidates, ²< (do)(peed) > and ²< do(peed) >, both appear to be possible foot structures of <dopeed>, and this is not surprising, given the two pronunciations of proceed. The noun is < (pro)(ceed) > and the verb is < pro(ceed) >, i.e. /ˈprɔʊsiːd/ and /prəˈsiːd/.

The two remaining constraints will help us to predict the more likely foot structure. Align-gFoot-Left means that every graphematic word should start with a g-foot. In English (and German, notes Evertz 2014: 179), this constraint is lower-ranked than Align-gFoot-Right. However, bi-syllabic words automatically end in a foot (whether the foot is mono- or bi-syllabic), so the tension between the two candidates only plays out in tri-syllabic words or longer. Hence, the final trochee of < Oc(to.ber) > is more likely than the dactyl < (op.ti.mal) >, which begins with a foot but doesn’t end in one (at least according to Evertz’s Minimality constraint (discussed above), which requires feet to have just one or two syllables). Some tri-syllabic words begin with a foot and end in a foot, as in < (kan.ga)’(roo) >. The final constraint is Parse-σ, which states that g-syllables are parsed by g-feet. Note how all three g-syllables of <kangaroo> fit neatly into well-formed graphematic feet. By contrast, the word aroma violates this
constraint, because its first syllable, <a>, can attach to a preceding foot. Parse-σ and Align-gFoot-Left are equally ranked (shown with broken lines) but dominated by all of the other five constraints (shown with an unbroken line). Table 5.11 completes the analysis begun in Table 5.10.

<table>
<thead>
<tr>
<th>&lt;dopeed&gt;</th>
<th>Non head RGS</th>
<th>Trochee</th>
<th>Mini Mality</th>
<th>WHP</th>
<th>Align gfoot right</th>
<th>Align gfoot left</th>
<th>Parse σ</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. ☞ (d)o(peed)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| b. (do.eed) | | | | | | | *
| c. do(peed) | | | | | | * | *
| d. do.eed | | | | | * | | *

Table 5.10
Full OT-analysis for the foot-structure of <dopeed>

This time we see that Candidate c., <do(peed)>, does not begin with a foot, nor are all of its syllables parsed by feet. Thus it is a less likely reading of the input spelling <dopeed>.

<table>
<thead>
<tr>
<th>&lt;doped&gt;</th>
<th>Non Head RGS</th>
<th>Trochee</th>
<th>Mini Mality</th>
<th>WHP</th>
<th>Align gfoot right</th>
<th>Align gfoot left</th>
<th>Parse σ</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. (d)o(ped)</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>*</td>
<td></td>
</tr>
</tbody>
</table>
| b. ☞ (do.ped) | | | | | | | *
| c. do(ped) | * | | | | | * | *
| d. do.ped | | | | | | * | *

Table 5.11
Full OT-analysis for the foot-structure of <doped>

We can now determine the likely foot structure of the input spelling <doped>. Table 5.12 shows the following. Candidate a., <(d)o(ped)>, whose foot structure is like the noun moped, breaks the highest-ranked constraint, NonHeadRGS, which means that reduced graphematic syllables are not heads of feet. Candidate c., <do(ped)>, has the same problem and it also violates the final two constraints (although that is redundant, and the latter asterisks do not need to
be marked in, since the candidate has been ruled out by for its fatal violation of NonHeadRGS). Words with this prosodic shape (iambs) frequently occur in §Latinate words with fossilised suffixes (e.g. refer, serene, supply), but this shape does not appear in English inflections. Candidate d. is not a possible English word as it has no stressed syllables (again, redundant violations are marked in for illustrative purposes). Candidate b. is not ideal, but it is clearly the best available option. We might expect < (do.ped) > /dəʊpɪd/, much like < (stu.pid) > /st(j)u:pid/, but of course this reduced vowel was dropped from speech long ago (Lass 1992), leaving us with spellings that indicate morphemic information more faithfully than the phonological form. Incidentally, the reduced vowel is retained in adverbs such as supposedly, markedly, assuredly, etc.

![Table 5.12](image)

Table 5.12
Full OT-analysis for the g-foot structure of <boped>

We can now apply the same OT analysis to the possible spellings ?<boped> and ?<bopped>. Unsurprisingly, ?<boped> undergoes the exact same analysis as ?<doped>, but with <b> replacing <d>, a change that does not affect the foot structure. Assigning foot structure to ?<bopped> is also quite similar to these two, as the candidates all violate the exact same set of constraints (see Table 5.14). Hence the model predicts that the spelling <bopped> will be read as the trochee < (bop.ped) >, exemplified by Candidate b.
At this point, we have assigned foot structures to the candidates in question, and the evaluations show that `<dopeed>` suggests the spondee, `<dopeed>`, while `<bopped>` suggests the trochee `<bopped>`. Similarly, we have seen that the amended spelling `<bopped>` suggests the trochee `<bopped>`, with a short or unary vowel, while the IdP draft spelling `<bo.ped>` suggests `<bo.ped>`, with a long or binary vowel.

Armed with this information, we can soon compare each pair of these spelling-cum-foot structures against the known phonological forms /dəʊpt/ and /bɒpt/, respectively and, in doing so, we can determine which spelling is the better phonographic match. But first, we must examine how the decoding process is theorised within Evertz’s framework.

**Coding and decoding in OT**

Evertz (2014: 181–207) enumerates six mapping constraints which work for both decoding and coding, and one extra constraint for the latter, although this will be dispensed of in the present analysis. The first constraint, **Head-Match**, is a faithfulness constraint which requires the head of a p-foot (i.e. the stressed syllable) to map onto the head of the g-foot, and vice-versa. For example, the spelling `<a.greed>` is a suitable spelling for its phonological form because the head g-syllable `<greed>` suggests a stressed p-syllable, whereas `<a>` suggests a reduced syllable. The constraint **Unary-Binary** is Evertz’s term to explain the fact that we can read the difference between `<pl.ning>` and `<pla.ning>`. In the first word, the single g-vowel in the closed g-syllable `<pl.>` represents a short, or unary, p-vowel; whereas in an open syllable, `<pla.>`, it represents a long, or binary, p-vowel. While this may appear to be two constraints, one implies the other, so Evertz combines them (2014: 182). In our examples, we can expect the
<bop> of ?<bop.ping> to have a short phonological vowel but the <bo> of ?<bo.ping> should have a long p-vowel. The constraint Geminate explains the need, or otherwise, for double consonants in g-trochees. While Unary-Binary can function in monosyllabic words (for example, the binary p-vowel of the open monosyllable in the word me versus the unary p-vowel in met), Geminate only applies to bi-syllabic feet and it is useful for explaining the encoding and decoding of ambisyllabic p-consonants, which tend to be spelt with double consonant letters. This constraint applies to both CLD and historical doubles (a distinction that Evertz does not explicitly make). In the coding direction, Evertz adds an extra constraint Geminate\textsuperscript{can} which helps to explain the presence of <mm> in canonical trochees (e.g. hammer) but also its frequent absence in non-canonical trochees (e.g. lemon, comic). However, this distinction has not been included in the present work in light of Berg’s (2016) observations that the spellings of final syllables in monomorphemes often provide clues about the grammatical and phonological structure of the word (Section 5c.2.4 above). A side issue is that removing the constraint Geminate\textsuperscript{can} removes the need to have a different set of rules for both coding and decoding.

The next pair of constraints are faithfulness constraints. Max says that there should be no deletion going from input to output, while Dep says that there should be no insertion. The third pair of constraints are well-formedness constraints, one working in either direction. GSyll-Wellformedness requires graphematic syllables to be well-formed, and this applies only to encoding. That is to say, when we form a spelling, it must look like a normal English spelling. However, GSyll-Wellformedness is in fact a bundle of constraints and Evertz (2014: Ch. 3) devotes an entire chapter to its explication. The details are too great for the present discussion but suffice it to say that most problems would appear self-evident to the experienced literate, and I will point out any details which deviate from the obvious. PSyll-Wellformedness is the converse constraint which is applicable to decoding, meaning that ill-formed phonological words cannot be constructed from the spelling. We discussed Giegerich’s definition of well-formed p-syllables in Section 5b.1 above and the definitions from there will suffice, even though Evertz’s (2014) definitions may vary on certain details (e.g. whether coda is part of the syllable nucleus or not).

Table 5.14 shows the evaluation of six spelling candidates for encoding the p-word /lætʃ/, and the tableau is taken from Evertz (2014: 187). The choice of candidates is very different from the conditions we met in Tables 5.9-5.13, where the input was a spelling and the outputs were possible g-foot structures. In Table 5.14, Evertz’s analysis omits a discussion of how g-foot structure is assigned, ‘For
the sake of simplicity’ (2014: 187), and all of the candidates, are assumed to represent trochees. He therefore includes three distinct spellings: ?<later>, ?<latter> and ?<latere>, and these are split into some possible g-syllable patterns, including ?<la.ter> and ?<lat.er>, as well as ?<lat.ter>, ?<lat.ter.e> and ?<la.ter.e>. Only ?<la.te.re> is posited for that letter sequence but ?<lat.e.re> or ?<la.ter.e> etc. might well have been included as output candidates. Evertz does this because the p-foot structure is included in the input form /lætə/, a situation that makes sense when creating spellings from phonological forms (i.e. coding).

Candidate a. is the chosen candidate because it violates a low-ranked constraint, Dep, which means that there is insertion between the input form, with a single phoneme /t/, and the output form, with a double consonant <tt>. Candidate b. is a not a well-formed graphematic word because <tt> is not a well-formed g-syllable onset. This candidate violates several other constraints although, strictly speaking, it is not necessary to demonstrate this. Candidates c. and e. do not adequately indicate phonological ambisyllabicity, and thus they violate GeminateCanon, among other constraints. Candidate d. violates Unary-Binary is because the g-syllable is open <la>, while the input calls for a closed syllable and short vowel, /læ/, which ought to be spelt <lat> or even <latt> (as in Candidates a. and c.). Candidate f. shows a mismatch between the number of phonological syllables and graphematic syllables, and thus it violates HeadMatch.

<table>
<thead>
<tr>
<th>/lætə/</th>
<th>Head Match</th>
<th>GWell Form</th>
<th>Un-Bin</th>
<th>Gem-Canon</th>
<th>Max</th>
<th>Dep</th>
<th>Gem</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. lat.ter</td>
<td></td>
<td></td>
<td></td>
<td>*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. latter</td>
<td></td>
<td>*</td>
<td></td>
<td>*</td>
<td>*</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>c. latt.ter</td>
<td></td>
<td></td>
<td></td>
<td>*</td>
<td></td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>d. la.ter</td>
<td></td>
<td>*</td>
<td></td>
<td></td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>e. lat.ter</td>
<td></td>
<td></td>
<td></td>
<td>*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>f. la.te.re</td>
<td>*</td>
<td></td>
<td></td>
<td>*</td>
<td></td>
<td>*</td>
<td></td>
</tr>
</tbody>
</table>

Table 5.14
OT-analysis for the spelling of /lætə/, taken from Evertz’s (2014: 187)

5.d.4.4 Choosing between candidate spellings
We are now ready to determine which is the optimal spelling for a particular inflected form. Tables 5.15 and 5.16 evaluate the best spelling for {dope+ed} and
{bop+ed} respectively, using the same set of decoding constraints proposed by
Evertz for mapping from phonological forms to spellings (c.f. Table 5.14). One
very important addition is to include Complex-IdP as the highest-ranked
constraint in the evaluation. The purpose is that this immediately weeds out any
other candidates which adhere to PhM. I have included a third candidate in order
to demonstrate its purpose.

<table>
<thead>
<tr>
<th>/dəʊpt/</th>
<th>Complex IdP</th>
<th>Head Match</th>
<th>G Well Form</th>
<th>Un-Bin</th>
<th>Gem-Canon</th>
<th>Max</th>
<th>Dep</th>
<th>Gem</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. (do)(peed)</td>
<td>*!</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. (do.ped)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>c. (doapt)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>*!</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 5.15
Full OT-analysis for the spelling of {dope+ed}

Firstly, Candidate c. does not obey Complex IdP, and is immediately
excluded. The same would apply to any other similar candidate (e.g. < (dowpt) >,
< (dopte) >, < (dohpt) >. The spelling < (do)(peed)> violates HeadMatch because
the single, stressed p-syllable, /dəʊpt/, maps onto two stressed g-syllables, <do>
and <peed>. The spelling < (do.ped) > does not violate HeadMatch because its
second syllable is reduced. Both spellings violate Dep because they ‘insert’ an
<e>, or indeed <ee>, where there is no equivalent p-vowel in the input /dəʊpt/.
Since < (do.ped) > violates a lower-ranked constraint, it is the chosen candidate.

<table>
<thead>
<tr>
<th>/bɒpt/</th>
<th>Complex IdP</th>
<th>Head Match</th>
<th>g- Well Form</th>
<th>Un-Bin</th>
<th>Gem-Canon</th>
<th>Max</th>
<th>Dep</th>
<th>Gem</th>
</tr>
</thead>
</table>
| a. (bop.ped) | | | | | | | | *
| b. (bo.ped) | | | | *! | | | | *
| c. (boapt) | | | | | | *! | | |

Table 5.16
Full OT-analysis for the spelling of {bop+ed}
In Table 5.16, < (bop.ped) > is the chosen candidate. Dep is the only constraint it violates, due to the insertion of a second <p> and an <e>. Candidate b. is eliminated because the open syllable <bo.> indicates a long, or binary, phonological vowel, and thus violates Unary-Binary, as the input form has a unary p-vowel /ɒ/. Candidate c. is of course eliminated for not adhering to Conplex-IdP. The fact that this candidate does not break any other constraints is an indication that it is a good spelling otherwise, and in circumstances where no form of IdP apply, <bopt> would be a good spelling of a monomorphemic word /bnpt/. By contrast, <boppt> would violate Dep, due to the insertion of an unwarranted <p>, and it is thus less likely to be chosen.

The other important point to emerge from this final set of examples is that a phonological input form is only relevant in the latter stages of spelling formation. In the first evaluation, the morphemic information forms the basis of the input forms. In this way, we see how the spelling formation of English inflections appears to be an ordered process. Morphological information is represented first and the role of PhM amendments is to help choose between possible candidate. In dialect spelling or creative spelling, by contrast, morphological information may be ignored and thus the input form is phonological.

5d.4.4 Conclusions
The analysis above shows that it is possible to formalise our intuitions about what makes a good or a bad spelling. The crux of the problem boils down to the fact that phonographic matching (PhM) is an extremely complicated process which involves the harmonisation of spelling and sound at several levels of analysis, and not just the matching of graphemes and phonemes. So while several studies have recognised the importance of the morpheme, as well as the phoneme, it should be clear that we also need to recognise the role of feet, syllables and rhymes, in English spelling. By contrast, IdP is a rather simple concept. In Chapter 4, IdP was defined as the simple concatenation of input spellings. However, in order to explain it within an OT framework, it has been necessary to redefine it as Complex-IdP, whereby the preservation of a spelling’s identity can include either reduplication (e.g. CLD), or the loss of existing letters (<e>-dropping). Notably, such changes are made in response to the need to match letters against phonological structures. Other examples of Complex-IdP may have other prosodic functions (e.g. <heeeeeey>).

Conceptualising PhM as a single OT constraint is of little use because that makes it impossible to decide between a bad spelling and a worse one (see Table
5.8). It was therefore necessary to break PhM up into a bundle of constraints, several of which can be used to assign g-foot structure to a spelling, and several more to determine lower-level mappings. The first of these is a highly innovative approach, introduced by Evertz (2014), while the second part involves using his OT-framework to explain positionally-determined mappings discussed elsewhere in the literature, such as whether a vowel-letter ought to be decoded as a short or long vowel, or whether consonant letters indicate ambisyllabicity (c.f. Rollings 2004). In the analysis presented here, we have not looked at simple letter-to-phoneme mappings (of the kind discussed by Carney (1994), *inter alia*), and it has been assumed the readers can fill in these blanks (e.g. `<b> ≡ /b/`, etc.). Nevertheless, a fuller description of the mapping relations between spelling and sound would require us to introduce constraints to explain all levels of the hierarchy. We might then expect decoding to be reduced to three phases of analysis. First, a spelling is assigned a likely g-foot structure, then the g-syllables can be analysed according to whether they represent stressed or reduced p-syllables, and, finally, mappings between graphemes and phonemes can be determined.

This three-step procedure could be developed to decode any given spelling, but it also has the secondary function of allowing us to decide between two (or more) candidate spellings in spelling formation. That is the major innovation of this entire subsection. In the language of Chapter 4, we can decide between a draft and an amended spelling. And thus we can see exactly why `?<doping>` is a better spelling than `?<dopeing>`. The case of `?<doped>` and `?<dopeed>` is more problematic because neither is an ideal phonographic match for the known phonological form `/dəʊpt/`. However, our analysis has shown that the g-trochee `< (do.ped) >` is a better phonographic fit than the g-spondee `< (do)(peed) >`. This makes sense from the viewpoint of historical phonology, because the contraction of a p-trochee into a monosyllable seems far more likely than the loss of an entire stressed syllable in a p-spondee.

In theory, the three-step procedure could be used for any set of candidate spellings, although we might expect some slightly different constraint rankings for different kinds of words. In the examples of inflection presented here, it was necessary to set Complex-IdP as the highest-ranked constraint.

Finally, it is not fully clear how reliable the constraint ranking is. For example, the spelling `<proceed>` is most likely, according to Evertz’s methods, shown in Table 5.11, to be assigned as the spondee `< (pro)(ceed) >`, which is the structure of the noun spelt that way. However, the verb takes the iambic form `< pro(ceed) >`, which is the second most optimal structure. Such examples show
that the OT framework put forth cannot account for all details of the writing system, and while the theory discussed above shows great promise, it may need to be augmented with other evidence. Likely options include statistical analyses of words with respect to the constraints. It may be the case that optimal candidates are chosen with a certain degree of statistical predictability, and there may be a predictable number of semi-regularities. It may also be the case that the spelling subsystems, and their concomitant decoding patterns, may correlate strongly with a different set of constraint rankings or, as in the example of *proceed*, different grammatical categories may require different rankings. The decoding patterns may also have changed slowly over the centuries. Furthermore, readers’ intuitions may not tally with the distribution of pronunciations of a particular spelling pattern. We shall see more of this in the remainder of the chapter through the phenomenon of spelling pronunciation, whereby words are given a new pronunciation due to the re-analysis of how their spellings ought to be decoded.

***
5e. The re-coding process: constructing phonological structure out of spellings

Up to now, we have only looked at how spellings are formed when the phonological and morphological forms are already known, but the G↔P trees can also be used to construct phonological forms out of the existing spelling. This has two ramifications. The first of these (discussed in Section 5e.1) is that a greater number of spellings can be accurately decoded if a foot-level of analysis is applied. This is because the foot structure helps to subdivide long and complex words into stressed and unstressed syllables, so we can more easily predict the pronunciation. This is particularly useful for predicting the palatalization of consonants, one of the more complicated mappings from spelling to sound which previous theorists have struggled to model. Mapping from spelling to sound in this way will be known henceforth as metrical decoding, in parallel with metrical phonology, the study of phonological stress.

The other important ramification is that we now have a path by which spelling pronunciation can be both explained and predicted (Section 5e.2). We shall see several pronunciations of the word omega, all of which are predictable from the letter-patterning, although some are more likely than others. This will be demonstrated with a combination of the metrical-decoding notation and the visual model, but it will be shown how OT can be used to predict what variation is likely to occur.

5e.1 Metrical decoding

5e.1.1 Consonant palatalization

Carney (1994: 21) discusses the difficulty of isolating the sound-spelling correspondences in words which have undergone palatalization. In his analysis, a word such as picture appears to have the correspondence <t> ≡ /tʃ/ (with <c> ≡ /k/ and <ure> ≡ /ə/). This same correspondence occurs with regularity in many such words: structure, virtual, perpetuate, combustion, etc. (Carney: 1994: 226), but the problem is that the analysis requires the reader to predict palatalization without any auxiliary evidence. The same problem arose in Venezky’s example of social, where the reader is simply expected to apply a palatalization rule on the second level of analysis, mapping from the intermediary morphophonemic level to the actual pronunciation. If we apply a metrical-decoding analysis it becomes easy to see a pattern that applies across all words with a similar prosodic shape.
/tʃ/ often occurs in dactyls, after the stressed syllable, as in < (vir.tu.al) >, < (per.pe.tu.al) > and < (vo(lup.tu.ous) ) >. Palatalization also occurs in words where there is a g-dactyl but just a p-trochee. Examples include < (ques.ti.on) > and < (sug.ges.ti.on) >, or < (pic.tu.re) >, < (rap.tu.re) >, < (struc.tu.re) > and < (vir.tu.e) >.

The same metrical-decoding analysis also applies to several other palatalised consonants: to the /ʒ/ of < (col.li.si.on) > and < (lu.xu.ry) >; to the /ʃ/ of < (na.ti.on) >, < (pres.su.re) > and < (dis(ssi.on) ) >, to Venezky’s example of < (so.ci.al) >, and to the /dʒ/ of < (re.gi.on) > and < (sol.dier) >. The model can even be extended to made-up words such as *exapture, which could be analysed as < e(xap.tu.re) >, before being mapped onto the phonological form /ɪgˈzæptʃə(r)/. In order to recognise the g-dactyl, there must be three g-syllables, where the final two are weak or reduced, marked here in bold: < (struc.tu.re) >, < (ver.si.on) > and < (vir.tu.al) >. In some cases, there are pairs of related words whose very different pronunciations can be seen clearly through metrical decoding. Compare palatalized < (dis(cu.ssi.on) ) >, < (so.ci.al) > and < (anxi.ous) > against non-palatalized < (dis(cuss) ) >, < so(ci.e.ty) > and < an(xi.e.ty) >, where the same letter now appears at the beginning of a stressed syllable, so there is no palatalization. Some of these words may have once been p-dactyls, like < (bas.ti.on) > ñ /ˈbæstən/, but have since been reduced to p-trochees, leaving behind them a predictable mapping from spelling to sound. This sound change, known as syncope, is common in English (see Wells 2000: 165–6).

Metrical decoding is in fact consistent with Venezky’s (1970) lower-level mappings from spelling to sound, and the only difference here is that further information about the stress and syllabification patterns are supplied. Palatalization is not indicated explicitly in the present model and it has to be supplied by the reader who must make a deduction of the following kind: palatalization occurs whenever <t d s s c g x > appears after a stressed syllable and at the beginning of two unstressed g-syllables. And of course that will vary across accents and even among individual speakers, as in words such as tissue and sexual (OED 202513 / 177084). It is of course impossible to map the entire path from spelling to pronunciation — all we can ever do is map from spelling to some abstract level (e.g. phonemic, morphophonemic, underlying form, etc) and the rest of the information must be supplied by the reader / speaker. The rough outline presented here shows how metrical decoding of palatalized consonants simply adds a greater level of predictability between spelling and sound. It is perfectly reasonable to imagine that a more detailed, OT account could provide more. However, since Evertz’s (2014) model does not accept g-dactyls, there does not
yet exist the theoretical machinery for performing this analysis, and doing so is beyond this project.

Before moving on, it is worth noting that higher-level phonological changes have in fact given us some of the most important spelling patterns. CLD ultimately derives from the reduction of phonological geminates (Britton 2011), so that <(ap.ple)> would have had a phonological equivalent [ˈappə], while mute final <e> was once pronounced, giving mappings such as <(na.me)> ≡ [ˈnarna] (Lahiri and Fikkert 1999). (Square brackets are used to indicate phonetic forms as the phonemic inventories of reconstructed phonologies are far from certain; see Lass and Laing 2008–13).
Figure 5.25
G ↔ P trees showing the mapping < (so.cι.al) > ≡ /ˈsəʊʃəl/
5e.1.2 Ambiguities in the decoding process

Despite all of the above examples where palatalization can be predicted from an g-dactyl, not all words with a <V(C).CV.CV(C)#> spelling are dactylic. Examples include < a(ro.ma) >, < to(ma.to) > and < po(ta.to) >. Wells (20/01/07) writes that:

Ordinary spelling, of course, doesn’t indicate vowel length in <VCV(C)#> words. It’s one of the biggest shortcomings of our spelling system.

The reader’s job is as follows, he continues:

Faced with an unfamiliar word ending <VCV(C)#>, we crucially need to know whether the penultimate [phonological vowel] is long or short. If it is long, it will be stressed; if it is short, it won’t be, and indeed will undergo vowel reduction.

Examples of the long, stressed vowel include < to(ma.to) > and < po(ta.to) > (regardless of the quality of the stressed vowel which varies across accents).

Examples of the short, unstressed penultimate vowel include the dactyls < (vir.tu.al) >, < (a.ni.mal) >, < (e.ne.ma) >.

In Section 5e.2, below, we shall see that this ambiguity over vowel length can result in the emergence of two or more variant pronunciations, and the explanation for this seems to be that the spellings are decoded according to different metrical analyses.

5e.1.3 Unstressed p-vowels

It should be also noted that the penultimate syllable in virtual is (a short) /u/ rather than /ə/ or /ʊ/, despite the fact that this syllable does not carry the primary stress in the word. Evertz (2014: 79) assumes that all reduced syllables in English are either /ə/ or /ʊ/ (or sonorant consonants, as in bottle and button). This means that full vowels always correspond to stressed syllables and reduced vowels to reduced syllables. However it is important to state that there are three kinds of vowel quality in English, whatever way they may be grouped with respect to stress and foot structure. The word virtual has all of them, in the following order: (1) full stressed vowels which occur in syllables that receive primary word stress; (2) full vowels which do not receive primary stress; and (3) reduced vowels, often transcribed as /ə/ or /ʊ/, which cannot carry any stress. Ladefoged (2006: 96) provides examples of this three-way distinction for eleven English vowels. Two sets of examples are: ap’preciate, cre’ation and ’deprecate as well as con’front, um’brella and confron’tation. He notes that the five vowels /ɔ u əʊ ɔɪ/ are never
fully reduced, giving the following pairs of examples: *cause, cauˈsality;ˈhoodwink,ˈneighbourhood; aˈcoustic, aˈcousˈtician, outside, ouˈsider and exˈploit, exploiˈtation*. The fact that these vowels are generally spelt with a digraph is irrelevant from the viewpoint of phonology, but it may be of importance in a foot- or syllable-based account of English spelling.

Wells (25/03/2011) notes that there have been different ways of dealing with full vowels which do not receive primary stress (group 2, above). In the British tradition, they tend to be analysed as unstressed while in the American tradition they are considered to carry secondary stress, and this would make them the head of a foot. If these theories were applied to a metrical decoding of *virtual*, the British tradition would render <vir.tu.al> and the American tradition would give <ˈ(vir)(tu.al)> . I have chosen the British-influenced path, at least where the full, unstressed vowel occurs immediately after the stressed vowel, in a dactyl. In spondees, however, I have taken the latter approach in analysing certain words, including <ˈ(pro)(tein)> and <ˈ(de)(tail)> . In words where the first syllable does not receive primary stress, I have left it unparsed, as in <Oc(to.ber)>, rather than treat it as a single foot <(oc)> . These issues will be important for a foot-based theory of English spelling which goes into greater detail than the present study.

Regardless, there is one thing we can be sure of: reduced vowels are not expressed explicitly in English spelling. This has greater implications for encoding than decoding. Since /a/ and /i/ can be spelt in so many different ways, phonological information is not a very good aid to spelling, in such situations. Writers must rely on auxiliary information such as morphologically-related words or etymology. If we cannot tell what vowel letters are missing in <democr_cy> and <hypocr_sy>, we can refer to *democratic* and *hypocritical* for help, despite the fact that their stress patterns are different. However, if the reader knows that the former words can be analysed as <de(mo.cra.cy)> and <hy(po.cri.sy)> they can predict the reduced vowels from their position within the g-dactyl. In this way, reduced p-vowels are predictable from foot structure, and it doesn’t matter which vowel letters are in the reduced g-syllable.

***
5e.2 Spelling pronunciation: building P-structure from O-structure

Spelling pronunciation is a phenomenon whereby the pronunciation of a word is changed under the influence of its spelling. For example, the name Anthony may be pronounced with /θ/ on account of the <th> spelling, rather than with a traditional /t/. Similarly, /h/ is nowadays pronounced in Hebrew and hermit where once it wasn’t (Neuman 2009). The phenomenon has been well-known for a very long time but it has not received a great deal of attention from linguists, with Householder (1971) being a notable exception. Most attempts to explain spelling pronunciation focus entirely on examples of segmental spelling pronunciation (e.g. changes from /t/ to /θ/), but here we shall see how the phenomenon can occur at higher-phonological levels, and this arises from a re-analysis of a word’s metrical structure.

The argument presented is that spelling pronunciation involves the construction of phonological forms out of graphematic forms. While there is often a predictable mapping in both directions, ambiguity in the decoding can lead to two readings and thus two pronunciations. One of these may have been the traditional and/or prestige form. Others may begin by having lower social value before becoming the norm (and possibly even the prestige variant). Our concern is solely with the actual mappings between spelling and sound: issues such as normativity, prestige and, to a lesser extent, frequency of usage, are not relevant. The aim of this section is not simply an argument for the existence of spelling pronunciation in this isolated case but instead to provide a theoretical path which explains how the phenomenon can occur, with phonological structure being built out of graphematic structure. We shall look first at the relevant literature, before showing how several pronunciations of the word omega can all be explained as valid mappings from spelling to sound. This will be modelled using a combination of metrical-decoding notation and the visual model. At the end, we shall see how the variation is somewhat predictable within an OT-based model.

5e.2.1 Background literature

Giegerich (1999: Ch. 5) provides evidence of spelling pronunciation in derivational morphology. He essentially assumes that words such as atomic, systemic and managerial all require knowledge of the spelling in order to be pronounced ‘correctly’ (by literate speakers). His main observation is this: if atom, system and manager all have /ə/ in their base form, then how can their derived forms, atomic, systemic and managerial, have three different phonemes,
respectively /ɒ/, /i:/ and /ɪə/? The only reasonable answer is that these pronunciations are corroborated by the spelling. To check the influence of spelling on pronunciation, Giegerich tests a sample of ten students on their pronunciation of some names and their derivations. He analyses ten names, including Anderson, Mendel and Pentzl, which all have a reduced vowel (or syllabic consonant) in their final syllable, and asks the students to derive adjectives with {ian}, a derivational suffix that attracts primary stress to the preceding syllable. The spellings of these derivations is obvious: simply add <ian> to get <Andersonian>, <Mendelian> and <Pentzlian>, but the pronunciations are all derived from the spelling, where possible. All of the students have /ændəsəʊnɪən/, but Mendelian results in a mix of /menˈdiːlɪən/ and /menˈdeɪlɪən/ (six of the former, four of the latter). Most students fail to answer for Pentzlian. Since the spelling has no vowel letter in the appropriate place, there is no guide to the quality of the (expected) vowel sound.

Giegerich does not discuss how the students attach the stress patterns to the spelling — it seems they just know how to organise the stress around the stressed syllable peak they have chosen. It will be assumed here that their method involves metrical decoding and, in the notation introduced here, these words would be < (An.der)’(so.ni.an) >, < (Men)’(del.i.an) > and, perhaps, < (Pent.zli.an) >. Clearly a problem arises with the occurrence of <zl> in syllable-initial position, as it is not a valid onset. < (Men)’(del.i.an) > does not specify what vowel sound the (second) <e> should correspond to. This is a reminder that: a) the representation of English p-vowels is highly defective in the writing system, but also that b) the phonemic correspondence of this vowel letter bears no effect on the word’s foot-structure (the same issue arose above with tomato).

Building on the work of Giegerich (1999) and Giegerich’s student Montgomery (2001, 2005), Neuman (2009) sees writing and speech as semiotic systems in contact, with one influencing the other through reciprocal transfer. Not only are spellings formed with some reference to the phonological form but the phonological form is often constructed out of the spelling. This is made possible by what he calls ‘graphophonemic operators’, mappings from letters to sounds where readers fill in the supposedly missing or incorrect details. For example, the spelling <often> suggests /ˈɒfən/ and indeed this pronunciation has gained ground on the traditional pronunciation /ˈɒfən/ (see also Wells 3/11/10). The presence of such variant pairings may seem marginal at any given stage in a language’s history but such changes have a cumulative effect over time, slowly bringing writing and speech closer together. In other words, decoding and encoding become more bi-directional, despite the better-documented opposing
force whereby phonological processes make the two systems diverge over time (something that will be discussed at length in Chapter 6). Neuman provides copious evidence of such changes throughout the history of English (and in several other languages). Examples include the introduction of /h/ in words spelt <h> (hostage, humour), and the change from /t/ to /θ/ in words spelt <th>, especially in §Greek words (theatre, lethargy) but sometimes in words mistakenly thought to be of Greek origin (author, anthem; see also Scragg, 1974: 57).

Neuman also argues that there has been a long tradition of inventing English pronunciations from the spelling of erudite words found in books. Renaissance English speakers introduced Latin and Greek words into the language and the pronunciations often had to be constructed from the spelling. The process was even extended to established words in the language so, for example, perfite was re-modelled on Latin perfectus, and its <c> instigated the use of /k/ in the word (see the notes at OED 140704). Further evidence of spelling pronunciation comes from a word such as data which have two variant pronunciations, /deɪtə/ and /dɑːtə/ (in RP), or /dɛɪtə/ (in GA). Neuman (2009) suggests the latter pair of readings follow a Latin-derived pronunciation (compare the phrase pro rata). Evertz (2014) would classify this as a non-canonical trochee. The pronunciation /dɛɪtə/ may be on analogy with the §Basic decoding of <(da.ta)>, rather like the acronym NATO.

It is easy to see why spelling pronunciation has never been fully accepted in linguistic theory. Once a spelling pronunciation becomes the established pronunciation, succeeding generations can have little or no awareness of the existence of the older pronunciation, partly because they have not heard it themselves, and partly because there may no recordings of the old pronunciation. This means that spelling pronunciation can exist in a perpetually marginal state and, while the examples of words undergoing the process change over the generations, no one generation can be aware of the phenomenon’s long-standing influence on the language. Neuman’s use of historical evidence, dating back to the late medieval period, makes a compelling argument for the continued presence of the spelling pronunciation over many centuries. Jesperson (1909) also provides numerous examples, e.g. nephew, falcon, soldier, as does Householder (1971: 252–3), including yesterday, diphthong and harass. Cummings (1988: 22–24) claims that the phenomenon is more common in American speech than British, although he does not provide any evidence.

Spelling pronunciation can also be disregarded somewhat because instances of it, such as those exemplified by Neuman, have generally been confined to changes at the segmental level, which means that any changes do not
affect the rest of the word. However, if spelling pronunciation were to affect the entire word, or even the foot structure, then it would be very hard to ignore its relevance. Giegerich’s examples of Andersonian and Mendelian provide a step in this direction because the changes affect the rhythmic shape of the entire word (thanks to the stress-shifting suffix {ian}).

5e.2.2 Metrical decoding and stress-shifting suffixes
A metrical-decoding analysis can in fact be applied to any kind of stress-shifting suffix, and this makes it much easier to predict the pronunciation from the spelling. For example, the suffix {ity} also attracts stress to the preceding syllable, and it forms a dactyl with a lax antepenultimate vowel. This occurs in

\(<\text{nor.mal}>\) and \(<\text{nor(ma.li.ty)}>\), where the stress shifts from the first to the second syllable of the base. A similar kind of analysis can be applied to any stress-shifting suffix. Compare the changing stress patterns of \(<\text{o.(ri.gi)n}>\), \(<\text{o(ri.gi.nal)}>\) and \(<\text{o(ri.gi)’(na.li.ty)}>\). English spelling does not indicate this information explicitly but its inherent conservatism and refusal to indicate reduced vowels explicitly means that many vowel letters are effectively inert letters whose pronunciation is triggered in related forms. In the following section, we shall see that the occurrence of spelling pronunciation can only follow certain available paths, and this provides further evidence of its existence.

5e.2.3 Modelling spelling pronunciation
The many pronunciations of the word omega suggests that spelling pronunciation can result from regular and systematic re-analysis of spellings at the level of the foot. Wells (2000) notes the striking differences between British and American pronunciations of the word. British English traditionally has /ˈəʊməɡə/, with a dactylic stress pattern, one which has followed the natural path from Greek to English via Latin (see Wells 05/02/07). American English, by contrast, has three pronunciations: /ˈoʊməɡə/, /ˈoʊ meɡə/ and, increasingly, /ˈəʊ miːɡə/ (see also OED 131185). These pronunciations, with stress on the second syllable, have since become more common among younger British people, according to Wells’s more recent survey (16/07/07), and there is only one way to explain it:

[T]his historical coherence is now being superseded by arbitrary think-it-up-yourself spelling pronunciations.

Wells (05/02/07)

The following argument shows that spelling pronunciation is not arbitrary, but instead, it is possible to think up a predictable path from spelling to

212
pronunciation, using metrical decoding. Where there is ambiguity in the
 correspondences, it can be manifested through variant pronunciations. In the case
 of *omega*, there are at least three kinds of ambiguous correspondences, leading to
 four documented pronunciations. A fifth pronunciation, undocumented in the
 major pronunciation dictionaries, is also predicted by the model.

The dactylic British pronunciation /ˈəʊ.ɪ.ɡə/, maps onto a dactylic g-foot,
< (o.me.ga) >, much like < (e.ne.ma) > or bimorphemic < (o.pe.ning) >. The
reason for this antepenultimate stress, notes Wells (17/07/07),49 is because the
penultimate syllable is light in Latin so the stress is pushed to the antepenultimate
syllable. The same goes for *cinema*, whose metrical analysis is < (ci.ne.ma) >.

In the first G↔P tree (Figure 5.26), the g-syllable <o.> maps onto the first
p-syllable, which is stressed. Since the g-syllable is open, it maps onto a binary
p-vowel, with a branching nucleus, which is the minimum requirement to fill
a(ny) stressed syllable. The second and third g-syllables have <CV> spellings and
each of these maps onto a light p-syllable, with a non-branching nucleus. As ever,
the spelling gives no hint as to whether the reduced vowels are /ɪ/ or /ə/. In the
figure, brackets and baseline dots have been included in order to indicate foot
structure and syllabification. This should make it easier to distinguish the three
analyses.
Figure 5.26
G ↔ P trees showing the mapping < (o.me.ga) > ≡ /ˈəʊmɪɡə/
Figure 5.27 shows that the visual model also predicts the pronunciation '/ɒmɪɡə/' as a possible reading of *omega*, even though it is not recorded by Wells (2000, 2007) or the OED (131185). We saw in Section 5c.2.4 that dactyls often have a lax stressed vowel due to tri-syllabic shortening. The reader might
therefore decode the stressed first syllable of \textit{omega} as /ɒ/, as it is in the name \textit{Oliver}. This reading is only possible because ambisyllabicity is not marked in the spelling of dactyls, and the variant readings are the tri-syllabic equivalent of the ambiguity between \textit{lemon} and \textit{demon}.

The other three attested pronunciations of \textit{omega} all involve the metrical analysis < o(me.ga) >, with stress on the penultimate syllable. This analysis can be decoded in three ways, two of which are related. The related pronunciations are /əʊˈmiːɡə/ and /əʊˈmeɪɡə/, where the <e> is interpreted as a binary phonological vowel, just like the <o> of < a(ro.ma) > or the <a> of < po(ta.to) >. The ambiguity between /i:/ and /eɪ/ also occurs in the word \textit{Mendelian}, above, and it is a reflection of underspecification in the writing system. In §Basic and §Latinate words, <eCV> generally corresponds to /i:/, as in \textit{veto, Peter, compete, these}, etc. In §French words, it is often /eɪ/, as in \textit{debut} and \textit{bouquet}. This leaves an unparsed extrametrical syllable at the beginning of the word and the phonological status of this vowel, as an unstressed full vowel /əʊ/, is rather like the short /u/ of \textit{virtual}. As can be seen from Figure 5.28, the mappings from < o(me ga) > to /əʊˈmiːɡə/ and /əʊˈmeɪɡə/ are identical at all levels, except for one phoneme. This pair also provides a good example of how a long p-vowel and a diphthong are structurally interchangeable. In this figure the foot-level is distinct from the word-level and their relations are shown in the figure.
Figure 5.28

G $\leftrightarrow$ P trees showing the mapping $< \text{o(me.ga)} > \equiv /\text{o}u\text{e}m\text{e}\text{g}a/$. The pronunciation $/\text{o}u\text{m}i:g\text{a}/$ has an identical set of mappings, except for $<e> \equiv /i:/$.
The final pronunciation /əʊˈmeɡə/ also has an unparsed syllable followed by a trochee, but this time the stressed syllable has a unary phonological vowel followed by an ambisyllabic consonant which is not marked in the spelling. The same pattern occurs in the trochaic words lemon and feta. If the spelling were §Basic, then we might expect it to be spelt *<omega>. In Figure 5.29, we can see how the absence of CLD causes a rift which runs through the entirety of the phonological and graphematic trees.

The OED (131185) note that /əʊˈmeɡə/ used to be a British variant and this pronunciation was listed by the 1902 New English Dictionary (the first edition of what became the Oxford English Dictionary, c.f. OED 124731). It is not clear exactly why this particular pronunciation has receded but it cannot have been helped by the unclear mapping from the pronunciation to the spelling, especially given the number of competing variants. Thus the pronunciation may be obsolescent due to its being a poor match for the spelling.

**Summary**

In the decoding of the word omega, there are three sets of ambiguity in the correspondences, and these give rise to four attested pronunciations, and one other pronunciation predicted by the model. There is ambiguity over whether this spelling ought to represent a) a g-dactyl, < (o.me.ga) >, pronounced /ˈəʊmɪɡə/ or, potentially /ˈɒmɪɡə/; or b) an unparsed syllable followed by a trochee, < o(me.ga) >. If it is the latter then it could be pronounced /əʊˈmeɡə/ or /əʊˈmiːɡə/, based on <me.> being analysed as an open g-syllable. If not, it suggests the obsolescent pronunciation /əʊˈmeɡə/, which doesn’t match the spelling very well.

If we do an OT-analysis the spelling <omega> we can see what the most likely stress patterns are.
Table 5.17

OT analysis of the g-foot structure of &omega;  

Table 5.17 shows eight potential stress patterns for the spelling &omega;, assuming as always that any given syllable can be either stressed or unstressed. Candidates a. and c. violate WHP because they both assign stress to a &lt;CV&gt; syllable. Open &lt;CV.&gt; g-syllables cannot occur at the end of a word (such stress patterns might be better spelt as &lt;(o.me)gah&gt; (see Evertz and Primus 2013). Candidates e-h. are all disqualified for not beginning with a g-foot and, related to this, their syllables are not parsed by feet. These candidates help to resolve the phonological status of full vowels in syllables that are not stressed, at least when they occur before the primary stress. They ought be considered as heads of feet, rather than unparsed syllables. Candidate b. may be assumed to be the default pattern as it violates no constraints. However, Candidate d. is a lot more interesting as it raises several questions about Evertz’s analysis and, more importantly, our theory of decoding. Candidate d. violates the Minimality constraint which requires g-feet to have just one or two syllables. However, if we consider g-dactyls to be valid g-feet, then this constraint is de-validated, and the word would also obey Align-GFoot-Right. Evertz’s analysis does not allow dactyls but the reality is that they exist in English, even though they are not the most common pattern, especially among monomorphemic forms. A more nuanced analysis than OT might examine the statistical likelihood that the word be analysed as &lt;(o)(me.ga)&gt; or &lt;(o.me.ga)&gt;. Instead of insisting that decoding is due to a set of ranked constraints, the analysis might factor in the relative
prevalence of each stress pattern. Either way, the fact that variant pronunciations can be predicted from the spelling provides strong evidence that spelling pronunciation exists in English. Perhaps even more importantly, however, is that we can use the same theoretical machinery to explain spelling pronunciation and the decisions required to choose between two possible spellings. This suggests that spelling pronunciation is a design-feature of the writing system and not just a peripheral phenomenon. The fact that its existence is only noticeable in a few extreme cases (e.g. *often*) simply highlights how powerful the bond is between spelling and pronunciation.
Figure 5.29
G ↔ P trees showing the mapping \(< o(mega) > \equiv /əʊˈměɡə/\).
5f. Conclusion

In this chapter, we saw how mappings between spelling and sound can occur at all levels of the phonological and graphematic hierarchies, including the foot, the syllable, sub-syllabic elements and the phoneme. This knowledge allows us to explain several things better than was previously possible. Firstly, we have more rigorous criteria for deciding which spelling is the best phonographic match for a known word, and thus we can decide between a draft and an amended spelling (or any two or more spelling candidates), rather than relying on our intuitive understanding of what counts as a good spelling. Secondly, we can decode spellings with a greater degree of predictability. Thirdly, it has been shown that there exist predictable paths by which spelling pronunciations can arise, due to a re-analysis of the metrical structure encoded within a spelling.

These advances have only become possible thanks to the recent work of Evertz (2014) who argues convincingly that the foot is an important structural level within English spelling. He provides a detailed discussion of the entire graphematic hierarchy, including segments (letters), graphemes (units of spelling such as <s>, <h> and <sh>), sub-syllabic elements (onsets, nuclei and codas), syllables (full and weak), and feet (canonical and non-canonical trochees). The hierarchy is defined in parallel with the phonological hierarchy and thus the early parts of this chapter provide the necessary background phonological theory (taken from Giegerich 1992), showing how well-formed English syllables must adhere to strict criteria (hence /lkæpm/ is not a possible English syllable). The same constraints apply to polysyllabic words, except that unstressed syllables, with non-branching nuclei, alternate with stressed syllables. Syllables are parsed by the syllable-boundary rule which requires onsets to be maximal, as long as they are legal. Hence ma.tron syllabifies thus, because /tr/ is a legal onset. However the /t/ of at.las must remain in the first syllable, which it closes, because /tl/ is not a permissible onset. The distinction between closed syllables (with a vowel plus a consonant) and open syllables (with a long vowel and no consonant coda) is very important and is deeply embedded in the English writing system.

Rollings (2004) recognises that <plan.ning> and <pla.ning> have, respectively, a closed and open first syllable, and this is marked with CLD. Evertz (2014) shows how this structure can be explained with reference to the foot and he goes so far as to argue that <pla.ne> has two g-syllables but only one p-syllable. Thus we recognise that <pla.> is an open syllable and thus a long, or binary, vowel. The syllable <ne> is a minimal graphematic syllable, which has no phonological correspondence. Full, or stressed, graphematic syllables must have a
branching nucleus, as in <sat>, <say>, or <hear.say>. However, in g-trochees, open syllables can be spelt <CV.>, as long as they are followed by a reduced syllable, as in <pla.nar> and <ta.ble>. Evertz distinguishes between canonical and non-canonical g-trochees, although this distinction was abandoned here, in light of Berg’s (2016) observations, and it is believed that the two may be better explained with reference to Carney’s spelling subsystems. Canonical trochees are mostly §Basic, while the less regular non-canonical trochees (e.g. lemon, feta) often belong to other subsystems. Evertz also disregards the graphematic dactyl but it was shown that this can be very important in helping to decode English words.

The graphematic hierarchy is very useful in decoding English spelling. We can first work out a word’s most likely g-foot structure, using an OT-based analysis, and analyse the spelling into stressed and unstressed syllables. This allows us to break polysyllabic words into a series of more manageable monosyllables and thus compute the phonological form with a greater degree of accuracy. The focus in this chapter has not been on the segmental-level mappings, but rather on how foot and syllable structure can help explain phenomena such as CLD, mute final <e>, consonant palatalization and spelling pronunciation.

However, the most relevant innovation of this chapter has been the introduction of criteria by which we can test if a draft or amended spelling is a good phonographic match for a known phonological form. We looked at pairs such as ?<dopeing> and ?<doping>, two possible spellings for the {ing} form of the verb dope. The first analysis involved the establishment of a visual model which allows us to compare the phonological and graphematic hierarchies. We could see that ?<dopeing> is not a good match for /dəʊŋ/, because <ei> is not a good spelling of /i/, and thus <eing> is not a good spelling of /ɪŋ/. The spelling ?<doping> matched all levels and was thus considered to obey PhM and be a good spelling. However, problems arose for ?<dopeed> and ?<doped>, two possible spellings for the {ed} form of the same verb. Neither looked like an ideal match for the monosyllabic phonological form /dəʊpt/, and thus an OT analysis was introduced to help decide between two ostensibly bad spellings. It quickly became clear that simple IdP and PhM were not adequate constraints to help choose the right spelling. The OT analysis thus involved the following. First, the highest-ranked constraint was defined as Complex IdP, a subtle development of IdP which allows for either IdP-reduction (e.g. <e>-dropping) or IdP-extension (e.g. CLD). This constraint weeds out all other spelling candidates such as <doapt> or <dohpt>, which only obey PhM. It was found that PhM is too general a constraint and it did not allow us to decide between two unsuitable spellings. Thus Evertz’s decoding analysis was applied to the words. The first job was to assign the most
likely foot structure to the words, rendering < (do)(peed) > and < (do.ped) >, a spondee and trochee respectively, and these could then be compared to the known phonological form, /dəʊpt/. The notation, introduced here, allows us to distinguish feet, marked in round brackets, and syllable structure, marked with a dot. Only foot-initial syllables are stressed. The OT analysis showed that < (do.ped) > was the better candidate, and is thus chosen. In this way, it was possible to confirm that the amended spelling is simply the best of a bad lot. A more elaborate analysis would be necessary to explain some of the trickier inflectional spellings, such as why <c> is retained in <eyeing> but not in <owing> or <suing>. Such examples are reliant upon a deeper understanding of sub-syllabic mapping constraints, something that there has not been space to discuss here. Despite the absence of such details, it should be clear now that the above method is useful for an initial testing of the suitability of a spelling, whether it is a draft spelling, an amended spelling, or otherwise.

It has also been shown that the graphematic hierarchy can be used to explain palatalization and some other details in the decoding process which have, up to now, defied analysis. For example, Venezky’s model for decoding the word social included several phonological processes that had to be supplied by the reader, including stress, vowel reduction and palatalization. If social is instead analysed as a graphematic dactyl, as in < (so.cial) >, then we can predict palatalization occurring in this position, just as it does in < (na.ti.on) >, < (re.gi.on) > and several other instances. (The underlining indicates that two g-syllables correspond to one p-syllable, a very common occurrence). A very different reading occurs in < so(ci.e ty) >, where the <c> now occupies the onset of the stressed syllable, and thus represents /s/. Such metrical decoding, as it has been called here, can increase the predictability of mappings between spelling and sound, thus reducing the amount of knowledge that must be supplied by the reader. It should be noted that Evertz’s OT model for assigning foot structure cannot be used to predict < (so.cial) >, as his model assumes that graphematic feet must have either one or two syllables. However, a study into English dactyls could easily rectify this problem and enrich the model. Metrical decoding can be applied to any complex word, and is especially useful in explaining stress-shifting suffixes because the reader can often locate the stress by counting syllables backwards from the suffix.

The final part of the chapter argued that spelling pronunciation can be viewed as a generalisation of the decoding process to words whose spelling may include ambiguities. For example, the word omega can be analysed as either < (o.me.ga) > or < (o)ˈ(me.ga) >, rendering up to five different pronunciations.
Evertz’s OT-analysis suggests that <(o)ˈ(me.ga)> is the default stress pattern, but the traditional pronunciation is a dactyl, represented by <(o.me.ga)>.

The variable pronunciations, all of which are predicted by decoding these two analyses, suggest that readers have a tendency to pronounce words according to the patterns they find in the letters. In this one, one pronunciation can slowly supersede another. No effort has been made here to show the sociolinguistic distribution of the pronunciations or how they have changed over time. The aim has simply been to show that pronunciation can be predicted from spellings. Where there exists ambiguity, variable pronunciations may emerge. It may be the case, therefore, that pronunciation has been streamlined over the centuries in order to fit spelling. Amassing evidence of this is not easy, due to the difficulty of reconstructing ancient pronunciation, although Neuman (2009) provides a useful step towards this goal. More important, for the present study, is that the formation of new spellings may involve a system of checks between phonology and spelling, and that there appears to be an explicable set of mappings between the two systems, operating at all levels of the hierarchy.

It is possible to bring all of this together. The first step is to map from the draft spelling to the decoded phonological form (DPF). If this is identical to the known phonological form (KPF), then the spelling is good. If the DPF and KPF differ, then we can look to see if amendments are possible. Thanks to IdP restrictions, a very small number of amended spellings are possible. Each of these is decoded in turn and the one that best matches the KPF is the chosen spelling. Spelling pronunciation, involves creating a DPF which differs from the KPF.

Nevertheless, the KPF remains elusive. It is not clear exactly what kind of phonological forms English spellings attempt to represent. In the next chapter, I shall argue the English spelling seems to function as a pan-lectal phonology of the language, which means that the spelling encodes a deep phonological form from which various accents can often be derived. This is due to the extensive phonological changes that have taken place in all varieties of English since spellings became fixed in the Early Modern period.

***

***

225
6. Invariant spellings, changing phonologies

If spellings never change but phonologies constantly do, however slowly, then the gap between spelling and phonology will widen over time. This is essentially what DeFrancis (1989) calls ‘the dead hand of the past’. One important result is that the writing system can change, even though spellings stay the same (Scholfield 2016). More specifically, the systems of relations between spellings and phonology change, and these changes will naturally vary across accents. This chapter focuses on how various types of phonological change affect the writing system and, following on from that, which changes affect readers more than writers, and vice versa. The chapter also makes steps towards explaining how English spelling has managed to function for so long and across so many varieties, despite the changes in phonology and the writing system.

The chapter begins by looking briefly at the four ways in which accents vary and then Wells’ (1982) concept of lexical sets is introduced (Section 6a). Section 6b analyses what needs to be studied and why. The bulk of the chapter (Sections 6c-f) explores how phonological rotations, phonemic splits and phonemic mergers all affect the writing system in their own particular way. Some of the changes discussed (e.g. the Great Vowel Shift, the MEET-MEAT merger) have affected all, or almost all, accents of English, while others have only affected certain subsets of the language (e.g. the loss of /r/ in certain phonological environments). The focus is primarily on RP English, but also GA, and occasionally other accents where the changes may not have taken place. This latter case is of interest where the spelling reflects the phonology of one accent more than another. The chapter concludes (Section 6g) by looking at what the discussion can tell us about the nature of writing systems in general.

It should be noted that this chapter is a counterpart for Chapter 7, which deals with certain kinds of non-standard spelling. While these two chapters may seem like a strange pairing at first, it should become clear during the following discussion how the existence of changing phonology amidst invariant spelling gives rise to different spelling units representing the same sounds, a situation that is best exemplified by homophones such as know and no or beet and beat. Accordingly, we might expect certain spelling units (e.g. <ee> and <ea>) to be interchangeable in creative spelling, where the Invariance principle does not have to be applied.
6a. Discussing accent variation

6a.1 Four ways in which accents vary

There are four major ways in which accents can vary (Wells 1982, Allerton 1982 [1967]): differences of phonetic realisation; of phonemic system; of phonotactic distribution; and of lexical distribution.

Differences of phonetic realisation may be the most obvious difference between accents but they are not recorded in spelling. Writing does not attempt to record every kind of phonetic variation, including predictable variation such as allophony. A simple example of this comes in the words *lid* and *dill*, where the respective allophones of /l/ have two rather different realisations (in RP and many other accents), ‘clear’ or alveolar [l] in *lid*, and ‘dark’ or velarized denti-alveolar [ɿ] in *dill*. The writing system does not attempt to record processes that affect entire classes of sounds, such as the nasalisation of vowels before nasal consonants, as in [hʌm], [hænd] and [hʊŋ] for *ham*, *hand* and *hang*; or the aspiration of unvoiced stops /p t k/ which are realised as [pʰ tʰ kʰ] in *pin*, *tin* and *kin*. This means that further localised variation, such as affrication of this series in Liverpool English (Watson 2006), can go unmarked in spelling; and as long as the phonemic distinctions between the sounds are maintained then there is no need to change the spelling.

However, sometimes changes affect the phonemic contrasts between words and this leads to differences of phonemic system. For example, some accents have no distinction between the words *cot* and *caught*, while others do. The same goes for *put* and *but*, or *spa* and *Spar*, or *girl* and *hurl*. Cases such as these will be studied in detail throughout this chapter, and the resulting effects on the writing system.

Differences of phonotactic distribution also affect the writing system. For example, non-rhotic accents have lost historic /r/ in syllable rhymes, as in *car park*, although has not been lost in syllable onsets, as in *rock* and *railroad*.

Differences of lexical distribution may be somewhat random or isolated, so that the word *derby* may have the same vowel as *start* or *nurse*, depending on the accent. The first vowel in the word *vitamin* is strikingly different in GA and RP. The former has /aɪ/ while the latter has /ɪ/.
6a.2 Lexical sets

Lexical sets are a useful tool for comparing accents. A lexical set is a set of words which: a) contain the same stressed vowel sound; and b) behave in predictable patterns across accents. As a result, lexical sets are a useful tool for comparing accents, much more so than normal phonetic variation. Wells (1982: 127–168) introduces the ‘standard lexical sets’ to compare RP and GA. More often than not, these accents coincide. For example, the KIT lexical set comprises all words with stressed /ɪ/. Most of these words are spelt with <i> (kit, sit, sick, sister), some are spelt with <y> (myth, symbol), and some are totally irregular (pretty, English, women, breeches). Other lexical sets do not correspond to the same set of phonemes across accents. For example, in GA, the vowel of TRAP is identical to BATH (so cat and giraffe have the same vowel), but PALM is different (e.g. father). In RP, TRAP is not like BATH (so cat and giraffe differ) but BATH is the same as PALM and START (hence giraffe, father and far all have /ɑː/).

Lexical sets do not explain all variation between accents (Wells 1982: 78–80). They are no help in explaining phonetic differences — in fact they are designed to do the opposite, seeing structural similarities across accents in spite of phonetic differences. Differences of lexical distribution are also inexplicable with relation to lexical sets. Hence the word vitamin falls between two lexical sets. In RP it will be in the KIT lexical set but in GA it is in PRICE.

The three tables below show how the lexical sets correspond to the phonemic systems in the two reference accents. Notice how KIT, DRESS and PRICE all map onto the same phoneme in both accents, but TRAP, BATH, PALM, LOT differ. Words in small caps are the headword for each lexical set. Words written in lower case (e.g. via, fire, Noah) are merely examples of words containing a particular phoneme or phonemic sequence. Another major difference is the effect of non-rhoticity on RP, where many words with historical //r// now have a long vowel or diphthong. In some cases, this has led to new phonemic distinctions (e.g. NEAR, SQUARE), but not always (fire, sour, soya).

Further details will be discussed throughout the chapter as we explore the phonological history of English vowels, and their relationships with the spelling. At the end of the chapter, the lexical sets are presented again, but with examples of the different spellings occurring in each set.
<table>
<thead>
<tr>
<th>Simplex vowels</th>
<th>Complex vowels</th>
<th>Vowels +/r/</th>
</tr>
</thead>
<tbody>
<tr>
<td>i</td>
<td>ai</td>
<td>air</td>
</tr>
<tr>
<td>e</td>
<td>i: FLEECE</td>
<td>ier</td>
</tr>
<tr>
<td>æ</td>
<td>æ: PALM</td>
<td>ær</td>
</tr>
<tr>
<td>BATH</td>
<td>LOT</td>
<td>START</td>
</tr>
<tr>
<td>æ</td>
<td>ð: THOUGHT</td>
<td>ð: North</td>
</tr>
<tr>
<td>CLOTH</td>
<td>GOAT</td>
<td>FORCE</td>
</tr>
<tr>
<td>a</td>
<td>a: GOOSE</td>
<td>a: Nurse</td>
</tr>
<tr>
<td>o</td>
<td>o: CHOICE</td>
<td>o: Cure</td>
</tr>
</tbody>
</table>

**Table 6.1**

GA Stressed vowels

<table>
<thead>
<tr>
<th>Simplex vowels</th>
<th>Complex vowels</th>
</tr>
</thead>
<tbody>
<tr>
<td>i</td>
<td>ai</td>
</tr>
<tr>
<td>e</td>
<td>i: FLEECE</td>
</tr>
<tr>
<td>æ</td>
<td>æ: BATH / PALM</td>
</tr>
<tr>
<td></td>
<td>START</td>
</tr>
<tr>
<td>ð</td>
<td>ð: THOUGHT</td>
</tr>
<tr>
<td>CLOTH</td>
<td>ð: North-FORCE</td>
</tr>
<tr>
<td></td>
<td>ð: FORCE</td>
</tr>
<tr>
<td>a</td>
<td>a: GOOSE</td>
</tr>
<tr>
<td>o</td>
<td>o: CHOICE</td>
</tr>
</tbody>
</table>

**Table 6.2**

RP stressed vowels

<table>
<thead>
<tr>
<th>GA unstressed vowels</th>
<th>RP unstressed vowels</th>
</tr>
</thead>
<tbody>
<tr>
<td>æ</td>
<td>len</td>
</tr>
<tr>
<td></td>
<td>lenIn</td>
</tr>
<tr>
<td></td>
<td>commA</td>
</tr>
<tr>
<td>æ</td>
<td>len</td>
</tr>
<tr>
<td></td>
<td>i</td>
</tr>
</tbody>
</table>

**Table 6.3**

Unstressed vowels in GA and RP
6b. Categories for inspection

If we take any two existing words with different meanings, there are four sets of relations that can arise between their spellings and sounds:

1. Different spellings, different sounds.
2. Same spelling, same sound
3. Same spelling, different sounds
4. Different spellings, same sound

We shall look at each of these categories in turn in order to work out which issues are most relevant in this chapter and later on in Chapter 7.

6b.1 Different spellings, different sounds (heterographic heterophones)

This is just a complicated way of saying the most trivial of things about spelling, which is that different words are spelt differently. Obviously bank and electricity are spelt differentially. There is one other case that needs to be removed from the discussion, namely ‘pronunciation respellings’ (see Section 7b.3), which use a different spelling to indicate a different sound without affecting the semantic value of the word (e.g. <wanna>, <gonna>, <wint> in place of the normal spellings <want to>, <going to> and <went>). The present discussion is only concerned with the spelling and pronunciation of words that mean different things.

6b.2 Same spelling, same sound (homophonous homographs, or homonyms)

It is not that rare for two different words to have the same spelling and the same sound. This is especially the case when words undergo semantic changes which affect neither the spelling nor the phonology. The spelling <bank> has, over the centuries, taken on several different meanings, as in river bank, Bank of Scotland and even ‘a row of keys on an organ’ (OED 15237). Carney (1994: 399–400) lists twenty-eight examples of such homonyms but notes the difficulty of making an exhaustive list since ‘it depends on what constitutes a difference of meaning’. Is there good reason to maintain an orthographic link between the body part back and its ancient relationship to the adverb? Would it be better to have <back> and *<backe> as well as <nail> and *<nale>. In Chapter 7, we shall see that there are many cases where such differentiation has occurred, although not always systematically.
In Chapter 4 we saw several cases where the spelling of polymorphemic forms coincided with existing monomorphemic forms. Such cases are strong candidates for distinctive spellings. Yet within the English writing system, there is no other way to spell <even-ing>, <mope-d>, <add-er> (‘a snake’ or ‘a person who adds), nor <jump-er> (an ‘item of clothing’ or ‘a person who jumps’). In theory a hyphen could be used in these words, as I have done here to show their structure, but this is not the normal practice, perhaps because its unsightliness outweighs its practical use, and the fact that these pairs of words do not normally appear beside one another (a high jumper and a woolly jumper are seldom bedfellows). A hyphen would be of no use in the case of rose, which can be a flower or the past tense of rise. We did see a few cases where DSTNCT was possible among polymorphemic spellings, including dying – dyeing, singing - singeing. Such cases are only possible because the spelling of the base includes a final <e>, and IdP could be applied without the expected amendment, dropping <e>.

6b.3 Same spelling, different sounds (heterophonous homographs)
The occurrence of spellings with two different pronunciations is very common in §Latinate spellings such as <conflict>, <import>, <record>, etc. which all have a noun form (with initial stress) and a verb form (with end stress). Carney (1994: 397–8) lists a large number of these spellings. There is undoubtedly an argument for using an accent on these words to mark the stress pattern, giving us <convért> for the verb and <cónvert> for the noun, etc. (c.f. Spanish and Portuguese writing). It would of course be possible to have <convért> for the verb <convert> for the noun, or vice versa. The need for an accent breaks down in converted, where the stress naturally falls on the heavy penultimate syllable. Hence, any accented solution to the problem would need to take morphologically related forms into account. It is not clear whether these stress patterns have heavily impacted the English writing system or not and they will not be studied in any detail here as the focus is on changes occurring at the segmental level.

There also exist a handful of cases where a spelling has two unrelated meanings, as in <lead>, <resume>, <buffet>, and we have also seen that this can arise through the spelling of a polymorphemic word, as in <hops>, <does>, <axes>, <rose>, <evening>, etc. Such examples are graphotactically motivated, resulting from either a lack of orthographic resources for making DSTNCT spellings (e.g. <rose>, <axes>), or from an unwillingness to use those resources (e.g. <even-ing>, <hop-s>).

231
There are some grammatically-related §Basic homographs, including the past and present forms of <read>, or the spellings <house> and <mouth> which can be either verbs or nouns. These latter two occur due to under-specification in the writing system for indicating fricative voicing, owing to the historical lack of phonemic distinction between /s/ and /z/, /f/ and /v/ and /θ/ and /ð/, until the late medieval period (see Scragg 1974: 13–14n). However, much of the homography has been dealt with in the writing system. For the latter pair, the solution is often to add an <e> to the verb form, as in bath – bathe, teeth – teethe, breath – breathe. Carney (1994: 129) notes that the verb mouth ought to be *<mouthe>. For /f/ and /v/, the solution is sometimes to violate IdP and obey PhM, as in half – halve or knife – knives. For /s/ and /z/, the solution is more complex. A pair such as advice, advise is relatively simple, with /s/ and /z/ respectively, but practice and practise both have /s/. The problem is amplified by the fact that <s> represents both /s/ and /z/ in plural and third-person singular forms. Carney (1994: 235–7) notes that the spellings <se> and <ce> are often used to distinguish monomorphemic forms from polymorphemic, as in tense, tens; dense, dens; since, sins and peace, peas.

The main interest in this chapter will be on cases where phonological changes cause spelling units to develop two or more pronunciations. There are two cases for examination — phonological rotations and phonological splits, and sometimes these can occur together.

The most important phonological rotation is known as the Great Vowel Shift (GVS), where long English vowels underwent radical phonetic changes with almost no changes in their phonemic contrasts. Where the phonemic distinctions were maintained, the writing system could cope and the structural relations stayed the same. It doesn’t really matter whether <i> represents [i:] or [ai] or anything else, what is important is that it is not confused with <e>, which used to represent [e:] before that sound changed to [i:]. If the changes take place in tandem then the structural distinctions are preserved.

Phonemic splits tend to affect sequences of letters within words more than entire words. For example, the fact that put and but don’t rhyme (in most accents) is due to a phonemic split called the FOOT-STRUT split, which resulted in one set of words being pronounced like foot and put, while the other ended up like but and strut (Wells 1982: 196–9). In the North of England, and in many Irish accents, this change never took place, so put and but still rhyme. We shall see many such cases below.

Something of a halfway situation between splits and rotations occur in morphological alternations such as electric – electricity – electrician where, once again, the spelling does not change but the pronunciation of the allomorph is
phonetically conditioned (which means that it changes according to its phonetic environment). The phonemes /c/, /s/ and /ʃ/ are phonemically distinct nowadays, although all of them can be spelt <c>, at least in the relevant graphematic environment, and readers have to work out the pronunciation. The notation introduced in the last chapter shows the process: the spellings can be analysed as: <e(lec.tric)>, <(e.lec)ˈ(tri.ci.ty)>, and <(e.lec)ˈ(tri.ci.an)>. The stress patterns are predictable from the suffixes, in the latter two words at least, and the mapping for each incidence of <c> is predictable from the metrical decoding analysis. But in all three words, the same spelling, or CPU, <electric> is used for different sounds.

6b.4 Different spellings, same sound (heterographic homophones)
This frequently occurs when two (or more) phonemes merge. This is where a phonological distinction has been lost but the difference is maintained in the conservative spellings. Examples include meat – meet or wait – weight. There have been many such mergers throughout the history of English and they are explored below in great detail.

There also exist a few homophones that emerge from different subsystems, as in shoot, chute; fisher, fissure; and catarrh – Qatar. These pairings may not be homophonous for all speakers. In Chapter 7 we shall look in detail at spellings which are deliberately constructed in order to create homophones with DSTNCT spellings (e.g. OutKast in place of outcast).

Finally, we need to rule out variant spellings from the discussion. While pairs such as <judgement> and <judgment> or <zincking> and <zincing> may both occur, they do not represent different lexemes, and thus they are not under discussion in this or the next chapter.

6b.5 Summary
The analysis above shows that the most important categories for examination in this chapter are phonological rotations, splits and mergers, and their respective effects on the writing system will be discussed in that order. We shall also see that these can occur in combination, with splits feeding mergers, causing further disorder in the spelling-to-sound correspondences.
6c. Phonological rotations

6c.1 Chain shifts
A vowel chain shift occurs when two or more vowels move in unison through phonological space. One vowel might begin to move into the space of a nearby vowel and then that vowel is pushed into the next space. This is called a push-chain shift, in contrast to a pull-chain, where one vowel moves into a new space, leaving a gap in the phonological space for another to fill. According to Labov (2001: 116), long vowels always get raised and short vowels fall in a chain shift. Meaning, however, is not affected in such a phonological rotation. To appreciate this, we might imagine a set of traffic lights were to undergo a positional shift without changing colour. Imagine the green light moved into the space of amber, and amber pushed into the space of red. The red light, having nowhere else to go, is forced off to one side. Despite the positional shifts we would still understand the meaning of the colours.

Similarly, chain shifts cause vowels to change their phonetic quality but not their abstract phonological contrasts, so there is no need for the spelling to change. However, the quality of the phones change. The relations are shown in Table 6.4. As we shall see, the amount of phonetic change the writing system can cope with is surprisingly large.

<table>
<thead>
<tr>
<th>Spelling</th>
<th>Unchanged phoneme</th>
<th>Phone before</th>
<th>Phone after</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;α&gt;</td>
<td>⇔</td>
<td>/X/</td>
<td>[X]</td>
</tr>
<tr>
<td>&lt;β&gt;</td>
<td>⇔</td>
<td>/Y/</td>
<td>[Y]</td>
</tr>
</tbody>
</table>

Table 6.4
Unchanged relations between spellings and phonemes, despite phonetic changes

A simple example of a modern chain shift has occurred in certain kinds of Dublin English over the past generation (Hickey 2005). The diphthongisation and fronting of the GOAT vowel (from [ɔː] or [ʌɔ] to [æʊ]) has left a gap for the THOUGHT vowel to rise (from [ɔː] to [ɔː] or even [oː]). The GOOSE vowel (/uː/) has also fronted somewhat, and the overall effect is that the accent has become remarkably similar to Estuary English in the relative positions of these three back
vowels (Hickey 2007). Further examples of chain shifts are legion (see Labov 2001; Labov et al. 2006).

### 6c.2 The Great Vowel Shift

<table>
<thead>
<tr>
<th>Front vowels</th>
<th>1400</th>
<th>Back vowels</th>
</tr>
</thead>
<tbody>
<tr>
<td>bite</td>
<td>i:</td>
<td>u:</td>
</tr>
<tr>
<td>meet</td>
<td>e:</td>
<td>o:</td>
</tr>
<tr>
<td>meat</td>
<td>ɛ:</td>
<td>ɔ:</td>
</tr>
<tr>
<td>mate</td>
<td>a:</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Modern</th>
</tr>
</thead>
<tbody>
<tr>
<td>meet/meat</td>
</tr>
<tr>
<td>mate</td>
</tr>
<tr>
<td>bite</td>
</tr>
</tbody>
</table>

**Table 6.5**  
The Great Vowel Shift (adapted from Lass 1999: 72)

The Great Vowel Shift is undoubtedly the biggest and most important set of chain shifts in the history of English and yet its effect on the actual spelling of individual words has been negligible. Affecting all of the long monophthongs, the GVS involved a re-organisation of both the front and back phonological spaces, during the transition from Late Middle to Early Modern English (Lass 1999). Table 6.5 shows the changes which took place, with the vowels of the old system shown in the top half, and the vowels of the new system shown on the bottom. The middle sections represent the phonological space. The high vowels (of *bite* and *out*) diphthongised and dropped, while the mid and low vowels rose, with *meat* rising two spaces and merging with *meet* (see Section 6e.2.1 below). Similar processes affected both the front and back phonological spaces and the new system is more symmetrical than the old, thanks to the single merger. The order in which this happened has been the subject of much debate (see Lass 1999 and Stockwell 2002, for opposing views). The issues are of great phonological interest but the changes scarcely affected the writing system because they occurred relative to one another. *Videlicet*, like the traffic lights discussed above, the
vowels retained their meaningful contrasts, despite often moving into one another’s pre-existing phonological space, so the mappings to the spelling were kept intact, and the system of relations is almost unchanged. The one exception is that the merger between *meet* and *meat* resulted in the emergence of two spellings, <ea> and <ee>, for one sound /iː/. The effect of mergers on the writing system will be discussed throughout Section 6e.

Despite the lack of structural changes among the historically long vowels, the inter-relationship of long vowels with their short counterparts complicated matters further. We have seen that there exist regular alternations for ‘underlying’ or historical //a/, //e/, //i/ and //o/.

<table>
<thead>
<tr>
<th>Underlying form</th>
<th>Spelling</th>
<th>Modern phoneme</th>
<th>Spelling</th>
<th>Modern phoneme</th>
</tr>
</thead>
<tbody>
<tr>
<td>//…//</td>
<td>&lt;…&gt;</td>
<td>/…/</td>
<td>&lt;…&gt;</td>
<td>/…/</td>
</tr>
<tr>
<td>a</td>
<td>inane</td>
<td>əɪ</td>
<td>inanity</td>
<td>æ</td>
</tr>
<tr>
<td>e</td>
<td>serene</td>
<td>iː</td>
<td>serenity</td>
<td>ɛ</td>
</tr>
<tr>
<td>i</td>
<td>divine</td>
<td>əɪ</td>
<td>divinity</td>
<td>i</td>
</tr>
<tr>
<td>o</td>
<td>cone</td>
<td>əʊ</td>
<td>conical</td>
<td>ɒ</td>
</tr>
</tbody>
</table>

**Table 6.6**

Vowel alternations of modern English

However, if the Great Vowel Shift had not occurred, and all vowels had kept their late ME English values, contrasting for length but not quality (as per Lass 1999), then the table would simply look like this:

<table>
<thead>
<tr>
<th>Underlying form</th>
<th>Spelling</th>
<th>Putative modern phoneme</th>
<th>Spelling</th>
<th>Putative modern phoneme</th>
</tr>
</thead>
<tbody>
<tr>
<td>//…//</td>
<td>&lt;…&gt;</td>
<td>/…/</td>
<td>&lt;…&gt;</td>
<td>/…/</td>
</tr>
<tr>
<td>a</td>
<td>inane</td>
<td>ə:</td>
<td>inanity</td>
<td>a</td>
</tr>
<tr>
<td>e</td>
<td>serene</td>
<td>e:</td>
<td>serenity</td>
<td>ɛ</td>
</tr>
<tr>
<td>i</td>
<td>divine</td>
<td>i:</td>
<td>divinity</td>
<td>i</td>
</tr>
<tr>
<td>o</td>
<td>cone</td>
<td>o:</td>
<td>conical</td>
<td>o</td>
</tr>
</tbody>
</table>

**Table 6.7**

Vowel alternations re-imagined

The system of relations would remain identical and mute <e> would still need to mark vowel length (but not quality, as in /iː/ v. /ɪ/, etc.). Furthermore, reader would still need to be aware of the complex relationships between the spelling and the stress in the poly-morphemic words.
Despite the major phonetic changes that were precipitated by the Great Vowel Shift, the effect on the writing system has been minimal because the relations between letters and abstract phonemes have remained the same. Where once there were phonemic distinctions indicated solely by vowel length, those same distinctions are nowadays expressed through both length and quality. Indeed the qualitative differences vary enormously across English accents and yet the letters still preserve the same set of abstract contrasts. The one exception to this was the merger between MEET and MEAT words, where a distinction was lost between two long vowels, resulting in a complication to the system of correspondences. In the next section we shall see how English spelling has long functioned as a pan-lectal analysis of English phonology, one which indicates long-standing, ‘deep’, phonological forms which remain impervious to phonetic changes.

6c.3 English spelling as a pan-lectal phonology

To appreciate the phonetic ‘reach’ of English spellings, and its ability to cope with phonetic change, it is instructive to look at the recent history of symbols used by phoneticians to describe the language. That way, we can fully appreciate how spelling has continued to represent radically different accents over such a long time, as well as radically different allomorphs such as serene – serenity and cone – conical.

The current standard system of notation for contemporary English indicates both vowel quality and length, as seen in Tables 6.1 and 6.2, and throughout the present work. The notation is faithful to RP from a phonetic point of view but it is a phonologically redundant system compared to its predecessors. The system was popularised by Gimson’s (1970) English pronouncing dictionary. It was a synthesis of two existing notation systems: Daniel Jones’s length-marking system and David Abercrombie’s value-marking system, both of which were more phonologically abstract. Jones’s system was much like Table 6.7, so bead was /biːd/ and bid was /bid/. Since vowels can be sorted into pairs like these, the phonetic value ought to be predictable. This level of abstraction is phonologically elegant but it causes problems for language learners. Anyone encountering Jones’s transcription would need to know that bid in fact has a more lax vowel [ɪ], a potential problem for learners of the language who may not be able to even hear the difference between [bid] and [bɪd], let alone pronounce it (native speakers of Romance languages often have this difficulty). The same goes for caught and cot which Jones transcribed as /kɔːt/ and /kət/, rather than /kɔːt/ and /kət/. Abercrombie (1965), on the other hand, only marked vowel quality, so
bead was /bɪd/ and bid was /bɪd/. This system provides greater phonetic specificity but it requires users to learn more symbols, including some unfamiliar ones (a task that many students — and language teachers — find most unappealing). It also lacks a level of phonological abstraction, as it does not explicitly mark vowels that are pairwise connected. Gimson’s compromise notation is highly redundant and inelegant (as a system of minimal phonological contrasts) but it has swept the (RP-influenced) English-language teaching world, perhaps because its greater phonetic precision is, all things considered, preferable in the classroom, and the spelling is such a poor indicator of the phonetic forms of the large set of vowels in English. Notation influenced by Gimson is increasingly the standard used in English dictionaries which actually employ IPA transcription. This includes the major pronunciation dictionaries, including the several editions of the English pronouncing dictionary that have appeared since Gimson took over the editorship from Jones, and also Wells’s three editions of the Longman pronunciation dictionary (Wells 1990, 2000, 2007).

There is a major drawback to Gimson’s notation. The greater the phonetic precision, the smaller the range of accents it is possible to describe. Gimson’s notation is useful for RP and closely related accents but a description of other accents requires significant amendment, and even Gimson’s notation has required some updates, most significantly in Wells’s (passim) introduction of /ɪ/ and /ʊ/ for indicating reduced vowels (e.g. happy /ˈhæpi/ and visual /ˈvɪsuəl/). For a long time, Wells struggled to apply Gimson’s notation equally to both RP and GA. In Accents of English (Wells 1982) the DRESS vowel is given as /e/ in RP but /ɛ/ in GA. PALM is /ɑ:/ in RP but /ɑ/ in GA — the same quality but with no length marker, while GOAT is /ɒʊ/ in RP but /o/ in GA. This variation is manageable in a book whose primary aim is to describe different accents but it is more problematic in the Longman pronunciation dictionary (Wells 1990, 2000, 2008) which aims to tell readers how to actually pronounce words. As much as possible, Wells tries to express an underlying unity between RP and GA so his more recent transcription has been streamlined at the expense of some phonetic precision. In the latest edition of LPD, DRESS is /e/ for both accents and PALM is /ɑː:/ (with an explicit length marker), although GOAT remains different: /əʊ/ in RP and updated to /ʊ/ in GA. All of these can be seen in Tables 6.1 and 6.2.

Giegerich’s (1992) phonological analysis is much more abstract and closer to English spelling. His aim is to discuss and compare the phonologies of three reference accents: RP, GA and Scottish Standard English (SSE), which is structurally rather different from the other two. One of his main reasons for choosing these three accents (pp. 43–4) is that so many English accents
worldwide are historically related to at least one of them (e.g. Australian and New Zealand English are structurally similar to RP, Canadian is related to GA and Ulster English is like SSE). Giegerich uses Abercrombie’s qualitative notation, to allow for a compromise between phonetic precision and phonological abstraction. For example, the GOAT vowel (Giegerich does not use lexical sets) is transcribed as an underlying monophthong /o/ (with no length marker), from which its phonetic realisations in all accents are derivable. SSE will indeed have [e] and [o], with conditioned lengthening in accordance with the Scottish Vowel Length Rule (ibid: 229–34), while RP will be [ɔʔ] and GA [oʔ], or indeed however a particular speaker happens to pronounce it. Thus the broader transcription used by Giegerich is applicable to a much wider range of accents than Gimson’s narrower transcription. Giegerich’s abstractions are suitable for phonologists while Gimson’s are more applicable to the teaching or learning of one particular accent. English spelling, inasmuch as it is a kind of transcription, provides an even broader transcription than Giegerich’s. Hence its applicability to so many accents, and we can think of the writing system as a highly defective kind of ‘pan-lectal’ phonological analysis of the language, a term that will be defined shortly.

Giegerich uses symbols that represent a phonological form underlying all three accents and, following his logic, all standard accents of English. However, the notation cannot express morphophonological alternations seen in cone – conical. While the notation can express the vowel of cone, it is necessary to have a different symbol for conical, because phonemic notation cannot capture the morphophonemic unity below this particular set of surface variations: RP /kɔnɪkəl/, GA /kʌnɪkəl/, and SSE /kənɪkəl/. English spelling, on the other hand, does try to represent an abstract form //kon// underlying both cone and conical, and that’s why they are both spelt with <co>. Of course, mute <e> is required to clarify the value of vowel (compare <con>).

There are two reasons why invariant English spellings have been able to represent such different sets of pronunciations. One is the use of mute <e> which functions as a diacritic, indicating the vowel quality in the spelling of one allomorph (e.g. cone) but which is dispensable in another (conical). This allows the same letter, in this case <o>, to represent two phonemes which are historically related but phonetically very different, even across different varieties of present-day English. Since all of these varieties can be traced back to a putative original pronunciation, the spelling functions as a pan-lectal representation of the morphophonological form. Somewhat independently of this, similar changes have occurred to other vowels but since they have all moved in unison since the late
Medieval period, without affecting the structural relations, the writing system continues to represent the relevant structural contrasts.

**Summary**

Phonological analyses that only describe one accent, such as Gimson’s RP, have the advantage of being able to provide a large amount of phonetic precision, a situation that has a useful application in the teaching of RP English to non-native learners. It is problematic in the teaching of accents which are not closely related to RP (e.g. GA, SSE). Phonological analyses that cover a wider range of accents are known as *pan-lectal* phonologies (Wells 1982: 69–72; Honeybone 2011) and they have to be more abstract in order to capture unities underlying the accents. Even then, though, it is seldom possible to account for all the derivable facts. Giegerich’s analysis cannot capture succinctly the various pronunciations of *conical*, while English spelling requires an *<e>* in *cone* to mark the vowel. Despite these complications, English spelling can be seen as something of a pan-lectal phonology of all English accents in at least two ways. It maintains phonological distinctions between vowels which have undergone radical phonetic changes over the past half-millennium and which continue to diverge. Secondly, it continues to represent morphophonological alternations whose ‘underlying form’, as it is often called in phonology, is represented in core lexical spellings such as the *<con>* of *cone, conical* and the *<seren>* or *serene, serenity*. These pronunciations can be disambiguated through the use of final *<e>* or consonant-doubling.

In short, the spelling is a good guide to the pronunciation in cases where phonemic and morphophonemic distinctions have been maintained without any disruption to the systems of relations between spelling and sound. In theory, it doesn’t matter how much phonetic change has occurred in the accent, as long as the structural relations are maintained. However, phonological rotations are the only phonological change which maintain structural relations between spelling and sound. Phonemic splits and mergers cause major problems for the writing system.
6d. Phonemic splits

Phonemic splits which have taken place since English spelling became fixed pose a major problem for the writing system, specifically for readers who may not know how to pronounce a particular spelling. Why, for example, does fatter have /æ/ in RP while father has /ɑː/? We shall look here at some of the most important splits to occur in recent centuries. Unfortunately, each of the splits has been complicated by some differences of lexical distribution.

6d.1 The TRAP-BATH split

Phonemic splits arise when a phonetically conditioned allophone becomes a distinct phoneme. There was a time when almost all words spelt with <a> in a closed syllable had /æ/, just as words spelt <i> have /ɪ/ (kit, kitten, kissed), and words spelt <e> have /e/ (dress, deck, wrecking). In Southern British and related accents, the phoneme /æ/ underwent lengthening before the unvoiced fricatives /f/, /s/ and /θ/, in a process called ‘pre-fricative lengthening (Wells 1982: 203–6). Later, this long, conditioned allophone changed its quality and became a distinct phoneme of English. This is why words such as staff, pass and bath have /ɑː/ in these accents (Wells 1982: 133–6). The split also occurred before clusters beginning with /n/, as in dance and grant. The TRAP-BATH split never happened in the north of England so the situation there is simple — all words in both of these lexical sets have /æ/ and are spelt <a>.

Despite the changes, the writing system ought to remain relatively simple in one direction, even for those with the split: all words with either /æ/ or /ɑː/ should be spelt <a>. Splits are harder for the reader than the writer, as it is impossible to see which vowel sound should be attached to the one letter — why does plaster have /ɑː/ while plastic has /æ/? Labov (2001: 334) provides a characteristically in-depth, if unusually funny, analysis of the facts (using some rather different notation):

BRITISH BROAD a

The classic example of such a lexical split in English is the British broad a class. Broad a is found before nasals in borrowings from French words originally spelled with au like dance and chance, or words spelt with a like lance, France, and plant — in the majority of monosyllabic words ending in clusters of /n/ plus a voiceless obstruent. It also occurs before front voiceless fricatives in half, laugh, bath, pass, past, with a pattern of phonetic conditioning independent
of etymology. Grammatical insights may help the second dialect learner formulate the rule that gives broad $a$ in bimorphemic *can’t* but not monomorphemic *cant* (*slang* or ‘tilt’), monomorphemic *class* but not derived *classic*. But no phonological, grammatical, or etymological rule will tell the second dialect learner that broad $a$ is used in *class* but not *mass* ‘substance, crowd’ both of French origin, though it is variable in ecclesiastical *mass* of Old English origin; in *pastor* but not *pastern*, both of French origin; in *plaster* from Old English but not *plastic* from Latin. A very adroit language learner may attack the situation before nasals by deducing that broad $a$ occurs only before clusters, never before a single final nasal. If the cluster begins with /m/, it must further be specified that as /mpl/, which yields broad $a$ in *sample*, *example*, but not *camp*, *lamp*, etc. Unfortunately, the learner must then learn to say *ample* as /æmpl/ and not /ahmpl/. The /n/ clusters are even more difficult, with broad $a$ in *demand*, *command*, *slander*, but not in *land*, *grand*, *pander* (Jones 1964). In each case, there is a tantalizing suggestion of an explanation — in some cases a hint of high style for broad $a$ versus low style form short $a$, old patterns for broad $a$ versus new ones for short $a$ — but these tendencies are only tendencies, explanations after the fact. The one true path for learning the broad $a$ class is to absorb it as a set of brute facts as a first language learner, or failing that, to be enrolled in a British public school in early childhood.

Labov (2001: 334)

6d.2 The FOOT-STRUT split

The FOOT-STRUT split is a little simpler, at least in terms of its consistent distribution across accents, if not in its lexical inputs. The split explains why *put* and *but* don’t rhyme in most accents (/pʊt/ v. /bʌt/) despite both words having a closed syllable spelt with <$u$>. The split may have been phonetically conditioned, at first, but even then it is hard to tease out the details, especially as both sets absorbed words from other sources, as can be seen in the table below where the spelling is a semi-reliable guide to the pronunciations before the split happened. This is why the sets contain several words spelt <$oo$>, mostly in the FOOT set (*foot*, *good*, *wool*), but some ended up in STRUT (*blood*, *flood*). In certain accents, a few words (*book*, *cook*, *hook*) did not move and have ended up in the GOOSE lexical set. Another point to note in Table 6.8 below is that /ʌ/ occurs before a much wider range of consonant phonemes than /ɜː/, at least in monosyllabic
words, hence the likelihood of the split being phonetically conditioned in its genesis.

The table below shows the orthographic complexity arising from the FOOT-STRUT split and the ensuing changes of lexical distribution. All four spellings occur for each of the sets and only certain sub-patterns can be deduced (see Carney 1994: 144–149).

<table>
<thead>
<tr>
<th>Spelling</th>
<th>STRUT lexical set</th>
<th>FOOT lexical set</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;u&gt;</td>
<td>cup, cut, suck, much, snuff, fuss, rush, rub, bud, jug, budge, buzz, hum, run, lung, dull, pulse, bulge, punch, lump, hunt, trunk; butter, study, punish, number, mustn't, Guthrie,…;</td>
<td>put, puss, bush, full. cuckoo, butcher, cushion, pudding, bullet,…;</td>
</tr>
<tr>
<td>&lt;o&gt;</td>
<td>done, come, love, mother, stomach, monk, tongue, onion, money, front,…;</td>
<td>woman, wolf, bosom,</td>
</tr>
<tr>
<td>&lt;ou&gt;</td>
<td>touch, enough, young, double, southern, country…;</td>
<td>could, should, would, couldn't…</td>
</tr>
<tr>
<td>&lt;oo&gt;</td>
<td>blood, flood.</td>
<td>good, stood, wood, wool; cook, look, book.</td>
</tr>
</tbody>
</table>

Table 6.8
The FOOT-STRUT split and the related spellings.
Examples taken from Wells (1982)54

6d.3 The LOT-CLOTH split
As with the TRAP-BATH split, this split was triggered by pre-fricative lengthening, and, this time, /b/ lengthened before /f/, /s/ and /θ/, as in the words off, Australia and cloth. However, the LOT-CLOTH split is rather different because it did not result in a new phoneme. Instead, CLOTH words adopted the same phoneme as the THOUGHT lexical set. In RP, this split did not occur to all words in the appropriate phonological environment and the split has been reversed for most present-day speakers. The split was much more important in GA, where a larger group of words changed, including Boston, coffee, wash, as well as words spelt <or>, as in origin, Florida, majority. In many American accents, THOUGHT and CLOTH words have since merged with PALM and LOT (see Table 6.1). Fuller lists of words in each group can be seen in Table 6.16.
6d.4 NG-coalescence: the *singer* – *finger* split

Not all phonemic splits occur to vowels. Present-day English /ŋ/ began as a conditioned allophone of /n/, occurring before /ɡ/ in words such as *sing*, *singer*, *singing* (Giegerich 1992: 297–301). Eventually, /ɡ/ was lost and [ŋ] became a distinct phoneme, albeit a phonotactically restricted one which cannot occur in syllable-onset position. The change occurred morpheme-finally, so monomorphemic words such as *sing*, *hang* and *belong* have a simple correspondence: `<ng#> ≡ /ŋ/`, as do polymorphemic words such as *singer* and *hangman*. However, in polysyllabic, but monomorphemic, words such as *finger*, *anger*, *(fish)monger*, /ɡ/ was not lost. In such cases, `<ng> ≡ /ŋɡ/`, and [ŋ] is functionally no different from a conditioned allophone of /n/. That is to say, the sound is simply assimilated to a succeeding velar consonant, just like the /ŋk/ of *ink*, *sphinx* and *angst*. The split is sometimes called the ‘*singer* – *finger* split’. The change did not occur in certain parts of Northern England and in these accents, [ŋ] remains an allophone of /n/ (Wells 1982: 188–189).

Complicating the picture is the fact that three comparative words, *longer*, *stronger*, and *younger*, never lost /ɡ/. This has led to some debate around the pronunciation of the rare word form *wronger*, synonymous with ‘more wrong’. Should *wronger* follow the pattern of the small number of comparatives which retain /ɡ/? Or should it follow the pattern of the derivational affix {er} occurring in *singer* and *hanger*? Both Pullum (Language Log 12/02/04) and Wells (18/01/07) have the intuition that the word should be pronounced /rɒŋər/, not /rɒŋɡər/. Further evidence, I suggest, comes from *cunninger* meaning ‘more cunning’, which surely has no /ɡ/. No pronunciation is suggested by the OED (45866).

Spelling /ŋ/ thus poses little trouble for the writer (see Carney 1994: 249–50). It is always spelt `<ng>` unless followed by /k/ or /ɡ/, in which case it is spelt `<n>` (as in *wink*, *anxious*, *income*; *finger*, *Angola*). However, decoding the phonemic correspondence of `<ng>` requires a deep knowledge of English morphological patterns, hence the split between bimorphemic *winger* and monomorphemic *finger*. Things are complicated even further due to words with the correspondence `<ng> ≡ /ndʒ/`. In monomorphemic words, the correspondence is often marked with final `<e>`, `<i>` or `<y>`, as in *singe*, *tangible* or *stingy* (loc. cit.). In polymorphemic words, however, this marking function may be compromised. The spelling gives no clue to the difference between *stranger* and *hanger* or *dingy* and *clingy*. The one allowance for this confusion, as we saw in Chapter 4, is that the retention of `<e>` in *singeing* and *bingeing* marks `<ng> ≡ /dʒ/` (as well as creating DSTNCT spellings).
Summary
The various phonemic splits discussed above are only an introduction to the large number of phonemic splits that have occurred across different varieties of English, primarily among vowels, but sometimes among consonants too. It should also be clear that splits may only affect certain varieties. NG-coalescence did not occur in parts of Northern England, and the same for the FOOT-STRUT split. The LOT-CLOTH split has had greater impact on GA than RP. The opposite is the case for TRAP-BATH. Regardless, the examples above show that phonemic splits cause much greater trouble in decoding than in spelling because one spelling unit (e.g. <a>, <u>, <o>, <ng>) now has two or more readings, depending on various, often indeterminable, factors. Towards the end of the chapter we shall see how conditioned splits occurring in non-rhotic accents, such as RP, have paved the way for further phonemic mergers. But before we can discuss such complex changes, we must first look at mergers.
6e. Mergers

A phonological merger occurs when two (or more) sounds converge on the one phonological space and a phonemic distinction is lost. Sounds that were once distinct become indistinguishable, a process that can happen to vowels, as in *meat* and *meet*, or to consonants, as in *three* and *free* (Foulkes and Docherty 1999; Stuart-Smith et al. 2007). Phonetically, mergers can happen in three ways. One phoneme might move into the phonetic space of another (merger by transfer); or they may 'meet halfway' (merger by approximation); or they may continue to occupy the same shared space but no longer be distinguishable (merger by expansion) (Labov 2001: Ch. 3). From the viewpoint of phonological systems, the most important thing is that mergers cannot be reversed by linguistic means. Once two sounds merge, entire speech communities have no way of reversing the distinction, except by taking recourse to the unreliable evidence of spelling. This is because mergers are ‘lossy’ while splits are ‘lossless’ (at least until they trigger further system changes). If your accent has a FOOT-STRUT split, it is easy to work out which words used to be in the old, unified set — all of the FOOT words plus all of the STRUT words (notwithstanding words such as *blood*, *good*, *book* which have jumped ship from a different source). Hence no information is lost in the split. However, if your accent doesn’t have the split, it is nigh on impossible to work out which words are in which set. It is much harder for a Northern English actor to learn the FOOT-STRUT split in a Southern accent than a Southern English actor to re-merge the split in learning a Northern accent. The same goes for TRAP-BATH and hence Labov’s explanation of the difficulties of knowing which words are in which set.

Mergers have a very different effect on the writing system to splits. For the reader, they are relatively easy — two different spellings map onto one sound, as with the homophonous trio *meat*, *meet* and *mete*. For the writer, however, they are problematic because different spellings are associated with the one sound and there may be no obvious reason why one word is spelt one way and an identical-sounding word is spelt another way: how is the writer to know if */miːt/ is to be spelt <meat>, <meet>, <mete>? It cannot be deduced (synchronically) from the workings of the current system and even a deep understanding of the language’s phonological history is not enough here, given that scholars still disagree on the phonemic status of these sets of words in Early Modern English (compare Lass 1999; Labov 2001: Ch. 13; and Stockwell 2002). The real advantage of maintaining ancient phonological distinctions in spelling is that words with different meanings remain visibly distinct and this allows for a more direct
relationship between spelling and meaning. Hence DSTNCT is a powerful force preventing reform on phonological grounds alone.

We shall begin this section by looking at several examples of mergers which have occurred since the stabilisation of English spelling, beginning with consonant cluster reduction (a special kind of merger), then vowel mergers and finally vowels before historic //r// which pose yet more challenges to the writing system. For ease of explication, the emphasis is on the homophones that arise from mergers.

**6e.1 Consonant Mergers**

**6e.1.1 Consonant-cluster reduction**

Some of the most abstruse English spellings are a result of consonant cluster reduction. Of the ones that occurred within English there is the KNOT – NOT merger, GNOME – NOME and WRAP-RAP. In each case, a consonant cluster has merged with a single consonant, and this has resulted in many homophonous pairs, which the spelling disambiguates. The <kn> group providing a large number.

<table>
<thead>
<tr>
<th>nap</th>
<th>knapp</th>
</tr>
</thead>
<tbody>
<tr>
<td>nave</td>
<td>knave</td>
</tr>
<tr>
<td>need</td>
<td>knead</td>
</tr>
<tr>
<td>Neil</td>
<td>kneel</td>
</tr>
<tr>
<td>neal</td>
<td>kneel</td>
</tr>
<tr>
<td>new</td>
<td>knew</td>
</tr>
<tr>
<td>nick</td>
<td>nick</td>
</tr>
<tr>
<td>nicker</td>
<td>knicker</td>
</tr>
<tr>
<td>night</td>
<td>knight</td>
</tr>
<tr>
<td>nit</td>
<td>knit</td>
</tr>
<tr>
<td>nob</td>
<td>knob</td>
</tr>
<tr>
<td>nock</td>
<td>knock</td>
</tr>
<tr>
<td>not</td>
<td>knot</td>
</tr>
<tr>
<td>no</td>
<td>know</td>
</tr>
</tbody>
</table>

**Table 6.9**

Homophones arising due to consonant cluster reduction\(^{58}\)

These mergers, it must also be noted, may only occur in word-initial position. In the case of /kn/ and /gn/, the velar consonant is retained in *acknowledge* and *agnostic*. Accordingly, the <k> and <g> are examples of inert consonant letters,
‘silent’ in one allomorph (knowledge, gnostic) and triggered in others (acknowledge, agnostic).

Other mergers have not taken place in all dialects, notably the whine–wine merger, which is complete in RP but not in Scotland, Ireland or America, although in many areas the distinction is highly recessive. The phoneme in whine words is transcribed either as the consonant cluster /hw/ or an unvoiced /ʍ/, a reminder that this peculiar phoneme is the lone survivor of a series of /hC/ clusters: Old English hlāf, hring and hnutu have become loaf, ring and nut in Modern English. The order of the <wh> spelling is anomalous. A small group of words with /h/ are spelt with <wh> as in who, whole and whore (on the history of <wh> spellings, see OED 228009).

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>whale</td>
<td>wail</td>
<td>what</td>
<td>watt</td>
<td>whales</td>
</tr>
<tr>
<td>wheeled</td>
<td>wield</td>
<td>whether</td>
<td>weather</td>
<td>where</td>
</tr>
<tr>
<td>whet</td>
<td>wet</td>
<td>which</td>
<td>witch</td>
<td>whine</td>
</tr>
<tr>
<td>whirled</td>
<td>world</td>
<td>whirred</td>
<td>word</td>
<td>whole</td>
</tr>
<tr>
<td>whoa</td>
<td>woe</td>
<td>whore</td>
<td>hoar</td>
<td></td>
</tr>
<tr>
<td>why</td>
<td>Y/ Wye</td>
<td>who</td>
<td>hoo(t)</td>
<td></td>
</tr>
</tbody>
</table>

**Table 6.10**
Homophones arising from the merging of /ʍ/ and /w/; three words are re-spelt with excrescent <wh>.

The occurrence of final <mb> is a strange mix. It occurs in some Old English words with final /mb/, such as comb and lamb, where the /mb#/ reduced to /m#/, leaving the final <b> as a residualism (OED 36737 / 105214). There is an inert <b> in bomb which is triggered in bombard. §Latinate plumb and plumber have borrowed the <mb> spelling but they retain the pattern of <mb#> → /m#/. The excrescent <b> of thumb has no etymological pedigree at all (OED 201489) unlike the more notorious examples of debt and doubt, whose original Latin form had /b/, before it was lost in Old French (OED 47935 / 57078). Unsurprisingly, an exception occurs, for some speakers: iamb has /mb/, along with its related form iambic.
Words like *lamb*, *know*, *gnaw* and *wreak* ought to pose no trouble to the experienced reader, who can simply ignore the ‘silent’ letter with no correspondence. However, these words are clearly of great trouble to the writer, who has no reason to expect silent letters to be tacked on, either at the beginning or the end.61

**6e.1.2 Imported §Greek consonant-letter clusters**

A rather different set of consonant-letter clusters have been imported in §Greek words such as *pneumonia*, *psychiatrist*, *pterodactyl*, and *mnemonic*. Unlike words with initial *<kn>*, *<gn>* or *<wr>*, these words have never had a consonant cluster in English speech — the spelling has simply been inherited from Greek via Latin transliteration.62 In theory, these words should pose no trouble to the experienced reader because there is a predictable mapping from spelling to sound, a fact that has often been forgotten in the literature. Albrow (1972) groups these spellings into his System 3, along with the likes of *<ch> → /ʃ>* , which is an ambiguous correspondence, since *<ch>* can also be /tʃ/ or /k/ (§Basic Chaucer, §Greek chaos and §French chauffeur). Nonetheless, predictability of correspondence must count for something in a systematic analysis, independent of its frequency of occurrence. These spellings are clearly of great trouble to the writer, and a knowledge of the etymology and/or the meaning of the constituent morphemes is a great help in spelling them. For example, the morpheme spelt *<psych>* re-occurs in *psychologist*, *psychotherapist*, *psyche*, *psychedelic*, etc. and it is clipped in *psytrance*; the word *pneumonia* is related to *pneumatic*; *mnemonic* to *amnesia*; and *pterodactyl* to *helicopter* and *archaeopteryx*. Again, the greater one’s knowledge of the lexicon and the spelling, the easier it gets, and the more information one can glean. That might not be of much consolation to the early learner but the writing system does reward experienced readers.

The final thing to take from the examples above is that once a unit gets into the writing system it can become productive, irrespective of its lack of phonetic transparency. This is especially the case for word-initial and word-final consonant clusters where an extra letter can reside *extra legem*, having no phonological correspondence. An example of this occurs in the name *Psmith*, P.G. Wodehouse’s fictional character, whose ludic name is homophonous with *Smith*. We shall see more examples of this kind of creativity in Chapter 7.
6e.2 Vowel mergers

Vowel mergers are more numerous and their histories harder to trace than consonant mergers, if only because of the usual problem in English spelling of having so many vowel sounds with so few vowel letters. Relying on orthographic evidence from earlier periods of the language is highly problematic (see Lass and Laing 2008–13 for an in-depth discussion of Early Medieval English) but there are more than enough examples occurring in the Modern period to illustrate the effect of vowel mergers on the writing system. The outline that follows is largely taken from Kruse’s (2016) account of instances where English spelling indicates distinctions that no longer occur in speech — beach, beech; bate, wait, weight; bird, heard, word, Kurd — but which may survive in isolated dialects. Examples are frequently lifted en masse from his work but augmented where necessary.

6e.2.1 The FLEECE merger

Part of the phonological restructuring known as the Great Vowel Shift was the merger of the mid-vowels /e:/ and /ɛ:/, or in terms of lexical sets, the merging of MEAT and MEET. The effect on the writing system was that <ea> and <ee> often took on equivalent value in what is now the /i:/ phoneme. This merger resulted in a huge number of homophones, often involving common words.

<table>
<thead>
<tr>
<th>beach</th>
<th>beech</th>
<th>peak</th>
<th>peek</th>
</tr>
</thead>
<tbody>
<tr>
<td>beat</td>
<td>beet</td>
<td>peal</td>
<td>peel</td>
</tr>
<tr>
<td>cheap</td>
<td>cheep</td>
<td>read</td>
<td>reed</td>
</tr>
<tr>
<td>creak</td>
<td>creek</td>
<td>sea</td>
<td>see</td>
</tr>
<tr>
<td>dear</td>
<td>deer</td>
<td>seam</td>
<td>seem</td>
</tr>
<tr>
<td>heal</td>
<td>heel</td>
<td>steal</td>
<td>steel</td>
</tr>
<tr>
<td>leak</td>
<td>leek</td>
<td>team</td>
<td>teem</td>
</tr>
<tr>
<td>meat</td>
<td>meet</td>
<td>weak</td>
<td>week</td>
</tr>
</tbody>
</table>

Table 6.11
Homophones arising from the merging of the MEAT and MEET lexical sets

This merger is complicated by the fact that many <ea> spellings are in the DRESS set (bread, leather, measure) and the spelling <lead> has a lexeme in both sets (follow the lead is FLEECE and lead pipe is DRESS). There are a few grammatical alternations which occur in both, e.g. mean, meant and read, read; and this is evidence of an earlier phonemic split between a long and a short
allophone of the one vowel. The alternation is spelt explicitly in \textit{keep} – \textit{kept}, \textit{sleep} – \textit{slept}. As a result, the correspondence between <\textit{ea}> and /i:/ is not as robust as that of <\textit{ee}> and /i:/.

Another group of spellings contributing to the FLEECE lexical set are those described by the notorious (and truncated) rule '<\textit{i}> before <\textit{e}> except after <\textit{c}>'. Words with <\textit{ie}> (\textit{field}, \textit{yield}) used to be in the MEET set while those with <\textit{ei}> were in MEAT (\textit{conceit}, \textit{receive}, \textit{seize}), along with §Latinate <\textit{eCe}> words (\textit{complete}, \textit{decent} and \textit{legal}). The modern spelling is not always a reliable indicator of which words used to be in which subset. For example, \textit{these}, \textit{Peter} and \textit{even} were in the MEET set, not MEAT (Wells 1982: 141).

One way or another, the reader is at an advantage over the writer because it is much easier to deduce an /i:/ pronunciation from the spellings <\textit{ee}, \textit{ea}, \textit{eCe}, \textit{ei}, \textit{ie}> than it is to remember which spelling is used in each word. A few words spelt with <\textit{ea}> ended up in the FACE lexical set — \textit{great}, \textit{break}, \textit{steak} and \textit{yea}, as well as the name \textit{Yeats} and sometimes \textit{Reagan}. This array of spellings is a result of the complex interactions of Middle English with French and Latin.

Adding to the complexity of the FLEECE lexical set are §French words such as \textit{machine}, \textit{police} and \textit{chic}, although this has not occurred through an English-internal phonological change.

\textbf{6e.2.2 The FACE merger and the WAIT-WEIGHT merger}

Words with <\textit{ai}> and <\textit{ay#}> spellings typically derive from a Middle English /ai/ diphthong and this set also includes some spellings in <\textit{ei}> and <\textit{ey}> (\textit{veil}, \textit{prey}). This was in contrast to words with /a:/, spelt <aCV> (as in \textit{lane}, \textit{lady}). The merger has resulted in another glut of homophones, as in:

<table>
<thead>
<tr>
<th>ale</th>
<th>aile</th>
</tr>
</thead>
<tbody>
<tr>
<td>bale</td>
<td>bail</td>
</tr>
<tr>
<td>made</td>
<td>maid</td>
</tr>
<tr>
<td>maze</td>
<td>maize</td>
</tr>
<tr>
<td>plane</td>
<td>plain</td>
</tr>
<tr>
<td>sale</td>
<td>sail</td>
</tr>
<tr>
<td>vale</td>
<td>veil</td>
</tr>
<tr>
<td>pray</td>
<td>prey</td>
</tr>
</tbody>
</table>

\begin{tabular}{|l|l|}
\hline
\textbf{Table 6.12} & \\
\hline
Homophones arising from the merging of earlier diphthongs with long front mid-vowels & \\
\hline
\end{tabular}
A subset of these words used to be followed by /x/, as in *straight, weigh, eight*, while others are spelt <gn> as in the §Latinate words *reign, deign, campaign*. Resulting homophones include:

<table>
<thead>
<tr>
<th>strait</th>
<th>straight</th>
</tr>
</thead>
<tbody>
<tr>
<td>wait</td>
<td>weight</td>
</tr>
<tr>
<td>slay</td>
<td>sleigh</td>
</tr>
<tr>
<td>rain</td>
<td>reign</td>
</tr>
<tr>
<td>Dane</td>
<td>deign</td>
</tr>
</tbody>
</table>

**Table 6.13**

Homophones arising from the merging of earlier diphthongs plus /x/, with front mid-vowels

At this point, we can appreciate the full length of the classic rule of spelling:

- <i> before <e>,
- except after <c>,
- and when sounding like ‘A’,
- as in *neighbour* and *weigh*.

<table>
<thead>
<tr>
<th>toe</th>
<th>tow</th>
</tr>
</thead>
<tbody>
<tr>
<td>moan</td>
<td>mown</td>
</tr>
<tr>
<td>no</td>
<td>know</td>
</tr>
<tr>
<td>nose</td>
<td>knows</td>
</tr>
<tr>
<td>road</td>
<td>rowed</td>
</tr>
<tr>
<td>groan</td>
<td>grown</td>
</tr>
<tr>
<td>sole</td>
<td>soul</td>
</tr>
<tr>
<td>throne</td>
<td>thrown</td>
</tr>
<tr>
<td>so</td>
<td>sew</td>
</tr>
</tbody>
</table>

**Table 6.14**

Homophones arising from the merging of /ɔ:/ and /ɔu/

**6e.2.3 The GOAT Merger**

Like the FACE set, there are two main input groups, both reflected in the spelling. The long Middle English monophthong /ɔ:/ has <o>-spellings (as in *rope, soap,*
noble, toe) while the diphthong /ɔu/ was usually spelt with <ow>, <oul>, <ol> or <ough>, as in blow, owe, soul, old, although.83

6e.2.4 Yod-dropping and the GOOSE lexical set
The GOOSE lexical set is more complex than others for two reasons. One is that it includes words with either /u:/ or /ju:/ in present-day English (moose, music) and the other is that many accents have lost that /j/ in certain phonological environments. The GOOSE set includes many §Basic spellings in <oo> (loop, shoot, boost, etc.), a few <oCe> (move, prove, lose, whose), some <o> (tomb, do, who, two) and some <ou> (group, youth, you). The other group all have a <u>: either simple <u..e> in §Latinate words (mute, rude, funeral) or <u..V>, as in (lunatic, Juno); while other spellings indicate two distinct rising diphthongs which have merged on /ju:/, spelt either <eu> (feud, neutral) or <ew> (newt, lewd, few) and <ui> (fruit, juice), <iew> ((re)view), and finally the <eau> of beauty, beautiful (Wells 1982: 147–9).

Yod-dropping affects certain classes of consonants: /j/ is generally retained in all accents after labials, velars and /h/ (as in mute, cute and human), except traditional accents in East Anglia (ibid: 206–8). In American accents, /j/ is dropped after all coronals (tube, duty, new, enthusiast, suit, Zeus, lute), giving rise to the homophones do and dew, loot and lute (Kruse 2016). Traditionally, British accents would have /j/ in all of these environments but less so nowadays, so it is retained in tube, duty and new but perhaps variably in enthusiast, suit, Zeus and lute.

6e.3 Loss of vowel distinctions before /r/
Some of the most remarkable changes in English phonology have involved /r/ in syllable rhymes. The loss of historic /r/ in syllable rhymes has resulted in many accents becoming non-rhotic (a slightly misleading term as /r/ survives in syllable onsets), and this has precipitated many further changes in the phonemic systems of such accents. However, another very important change, particularly from the viewpoint of the writing system, has been the significant reduction in the number of vowel contrasts before /r/. We will begin with these, as they have affected all accents of English.

6e.3.1 The NURSE mergers
Perhaps the most drastic change, from the viewpoint of the writing system at least, has been the convergence of Middle English /ɜ:/, /ɛr/ and /ɛːr/ as one single sound [ɔ] (known as ‘r-coloured schwa’). These distinctions remain indicated in the
spelling, as in *irk, jerk, earl, Turk*, and *work*, but the merger has resulted in only a few homophones, such as:

<table>
<thead>
<tr>
<th>birth</th>
<th>berth</th>
</tr>
</thead>
<tbody>
<tr>
<td>pearl</td>
<td>purl</td>
</tr>
<tr>
<td>serf</td>
<td>surf</td>
</tr>
</tbody>
</table>

**Table 6.15**
The Nurse Mergers:

homophones arising from the merging of */w/, */r/ and */ur/*

The nurse mergers have had no effect on reading but they make spelling a lot more problematic, as one has no way of knowing whether [ɚ] is spelt <er>, <ir>, <ear>, <ur>, as in *her, fir, heard, hurl*, etc.

### 6e.3.2 The FORCE mergers

The *NORTH-FORCE* distinction, between */ɔr/ and */ɔr/ is sharply recessive throughout the English-speaking world and Wells no longer considers it a feature of GA (compare Wells 1982 and 2008). It does survive in many conservative accents, irrespective of rhoticity, in much of Ireland and Scotland, the North of England, parts of the Caribbean (with further developments), and sporadically in Eastern parts of the US (Labov, Ash and Boberg 2006: 49–52).

The spelling <or> occurs for both vowels and while there is a strong tendency for it to appear before consonants in *NORTH* words (*short, orb, morpheme*) and syllable-finally in *FORCE* words (*before, wore, oral, Tory*), this is not a neat split. The words *or, for* and *nor* are all *NORTH* while *FORCE* includes *sport, afford, worn*. The confusion is best exemplified by the fact that *import* is in the *FORCE* set but *important* is in *NORTH*.

All words spelt <oar>, <oor> and <our> are part of the *FORCE* set, as in *board, door* and *mourn*, and they reflect distinct vowel sounds in Middle English which have since merged. *NORTH* contains all words in <aur>, such as *aura, aural, Laura* and *Taurus*. One result of this is that the words *aural* and *oral* ought to be homophones, although *aural* is often pronounced /ˈɔːrəl/, with the *MOUTH* vowel, presumably on account of its spelling.
**Table 6.16**
The NORTH-FORCE mergers: homophones arising from the merging of /ɔː/ /ɔːr/ and /ɔːr/

<table>
<thead>
<tr>
<th>NORTH /ɔː/</th>
<th>FORCE /ɔːr/</th>
</tr>
</thead>
<tbody>
<tr>
<td>border</td>
<td>boarder</td>
</tr>
<tr>
<td>for</td>
<td>four</td>
</tr>
<tr>
<td>horse</td>
<td>hoarse</td>
</tr>
<tr>
<td>morning</td>
<td>mourning</td>
</tr>
<tr>
<td>or</td>
<td>ore / oar</td>
</tr>
<tr>
<td>war</td>
<td>wore</td>
</tr>
<tr>
<td>warn</td>
<td>worn</td>
</tr>
</tbody>
</table>

A further development occurs in the ongoing merger of the CURE lexical set with these words. This is better detailed in the literature on non-rhotic accents, as we shall see below, but even for many rhotic speakers *poor* has become homophonous with *pour* and *pore* (Hickey 2014).

**6e.3.3 Vowels before historic //r//: in non-rhotic accents**
The loss of historic //r// in syllable rhymes has radically changed the vowel system in RP and other accents which have undergone this change. This has greatly increased the number of vowel phonemes in the inventory, although the total has been tempered somewhat by the ensuing merger of existing phonemes with these new, complex-vowel phonemes. Generally, the changes begin with a conditioned split before the resulting vowel becomes a distinct phoneme. In the START lexical set, for example, the sequence //ar// changed to the complex vowel /ɑə/ and later became a long, simplex vowel /ɑː/: In doing so, the vowel has merged with the BATH and PALM sets. This explains why *giraffe, father* and *far* all have the same vowel. Once again, this merger is manageable for the reader but difficulty arises for the writer who must remember that the name *Bart* has an <r>, unlike *bath*. NORTH and FORCE have long since merged (in RP) and their reflex /ɔː/ has gone on to merge with the /ɔː:/ phoneme of the THOUGHT lexical set, rendering the homophones *shore* and *Shaw*. The latest addition to this group is the CURE set, making *sure* homophonous with *shore* and *Shaw*, a merger that is all but complete in RP. The NURSE merger has had little effect on the vowel systems of RP. Words such as *irk, jerk* and *Turk* have converged on a long central vowel /ɜː:/ which has stabilised in the middle of the phonological space, and this does not impact on other vowels.
Some of the complex vowels have survived. Words in the near lexical set and (less so nowadays) the square lexical set still have centring diphthongs, as a reflex of underlying //ɪə//, and this is made explicit in their notation as /ɪə/ and /eə/. They are often considered distinct phonemes, on account of their merging with vowels of Ethiopia and Eritrea, respectively. A final group involves what Jones calls ‘triphthongs’ where a diphthong is followed by schwa [ə] within the same syllable (Jones 1918/1967: 30, §233, §234). Examples include layer /ɛɪə/, mower /əʊə/, via /aɪə/, sour /aʊə/ and soya /ɔɪə/ (see also Roca 2016). All of these are exemplified in Table 6.2.

6e.4 The weak vowel mergers

Even more drastic than the changes to vowels before /r/ has been the almost entire loss of distinction among reduced vowels. Many accents have just one reduced vowel but some retain two, so that abbot and rabbit don’t rhyme (/æbət/ and /raɪt/), and Lennon and Lenin and are not homophonous (/lenən/ and /lenɪn/). In the word remember, there are two unstressed vowels: in rhotic, GA accents, with no LennON-LenIN distinction, it is simply /raɪˈmembər/; in non-rhotic, RP accents, it is /rɪˈmembər/, with the first unstressed vowel being in the LenIN set, and the second being a special case of LennON. This is known as the letter lexical set, and it occurs before historic //r//, as in teacher, professor, humour, martyr. In non-rhotic accents, this is identical to the commA, lexical set occurring in non-§Basic words such as beta, sofa, saliva.

A final group is called the happy lexical set, distinct from commA, and it includes words such as sorry, taxi, movie, coffee. This used to be transcribed as /ɪ/, like KIT words, but Wells (2000, 2008) transcribes it as /i/, with no length marker (like the /u/ of virtual). This transcription partly reflects a tendency for tensing this vowel to [i], a process known as happy-tensing (Wells 1982: 165), and partly the neutralisation of the distinction between /i:/ and /ɪ/ in reduced vowels (as in vary and various, see also Wells 7/6/12).64

The loss of vowel distinctions can occur for any vowel sound making it impossible to reconstruct the spelling from the sound alone. For example, the <a> of democracy /dɪˈmɒkrəsɪ/ cannot be inferred from the sound of the word; one must either remember the spelling or observe the /æ/ in democratic /ˈdɛməˈkrætɪk/. Similarly, the <i> of hypocrisy can only be retrieved from memory or by comparing the second stressed /ɪ/ of the related word hypocritical. This situation is extremely problematic for writers who cannot rely on their speech as even a rough guide to spelling. Such vowel reduction can correspond to any vowel letter or digraph: from this paragraph alone we have alone,
problematic, distinction, second, etc. The problem is particularly acute among complex words (e.g. hydrometer and arbiter; physician and repetition, etc.) and Carney (1994: 417–442) examines the issue in enormous detail. An awareness of such connections between morphemically related words would be of huge help to learners of the writing system.

Finally, the existence of so much homophony among the last syllable or two of words is another reason why there is a general tendency for affixes to be spelt consistently across the writing system. For example the spelling <ese> is used for languages, including Chinese, Japanese, Portuguese, and by extension, for writing styles such as journalese, textese, etc. By contrast, the spelling <ees> occurs in plural nouns such as employees, fricassee, etc. Such a distribution undoubtedly helps experienced readers to establish meaning with greater ease, although the variation poses difficulties for writers who must remember which spelling to associate with which semantic group.

***
6f. Conclusion

Some kinds of phonological change affect the writing system in more manageable ways than others. Chain shifts among vowels tend not to affect the structural relations in letter-phoneme correspondences and this is why there was no need to change the spelling of words such as boat, boot and out, during the Great Vowel Shift. It also explains why oft-cited spelling pairs such as <profane – profanity>, <serene – serenity>, <divine, divinity>, <cone – conical>, have continued to function for such a long time and across so many varieties of English. However, the retention of the same spelling in these words relies on the function of final <e> to mark the value of the preceding vowel letter.

Unfortunately, splits and mergers cause radical discrepancies to the writing system’s sets of correspondences, and this can affect readers and writers in different ways. Phonemic mergers are more manageable for the reader than the writer. The situation is reversed for splits. Mergers cause two spellings to represent one sound and this often results in homophones, including meat, meet; way, weigh; wrap, rap; and knot, not. Such spellings cause little trouble to the experienced reader who can map both spellings onto the same sound and yet associate each spelling with its meaning. Writers, however, only have one sound (in their heads) and must remember which spelling to associate with which meaning. In theory, the reverse is the case for splits. This time, the writer should be able to use one letter (or set of letters) for two different sounds (e.g. <ng> for both /ŋ/ and /ŋɡ/) but the reader may not be able to tell which sound to associate with which spelling (e.g. singer and finger). An understanding of the morphological structure of the words is essential in such cases. Yet deep knowledge of the language may not always be enough to explain a split, as Labov (2001) explains with regard to the TRAP-BATH split.

These complications do have their advantages, however. When phonemic splits correspond with morphophonological variation, as in the case of electric, electricity and electrician, then there can remain accord between the spelling and morphological structure, despite variation at the phonemic level. Thus the CPU <electric> has the same basic meaning in each word even though its pronunciation varies. Phonological mergers create a sense of equivalence between two spelling units which represent the same sound, such as <ee> and <ea>, and hence these units can be swapped in creative or DSTNCT spelling. For example, The Beatles deliberately respell the word beetle by swapping <ea> for <ee> and, while this indicates no change in the pronunciation, it creates a new visual meaning for the
such spelling changes, many of which result of complications in the writing system, will be the focus of Chapter 7.

Every English accent will have its own set of complications due to its unique set of phonological changes which have taken place since English spelling stabilised, but the evidence provided here shows that the complications can be better understood when subcategorised according to each kind of phonological change. The examples presented here also provide systemic evidence of Venezky’s insistence (1970: 120) on the importance of letter position and phonological environment for the reader looking to decode English spellings. The occurrence of word-initial <gn> will have a different reading to word-final <ng>, and the same goes for <rV> and <Vr> words, etc.

Finally, we must look at the deeper implications of the foregoing discussion on writing systems research in general. The focus in this chapter has been on the gap that grows between invariant spellings and changing phonologies, what DeFrancis (1989) calls ‘the dead hand of the past’, what he considers to be a third universal of writing systems, separate from the need to represent sound and meaning. However, the dead hand of that past is not really separable from the other two; it simply accounts for the gap that grows between them. Furthermore, that gap only grows in one direction because spellings almost never change, once they are settled upon. Of course, we have seen a few examples of variable, and thus changing, spelling, including <gaol>/<jail> and <focussed>/<focused>. However, we shall assume for now that such examples are negligible in the wider scheme of things (they may not be — the present study has only investigated isolated examples, precipitated by observed patterns. A fuller, more systematic study could explain in much greater detail how much variation actually exists in English spelling). Thus, assuming English spellings are invariant, we can say that, over time, their degree of morphography stays the same but their degree of phonography can be significantly reduced. If we define a spelling’s degree of phonography as a measure of how predictable its spelling-to-sound correspondences are, in both directions, then we can comfortably say that phonemic splits and mergers reduce its predictability. Hence it is harder to know how to read <put> than <dress>, since <e> ≡ /e/ in closed stressed syllables, but <u> ≡ /ʌ/ or /ʊ/. The situation is best exemplified by the existence of so many homophones which have arisen due to phonological mergers. Numerous examples have been presented throughout the chapter, including knapp, nap; wrap, rap; gneiss – nice; feet – feat; groan – grown; berth – birth, etc. Many homophones are specific to certain accents or groups of accents (e.g. shore – Shaw – sure). In all of these cases, the words have lost a certain degree of phonological
transparency and yet none of them has lost their degree of morphography. That is because all invariant spellings of words are complex pleremic units and this is independent of their ever-diminishing ability to provide clear phonographic matches for their ever-changing phonological forms.

***

***
7. The Distinctiveness principle (DSTNCT)

This chapter examines an important ramification of having a complicated writing system – the fact that it becomes possible to have different spellings for two words or names which sound alike. We saw the example of *The Beatles*, where `<ea>` has replaced the `<ee>` of the lexical word *beetle*, allowing the pop band to have a distinct orthographic identity, and yet the spelling suggests the same phonological form as the lexical word. Such creativity is a manifestation of the strong tendency in the English writing system for there to be one distinct spelling for each distinct meaning. In most cases, this principle is so obvious that it is scarcely worth mentioning, and hence the words *cat* and *dog* are not spelt alike, nor are *giraffe* and *electricity*. The importance of DSTNCT becomes clearer in the non-trivial situation where two different words happen to sound the same. Having DSTNCT spellings allows us to distinguish the two words by sight, even though they can’t be distinguished by sound. We have seen many cases where this occurs due to accidental changes in phonology (e.g. *beat* and *beet*), but it can also happen by design. Hence the names *Kit-Kat* and *Snoop Dogg* are spelt differently from *cat* and *dog*. In this way, DSTNCT can be viewed as the converse of IdP. Instead of related words having related spellings (IdP), unrelated words must have DSTNCT spellings.

The focus in this chapter will be on the several ways in which DSTNCT is manifested in the writing system. The first part of the chapter looks at its occasional occurrence among lexical words. This can occur in homophonous pairs, such as *flour* and *flower*, where the spelling allows us to distinguish the two. Unfortunately, such distinctions are not applied systematically across the system.

DSTNCT is much more common among names because they frequently derive from lexical words and thus there is a practical need for the spellings to be individuated. The simplest example of this involves the capitalisation of all proper nouns. Yet there are several kinds of names which are particularly prone to DSTNCT spellings, including surnames, trade names and band names (meaning the names used by popular musical artists). These will be discussed based on available evidence provided in previous studies (e.g. Carney’s 1994 analysis of English surnames; Jacobsen’s 1966 and Praninskas’s 1968 analysis of trade names; and Ryan’s 2010 account of band names).

Unlike normal words, the spelling of a name is not always decided by use, consensus or dictionary compilers; often the spelling can be safeguarded by those
whom the name refers to. And the degree to which a name can be invented by the user seems to relate to their ownership of the name. If one invents a business or a band, one can choose its spelling. Spelling one’s surname differently would be a lot more contentious, as it also belongs to the rest of one’s family. Changing the spelling of a place name might require the consent of an entire community, so place names are akin to lexical spellings in that respect. Accordingly, we see huge creativity in the spelling of trade names and band names but much less in surnames, where the range of variant spellings is much smaller.

Throughout the chapter, we shall see that there are some very strong constraints on what spelling changes can be made in order that the same phonological form be recuperable from the new, DSTNCT spelling. For example, the band <Phish> is a viable respelling of fish unlike the infamous spelling <ghoti>. This is because spelling correspondences are so often reliant on their relative position within a word: <gh> ≡ /f/ only occurs word-finally (laugh, cough), and <ti> ≡ /ʃ/ occurs in graphematic dactyls such as < (na.ti.on) >. In the spelling of names, it is often helpful to have a DSTNCT spelling, in order to create a distinct orthographic identity. It is therefore not always necessary, or even desirable, to satisfy lDlP.

We shall see that changing the spelling of a word or name successfully requires a detailed understanding of how the writing system functions because changes must be confined to ones that readers can decode as corresponding to the same phonological form (i.e. Beatles and beetles indicate the same form). These changes also highlight which spelling patterns are productive and which aren’t. It is for this reason that names can be seen as an essential element of a theory of spelling formation, even though they have only even been peripheral in previous studies of English spelling. The discussion of trade names and band names is further contextualised by a section which provides an updated taxonomy of non-standard spelling. Here we shall see how some non-standard spellings introduce new linguistic information while others don’t. Names which derive from the lexicon but which have DSTNCT spellings form something of a bridge between standard and non-standard spelling. The spelling of the name may be the only spelling in use, but the orthographic techniques occurring therein are much the same as the techniques used in certain kinds of non-standard spelling.

It should be noted that recent findings on the spelling of affixes, outlined by Berg and Aronoff (2017), confirm the basic premise of this chapter, which is that each word (or morpheme) ought to have one DSTNCT spelling. Their argument follows the logic of that idea even further. Not only does each morpheme have its own spelling, but words with homophonous endings ought to
have a spelling which does not coincide with the spelling of the affix. Hence <ic#> is reserved almost entirely for adjectives (atomic, Germanic) while other spellings are used for non-adjectives (gimmick, limerick, buttock, eunuch etc.). This idea also tallies with Carney’s (1994: 130) observation that final <e> is used in words such as tense and browse in order to distinguish them from the plural nouns tens and brows. Carney devotes an entire chapter to the topic of homophones and homographs (1994: 395–442), and it would certainly be helpful to examine his data in light of Berg and Aronoff’s recent findings. However, that task has proved beyond the present investigation.

Before beginning, it should be pointed out that distinctiveness is not quite the same as uniqueness, even though there is a great deal of overlap between them. Unique spelling requires a strict one-to-one relationship between spellings and their referent, a situation that occurs in, for example, car registration plates or airport codes, where it is not permissible (for good reasons) to have two occurrences of the one written form. Distinctiveness simply requires the spelling to be different from an existing, often homophonous, spelling. For example <bye> is distinct from <by>, but <bye> remains ambiguous between a bye-law and the clipped form of goodbye. Distinctiveness is a practical solution for a functional writing system in a way that uniqueness is not. Yet uniqueness may be required in restricted domains of the writing system, such as the spelling of trade names, which must be spelt uniquely in America, thanks to a law passed in 1946 (Toulmin 1946). This law has precipitated some unique spellings which are formed in order to be memorable and decodable.
7a. DSTNCT among lexical spellings

7a.1.1 Semantic specialisation
We saw throughout Chapter 6 that phonological mergers can accidentally result in homophonous pairs of words with different spellings, as in meat, meet; way, weigh; knob, nob; wreck, reek, etc. There also exist a few homophones that emerge from different subsystems, as in shoot, chute; fisher, fissure; or perhaps catarrh and Qatar. However, the remainder of this chapter will involve looking at the rather different case where differences of meaning have precipitated deliberate changes in the spelling, despite the words continuing to have the same pronunciation.

There are several cases where a semantic split between once-related words results in a deliberate attempt to have two DSTNCT spellings. Carney (1994: 413–6) provides numerous examples, including metal and mettle; flower and flour; discrete and discreet; plain and plane; or check and cheque. A contemporary formation has occurred between bite and byte (in computing). Other cases have not settled down fully or gained sufficient traction, including set and sett, reflection and reflexion (loc. cit.). There are two ways this can happen – two variant spellings can become associated with two different meanings, as in <flour> and <flower> (Scholfield 2016: 153); or an entirely new spelling can be invented, as in <byte>. In some cases, the new spellings have given rise to different pronunciations, as in §Latinate ‘artist and §French ar’tiste, or ‘local and lo’cale. Both <artiste> and <locale> choose the spelling of the French feminine form (with <e>), spellings which helps to preserve both a distinct meaning and a distinct pronunciation (OED 11239 / 109551).

Such cases are isolated and there are no systematic principles for the application of DSTNCT to semantically unrelated homophones. As a result, the writing system continues to make do with a very large number of homographs with very different meanings, such as <nail>, <bank> and <lie>. Hence DSTNCT does not play a large role in the spelling of homophones in the core lexicon and distinctive spellings are haphazard rather than systematic. This is perhaps why DSTNCT has seldom been treated as a fundamental principle of English spelling. However, there are two cases where the principle is more carefully applied: in the so-called ‘three-letter rule’ and in the use of capital letters for proper nouns.

7a.1.2 The three-letter rule
The ‘three-letter rule’ of English spelling requires lexical words to have three letters but not grammatical or function words. The rule explains why me, we and
he can have just two letters but see, fee and pee have to have three (see Jespersen [1909] 1928: §4.96; Carney 1994: 131–4; Venezky 1999: 86; Cook 2004: 56–7; Evertz 2014: 161–6; Berg 2016: 456). A very useful by-product of this rule is the possibility of creating an orthographic distinction between pairs such as we and wee or by and bye. When stressed, the following pairs are homophonous, although some vowels may differ across accents (these are marked here with ‘%’).

<table>
<thead>
<tr>
<th>to</th>
<th>too</th>
</tr>
</thead>
<tbody>
<tr>
<td>we</td>
<td>wee</td>
</tr>
<tr>
<td>be</td>
<td>bee</td>
</tr>
<tr>
<td>by</td>
<td>bye</td>
</tr>
<tr>
<td>or</td>
<td>ore%</td>
</tr>
<tr>
<td>I</td>
<td>eye/aye%</td>
</tr>
<tr>
<td>in</td>
<td>inn</td>
</tr>
<tr>
<td>no</td>
<td>know</td>
</tr>
</tbody>
</table>

Table 7.1
The three-letter word rule, functional and lexical words compared

<table>
<thead>
<tr>
<th>ad</th>
<th>add</th>
</tr>
</thead>
<tbody>
<tr>
<td>od</td>
<td>odd</td>
</tr>
<tr>
<td>of</td>
<td>off*</td>
</tr>
<tr>
<td>as</td>
<td>ass*</td>
</tr>
<tr>
<td>aw</td>
<td>awe</td>
</tr>
<tr>
<td>ow</td>
<td>owe</td>
</tr>
<tr>
<td>pi</td>
<td>pie</td>
</tr>
</tbody>
</table>

Table 7.2
The three-letter word rule, non-function words and lexical words compared

There are other cases involving spellings which do not adhere to the division based on grammatical function. Some of the two-letter spellings are §Exotic or onomatopoeic or clipped forms of longer words. The asterisked cases differ for voicing, and again vowels may differ substantially.

Most of the three-letter spellings contain letters that are redundant from the viewpoint of PhM, thanks to consonant-letter doubling (inn, add, odd, off,
odd); vowel-letter doubling (bee, too); or final <e> (owe, awe, aye, bye, eye, pie, ore). Carney calls this ‘padding’ (1994: 454–458) and notes its frequency among the spelling of surnames, as a way of making the name have a distinct spelling (see Section 7b.2 below). Another way of explaining it is to say that existing spelling correspondences are redeployed in order to create spellings that obey both DSTNCT and PhM. However, the three-letter rule is not motivated solely by DSTNCT, and the details are explained in greater detail by Evertz (2017, 2014: 161–7).66

***
7b. DSTNCT in the spelling of traditional names

This section begins by looking at the use of capital letters in proper nouns, a simple and systematic way of making DSTNCT spellings. The next part looks at the spelling of surnames. These are useful to begin with because the small number of spelling variants occurring therein will allow us to see how creative spelling can be explained within the model of spelling-sound correspondences developed in Chapter 5.

7b.1 Capital letters: systematic DSTCNRNT spelling

If the purpose of writing were simply to record speech, as Daniels (1996) argues, then there would be no need for capital letters. Capital letters are graphic entities and they can be used to add a layer of meaning independent of the phonological form. Whether a word is spelt with Capital Letters, SMALL CAPS or in CaMeL CaSe, the mapping from spelling to sound is always the same, yet sometimes the meaning is different. Capitals have several uses, and they do not always indicate semantic information. They mark the beginning of a written sentence and they are used to mark various kinds of important words, including proper nouns, days of the week, months of the year and festivals such as Christmas, Hallowe'en and Hajj; they mark languages, nationalities, religions, titles and headings. Newspaper headlines and notices are often written entirely in capital letters, as are initialisms, as in EU, FIFA and UEFA. Content words are capitalised in the titles of books and sometimes Important Words are capitalised in High Culture, a survival of an outdated trend for all nouns to be capitalised, as they are to this day in German.

A useful by-product is that capitals can help us distinguish visually between homophones, a situation that arises in names derived from the lexicon. The English football club Arsenal is always capitalised, unlike an arsenal of weapons, the word which gave the club its name. Similarly, we can distinguish between turkey the bird and Turkey the country – because the spellings look different. Anderson (2007: 191–3) gives the example of The Channel where the capitalised form refers to the body of water between England and France, whereas the channel can refer to many things. This also occurs in surnames, where we can separate Cook and Butcher from their lexical counterparts. Nevertheless, the distinction between the name and the lexical word is lost at the beginning of a sentence.
Capitalisation can be applied systematically across the writing system because all letters have a lowercase and capital variant. This makes conditioned variation possible, such as capitalising all proper nouns, and it means the meaningful distinctions crop up as a by-product (hence *arsenal* v *Arsenal*). Such systematicity is not possible in the other cases of distinctiveness outlined in the rest of this chapter, because interchangeable units of spelling are fortuitous rather than systematic. For example, *<f>* and *<ph>* are somewhat interchangeable spellings of /f/, at least from the viewpoint of phonology, but other phonemes don’t have such pairings. Of course vowel-letter sequences do afford more potential substitutability than consonants but, even then, there is a lack of symmetry. For example, *<ee>* can be replaced by *<ea>* but *<oo>* can’t be replaced by *<oa>*.

### 7b.2 Surnames

Variation in the spelling of surnames, such as *Cook* and *Cooke*, is primarily a relic from an earlier time when English spelling was more varied, as in the time of William Shakespeare, whose surname might be spelt as *<Shaksper, Shakspear, Shakspeare, etc.>*. However, McKinley points out that there has been some further tinkering, especially in the Victorian period.

During the Victorian period minor changes in spelling were sometimes deliberately made to distinguish one branch of the family from another or to improve the appearance of the name to give it what was thought to be a more aristocratic ring. The fluctuations to be found at earlier periods are often much greater. It must be remembered that even as late as the nineteenth century there were still many people who were illiterate to the extent that they could not spell their own names. Entries in parish registers were made by parsons and parish clerks from what they knew, or could hear, of the pronunciation of any surnames, and not surprisingly this produced a variety of spellings, some of them eccentric.

McKinley (1990)

Carney’s (1994: 433–466) study of the spelling of names reveals that, among surnames, the same set of spelling changes occur repeatedly. The spelling is often fleshed out with CLD and/or an extra *<e>*, a practice that he calls 'padding' (1994: 454–8). Examples include *Webb, Wilde* and *Cuffe*, which appear to derive from variant spellings of the lexical words *web, wild* and *cuff*. Padding is by no means confined to such words, as can be seen from *Alsopp, Arnott, Askrigg.*
The other kind of change is substitution, which mostly involves using <y> instead of <i>, as in Traynor, or <ie#> instead of <y#>, as in Archie and Smellie. Substitution is often used in conjunction with padding and all three kinds of change take place in Thynne and Wynne, homophonous with the words thin and win.

Padding is not just a marker of distinctiveness but one of distinction. Families keep the spelling of their name to help preserve their identity and, as Carney wryly observes (1994: 44–53), ‘archaism has a value’, and a certain kind of prestige may be associated with having an unusual spelling and, especially, a peculiar set of spelling-sound correspondences. This is an idea we shall see throughout this section on names, because having an unusual or creative spelling gives you a distinct visual identity. Uniqueness is essential in trade names and band names, but among surnames it is neither necessary nor fully desirable. It is not necessary because one's surname is often rare enough to function uniquely within one's milieu, and it's not always desirable because if people can't remember how to spell your name then things can get annoying or impractical. The balance seems to come in having a spelling that looks different enough to be distinct but normal enough to be memorable.

7b.1.1 Constraints on the spelling of surnames
Among the changes that do occur in surnames, the position of the changes is highly restricted, and almost all of them adhere to the following principle:

- Changes to the spelling of a surname never suggest a new pronunciation.

7b.1.2 <i> → <y>, <y> → <ie>
This pair of changes is particularly easy to make: replace word internal <i> with <y> or, conversely, replace word-final <y> with <ie>. This gives Traynor, Taylor, Payne, Wynn, Archie and Smellie.

7b.1.3 CLD-padding
The distribution of CLD-padding is relatively simple, although the phonological environment is slightly broader than what we saw in the spelling of inflected forms throughout Chapter 4. CLD occurs in /VC#/ rhymes, but it can also occur in unstressed syllables:
Stressed /VC#/ rhymes:
Chipp, Chubb, Kitt, Dodd, Clegg, Grimm, Scragg, Flynn, Carr etc.

Unstressed /VC#/ rhymes:
Arnott, Beckett, Cockshott, Leggatt, Kellogg.

Carney suggests the frequency of padding in polysyllabic names with /t/ may result from efforts to avoid the name having a §French reading, a problem that the character Hyacinth Bucket, of the television comedy Keeping up Appearances, is not so fortunate to have. She continually has to explain that her name is not like the §Basic word, *bucket*, but like the §French word *bouquet*: /bu’keɪ/.

Iambic words, with an unstressed syllable before a stressed syllable, are characteristic of §French words, including < bou(quet) >, (with §French <et> ≡ /eɪ/). Lady Hyacinth would like her surname to be analysed as *< bu(cket) >*, again with §French <et> ≡ /eɪ/). Unfortunately, she misses the fact that the principle of onset maximisation cannot be applied to <bucket>, since <ck> is not a valid graphematic onset cluster: hence people read it as < (buc.ket) >. The use of final <tt> in surnames confirms that these words have §Basic correspondences. However, they do also suggest final stress, due to the long graphematic rhymes of the final g-syllables, and could easily be decoded as < (Ar)nott >, < (Bec)kett >, < (Cock)shott > and < (Leg)gatt >.

In the other cases of CLD padding, the extra consonant letter has no possible effect on the analysis of the spelling, and thus no effect on the pronunciation. What we don’t see is CLD-padding in names such as Jones, White and Mason, where *<Jonnes>, *<Whitte> and *<Masson> would suggest different pronunciations. Once again, the implication is that the pronunciation can be worked out from the spelling.

7b.1.4 <e>-padding
<e>-padding (in surnames) is more complicated than CLD-padding because there are more orthographic environments in which it can introduce an unwanted phonological function, and more where it can have no effect. The simplest problem case is changing <VC#> to <VCe#>: it would not make sense for *Wilson* to be spelt *(wil)so.nes* lest it suggest */wl'sən/ or */wl'sən/, with the analysis as < '(wil)(so.ne) or <(wil)'(so.ne)>. In other cases, <e>-padding might introduce an unwanted marking function: if *King* were spelt *(Kinge)* it might be rhymed with *binge* (/-'ndʒ/).
Working from Carney’s extensive list of examples (loc. cit.), <e>-padding can occur in the following contexts (italicised names do not have lexical counterparts):

<table>
<thead>
<tr>
<th>Spelling</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;VC₁C₁e&gt;</td>
<td>Webbe, Wynne; Legge; Cuffe, Radcliffe</td>
</tr>
<tr>
<td>&lt;VC₁C₂e&gt;</td>
<td>Wilde, Aske</td>
</tr>
<tr>
<td>&lt;Vy/we&gt;</td>
<td>Toye, Rowe, Bowe</td>
</tr>
<tr>
<td>&lt;Vy/wCe&gt;</td>
<td>Payne, Doyle, Lansdowne</td>
</tr>
<tr>
<td>&lt;VVCe&gt;</td>
<td>Freake, O'Keefe; Cooke, Goode</td>
</tr>
</tbody>
</table>

Table 7.3
Various kinds of <e>-padding in surnames

7b.1.5 <VC₁C₁e>
In <VC₁C₁e> spellings, there is a combination of CLD and <e>-padding. Phonologically, they all have /VC#/ rhymes which results in <VC#> lexical spellings, as in leg, web and win, but <ff> in cuff and cliff. These last two can have padded <e> without it marking the vowel as long, hence Cuffe and Radcliffe, analysed as <(Cuf.fe)> and <(Rad)(clif.fe)>. There is no ‘danger’ here of these being rhymed with spoof or wife, because CLD blocks the length-marking function of <e>. The other three examples, <(Web.be)>, <(Wyn.ne)> and <(Leg.ge)> all require CLD to create a buffer between the vowel letter and <e>, otherwise *<Le.ge>, <We.be> and <Wy.ne> would suggest the wrong phonological form.

Carney (1994: 457) also notes that there seems to be an informal four-letter rule for surnames, so three-letter lexical words require padding when turned into surnames, just as lexical words must be padded to at least three letters (inn, bye, odd, egg, see etc.).

7b.1.6 <VC₁C₂e>
At first glance, one might expect the spelling <Aske> to have a long vowel, as in waste and pastry. According to the principle of onset-maximisation, <Aske> ought to be interpretable as /ensk/, with an analysis of <(a.ske)>, similar to <(wa.ste)> and <(pa.stry)>.

However, /k/ cannot appear in this position (X₄) of a well-formed p-syllable as it is not a coronal consonant (see Section 5b.1.1). This
is different from the /t/ of asked /ɑːskt/, or indeed the /d/ of wild /wɔld/. By this rationale, /eɪsk/ is not a well-formed English p-syllable so the <a..e> sequence cannot indicate a (historically) long vowel.

Wilde is easier to explain. The rhyme of the word wild has four X-positions filled (/aɪld/), including the appended coronal /d/, so it is already a near-maximal English p-syllable. Adding <e> can have no effect on the decoding of the spelling.

7b.1.7 <Vye>, <Vwe>, <VyCe> and <VwCe>

The names Toye, Rowe and Bowe form spellings that are similar to eye and owe above, and they can be analysed as <(To.ye)>, <(Ro.w) > and <(Bo.we)>, with two g-syllables but just one p-syllable. The spellings *<Toie>, *<Roue> and *<Boue> are not permitted, because <i> and <u> do not usually function as consonant letters in the same way as <y> and <w>, and hence they cannot be in the onset position of a g-syllable. The reverse is not true of course, and <y> and <w> can replace <i> and <u>, when vowels are needed, as in Payne, Doyle and Lansdowne, all of which also have <e>-padding, making for the following analyses: <(Pay.ne) > and <(Doy.le) > and <(Lans)(dow.ne) >.

7b.1.8 <vvCe>

Freake and O‘Keefe both have super-heavy rhymes (/VVC/ /iːk/ and /iːf/) so no further changes can be made (again, except for coronal endings, as in the polymorphemic forms leeks and beefed). This means the <e> is ‘safely’ padded, which is to say that <freak> and <Freake> both suggest the same phonological form. Cooke and Goode are more problematic, since <oo> is used for both the FOOT vowel and GOOSE, and <e>-padding may confuse the decoding. The name Goode may be pronounced like good with the FOOT vowel (as suggested by LPD3 and EPD18) or it may be pronounced with the GOOSE vowel, like food. The problem is that, in most accents, <ood> can correspond to either /ʊ/ or /u:/, although <ook> only corresponds to /ʊ/, except for the words snooker and spook, notes Carney (1994: 343; see also Table 6.8). Thus the <e> of <Goode> adds to the confusion, as it suggests the name may not be homophonous with the lexical word good. This is a rare case where the padding makes for a genuinely ambiguous reading.

7b.1.9 What padding tells us about the nature of decoding

The difficulty in deciding how to pronounce Cooke and Goode leads us to the deeper question of how we ever know how to pronounce any given word or name. If we hear the word clearly then we can repeat its sound but if we only see it
written then we may not be able to reliably work out its phonological form. This problem is obviously more common among unfamiliar names, the more technical or erudite vocabulary, or indeed any word that is new to the reader. In such cases, the spelling is obviously not a (fully) reliable guide but it may be the only evidence at hand if one has never heard the word spoken.

The distribution of padding tells us a lot about the decoding process because padding only seems to occur in orthographic positions that do not suggest a new pronunciation. This suggests that users of the writing system are highly attuned to the need to keep creativity within intelligible bounds. For example, *<Wilson>, *<Jonnes> and *<Kinge> are not possible spellings because they suggest a new pronunciation. The implication, therefore, is that spelling does indeed function as a guide to pronunciation. This point is so obvious that it might at first not be worth stating, but the evidence shows the following:

If DSTNCT spellings must fulfil strict orthographic constraints in order to indicate the same phonological form then we can assume that spellings which don’t adhere to the constraints will steer readers towards a pronunciation which matches that new spelling. Hence *<Wilson> might be pronounced with an excrescent /əʊ/, and *<Jonnes> with an unwanted /ɒ/. This idealisation does not take into account the fact that some people (an in-group) may know that the spelling doesn’t match the pronunciation while others don’t (an out-group). For example, the spelling of the surname Davies suggests the pronunciation /ˈdɛvɪz/, rhyming with the lexical word gravies, and I frequently hear this pronunciation on Irish and English rugby coverage. In Wales, where the name is significantly more common, people tend to know that there is a mismatch between the spelling and the local pronunciation /ˈdɛvɪs/, rhyming with the name Mavis. The issue here is with the voicing of the final consonant, as well as with the quality of the vowel, which we would expect to vary considerably across dialects (see Allerton: 1982: 65; Wells 1982: 257). The <ies> ending generally indicates /Vz/, rather than /Vs/, where ‘V’ stands for an undetermined vowel. Carney (1994: 456) observes similar examples, including Gillies, Harries, Herries, Margulies. In such cases, two pronunciations may be found. One seems to have been precipitated and reinforced by the spelling while the other relies on its being perpetuated through speech. Thus the maintenance of an irregular spelling-sound correspondence may require a certain level of frequency in order for it to be maintained. Compare, for example, London and Coventry. The capital city continues to be pronounced with the STRUT vowel (like Monday or monkey) but the provincial city is nowadays pronounced with the LOT vowel (like poverty), presumably under influence of the spelling (compare cover and covet, which both have the STRUT vowel).
Undoubtedly, the fact that *London* is a better-known place helps to maintain its irregular, or at least semi-regular, correspondence of \(<o> \equiv /\alpha/\). The examples above do not provide concrete evidence that spelling pronunciation is widespread across the lexicon but the distribution of padded surnames (with \(<e>\) and/or CLD) suggests that people are aware that spellings not adhering to the usual mappings from spelling to sound are liable to be given an unwanted spelling pronunciation. In the latter parts of the chapter, we shall see how the same principle is applied to a wider range of substitutions, which is to say that the permissible changes made in a DSTNCT spelling are restricted to those which do not suggest a different phonological decoding of the spelling. Before proceeding, it will be useful to contextualise the analysis by looking at taxonomies of non-standard spelling, to see how creative DSTNCT spellings relate to other kinds of spelling.

***
7c. Categorising non-standard spelling

Categorising non-standard spelling is a very different task to categorising standard spelling because the question is no longer how should a particular word be spelt?, but instead, how should the established spelling be re-spelt? Hence non-standard spellings mostly involve making slight adjustments to existing spellings in order to indicate some new information. Squires (2016) observes that work on non-standard spelling in computer-mediated communication has mostly been sociolinguistic in focus. However, in recent years, there has been a notable increase in studies focusing on the writing system itself. Classifications of non-standard spelling (Preston 1982, 1985; Androustopoulos 2000; Sebba 2007; Jaffe 2012; Honeybone & Watson 2013; Picone 2016) have been based around two recurring dimensions: whether or not the new spelling should indicate a pronunciation that differs from the standard pronunciation; and, secondly, what kind of sociolinguistic meaning can be attached to the spelling. We shall look at these classifications in detail here and then I shall argue that it is also possible for non-standard spellings to include a semantic element, and there is some support for this in the literature. The spelling of names can be included in this group, making them something of a bridge between standard and non-standard spelling. The section finishes with an updated classification of non-standard spellings, and this will prepare us to examine the creativity in the spelling of trade names and band names.

7c.1 Dialect literature

There exists a sizeable body of research which studies how non-standard dialectal speech is represented in literature and there have been some classifications of the spelling strategies employed therein. Picone (2016: 331) uses the term pronunciation respelling to refer to new spellings which indicate a change in the phonological form of the word, and Preston (1985: 328) has two categories of these: allegro forms and dialectal spellings. Allegro spellings include <gonna>, <wanna> and <innerested> where the spellings ‘attempt to capture the sound of rapid, casual or informal pronunciations and elements of stylistic choice in delivery’ (Jaffe 2012: 206). Such pronunciations occur in both standard and non-standard varieties but the writing system does not normally include this kind of information. Dialectal respellings attempt ‘to capture phonological variation associated with regional or social dialect’ (loc. cit.). For example, the spelling <wint> for went indicates a merger of /en/ and /m/ in p-syllable rhymes (the PEN-PIN merger). The spelling <Hahvuhd> for Harvard indicates a loss of rhoticity.
The other major group of non-standard spellings involves changing the spelling without indicating any new linguistic information. Examples include `<sez>` for `<says>` and `<’rong>` for `<wrong>`. Nowadays, this group is known as eye dialect (Preston 1985: 328; Picone 2016: 331), although Picone (loc. cit.) notes that that the term has often been used to refer to all kinds of non-standard spellings used in dialect literature, including pronunciation respellings. Eye dialect spellings have ‘sociolinguistic meaning but not linguistic content’ (Jaffe 2012: 206), and this introduces us to a very different dimension of classification.

Both pronunciation respellings and eye dialect can have the effect of creating two social groups, the self-representation of an in-group and the other-representation of an out-group (Jaffe 2012: 206). Traditionally, all of the above-mentioned spelling strategies have been used to represent ‘other’ speech communities, in fiction which depicts dialectal speech and, latterly, in linguists’ efforts to transcribe dialects (see Preston 1982, 1985). The phonological function of pronunciation respelling is reinforced by the often-negative function of eye dialect which can remind the reader that the depicted speaker speaks in a stigmatized way, often in manner that is ‘boorish, uneducated, rustic, gangsterish’ (Preston 1985: 328).

Honeybone and Watson (2013) argue that it is possible to use non-standard spelling as part of a positive in-group self-representation. They observe its occurrence in what they call contemporary humorous localised dialect literature (CHLDL), which describes a very specific type of writing: funny fiction aimed at readers familiar with the characters’ accents. The humour of the respelling can only be fully appreciated if one is deeply acquainted with the spoken variety and those less familiar with it may struggle to reconstruct the accent from the written forms. The real problem lies in the impossibility of using existing spelling units to represent phonetic, but not phonemic, variation. This is a common theme in the literature on dialect spelling. For example, Hickey (2016: 325) points out the long-standing difficulty of capturing the apico-dental fricative allophone of /ʃ/ in Irish English. In phonetic transcription, which uses an open set of symbols, Hickey can introduce the symbol [tʃ] specifically for the purpose. In pronunciation respelling, this freedom does not exist, and the writer is confined to the redeployment of existing symbols. In this case, potential options include `<sh>` and `<th>`, but this leads to major ambiguities because these units are already in use for representing other sounds. Nevertheless, the impossibility of capturing the sound has its satirical advantages. In the Ross O’Carroll-Kelly novels, for example, Howard (e.g. 2003) re-spells the word right as `<roysh>`. This spelling suggests two phonological mergers: between PRICE and CHOICE, and
between TIN and SHIN. In reality, these are merely near-mergers, and the humour lies in the suggestion that certain speakers have completed the merger. The problem, as Honeybone and Watson (2013) note, is that such subtleties will inevitably be lost on readers not in touch with the specifics of the accent.

When we look at the creative spelling of names, which often occur in isolation and out of any literary context, it is thus possible that the spelling may indicate subtleties that are not perceptible to most readers. One familiar example is the use of <q> in Arabic-derived placenames such as Iraq and Qatar. The use of <q>, rather than <k> or <c>, for example, is to indicate a distinction in Arabic phonology which does not occur in English.

7c.2 Other non-standard spelling strategies

Androutsopoulos (2000) identifies a wider range of non-standard spelling practices in his corpus of subcultural magazines. Most of the observed spellings are structurally identical to the categories discussed above. His category of regiolectal spellings is identical to pronunciation respelling. One group is very different, however: prosodic spellings include <rePEAT> and <baaad>, spellings which use graphic resources to crudely represent prosodic functions. Homophone spellings are very similar to eye dialect but they include a wider range of letter alterations than tends to occur in dialect literature. He splits this category into two: there are graphemic substitutions, such as replacing <s> with <z>, as in <boyz>; and lexical substitutions, such as spelling you as <U> and skater as <sk8er> for skater. It is not clear why these are called ‘lexical substitutions’ as they involve letters and numerals being used for their syllabic value, rather than their lexical value. Carney (1994: 448) also observes this use of symbols in the spelling of trade names and band names but notes the absence of a term for such a device. I propose the term nymophone, since symbols take on the phonological value of their name, whether it be a single p-syllable, as in <U> or spread over two p-syllables, as in the <8> of <sk8er>. Androutsopoulos identifies yet another group which he calls interlingual spellings. These involve the respelling of words borrowed from other languages: in German writing, the English word action might be re-spelt as <äktschn>. Interlingual spellings are quite rare in English, notes Sebba (2007: 36), as such foreign spelling peculiarities have been so deeply incorporated into the writing system. Such spellings are fundamentally the same as eye dialect and homophone spellings since they do not involve any changes to the phonological form indicated by the spelling.

Sebba (2007: 33–41) also enumerates several non-standard spelling strategies in use and, on top of Androutsopoulos’s categories, he notes the
prevalence of archaic or pseudo-archaic spellings in English and Swedish (Sebba 2007: 38–9). He gives the example of <wedding fayre> for wedding fair and he links such spellings to the tendency for English surnames to be spelt in an archaic fashion, something we saw in Section 7b.2 above.

Aside from prosodic spellings and pronunciation respellings, all of the above-mentioned categories (eye dialect, graphemic substitutions, lexical substitutions, interlingual spellings, pseudo-archaic spellings), involve the deliberate re-spelling of words without the addition of any further linguistic information. These might all be put under Androutsopoulos’s umbrella category of homophone spellings.

The meanings attached to homophone spellings are purely sociolinguistic and they exist in what Sebba (2007: 34) calls the ‘zone of social meaning’. This space includes the small range of options whereby a word can be re-spelt in a manner that is recognisable to readers, but without changing the sound or meaning of the word. Such spellings are of interest in the present study, not for their social meanings, but because they provide evidence of which substitutions can be made to English spellings without affecting the phonological correspondence.

7c.3 Semantic functions of non-standard spelling

Honeybone and Watson (2013: 313) identify yet another strategy, which they call forced lexical reanalysis, whereby words are re-spelt in order to add wholly new lexical information. Sometimes this relies on pure homophony, as in the spelling <Chuck Doubt> for chucked out, and in other cases it relies on near-homophony (and a familiarity with the accent being represented), as in <Jamaica> for did you make her. They note that such spellings are rarer than simple eye dialect because they are more complex and rely on motivated, rather than arbitrary, substitutions.

Forced lexical reanalysis can also involve the admixture of two written languages. Kallen (2016: 385–6) gives the example of bóbó’s burger restaurant in Dublin, where the spelling indexes the Irish Gaelic word bó, meaning ‘cow’. The Irish-ness is reinforced with the use of superscript English <s>. Such a spelling relies on readers being familiar with the Irish language and its writing system, a reminder that interlingual spellings may only by intelligible to certain sub-groups, and they are thus a form of in-group representation.

7c.3.1 Constructed homophony

Yet another group of non-standard spellings, often occurring in the spelling names, has been referred to as constructed homophony (Ryan 2010, 2011).
Constructed homophones are identical in their construction to the various kinds of homophone respellings discussed thus far, but they differ in that they are designed to refer to a distinct word or name. The name OutKast sounds the same as the lexical spelling outcast but <OutKast> does not refer to the same thing as <outcast>. Instead, it is a band name, and it gives the musicians an orthographic identity which separates it from the lexical spelling. This also separates the band from all derivatives of the base form, including broadcast, podcast, castaway, etc. Constructed homophones occur when existing words are re-spelt distinctively, without affecting the phonological correspondence. Hence they obey both DSTNCT and PhM, but they necessarily disobey IdP, because they have to break the connection with the original spelling.

The concept of constructed homophony is not entirely new. We saw above that several seemingly different kinds of non-standard spelling can be grouped under the heading of homophonous spelling (e.g. eye dialect, graphemic substitutions, lexical substitutions, interlingual spellings and pseudo-archaic spellings). What differentiates constructed homophones from homophone respellings is that they are used with a semantic value in order to refer to something new, or at least distinct from the referent of the original spelling. Constructed homophony is thus common among names as it gives the name a new orthographic identity.

One reason why constructed homophony is so common in the spelling of names is that they do not carry the same morphological ‘responsibilities’ as lexical spellings. They do not (usually) undergo the same morphophonological alternations, so they do not need to adhere to the IdP. Take the word electric. Its two <c>s have very different potential roles, and this affects how we can alter the spelling. The first <c> always indicates /k/ and it could easily be <k>, and it often is in the name Electra / Elektra. The latter <c> in electric has further morphophonological functions in the derived forms electricity and electrician. This <c> is necessary because it has a greater phonological range than <k>. A band called Electrik would be perfectly functional, but <electrik> would not be a good spelling of the lexical word since *< electrikity> and *< electrikian> wouldn’t indicate the right consonant sounds. Perhaps unsurprisingly there is a trade name called Electrik Sheep and a film collective, Elektrik Zoo (note the different spellings of the first /k/). There is also a band called The Electrix, using <x> in place of <cs>. Problems would only arise if there were a need for certain derivations. If the Elektrik Zoo people wanted to refer to their staff as Elektrikians that would make the pronunciation hard to determine from the spelling because,
as we saw in Section 5e.1.1, <k> does not indicate a palatalised consonant in this position.

7c.4 A taxonomy of non-standard spelling

Table 7.4 summarises the similarities and differences between the various kinds of non-standard spelling discussed so far. All non-standard spellings involve some kind of deviation from the existing, accepted spelling of a word. Some of them indicate a new pronunciation for the word (pronunciation respellings), or they may provide paralinguistic information, such as volume, duration or intonation (phonetic respellings). Some non-standard spellings include new semantic information, although this comes in two very different forms: constructed homophones indicate a new name or a distinct new word, while forced lexical re-analysis may have two meanings at once: one based upon the sound of the original word, the other based on the spelling that is actually used.

Pronunciation respellings cannot tell us a lot about the nature of the normal, functioning writing system because they do not tell us anything about which spelling units are interchangeable. Instead, they tell us about phonological forms, or at least attempt to, since it is almost impossible to represent dialect speech using spelling units which are conventionally used to represent standard correspondences. Thus there is a lack of equivalence between the spelling units occurring in the original form and the spelling used in the pronunciation respelling – in the case of <roysh>, discussed above, <sh> is not an equivalent spelling unit to <t>, nor <oy> to <igh>. They are merely approximations at representing phonetic forms.

Homophone respellings do not add any linguistic information. They are simply variant spellings of the same word, indicating the same pronunciation. They can only have sociolinguistic meaning, information which can indirectly tell the reader about the social context in which the spelling exists. The spellings may stigmatize the speech of certain communities or, as Androutsopoulos (2000) argues, it may function as an in-group code. The present study is not concerned with the social merits of particular spellings. Instead, the focus will be on what spelling units can be changed and what they can tell us about the internal workings of the writing system. Homophone respellings can tell us about which units of spelling are interchangeable, which are predictable and which are productive.

Constructed homophones are a special case of homophone respellings. Both are formed based on the assumption that certain spelling units are equivalent and thus interchangeable, once IdP and the Invariance principle are relaxed. The
use of <k> instead of <c> is based on a direct equivalence between these units, at least in the orthographic context under discussion. However, constructed homophones provide new semantic information because they have a different referent to the original spelling, and that is why the phenomenon is so common in the spelling of names.

Furthermore, the concept of homophone re-spellings and, by extension, the zone of social meaning, can only exist if there are phonographic matching limits to what changes can be made. The word *says can only be re-spelt as <sez> because readers can construct the same phonological form from this new spelling. They cannot do the same with a spelling such as *<siz> or even *<ses> because these spellings don’t reliably indicate the same phonological form. The very existence of non-standard spelling provides evidence that spellings can be used as a guide to pronunciation. If the situation were otherwise, then pronunciation respelling would have zero effect on how people pronounce written forms. The importance of spelling in indicating pronunciation among these marginal, non-standard forms suggests that standard spelling may also affect the pronunciation. Of course, this is not necessarily the case – the spelling <says> does not require us to pronounce the words as /seɪz/, rather than /sez/. Instead, the evidence suggests that where changes are made to the spelling, then readers are likely to use the new spelling as a guide to pronunciation.

Finally, forced lexical reanalysis can include new semantic information and, potentially, new phonological information. The new semantic information is contained in the CPUs occurring in the new spelling e.g. <chuck>, <doubt>, <Jamaica>, although the few examples presented here tell us almost nothing about the degree to which the phonological form suggested by the spelling ought to adhere to the original phonological form.
<table>
<thead>
<tr>
<th>Label</th>
<th>Examples</th>
<th>New sound</th>
<th>New meaning</th>
<th>In/out group</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pronunciation respellings</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Allegro forms</td>
<td>gonna, innerested, jus'</td>
<td>✔</td>
<td>(predictable)</td>
<td>×</td>
</tr>
<tr>
<td></td>
<td>wint, Hahvahd l'il</td>
<td>✔</td>
<td>(dialect specific)</td>
<td>×</td>
</tr>
<tr>
<td><strong>Homophonous respellings</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eye dialect</td>
<td>sez. wuz, pix</td>
<td>×</td>
<td>×</td>
<td>out</td>
</tr>
<tr>
<td>Graphemic substitution</td>
<td>boyz</td>
<td>×</td>
<td>×</td>
<td>in</td>
</tr>
<tr>
<td>‘lexical’ substitution</td>
<td>U, sk8r</td>
<td>×</td>
<td>×</td>
<td>in</td>
</tr>
<tr>
<td>Interlingual respellings</td>
<td>bóbó', äktschn</td>
<td>×</td>
<td>×</td>
<td>in</td>
</tr>
<tr>
<td>Constructed homophones</td>
<td>U2, INXS byte</td>
<td>×</td>
<td>✔</td>
<td>in</td>
</tr>
<tr>
<td>Forced lexical reanalysis</td>
<td>Chuck, Doubt, Jamaica</td>
<td>×/✔</td>
<td>often near-homophonous</td>
<td>✔</td>
</tr>
<tr>
<td><strong>Prosodic spellings</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>rePEAT, baaad</td>
<td>✔</td>
<td>✔</td>
<td>paralinguistic meaning</td>
</tr>
</tbody>
</table>

**Table 3.1**

Different kinds of non-standard spelling

### 7e.5 Linguistic exaptation

It is worth asking how the concept of constructed homophony spelling relates to broader linguistic patterns. The idea of recycling redundant language patterns is discussed by Lass (1990) in relation to (spoken) language change. He borrows the concept of *exaptation* from evolutionary biology to show how linguistic ‘junk’ gets re-used for some unintended purpose. Lass explains the concept thus, talking about morphology, rather than spelling:
Say a language has a grammatical distinction of some sort, coded by means of morphology. Then say this distinction is jettisoned, prior to the loss of the morphological material that it encodes. This morphology is now, functionally speaking, junk; and there are three things that in principle be done with it:

(i) it can be dumped entirely;

(ii) it can be kept as marginal garbage or nonfunctional / nonexpressive residue (suppletion, 'irregularity')

(iii) it can be kept, but instead of being relegated as in (ii), it can be used for something else, perhaps just as systematic.

Option (iii) is linguistic exaptation. The point of course is that it IS an option: languages may operate 'wastefully', dumping material that no longer does anything useful, or in a 'conservationist' mode, by recycling.

Lass (1990: 81–2)

An example of exaptation in morphology occurs in the plural of the computing system called Unix. Sometimes pluralised to the regular, but unwieldy, form <Unixes>, it may also be pluralised as <Unices> or <Unixen>. The first case recycles the Latinate pattern of index – indices and appendix – appendices, and the second case recycles the native {en} plural of oxen, brethren and children. One might like to argue that this case is merely jocular but, once again, the interesting patterns often occur on the fringes of the system. Either way, these recycled patterns are not invented from nothing but instead taken from existing morphology. It makes sense to view the pattern that occurs in the creative spellings in the same way. They are only possible because the writing system contains so much ‘junk’. The system’s very complexity is what provides the raw material for intelligible creativity, and nowhere is this clearer than in the spelling of trade names and band names.
7d. The spelling of American trade names

7d.1 Jacobsen

American trade names in the mid-twentieth century show a remarkable amount of creativity in their spelling, prompting Jacobsen to write *Unorthodox spelling in American trademarks* (1966). He wonders whether these spellings might provide a pathway for spelling reform, on the assumption that they are more ‘phonetically motivated’ than existing spelling. The prevalence of such spellings, he observes, was provoked by changes in American trademark legislation which required all trademarks to have unique spellings (Toulmin 1946: 4). Jacobsen carries out a taxonomic study, with a view to showing how the spellings relate to pronunciation. His categories, with examples, are the following:
1. Spellings based on substandard pronunciation
   - Lectric Shave
   - S'nuf
   - Kilzum

2. Spellings based on regional pronunciation
   - Buttacup
   - Kleerwite
   - Tube-Tule

3. Spellings based on nonce or fancy pronunciation
   - Sil-vor-plate (silver plate),
   - CONVAIR (conveyor),
   - X-ACT (exact)

4. Spelling regularization and simplification
   - Bild-a-Clock
   - Dun Ezy
   - Sonarfone

5. Reduced number of graphemes
   - Breez-Air
   - Nun Better
   - Sav-A-Lot

Using the terminology discussed above, we might consider the first three groups as being examples of pronunciation respellings. Jacobsen’s subjective labels only indicate sociolinguistic information (e.g. prestige, variety) rather than purely linguistic information. However, there is a deeper subtlety here. The new spellings may represent a specific pronunciation, as with the pronunciation respellings we saw earlier (e.g. wint, gonna, etc), but they also represent a new semantic referent, the trade name.

Jacobsen’s latter two groups are effectively constructed homophones as they involve re-spellings which do not indicate a new phonological form, but do have a new semantic referent (the trade name).
Jacobsen makes no attempt to explain what his observed patterns tell us about the nature of the writing system as a whole. This may seem surprising nowadays, given that Jacobsen’s aim was to see what the spelling reform movement might learn from these new spellings. Yet it is a reminder that up until Hanna et al.’s (1966) study, and Venezky’s landmark (1970) treatise, there were no linguistically-informed descriptions of the entire writing system that Jacobsen could refer to (c.f. Section 2b). Hence, he lacked a framework against which he could compare his findings, a problem that did not hinder Praninskas.

7d.2 Praninskas

Praninskas’s (1968) *Trade name creation* is the most insightful early work on the relationship between creative spelling and the rest of the English writing system. The book has chapters on ‘phonemics’, ‘semantics’ and ‘graphemics’ and it is the latter which is of interest here. She notes that the entire study is done ‘out of curiosity’ in the spirit of ‘pure research’, and this frees her up to make observations which are not encumbered by an external agenda (such as spelling reform, Jacobsen’s ultimate higher goal). Along with some important observations, Praninskas makes some useful theoretical steps towards connecting the patterns in trade-name spelling to lexical spelling, thus creating a theoretical model for their co-analysis.

She provides an inventory of symbols in use (in her data) including twenty-six letters, ten numerals, two logograms (<&> and <%>) and three punctuation marks (the full stop, inverted comma and the hyphen). Next, she lists a set of default spellings for the phonemes she considers relevant and only then does she start to taxonomise, observing, for example:

- Elimination: *Protex*
- Substitution: *Kontakt*
- Simplification: *Tru-Blu*
- Haplology: *Travelodge*

Unlike Jacobsen, Praninskas’s categories are based on changes to the spelling, rather than on the relationship of the spelling to pronunciation. For example, one heading is 'Fluctuations of <i> and <y>, and other minor patterns' (e.g. *Kleen Brite*, with <ee> for <ea> and <ite> for <ight>). She goes on then to assess the use of punctuation, particularly how hyphens are used to indicate that a letter (or numeral) is to be pronounced as its name, not by its usual phonetic value. In this instance, her use of small caps is retained here, as it is not clear which letters were capitalised in the original trade name.
Praninska's claims that these spellings, containing what I have termed nymophones, are the one true innovation in her data, not observed in Pound’s remarkably early studies of trade-name spelling (Pound 1925). Praninska summarises her findings as follows:

All these kinds of change occur but a small number of patterns are dominant, indicating a specific drift toward a phonemic representational system for consonants, a simplified and regularized system of diacritics for vowel symbols, and a type of syllabary. There is another important subtlety to her work, not mentioned in the quote. She spots counter-trends to (almost) all of the patterns she observes and she is astute enough to accept and take note of them (unlike Jacobsen who makes no generalisations about the, admittedly few, counter-trends he observes). For example, Praninska spots a general tendency towards the loss of historical doubles, as in Hot Stuf but she sees the reverse process in Arrid and Slip Nott (homophones of the words arid and slip knot). This is similar to the CLD-padding we saw above among surnames and, just like in surnames, <e>-padding also occurs (Goodaire and Septic Aide). Further counter-trends occur including <kn> for <n> in Lads of Knote.

Connecting all of the patterns she observes is something deeper than the ‘drift toward a phonemic representational system’ which she reports. Instead, the patterns in her data are part of a more significant tendency for new spellings to have a consistently predictable mapping from spelling to sound. The difference is subtle but profound – as long as the new spelling has a predictable pronunciation, it is viable. It does not have to be predictable in the reverse direction, from sound to spelling. This tallies with the examples of padded surnames, discussed above: Aske, Freake, Cuffe and Payne involve the use of sub-regularities which are predictable in the appropriate orthographic environment. These examples do not simplify the writing system overall but instead they make it more manageable for the reader, who only has to decode the spellings, but not the writer, who must remember which spelling variant to use. This makes sense for trade names which appear on labels and elsewhere, and less frequently in people’s writing. Hence the
distinctive spellings of names are primarily there to be seen and understood by the reader, while difficulties experienced by writers are not a concern. The result is an abundant set of unambiguous many-to-one mappings, unlike the current writing system with its ambiguous many-to-many mappings.

Praninskas does not provide analysis of all examples of creative spelling in her corpus (of about two thousand examples). She omits cases of dialect pronunciation, typographic eccentricities (pHisoac, RICHardson's Root Beer) and 'aberrations which are obvious puns' which are dealt with in her chapter on the semantics of trade name creation. Each of these remains relevant to the study of spelling. Creative spelling can be used to indicate double meanings, and her (omitted) examples include Ab-scent (a deodorant), Kar-Go (rental trailers) and Sell-a-bration (an outlet store). Such spellings are examples of forced lexical re-analysis, as discussed above. Consistent with Honeybone and Watson’s (2013) observations, this type of creative spelling is much less common than simpler, homophonous spellings.
7e. The spelling of band names

The following discussion is an update of my earlier study of the spelling of ‘band names’ (Ryan 2010). The data examined was selected from the names used by the several thousand (sets of) musicians who featured in the British music charts between 1994 and 2009. This corpus was supplemented by the names of other musicians in my own music collection and some general knowledge and observation. The analysis was not statistically motivated but sought to find and account for both major and minor patterns of creativity in one category of spelling, looking to see what they reveal about the workings of normal, core spellings. Examples were not discarded on the grounds of low occurrence – these were taken as portals of discovery into new territory, and ultimately led to the present study.

7e.1 Changes in the spelling of consonants

The range of changes that can be made at the level of the phoneme is obviously restricted to the choices available from within in the existing writing system, a point that is trivial but fundamental. While /f/ and /i:/ can be spelt in several ways, the options are fewer for /m/ which is restricted to <m>, <M>, and <mm>, or perhaps <mb> or <mn> word-finally (e.g. womb, damn) (see Carney 1994: 244–6). As a result, changes to consonant representations (in particular) are highly restricted and the same few changes occur repeatedly, most commonly to /z/ and /k/. The word music, which appears frequently in band names, demonstrates the major patterns for both of these phonemes:

- The Music
- Musikman
- Salsa Musika
- Musiq Soulchild
- Musique
- μ -Zik
- Mu$ic

The variants for /z/ show the two fundamentally different changes that can occur. Replacing <s> with <z> is a change in the correspondence mapping from spelling to sound indicating that these two have equivalent phonological values, at least in certain contexts. For example, it is common for Plural {s}, as in Gorillaz, Metalheadz, Redz and <s> → <z> can occur word-medially, as in μ -Zik above. Replacing <s> with <$>, as in Mu$ic is a very different kind of alteration,
redeploying $<$ on account of its visual similarity to $<$s>, an issue discussed in Section 7e.7.

The phonologically-motivated substitutions are more numerous for /k/, reflecting its larger set of correspondences in English. In the examples above we see $<$c> changed to $<$k>, $<$q> and $<$que>. The default spelling $<$k> occurs frequently, as in Kode 9, Krafty Kuts and Kontakt, but $<$q> and $<$que> also seem to be mildly productive, recycling a predictable subsystem pattern: $<$q> is §Exotic and $<$que> §French. Notice that $<$que> probably marks a stress shift (/mju‘zi:k/), which one would expect in the pronunciation of §French words (c.f. bouquet and bucket, above), and this also occurs in Electrique Boutique, Other spellings of /k/ also shift to the default: $<$ch> becomes $<$k> in Hi-Tek, while $<$qu> is changed to $<$kw> or even $<$qw>, in Kwirky and Qwilo.

There were several kinds of consonant changes in the data shown in Table 7.4. These examples show further that the urge for a name to be recognisable and distinct is not confined to choosing default spellings or making spelling more phonemic. If a band likes the sound of a word such as /fats/ or /flaʃ/ then there is no way of making the spellings $<$fats> and $<$flash> more phonemic. The option is to use a minority variant which remains pronounceable, hence DJ Phats and Phlash etc. Further examples include Beatniqz, Blaque Ivory and Cequenza. None of these changes affect the metrical decoding analysis of the spellings: compare $<$ (beat.niks) > and $<$ (beat.niqz) >, etc.
<table>
<thead>
<tr>
<th>Phoneme</th>
<th>Spelling</th>
<th>Examples</th>
</tr>
</thead>
</table>
| /s/     | <s> → <c> | The Pharcyde  
|         |          | Cequenza |
| /dʒ/    | <g> → <j> | Jemini,  
|         |          | Imajin  
|         |          | The Suffrajets,  
|         |          | Soulja Boy  
|         | <dier> → <ja> | Smujji |
|         | <dgy> → <jji> | Smujji |
| /f/     | <f> → <ph> | Phuture  
|         |          | Phuturistix  
|         |          | Phuture Assassins  
|         |          | Phender  
|         |          | Phlash  
|         |          | Phender  
|         |          | Phlash  
|         | <ph> → <f> | Phat Controller  
|         |          | DJ Phats  
|         |          | Phatt B  
|         |          | Phrogz  
|         |          | Phixx  
|         |          | Filosofy |

Table 7.4
Creative changes to the spelling of some English consonant phonemes in band names

**7e.2 Expanding the English writing system**

The data shows that creative spelling can expand the range of correspondences in use. An example is the use of CLD for letters which don’t normally double (e.g. <z>, <j>, <x>, etc.). CLD is a particularly useful re-spelling strategy when there is no obvious spelling variant for any of the sounds. Among band names, we see M/A/R/I/S, Iceberg Slimm and Snoop Dogg, none of which suggest new pronunciations. Padded <zz> occurs four times in the data, with differing levels of need: Chazz, Topazz and Barthezz are padded, while Bizzy is a simple §Basic
respelling of the irregularly-spelt word *busy*, with <zz> indicated the ambisyllabic consonant /z/. The spelling *<buzy>* would not necessarily be read as a homophone of *busy*. This is a case of an irregular spelling become regularized, not just at the phonemic level, but across the entire foot.

CLD is extended to consonant letters which seldom or never normally undergo CLD, in normal spelling, including *Smujji, Bruvver* and (perhaps) *Lykke Li*. A very rare example of <jj> in English is *Hajj*, an Islamic festival (OED 83098), but there is no reason, other than a lack of precedent, for <jj> not to be used in place of <dg>, hence <smudgy> \(\rightarrow\) <Smujji>. The spelling <vv> is more common, but relatively new in English, occurring in slang words such as *navvy, savvy, bovver* (‘bother’) and *revving* (Carney 1994: 376). We might note that there are two different uses here: <bovver> is monomorphemic and involves a re-analysis of /ð/ as /v/; the new spelling indicates ambisyllabicity through <vv>. This is the same pattern that occurs in *Bruvver* (‘brother’). The other three all occur at a morpheme boundary, either before diminutive {y}, as with *Smujji*, or participle-forming {ing}, as we saw throughout Chapter 4 (*hopping, bobbing* etc.). The creative spelling of band names therefore involves the appropriation of patterns which are marginal in normal spelling, but applicable in new spellings, because they are extensions of intelligible, productive patterns.

*Lykke Li* is different. She is from Sweden\(^7\) and the name brings the <kk> pattern with it. Borrowings that come from outside the normal subsystems potentially pose problems for the stability of the overall system but they also bring in new correspondences (c.f. Cummings 1988: 19), enriching the range of options for creative spellings. In English, <kk> occurs in *Akkadian* and *trekking*, the {ing} form of the Afrikaans borrowing *trek* (OED 205462) neither of which pose reading problems. In the case of *Lykke Li* there is no precedent for a spelling of this kind and this makes it difficult to work out the pronunciation. There are no words that end in <icke>, let alone <ykke>. One might simply argue that *Lykke* is a foreign spelling and has nothing to do with English, but the difficulty of giving the name a pronunciation (without hearing it, or knowing the Swedish writing system) suggests that readers rely on existing knowledge of correspondences to work out the sound. They are able to understand a much broader range of spellings than the ones that happen to exist. Indeed, the creative spellings under discussion are evidence of this fact: literate readers can, by applying known patterns, provide pronunciations for far more words than they may have actually heard.

The occurrence of <xx> twelve times furthers this argument. CLD does not make much phonographic sense for <x>, unlike <j, k, v, z> because <x>
usually represents two phonemes, and this is why the <e> of *axe* can be discarded in American spelling. Final <e> is redundant after /C1C2/ sequences, hence <ax> and <axe> → /æks/ (compare <aches> → /eɪks/). The prevalence of <xx> in the data reveals a lot about the force of orthographic regularity. It has two functions: it marks the vowel as short within the g-syllable (e.g. *Hixxy*); and it can also provide graphematic weight for letters that have been replaced, usually <cks>, as in *Basement Jaxx*, *E-Traxx*, *Hatrixx*, and *Staxx*. The names *Maxx* and *Phixx* are just padding for *Max* and *fix*, perhaps adhering to the four-letter rule that Carney observes in surnames (1994: 457).

The removal of historical double consonants also occurs, in a process that Praninskas (1968: 28) calls *graphemic haplology*. Some of these provide better spellings than others. *Ilogik* removes the (phonologically) unnecessary historical double of *illogic*, but *Hatrixx* is not a good re-spelling of *hat-tricks*, because it might be read as < (Ha.trixx) >, like the word *matrix*. Similarly, *t.A.T.u.*, is supposed to be a homophone of *tattoo* but it suggests /ˈtɛtuː/. In keeping with Praninskas’s observation, every major trend is complemented by a certain amount of change in the opposite direction. While CLD padding is the dominant force, the removal of CLD is sometimes possible, although it may not always be a good option. Some instances of non-CLD can result in ambiguous mappings to spelling and that explains why it is less common.

### 7e.3 Changes in the spelling of vowels

The complex relations between spelling and sound for English vowels was discussed throughout Chapter 6, and the changes occurring in the data provide some insight into the relative importance of those correspondences, especially in terms of their productivity. The most obvious trend was the loss of unpredictable or ambiguous spelling units. For the DRESS vowel, we see <ea> regularised to <e>, as in *Megadeth*, *Mos Def* and *Led Zeppelin*, while *Def Leppard* regularises <eo> to <ep> (CLD for the ambisyllabic /p/). Among examples of the STRUT vowel, we see <o> changed to <u> in *Bruvvers* and *Booty Luv*, while the entire rhyme /ʌf/ of *tough* and *rough* changes to <uff> in *Ruff Driverz*, *Ruff Endz*, *Ruff Neck*, *Mark ‘Ruff’ Rider*, *Tuff Twins*, *Tuff Jam*, *Tuff Productions*. This indicates changes occurring across the spelling of the rhyme, rather than just at the phonemic level.

The data also provides evidence of which units are the default spellings for particular vowel sounds. Carney (1994: 159–64) spends six pages trying to work out the default spelling for /i:/ based solely on evidence from its existence in standard spelling. In constructed homophony, we get much greater insight into the
productivity of spelling units, and for /i:/ it is clear that <ee> is much more productive than the <e…V> unit which occurs in metre, compete and ecology. Carney suggests this unit is the default, on analogy with mate, mite, mote, mute, etc. However, in band names we see Dreem Teem, Lovefreekz and Mis-Teeq, displacing <ea> and <i.e>, word-internally. The only counterexample in the data is Breze, reduced from a word that already has <ee>, so the new spelling may be a recycling of a subregularity of the kind we saw with the <ph> of Phats and Phuture above. Furthermore, <breez> might be read as a plural. (The spelling also occurs in Jacobsen's data in Breez-Air, a collocation where the meaning of <Breez> is clarified by the word Air). <ee> also appears in syllable-final position in Dee Tah, Jay Dee, Alice Dee Jay, The Bee Gees, Gee Motion, E-Zee Possee (note both <ee>s), Skee-Lo, Skeewiff, Deee-Lite (with <eee>). The unstressed happy vowel occurs as <ee> in The Monkees, The Funk Junkees and Dizzie Rascal, so that <ee> is here ousting <y>, <ey> and <ie>, providing orthographic evidence of happy tensing, an ongoing sound-change we saw in Section 6e.4 (see also Wells, 1982: 165–8).

The data provides a smaller amount of encouraging evidence for the default spelling of GOAT and GOOSE vowels too. There are several examples of <o> replacing <ow>, another regularisation in keeping with changes in the phonology. We saw in Section 6e.2.3 that many words with <ow>, <ol> or <oul> used to have different vowel (i.e. soul and sole were not always homophones), so there is no present-day need for <w#> in Sisgo, DEVO, Sessamoto and Lasgo (which is a clipping of the band’s hometown of Glasgow).77

Among the GOOSE lexical set, the word ‘crew’ appears frequently enough to examine: B-Crew, Blackout Crew and Cabin Crew are unchanged. Corrupted Cru, Rufuge Cru and Mötley Crüe provide enough evidence to show that <oo> is not necessarily the default spelling in word-final position. Similarly, t.A.T.u removes <oo>. Word-medially, we see the irregular <ui> of juice changed in Rootjoose. Such examples demonstrate that while there is still no clear default spelling of /u:/, the minor units <ew> and <ui> are replaceable and perhaps less likely to be productive.

As in Praninskas’ and Jacobsen’s data, there is reciprocal transfer between <i(e)> and <y>: as in Dyverse, Vybe, White Noyz, and Zomby, as well as Cosbi, Bizzi and N-Joi.

7e.4 Removal of graphotactic constraints

It was noted in Section 2c.4 that there is a graphotactic constraint against English spellings ending in either <i>, <j>, <q>, <u>, <v> or even <z>. This constraint is
often relaxed in the data under discussion. For <i#> we see Cosbi, for <u#> we see Blu Peter, Blu Cantrell, Nu-Matic (homophonous with pneumatic), and Dru Hill (its <u> replaces the <ew> of Andrew). Final <v#> occurs Herv and Booty Luv and final <q#> appears in Musiq Soulchild. The change of <s> to <z> occurs so often that final <z#> ceases to be a constraint. The spellings <Blu> and <Cru> also suggest that the constraint against <CCV> spellings may be relaxed in non-standard spelling (c.f. Section 7a.2 above). In short, the relaxing of the Invariance principle allows for minor patterns to be regularized. Significantly, the reverse trend does not occur. There are no cases in the dataset of an unnecessary final <e> being added where it might lead to confusion over the pronunciation.

7e.5 Nymophones

Both Praninskas (1968) and Carney (1994: 447) note the existence of nymophones (without using that term) and Praninskas sees them as an important innovation that had not been previously observed. Yet neither of them mentions the fact that, in examples such as U2 and INXS, the letters do not function at any one phonological (or indeed morphological) level. <U> and <2> are possibly morphograms (of you and too) whereas the <X> of INXS represents /eks/, three phonemes split over two p-syllables; and the <S> represents a rhyme /es/. Furthermore, these symbols are mixed in with phonograms <i> and <n>. This redeployment of symbols, be they alphabetic letters or logographic numerals, is different from a(n ideal) syllabic writing system, where one symbol is used for each p-syllable.

Nymophones occur frequently in band names with wildly divergent phonological functions and it is not always possible to draw a clear line between them and letters used in abbreviations. So we see EZ, E-Z Rollers, Eazy-E, E-Zee Possee, as well as JZ, Jay-Z, Jay J, Jay Dee and Alice Dee Jay. The full range of possibilities is amply provided by the letter <x>, which can represent phonemes, rhymes, p-syllables and stand alone as a graphematic word. <X> appears alone in the band X and twice in The XX. It is combined with words in Pale X, Chanel X and X Funk. It is syllabic in Shy FX (‘Effects’), where it is also part of a pronunciation respelling. It cuts across a p-syllable rhyme in XTC (‘Ecstasy’), Junkie XL (‘Excel’), and INXS (‘in excess’). The band X-ite also combines different levels within one word. NOFX and INXS mix them without a space or hyphen, and the same goes for Xpansions and Xelcia. This latter case is ambiguous because its <X> may represent /z/ word- or syllable-initially (e.g. Xavier, Xerxes). Xscape solves this problem at the expense of elegance, something not lacking from the name Xzibit. The 'silent' <h> is dropped and <z>
marks the voiced /gz/ sound which precedes a stressed vowel (compare the /ks/ of < (sexy) > and < (expo)’ (nen.ti.al) > with that /gz/ of < e(xhi.bit) > and < e’(xag.ge)(rate) >. The notation used in the latter two examples cannot capture the fact that the <x> occurs across a foot boundary, as well as a p-syllable boundary (marked with the dot below).

The wide range of uses for <x> highlights the problem of categorising written symbols according to their dominant phonological function. The letter <x> does not usually adhere to the principle of one (or more) letters to one phoneme, so it seldom matches a single phonological unit such as the phoneme, nucleus, rhyme or p-syllable etc. (It does word-initially in Xavier and Xena, etc.). Instead, the letter can be used wherever it can do the job its user needs. This is the principle upon which nymophones are built. Letters are not restricted to being used for representing phonemes. Instead, they are just tools for use within a system. In non-standard spelling, their range can be greater than in standard spelling, but even there, letters can have a wide range of uses. We have seen several kinds of <e>-marking functions (e.g. teeth, tone, tense, notice, manage). Similarly, <h> is used in several digraphs, either altering the phonemic value of the basic letter (as in chat, that and shat) or having no phonological effect (as in dhal, bhuna and khaki). It can be used to mark vowel length (Shah, cheetah). All letters are used for several purposes in standard spelling so it should be no surprise that creative spelling should extrapolate from those possibilities. Crystal (2008) notes that the use of nymophones (a term he doesn’t use) in text messages, or other creative spellings, can be traced to the long-standing use of initialisms in English writing. Forms such as e.g., i.e., p.s. etc., all require the reader to pronounce the full name of the letter rather than to give it a segmental value (as in the words egg, eye). It is perhaps no surprise, therefore, that the range of uses for nymophones has been expanded and they can be used spelling-internally, as we shall see.

It is worth looking at how nymophones can be captured within the graphematic hierarchy, and how they fill gaps at the segmental level. In order to capture the use of nymophones in the linear notation, we can use word internal capital letters, as in < ‘(U)(2) > and < ‘(in)(XS) >. Again, the notation cannot mark the foot boundary inside the <X>. However, this example highlights a more fundamental limitation of the notation. It is only designed to explain foot structure and stress patterns. If the phonemes are represented by nymophones, the linear notation cannot grasp this. Nevertheless, the a-symmetry between the phonological and graphematic hierarchies can be seen in Figure 7.1 (with the foot level removed). We see how the <S> marks the entire phonological rhyme of the
final p-syllable, while the <X> is spread across two p-syllables and three phonemes. The lack of symmetry in this kind of writing is clearly visible in the graphematic hierarchy. Furthermore, the hierarchy shows that the ambisyllabic /n/ is not represented by <nn> (as it generally is not in g-dactyls). A more phonologically faithful spelling might be the cumbersome <INNXS>. This is marked in red in Figure 7.1. Of course this would not be a morphologically faithful spelling.

Such difficulties in reading this spelling perhaps explains why the band’s logo has <X> is a different colour, so that the word is not read /ŋks/, much like the word sphinx. Furthermore, the white star functions as a marker of the morpheme boundary, rendering <in*XS>, something that is useful when the logo is printed in black and white (see Figure 7.2).

![INXS logo](image)

**Figure 7.2**
The INXS logo
Perhaps more interesting than the rudimentary substitution of symbols, as in all the cases above, are cases of forced lexical re-analysis, where the spelling units are chosen to create extra meanings. Rare examples in the data include *The*
Beatles and Creedence Clearwater Revival.78 Such spellings involve the simultaneous use of the DSTNCT and IdP. The spellings distinguish <beatle> from <beetle> and <creedence> from <credence>, yet they connect the names to the words beat and creed, along with these words’ associated meanings. Similar examples occur among the examples omitted by Praninskas, e.g. pHisoac and Ab-Scent. Such examples demonstrate the possibilities for the semantic interplay for both the eye and the ear.

Some names appear to be spelt so as to indicate two pronunciations – one homophonous with an existing word, the other being a pun from the spelling. A spate of these revolve around the letter <E>, slang for ecstasy, the drug of choice in the rave scene which spawned so many creative spellings: E-rotic, E-motion, Smart E’s, Halo Hipp-E. These names rely on the fact that spellings are fixed but readings aren’t, and two pronunciations can be associated with each word, depending on the metrical decoding attached to them. For example, E-motion can be read as <e(mo.ti.on)>, just like the lexical word emotion, or it can be read as <ˈ(E)(mo.ti.on)>, with the nymophone <E> functioning as the head foot, and the word being connected to E and motion. Similar pairings could be constructed for the other examples. Spellings are usually designed to not remind us of the other possible pronunciations, but the use of capitals and/or hyphens in these examples, helps to maintain both possible readings. This is an example of the use of the graphic resources of the writing system to add further meanings to a spelling. In logos, letterforms can be changed so that their colour, shape and size add pictorial information with no bearing on the reading, or they can be used to indicate a particular reading, as with the INXS logo above (Ryan 2015). In normal spelling, however, the possibilities are limited to the symbols available, which includes letters (upper- and lower-case), numerals and punctuation marks. The examples in the next section demonstrate what can be done within this very limited set of graphic resources.

7e.7 Visual meanings: capitals and punctuation

There are several other ways that distinctiveness can be achieved without changing the sound-spelling correspondences. Capitals can be used, as in deUS, OutKast and U.N.K.L.E. Punctuation marks can be added, as in The Go! Team, Add N to (X) and Sunn O))). Finally, symbols with similar graphic forms can be substituted for letters, as in Ke$ha, Deadmau5, Fe-M@il.

Sometimes capitals and punctuation have a useful role to play in marking word- or morpheme-boundaries, as in: PaloAlto, InMe, Deee-Lite, The Night-Trippers and Mo-Ho-Bish-O-Pi. Praninskas (1968: 29) remarks that the
hyphen is “often not so much a uniter of words as a separator of letters.” *Mo Lester* and *Test Icicles* reverse this pattern so that the expected pronunciations block the vulgar puns, and provide further evidence of the tendency for patterns and counter-patterns, as well as the possibility of two or more readings of the same string of letters. Spaces, hyphens and capitals can thus function as a guide to pronunciation, beyond their usual roles.

Other punctuation marks feature in related ways, and this frequently occurs after the letter <D>. *De:Progression, D*Influence, D’Angelo, D’Jaimin, D:Mob, D:Fuse, D:REAM, Fund*Mental. Some of these examples may involve different solutions to the phonological need to represent <the> with /d/. In other cases, the punctuation marks are a purely visual phenomenon with no correspondence in speech. So, in a (somewhat) descending order of phonological significance, we see:

Stellastarr*
N*E*R*D,
S*M*A*S*H
M!A|R|R|S
? and The Mysterions
The Maybes?
Y? N-Vee
Buy Now!
BOOM!
The Go! Team
And And! And
Add N to (X)
Sunn O)))
!!!

The asterisk of *Stellastarr* is a visual reinforcement of their name (with <*> being a ‘star’). In the case of N*E*R*D the band is pronounced as either /nerd/ or /eni:ardi:. In the latter case, the asterisks function like full stops.
*S*M*A*S*H and M!A|R|R|S also blur the boundary between spellings and logos, with their repeated symbols being a kind of orthographic decoration. The brackets in *Add N to (X)* also have this function, and also in *Sunn O)))* where there are no open-brackets corresponding to the three closed-brackets. At this point, the spelling has effectively become a picture as well having a phonological representation. The brackets <*> are purely decorative.
The use of a dieresis, as in Motörhead, Mötley Crüe and Hüsker Dü, has become something of a ‘genre cue’ (Androutsopoulos 2004) indicating that the band plays metal music. The function of a dieresis in normal English spelling is to mark a syllable boundary between two vowel letters, and rare examples include naïve and, in the past, coöperate. The examples among band names appear to be German words (even though they are not) because its writing system uses diereses, or umlauts, to mark vowel-fronting. Hence these band names are often given new phonological interpretations by their German fans (Spitzmüller 2012). The use of such ‘rock dots’, as they have become known, has extended well beyond its use in either English or German. Spitzmüller (ibid) notes its use in Queensrÿche and parody rock band Spïnal Tap, and many more. In short, rock dots function as a kind of semantic determiner, indexing a particular style of music, but also German or broader Teutonic culture.

The exclamation mark has also undergone a kind of orthographic grammaticalization. The band Buy Now! employ the mark for its usual purpose of indicating an imperative. Other bands add it ornamentally, as in The Go! Team and And And! And. It appears here to have no prosodic role, although when we look at all the other cases of its use, many of the words are exclamatory:

Amen!, Boom!, Coma B!, Embargo!,
Encore!, FAB!, Frente!, Girls Aloud!,
The Go! Team, Jakatta!, Only Men Aloud!,
Phlash!, Pop!, Sash!, Snap!, Skream!,
Wham!, Wubell U!, Viva!;
3oh!3, Panic! at the disco

The <!> thus has no segmental correspondence, but may indicate some kind of prosody.

The trend for exclamation marks is amplified by the American band !!! where there is clearly no phonological correspondence between the ‘spelling’ and any phonological name they may have. The name cannot be interpreted within the English writing system and, while the band were aware that <!> is used to represents clicks in some South African languages,79 that is clearly too obscure for most readers of English. One option is to refer to them as ‘exclamation mark, exclamation mark, exclamation’, a situation that applies to the band ? and the Mysterions. Over time !!!’s name has morphed into Chk Chk Chk, allegedly coming from the noise made when checking a microphone (‘check, check check’), although the band recommend that their name be pronounced with any three
monosyllabic sounds (perhaps referencing the click sounds). Despite the lack of any obvious connection to speech, !!! do have a number of common orthographic features to their name. Firstly, there are three symbols, like in the majority of abbreviations (Cannon 1988), and this is common in bands too: e.g. AFX, REM, Run DMC and LCD Soundsystem. Furthermore, the exclamation mark is in common use, unlike the symbols used in the following band names:

†‡†
Gr†ll Gr†ll
ℑ ⊃ ⋐ ⊇ ℜ
GL▲SS †33†H

A big difference between !!! and these bands is that one can at least type <!!!> on a normal keyboard, although the symbol cannot be used in a URL, so the band’s website is chkchkchk.net. One cannot search for !!! on Google, but one can on YouTube. Such factors may have had a huge constraint on creative spelling over the past twenty years because searchability is very important, at least for most artists. Part of the purpose of the strange names listed above is the difficulty of finding them on the internet, something which provides a certain amount of welcome anonymity for some.

Some of these bands adhere to PhM more than others, but first it is necessary to convert the symbols into normal letters: hence <GL▲SS †33†H> can be converted into <GLASSTEETH>. In other cases, such as ℑ ⊃ ⋐ ⊇ ℜ, the symbols cannot be interpreted with respect to PhM assuming we discount the circuitous technique of naming each of the symbols. Nevertheless, the spellings are invariant and they do exist, which means that they obey the two most fundamental principles. They also obey DSTNCT. IdP is not fully relevant as their names do not derive from existing spellings. They are more like names that were formed directly from their sound (like The Pogues) but instead they were formed directly from symbols. In short, these names are purely orthographic creation and any phonological form has been attached after the fact. This is not the usual pattern for word, or even name, formation.

There is one spelling that remains unique within the dataset. The name 5ive! is an example of a multi-dimensional spelling, discussed in Ryan (2015) in relation to Google Doodles. Out of context, <5> does not look enough like <F> to replace it (in the way that <$> could replace any <s>), but in context <5> provides enough of a visual gesture to <F> to function as that letter. This spelling relies heavily on the pragmatic context: obviously <5> refers to the same word as
<five>, so the numeral’s partial similarity to <F> is reinforced by the meaning of the word five.
7f. Conclusion

The examples provided throughout this chapter, drawn from various relevant studies (Carney 1994, Jacobsen 1966, Praninskas 1968, Ryan 2010) show that there is a strong desire among users of English to have one distinct spelling for each distinct word or name, even though the writing system does not always facilitate that need (e.g. bank, nail, jumper). When the need to obey IdP is relaxed, and when there is a need for a word or name to have a distinct or unique spelling, then new spellings can be created from the available orthographic tools, namely the functional correspondences of the English writing system. In the examples presented throughout this chapter, we have seen default correspondences being used in place of less regular ones but also the reverse: the recycling of non-default correspondences, on condition that they are predictable from spelling to sound. The spelling formation patterns follow the same basic principles. In a constructed homophone, changes to the spelling never suggest a new pronunciation. Units of spelling can replace any phonological level and normal graphotactic ‘constraints’ may also be relaxed. The overall effect on the writing system is that decoding is more predictable for the reader but spelling may be more difficult for the writer who must remember two spellings for the one sound.

There are certain conditions under which distinctiveness is desirable, although it is not always possible. The need for lexical words to have three letters throws up a handful of distinctive pairs, including we and wee or in and inn. However, DSTNCT is much more common in the spelling of names. The principle can be systematically applied through the use of capital letters, so that names deriving from the lexicon are visibly distinct from their lexical counterpart, as with Arsenal and arsenal or Cook and cook. Padding is possible because CLD and final <e> can be skilfully positioned so as to not affect the phonological correspondences. Spellings such as Freake, Payne, Cuffe, Webbe, Wilde and Aske have identical readings to their lexical counterparts. In the spelling of trade names and band names, we see DSTNCT applied using a much wider set of spelling correspondences. This occurs for two very different reasons. The sociolinguistic condition is that names often have a need for a unique orthographic identity. The linguistic condition is that names have little or no need for morphophonological alternations to occur and thus IdP can be relaxed: Electrik is a possible name because there is no need for the related word *<Electrikian>, and thus no need for <k> ≡ /ʃ/.

The most notable pattern is the tendency for spellings to be regularized towards a seemingly more phonemic writing system. Productive spellings include
the use of <k> for /k/ and <z> for /z/, as in *Kode 9* and *Gorillaz*. Such cases provide some evidence of the productive spelling patterns in English, and a noteworthy example is that the <ee> of *meet* seems to be the default spelling of /i:/, rather than the <e...e> of *mete* (as Carney, 1994 argues). The evidence from the spelling of band names shows that <ee> is highly productive whereas <e...e> is not. Similar points were made about final <o#> versus <ow#> and final <u#> versus <ew#>, <oo#> (e.g. *Lasgo* and *Rufige Cru*). This use of default spellings is counter-balanced by the use of predictable, non-default spellings such as <ph> for <f>, as in *Phuture* (compare Filosofy). Praninskas (1968) notes the existence of counter-trends to all major trends, but I have argued that the conditioning factor is not the collapse of spelling towards a simplistic one-to-one writing system, but rather one where all spelling units are predictable in the direction from spelling to sound (although clearly not in the other direction). This is borne out by two other pieces of evidence. One is the loss of unpredictable and/or vestigial spellings – *tough* is changed to <tuff> and *blue* to <blu>. The other evidence comes from the use of nymophones, as in *Altern-8*, *X-pansions*, *NOFX* and *INXS*. Readers may have difficulty parsing the letters and providing a higher-level analysis of the spelling, but once they do, the correspondences are predictable from spelling to sound, even though the representation occurs at several phonological levels. Once again, some converse patterns still occur, albeit for good reasons. There are puns with double meanings that can be read two ways, disrupting the system of predictable many-to-one mappings (e.g. *E-motion*). These can be categorised as forced lexical re-analysis, just like *The Beatles*, where the sound of the word suggests one meaning (*beetle*) but the spelling suggests another (*beat*).

The chapter also included a new taxonomy of non-standard spellings, one which builds on the work of existing taxonomies (e.g. Preston 1982, Androutsopoulos 2000, Sebba 2007, Jaffe 2012, Picone 2016), but which also introduces the new category of constructed homophones, spellings which form a bridge between standard and non-standard spelling. Many labels attached to non-standard spellings can be subsumed under the category of homophone respelling, changes that introduce no new linguistic information, only sociolinguistic (a perspective that is beyond the present study). Constructed homophones are a special case of homophone respelling because they involve the same orthographic technique of changing spelling units without affecting the phonology. The two groups differ because constructed homophones are used to refer to a new entity, such as a surname, trade name or band name, and thus they include semantic information. Forced lexical reanalysis also introduces a semantic element, but one that is based on the meaning of existing words (as in *beetle* and *beat*).
Pronunciation respellings involve efforts to mimic regional pronunciation using the limited resources of standard correspondences. Forced lexical reanalysis may also have a phonological element but it is not clear to what extent a spelling can indicate a different phonological form and a different referent without being considered a totally new word.

Including non-standard spellings and names in a theory of English spelling has seldom been attempted in the field and studies of these kinds of spelling have been confined to either studies that do not engage with standard spelling or marginal sections of major studies of standard spelling. However, it should be clear that non-standard spelling and the spelling of names is dependent on the workings of the standard system, and thus the core system constrains the possibilities of the marginal spelling practices. It was Jacobsen’s (1966) belief that the non-standard spelling would provide insight to reformers looking to change the standard writing system, and this hunch is borne out by the findings presented here that marginal spelling practices tend towards the use of spellings which can be decoded predictably, a situation that advantages the reader/decoder over the writer/speller. This chapter has shown how patterns of creative spelling emerge from the re-distribution of existing patterns, and how linguistic junk, as Lass (1990) calls it, can be re-used to generate new kinds of meanings, allowing for more words and names to have DSTNCT spellings, at least when the need to adhere to IdP is relaxed. Thus constructed homophones can occur because people are able to understand a much broader range of spellings than the ones that happen to exist, and a lot more spellings than tend to be listed in lexical dictionaries.

***

***
8. Different kinds of spelling for different kinds of words

The aim of this chapter is to apply what we have learnt about spelling formation (Ch. 4, 5, 7) to several kinds of word-formation category, including affixed derivation, compound words, neo-classical compounds, blend words, abbreviations and clippings. While analysis of each group is rather brief, with the notable exception of affixed derivation, the primary goal is to show that each category may be subject to its own orthographic peculiarities, and is therefore worthy of greater attention in future studies.

Affixed derivation will be examined in the most detail, as its spelling patterns are broadly similar to the spelling of inflections, except that there is often a significant etymological input, making the spelling formation more complex. This will allow us to see how the model developed in Chapters 4 and 5 can have further elements added. The focus will be on the spelling of two suffixes, words with adjective-forming {al} (e.g. functional, hexagonal, maternal, tonsillar) and words with adjective-forming {-ble} (e.g. possible, probable, noticeable). The first group are heavily influenced by Latin spelling and morphophonology (note, for example, the <ar> ending in tonsillar), while the second group are influenced by both Latin and French, which can either reinforce one another or provide contrasting models for English spelling. This data provides excellent evidence of how the need to represent etymology interfaces with the usual needs of representing morphological and phonological information. But first, it will be necessary in Section 8a to account for the exact status of etymology in English spelling formation, something that has remained elusive in the literature.

***
8a. Integrating etymology into a description of English spelling

English has no shortage of words borrowed from other languages and these have taken many shapes within the writing system over the last millennium. Some words were borrowed so long ago that their spelling has been completely integrated into the §Basic subsystem of spelling. Thus sheep and beef both contain the correspondence <ee> \equiv /i:/, even though sheep is from Old English and beef was borrowed from Old French (OED 177773/16953).

In other cases, borrowed words have had their spellings adapted slightly so that they fit into the English writing and morphological systems. Cummings (1988: 17–20) calls these adoptions with adaptation. The word philosophy, for example, has a spelling that largely reflects both French <philosophie> and Latin <philosophia>, which themselves were influenced by Greek spelling (OED 142505), but the English use of <y> is in keeping with other English words with this suffix (e.g. biology, orthography and masonry; OED 231080). Nevertheless, the plural form <philosophies> is identical to the spelling of the French plural, while the <ia> spelling survives in pairs such as Italy, Italian. Words derived from (older) French and/or Latin will be referred to as ‘Latinate’ in keeping with Carney’s §Latinate subsystem of spelling (1994: 96–101).

Other words have been adopted without adaptation (Cummings 1988: 17–20) meaning that the spelling has been taken directly from the borrowed language. This has become very common since English spelling became standardized (Scragg 1974, Upwards and Davison 2012, Scholfield 2016). Adoption without adaptation often occurs in specific domains of the lexicon: cuisine is associated with a lot of modern §French words, including chef, meringue and gateau; while coffee has a lot of §Italian words: espresso, macchiato and cappuccino.

A final group of borrowed words, such as kangaroo and hickory have §Basic English spellings (OED: 157186/86699), by default, simply because they come from languages which used to have no writing system.

8a.1.1 Previous attempts reviewed

It has long been known that English spelling is both morphemic and etymological but the two have never been connected formally. Taking etymological information into account allows for a much simpler description of existing spellings. Rollings grasps this:
Reference will frequently be made to etymological criteria for determining or explaining spellings, since they very often constitute the only real reason for the way a word is written.

Rollings (2004: 72)

Despite stating this, Rollings struggles to reconcile this fundamental design feature of the writing system with his synchronic, generative-phonology rules, and thus he ends up producing overly complex re-write rules for explaining correspondences which are regular within a particular subsystem (e.g. the use of <y> in Greek-derived words such as synonym and synchronise (c.f. Section 2b.3.4). Carney (1994) recognises the role of etymology in the existence of spelling subsystems but he only partially reconciles that with his otherwise segmental analysis. As we have seen (Section 2c.3; Section 5a), he argues that subsystem correspondences apply across an entire word. Thus a spelling such as <chlorophyll> is $\S$Greek throughout, so the presence of <ph> $\equiv$ /f/ and <y> $\equiv$ /v/ corroborate <ch> $\equiv$ /k/.

The distribution of correspondences for <ch> is a perfect example of where reference to etymology can be very useful in explaining English spelling, although this explanation is frequently avoided. For example, Venezky (1970: 66–67) simply lists examples of seemingly irregular correspondences. He makes no formal distinction between his short list of cases where <ch> has no correspondence (e.g. fuchsia, yacht, drachm), and the forty-three cases where <ch> $\equiv$ /k/ (e.g. chameleon, chaos, character, eucharist, parochial, etc.). He does not include in his description the fact that most of the latter group have Greek roots, nor that the thirty-seven cases (in his dataset) of <ch> $\equiv$ /ʃ/ are all of modern French origin (cache, chandelier, parachute, sachet, etc.). Clearly, Venezky was aware of the words’ origins but his descriptive concern was simply with letter sequencing. As a result, he considers chlorine, Christmas and technology to be regular on the grounds that <ch> $\equiv$ /k/ before <l>, <r> and <n>. But the rest are irregular, within the limitation of his theory, despite the obvious subgroupings. Carney, by contrast, uses the subsystems as a safety check for spelling/decoding. If a reader/writer recognises that a word is $\S$Greek, then the appropriate correspondences should apply. The same goes for $\S$French, $\S$Italian, $\S$Latinate, and so on. This is why contestants in Spelling Bee competitions are allowed to ask for a word’s origin.

Scholars working within a graphematic framework have also struggled to integrate their synchronic approaches with etymological spelling. Their argument
is made on the grounds that ‘speakers can only resort to etymological information in a very limited way’ (Berg et al. 2014: 285). This means that the way people learn to use the writing system is at odds with how the writing system itself has developed. The mistake is to assume that a description of the system should be based on learners’ experiences of it. At the core of their argument seems to be the following unstated assumption: as it is not reasonable, nowadays, to expect all learners of English to understand Greek, Latin and French writing, it does not make sense to refer to these writing systems in a description of the English writing system. Instead, their descriptions are based solely on phonological, morphological and letter-patterning information. This is in keeping with the prioritisation of synchronic approaches within linguistic theory, but it fails to recognise that the study of spelling formation can benefit from a mixed, synchronic-diachronic approach, as occurs in word-formation studies (c.f. Ch. 3). While graphematic approaches have yielded some remarkably useful results in recent years, the theory remains stymied by the refusal to accept some of the simpler facts about English spelling. However, it should be noted that these approaches are not interested in spelling formation nor how the writing system develops over time. They are only looking at a contemporary snapshot of the language, based on a sizeable corpus of words (e.g. Baayen et al. 1995, Thorndike 1941).

8a.1.2 Etymological spelling as simple IdP

Etymological information is easy to account for if we approach the topic from the viewpoint of how the spellings were formed, rather than accepting spellings as they are and trying to explain their subcomponents. Adopting the spelling of a borrowed word, without adaptation, can be viewed as an instance of simple IdP. The only difference is that the identity preservation occurs between English and another writing system, rather than within English. Thus we can expand our definition of IdP to include borrowed words whose spelling has been adopted without adaption.

Viewing etymological spelling as a case of IdP allows for a much broader definition of what counts as English spelling, and we no longer have to make arbitrary decisions about what is to be omitted. We noted earlier that Carney excludes a sizeable number of foreign words from his corpus. He writes:

Unassimilated words of foreign origin were excluded. The choice made under this heading is inevitably idiosyncratic, but can be illustrated by the following small sample of excluded words:
bunt, capita, contra, da, de, der, des, donna, du, duc, ecole, exeunt, fait, faubourg, frère, habeas, infinitum, livre, maître, mardi, memoriam, noir, noster, notres, obscura, oyez, padrone, princeps, priori, poilu, prix, sic, siècle, stadholder, summa, supra, tonneau, un, und, etc.

Carney (1994: 106, my emphasis)

Accepting these spellings as examples of etymological IdP removes the need for idiosyncrasy in delimiting what counts as English spelling. One may, of course, wish to discount these spellings on other grounds, such as their low frequency, or the difficulties readers may have in decoding and/or pronouncing them. But it is undoubtedly useful to have a descriptive model that can explain the presence of established ‘English’ words alongside ‘foreign’ words which may only crop up occasionally.

Adoptions without adaptation are so simple that there is surprisingly little to discuss about their actual formation (within English at least — how they were originally formed in the donor language is another matter). The interesting question revolves around how borrowed spellings are pronounced, and such words provide further evidence of the role of spelling on pronunciation. We shall look at this issue briefly before examining how adaptations occur in affixed derivatives.

8a.1.3 Pronouncing adopted spellings

There are two basic strategies for pronouncing borrowed words. They can have a pronunciation that approximates the word in the original language — this is why there is no /t/ at the end of the §French word dufet — or they can be given a spelling pronunciation, often making them a better fit for English phonotactics. A slightly different way of stating these two options is to say that a French borrowing (for example) can be decoded through the §French subsystem or it can be decoded as though it were a §Basic spelling.

Everything relevant to the present study can be explained by various pronunciations of the §Italian word macchiatato. The pronunciations /mæˈkjaːtəʊ/ and /maːˈkjaːtəʊ/ suggest attempts (in RP and GA, respectively) to render the Italian pronunciation /makˈkjaːtəʊ/, but changes are naturally made by English speakers. The Italian /o/ is diphthongised in both English accents, and other phonemes are Anglicized accordingly, but the basic phonological shape of the Italian word is preserved, or at least mimicked. Most notably, <(c)ch> ≡ /k/, as it often is in §Italian words, including zucchini and Machiavellian (OED 233097 /
Yet the spelling seems to affect even these Italianesque pronunciations. The first <a> is interpreted, in RP, as /æ/, while it is /a:/ in GA. The ambiguity is clear from the metrical analysis as <mac(chia.to)>, where the first <a> remains ambiguous, although the absence of <a> ≡ /e/ suggests that a syllable boundary is placed between <c> and <ch>, thus making <cch> an historic double in the vein of <tch> (compare Mitchell and Michael). Despite these ‘differences of lexical distribution’ (Wells 1982, c.f. Section 6a.1), there exists a much greater phonotactic problem in that the syllable /kja:/ is difficult for speakers to say. One solution is to re-analyse the spelling in order to make a more suitable pronunciation. This renders the word as two trochees and the spelling is analysed as <(mac.chi)ˈ(a.to)>. In RP, this generates /ˈmækiˈaːtsəʊ/, while in GA it is /ˈmaːkiˈaːtʊs/ (LPD2 454; OED 251771). Most importantly, a syllable boundary is located between <i> and <a>, and this in fact corresponds to a foot boundary. Roca (2016: 79) notes similar syllable boundaries being introduced in the English pronunciations of Spanish-derived words such as siesta, Santiago, adios, and these can be metrically decoded as <si(es.ta)>, <(San.ti)ˈ(a.go)> and <(a.di)ˈ(os)>, with the similar prosodic shapes to the words barista, Colorado and adipose.

The pronunciation of an adopted spelling is thus subject to two, often opposing forces: an English attempt to mimic the sound of the foreign word and a re-analysis of the word in keeping with an English decoding of the spelling. The first choice may render the more prestigious pronunciation but the second choice may often be the more pragmatic solution, especially if the decoding results in a pronunciation that is more in keeping with English phonotactics. We might expect one pronunciation to win out eventually but it also makes sense that spelling pronunciation would gain traction over time, especially as borrowed words become established in English and lose their connection with the donor language (see also Neuman 2009).
8b. Affixed derivation

Affixed derivation involves the formation of a new word by adding a prefix or suffix to a base form. In his discussion of affixed derivation processes, Bauer (1983) provides the following paradigm:

- nation
- nation.hood
- nation.al
- nation.al.ize
- nation.al.ist
- nation.al.ist.ic
- nation.al.ity

Bauer's examples are often chosen with one eye on the spelling, and these derivatives all display simple IdP, with the spelling of the stem <nation> being added to the spelling of the various suffixes. Note also that the spellings are unaffected by differences of stress or vowel quality, as exemplified by 'nation – 'national – natio'nality.

Not all affixed derivatives have such neat spellings (i.e. adhering to simple IdP) and amendments can be made at the morpheme juncture, just like in the spelling of inflections. The fact that inflection and affixed derivation are very different from the viewpoint of word-formation (Plag 2003: 18–23) is not relevant to spelling. Both of them involve suffixation and this results in similar spelling formation patterns. However, affixed derivatives are often subject to greater etymological influences because, unlike inflection, which is a morphological process occurring solely within English, many affixed derivatives may have been formed in French or Latin before being borrowed into English.

Many affixed derivatives have orthographic histories of spectacular complexity. The {-ble} suffix is a case in point. It is in no way obvious why, for example, possible is spelt with <ible> but probable with <able> (see 8b.2.1 below). In other cases, such as <emittable> the presence of CLD is a strong indicator that the spelling was formed within English, according to the same patterns of spelling formation as <emitted>. Before we look at the intricacies of how words containing the {-ble} morpheme are spelt, we will look at words with {al}. Its formation patterns are simpler, showing similarities to those seen in simple inflections and, while there has been quite a lot of influence from Latin, there has been much less from French, making the picture much simpler.
8b.1 Adjective-forming {al}

The OED (4478) notes that English {al} forms have all been streamlined to take an <al> spelling in keeping with classical Latin -alis, regardless of the suffix’s natural development as <el> in French and its consequent influence on earlier forms of English. Thus, *mortal* was often spelt <mortel> in Middle English, as it is in modern French. Compare also English *functional* with French *fonctionnel*.

8b.1.1 Simple IdP

The basic spelling-formation patterns of {al} words follow simple IdP, with the spelling of the affix being concatenated to the spelling of the base. There are many examples of this:

- renew – renewal
- function – functional
- hexagon – hexagonal
- baptism – baptismal

8b.1.2 <e>-dropping

As with the spelling of inflections, simple IdP is not always enough and amendments often need to be made, if possible. One simple, familiar change is <e>-dropping:

- fate – fatal
- mode – modal
- culture – cultural
- homicide – homicidal

8b.1.3 Germanic and Latinate pairs

In some cases, the base form is Germanic but the complex form is Latinate, a situation that often arises in learnèd vocabulary:

- mother – maternal
- mouth – oral
- tooth – dental

The words *maternal, oral* and *dental* contain two bound morphemes. All of them have the suffix {al}, while {matern-}, {or-} and {dent-} are bound forms, which cannot appear alone. They do appear with other suffixes, as in *maternity, orality* and *dentist*. As noted in Section 3e.1.6, it is not always possible to make a clear distinction between monomorphemic and polymorphemic forms in English, as
polymorphemic Latin or French words may come directly into English, with no base form. This applies to words such as apply, affirm, offer, suffer, suggest, serene, etc, which all have fossilized Latin prefixes (e.g. \{ab+plicare\}, \{ob+ferre\}, OED 9724/130620).

8b.1.3 The influence of Latin morphophonology
In other cases, both the base and complex forms are Latinate but it may not be immediately obvious that they are related.

- species – special
- distance – distal
- anus – anal

The connection between <anus> and <anal> is rather like the problem with <aging> because only two letters of the base CPU are retained, making it a violation of the three-letter rule (in its simplest definition). With <aging>, however, there does at least exist the possibility of <ageing>, clarifying the base form. No such solution is available for <anal> because the spelling has been borrowed from Latin/French (OED 6994). Words borrowed from other languages can introduce new patterns into the writing system and, in this case, it is harder to identify the connections between related words.

The following examples provide further evidence of how Latin morphophonology can influence English spelling patterns, although more systematically:

- larva – larval
- angina – anginal
- vagina – vaginal

The word larval comes from Latin larvalis (OED 105915), so the spelling has been imported directly into English (with the final Latin syllable dropped), but anginal and vaginal seem to have been formed within English, by adding \{angina\}/\{vagina\} to \{al\} (OED 7516/221018). The spelling of these complex words involves <a>-dropping, an amendment pattern much like the <e>-dropping of *<hopeed> and *<tenseed>, etc. The recurrence of this kind of amendment suggests it is a design feature of the writing system. Carney (1994: 106) calls it dovetailing, and notes its occurrence in eighth, adapted from <eight + th>. The process is another case of IdP-reduction, and it allows for a practical compromise.
between the competing needs to represent the base spelling, the suffix spelling, and the phonological form.

A further reflex of Latin morphophonology is that there is an allomorph, spelt <ar>, which occurs after /l/, as a way of avoiding phonological haplology.

- pole – polar
- molecule – molecular
- angle – angular
- table – tabular
- tonsil – tonsillar

Some of these also have no stem in English:

- moon – lunar
- star – stellar

And again, we see dovetailing:

- fibula – fibular
- patella – patellar

The <ll> of tonsillar and patella(r) are historical doubles, imported from Latin, and not the result of CLD within English (OED 203228 / 138722; for more examples of such spelling alternations, see Cummings 1988: 170–2). Nevertheless, we can consider the <ll> of tonsillar to be an example of IdP-extension, this time sanctioned by etymology, rather than being a case of CLD.83

8b.1.4 Semantic specialisation

The presence of two allomorphs leads to the possibility of semantic specification, and this occurs with molar and molal (OED 120822/120819). The former, a dental term, comes from Latin molaris, while the latter is used in chemistry and is formed within English, from mole. This spelling has replaced the earlier spelling <molar>. This pair of words is not an example of DSTNCT spelling but rather a case of distinctiveness occurring in speech, suggesting that there is a deeper need in language to attach to each meaning a single form (spoken or written). Another pair is familiar and familial. The former has been an English word for some time and its meaning has diverged from its original sense of ‘pertaining to the family’, as one might expect. That gap has been replaced somewhat by familial, allowing for a semantic distinction to be made (OED 67957/67956).
8b.1.5 Stress and CLD

Examples of CLD are rare among {al} forms but this is because the conditions are seldom met. CLD does occur in the rare cases of Latinate words with /VC#/ $\equiv$ <VC#>, as in:

- re'fer – re'ferral
- de'mur – de'murral
- re'but – re'buttal

The spelling <referral> was formed within English, although its <rr> might be explained as a historical double, deriving from Latin referre (OED 160839) much like the <ll> of tonsillar (above). However, the <rr> of demurral and the <tt> of rebuttal were undoubtedly formed within English (OED 49887 / 159308), and this provides good evidence that {al} forms are subject to the same spelling formation patterns as basic inflections such as {ed} and {ing}.

In all other cases, stress is not marked. Compare:

- 'hexagon – hex’a’gonal
- 'baptism – bap’tismal
- 'homicide – homi’cidal
- ‘molecule – mo’lecular
- 'origin – o’riginal

While the stress pattern is not marked explicitly, the stress (and accompanying vowel quality) can usually be predicted from a combination of the letter patterns and the phonological patterns, just like we saw in Chapter 5. Words ending in {al} end in a phonological foot (trochee or dactyl), with the affix being unstressed (Plag 2003: 119). Thus <bap(tis.mal)> ends in a trochee, and has a closed penultimate syllable; <(ho.mi)’(ci.dal)> also ends in a trochee but has an open penultimate syllable. <hex(a.go.nal)> and <mo(le.ru.lar)> both end in dactyls, and thus the, ante-penultimate syllable has a stressed lax vowel. The base form <(o.rig.in)> is itself a dactyl, so the stress shifts to produce a new dactyl, counting syllables from the end of the word, hence <o(ri.gi.nal)>. Otherwise, the word would be a quartus paeon, with one stressed syllable followed by three unstressed <(o.rig.in)>.

Stress-shifting is common among Latinate suffixes in English (Plag 2003: Ch. 4; Giegerich 1999: passim). Compare also how the {ic} suffix shifts the stress in ‘atom – a’tic and ‘totem – to’emic.

Occasionally, the stress can shift in the other direction:
an’gina – ‘anginal / an’ginal (i.e. /ˈændʒɪnəl/ and /ənˈdʒɪnəl/)
va’gina – ‘vaginal / va’ginal (i.e. /ˈvædʒɪnəl/ and /vəˈdʒɪnəl/)

The presence of this variation between final dactyls and final trochees echoes the variation in the pronunciation of omega (c.f. Section 5e). Similarly, the words in question have the <VCV(C)#> ending which causes ambiguity in the decoding. The dactylic forms require a metrical decoding of <(an.gi.nal)> and <(va.gi.nal)>. Yet there is a tendency for these words to be reconstrued with final trochees, based on the reanalysis of the spellings, as <an(gi.nal)> and <va(gi.nal)>

8b.1.6 <y> ↔ <i> and Anglicising §Latinate borrowings
As ever, word-final <y> is changed to word-medial <i>, as in:

- colony – colonial
- deny – denial

This alternation can be viewed the other way around. The Latin word *colonia* is the root of both *colony* and *colonial* (OED 36524), and therefore the change has taken place in the Anglicisation of the base form <colony>, not in the related forms *colonial, colonise, colonialise*. Nonetheless, if the base form is formed within English, it will have <y> and will be changed to <i>, as in deny – denial (OED 49965). This means that the rule is consistent in both directions, showing that the spelling patterns have been fully integrated into the English writing system.

8b.1.7 Summary
The analysis above shows that the spelling-formation patterns for {al} derivatives are broadly similar to the spelling of other kinds of suffixation (e.g. {ed} and {ing} inflections). We see CLD when the usual pair of conditions are met (re’but – re’buttal), although CLD overlaps with the retention of historical doubles (refer – referral). We also see <y> ↔ <i> (deny – denial) and two kinds of IdP-reduction: <e>-dropping (culture – cultural) and <a>-dropping (larva – larval). Many {al} forms have been imported from Latin, along with its morphophonology, and this explains why we have pairs such as *patella, patellar* and *species – special*. While the distribution of <al> and <ar> spellings is usually conditioned by Latin morphophonology, the presence of two allomorphs allows for semantic specialisation to occur (molar – molal). These are not DSTNCT spellings per se — instead the spelling reflects distinctions that occur in spoken
language, suggesting that distinctiveness may be a principle of the language in general, and not just the writing system.

The spelling patterns occurring for the {al} affix may not apply to all suffixes. One example that jumps out from the text above is the word *dactylic*, where one might expect to see CLD, on the grounds that dactyl has a /VC#/ \equiv <VC#> ending. Compare *metal* – *metallic*. However, there is an historical double in the Latin *metallicus* (OED 117243), whereas there is no equivalent <ll> for <dactyl>. It may be the case that {ic} does not trigger CLD and, extrapolating further, each suffix may have its own peculiarities. A fuller study of the spelling patterns of affixed derivatives must thus examine the interlocking roles of etymology, phonology and morphology.

One thing that does seem to be certain, however, is that it is impossible to divorce etymological considerations from a discussion of the spelling of polymorphemic English words and, furthermore, that the act of retaining spellings from other languages is a special case of IdP. The only difference is that the preservation occurs *between* languages rather than between related words *within* English.

**8b.2 Derivational affixes 2: {-ble} suffix**

The {-ble} suffix has not been streamlined as neatly as {al} in the transition from French and Latin into English, and examining the variation therein will give us greater insight into the complex, and often inconsistent, interactions of etymology, morphology and phonology in the formation of derived forms. As with the {al} suffix, we shall see that some {-ble} words have been formed within English while others have been borrowed directly from French or Latin or both, and this has produced some hybrid results.

**8b.2.1 The hybrid spelling of {-ble} forms**

Below is a lengthy quote from the OED (20035) outlining the turbulent paths by which this suffix has entered English and how it is represented in the writing system. Afterwards, the quote is analysed so that it can be better understood in terms relevant to English spelling formation. The OED conventions are retained throughout the quote, but the most important lines have been emphasised here in bold, and the paragraphs have been numbered.

1. Old French *-ble*: [derives from] Latin *-bilem*, nom[inative]-*bili*-s, suffix forming verbal adj[ective]s, with the sense ‘given to, tending to, like to, fit to, able to’; as in sta-*bili*- ‘like, fit to stand.’ *After*
consonant stems, -ībili-, as vend-i-bili-, vinc-i-bili-; after stems ending in a, e, i, o, u, -ābili-, -ēbili-, -ībili-, -ōbili-, -ūbili-, as honōrābili-, delē-bili-, sepelī-bili-, nō-bili-, volū-bili-.

2. Some of these Latin words lived on into Old French, e.g. noble, amable, fleble, meuble; later, these became models for the extensive adoption of others from the Latin of literature.

3. Many of both sorts were from 12th to 15th century adopted in English from French, and here served as models for the direct adoption or formation of others from Latin, a process which has gone on to the present day. By far the most numerous of the -ble words are those in -able. In Latin, -ābili- adjs. arose only from verbs in -āre; but, in French, all pres[ent] p[articiples in -ant may give rise to an adj[ective] in -able, and as -ant is now the universal form of [the] pres[ent] p[articiple],

4. -able is the universal form of the adj. suffix as a living element; -ible being only a fossil survival in words from Latin like horrible, possible, visible, not directly attached to a living French verb. When the verb lives in French, a modern adj[ective] in -able has always taken the place of the earlier -ible form, as in vendable, croyable, préférable, for Latin vendibilis, credibilis, *préferibilis.

5. But in English there is a prevalent feeling for retaining -ible wherever there was or might be a Latin -ibilis; while -able is used for words of distinctly French or English origin, as conceivable, movable, speakable. Hence, where there is a verb in French and English, as well as in Latin, English usage is distracted by conflicting and irreconcileable analogies. Thus in the compounds of -fer, Latin fero, Latin analogy requires preferible, referible, sufferible; French example gives préférable, référrable; French and English analogy sufferable; English analogy gives referrable, conferrable, deferrable.

6. There is also a mongrel spelling referrible, sanctioned by Dr. Johnson, but defensible on no analogy. So with the variant spellings admissible, -ible, tractable, contractible, partable, -ible.
7. **These discrepancies** no mere etymological grounds can settle; though their number **might be lessened by following French precedent**, and extending -able to all words having a verb (with the same accent) in English, thus admissible, contractable, corruptable, exhaustable, vendable, etc., leaving -ible in credible, intelligible, legible, possible, etc.

8b.2.2 Analysis

The modern English story begins in Paragraph (3), where the French forms ‘served as models for the adoption or formation of others from Latin’. This is very important because the English spellings are chosen not because of how the words are (or were) pronounced, but because there is a precedent for their spelling, either in French or Latin. In other words, IdP is being obeyed, irrespective of the pronunciation. This process, it is noted, ‘has gone on to the present day’.

In (4) we see that the spelling <able> is the only productive form nowadays and the spelling <ible> only occurs in fossilized forms. A good example of this can be seen in noticeable. If English spelling were looking for maximal efficiency, then *<noticible> would be a better spelling, much like <noticing>, with <i> performing the marking function on <c>. Instead, the unwieldy <eable> sequence is maintained, preserving IdP, but resulting in a (relatively minor) PhM conflict, <ea> ≡ /ə/ (disregarding the morpheme boundary). Because notice is not a direct borrowing from Latin, but comes via French (OED 128591), it does not pattern like reduce – reducible nor does it vary like <seduceable, seducible> (OED 174727 / 160512).

The emergence of <able> as a dominant form over <ible> tallies with Berg and Aronoff’s (2017) convincing argument that, where there are competing variants, one spelling tends to win out eventually, and this applies to affixes as well as stems (c.f. Section 3d). It may be the case that <able> and <ible> are still going through this phase, with <able> the clear favourite to take over. However, there is a need for the suffix to also be a marker of consonant quality, at least after <c> and <g>, which means that there ought to be some functional demand for <ible> over <able> (e.g. invincible, incorrigible).

Beyond case-by-case concerns about phonographic matching, or even the higher-level concerns over keeping the spelling of the morpheme constant, English spelling is often subject to a reverence for Latin over the need for system-wide stability. That is why ‘there is a prevalent feeling for retaining <ible> wherever there was or might be Latin <-ibilis> ’ (5). The solution in (7) is very revealing. The OED are saying, without using the linguistic terms, that <ible>
spellings are permissible when attached to bound morphemes but inadvisable when attached to free morphemes. Hence, <possible>, <legible> and <credible>, etc. have spellings that reflect their Latin roots, as do <probable> and <tractable>, while <contract-able> and <refer-r-able> are formed within English, making possible amendments such as CLD. The latter spelling patterns identically to referral and rebuttal in the last section, as it meets the necessary conditions for CLD.

It might appear that etymological spelling and morphemic spelling are potentially at odds but there remains one tendency that underlies all of these examples — complex pleremic units cannot be changed on the inside. Johnson’s (1755) ‘indefensible’ sin was to use an English-only pattern (CLD) within a complex pleremic unit <referib>. The sequence <referib> cannot be changed in the middle, so *<refer-R-ible> is not acceptable, even though <referib-LE> and <referib-ILITY> are both fine. Hence <ible> should only occur in fossilized borrowings (such as <referib>). It is for the same reason that another <de> cannot be added to edit, making it §Basic *<eddit>, and yet <t> can be added in <emit.t.ed>, when the changes occur at the edge of the CPU. Such changes on the margins of CPUs are necessary in order to make English spellings function and to simultaneously obey PhM. This is why we can have IdP-reduction and IdP-extension, but not CPU-internal changes. In short, <ible> is not considered to be a productive unit of English spelling — it only occurs as part of a borrowed complex pleremic unit such as <referib>, <horrib>, <visib>, etc. By contrast, *<noticib> does not occur in Latin and this is why we must continue with <noticeable> rather than *<noticible>. In other words, the spelling of affixed derivations (as well as inflections and compounds) can only be built out of existing strings of letters, new strings cannot be invented.

The foregoing discussion of the spelling of words containing the {-ble} suffix shows that the processes by which spellings actually enter the English writing system are often more complicated than the way in which complex words are spelt when formed within English. And while it is not always possible to make a rigid distinction between what is English and what is French or Latin, one deeper principle has emerged from the discussion. Videlicet, English spellings are not just purely morphemic or purely etymological. Instead, they are formed from CPUs that may develop within English or which may be adopted from other languages, and these units have given rise to the subsystems of English spelling. Changes to CPUs, such as phonographic amendments, can only occur at the margins such as morpheme boundaries.
8c. Compound words

Perhaps the most basic kind of word-formation process is compounding, which is, ‘roughly speaking, the process of putting two words together to form a third’ (Bauer 1983: 11). For example, girlfriend is intuitively one word yet it is combined of girl and friend (Plag 2003: 5). The spelling of compound words simply involves concatenating the spelling of the two input morphemes — in other words, obeying IdP without amendment. The only variation is whether those spellings are written solid, with a space or hyphenated. This is an orthographic, rather than a linguistic issue, and the presence of a space does not indicate ‘two words’. Thus <bar tender> and <bar fly> are established compounds even though they may or may not written with a space. It is also possible to invent compound words afresh (e.g. morphemic spelling, writing system, phonographic-matching principle) in order to label a concept, and there remains some degree of choice over whether these should be hyphenated, written with a space or written ‘solid’, making it one orthographic word.

Bauer (ibid) lists out the following examples of compounds, all involving the word paper.

- oil-paper
- paperclip
- paper aeroplane
- paper thin
- (to) wallpaper
- wastepaper
- wastepaper basket

Again, Bauer’s examples are elucidated somewhat by the spelling, which is useful for present purposes. Here we see <paper> being attached in various ways to other words. <oil-paper> is hyphenated whereas <paper aeroplane> is separated by a space and <paperclip> is solid. The distribution of hyphens and spaces is notoriously complicated, although not entirely capricious. Burchfield and Fowler (1996) call hyphenation a ‘stretchingly difficult subject’ and even Carney (1994: 48–49) has trouble making any major pronouncements on how hyphens, spaces and solid spellings can represent relations of semantics, syntax and prosody between two or more connected words. Familiarity is surely another factor: to-day and to-morrow were once hyphenated but they are always solid nowadays (OED 212816 / 2013126), so it would appear that some compounds gradually change from being spaced to hyphenated to solid. More recent studies suggest that
morphological complexity can play a role (Rakic 2009, Schmid 2005, Berg et al. 2014). This is why *barman* is less likely to be spaced or hyphenated than *bartender*, because *tender* is morphologically complex, whereas *man* is simple. This issue becomes particularly noticeable in three-way compounds such as *armchair critic*, where the complex form *armchair* is a compound unto itself, separate from *critic*. The spelling of compound words is thus of interest to anyone concerned with the distribution of hyphens and spaces, something that is beyond this study.

From the viewpoint of letter sequencing (i.e. spelling), it is not important whether spelling *<A>* + spelling *<B>* is spelt as *<AB>* or *<A-B>* or *<A B>*. However, there are two problems that can arise, one orthographic, one phonographic. In the first case, the concatenation can lead to the co-occurrence of unwanted letter sequences, *still life* would not be easily parsed if it were spelt *<stilllife>*. However, *pondlife* may well be spelt without a space, as its letters can be more easily parsed, due to the absence of words beginning in *<dl>* or ending in *<ndl>*. Phonographic issues arise in words such as *uphill* and *cathouse* where the *<ph>* and *<th>* sequences might be misinterpreted as /f/ and /θ/, respectively, unless the morpheme boundary is identified (compare *euphemism* and *catholic*; see Venezky 1967: 90).

The spelling of compounds has long been grouped with the spelling of affixed derivation in the belief that both kinds of words are subject to morphemic spelling. This idea dates back to at least Venezky’s (1970: 120) conclusions about the writing system. Yet there are some very important differences between the groups, despite the overlaps. Problems arising in the spelling of compounds can only be solved *between* the spelling of the morphemes, using a space or a hyphen. Problems arising in the spelling of affixed derivation and inflections are solved *within* the spelling of a single orthographic word. However, this can only happen if a solution is available. The confusion between compounds, on the one hand, and inflections and affixed derivatives, on the other, arises because many of the latter group are spelt using simple IdP. Sometimes this makes good sense (e.g. *edited, editing; editor, editable*); while other times, there are simply no available solutions (e.g. *signed, signing; signal, signify*). The existence of so many amended forms (e.g. *rebutted, rebutting; rebuttal*) tells us that these groups are subject to different spelling-formation patterns from compound words.
8d. Neo-classical compounds

Neo-classical compounds are words formed using Greek and/or Latin combining forms, as seen in ortho-graph-y, tele-vision, tele-phon-e and chloro-form. Such words did not exist in either Latin or Greek and have been formed in modern times within English or French.

Once again, IdP seems to be the primary force behind the spelling of these words, although again, PhM plays a very important auxiliary role. For example, neo-classical compound sociology was formed by merging the CPUs <socio> and <(o)logy>. The latter unit appears in biology, anthropology and mineralogy, etc. (OED 109868). In all of these cases, an amendment has been made, with <o>-dropping occurring in all cases, in order to make a better phonographic match. In the case of <mineral+ology>, an <l> has also been dropped as a way of avoiding haplology. A similar pair of changes occurs in adverbs such as singly, doubly and wobbly, with both <e>-dropping and <l>-dropping. It seems that such dovetailing can occur for any kind of letters, providing further evidence of the need to check draft spellings (e.g. *<mineralology>) against phonological forms. Combining forms tend to recur in several words. The sequence <socio> begets sociolinguistics, sociophonetics, sociobiological, sociopath, sociometric and socio-economic.

From the viewpoint of spelling formation, neo-classical compounds are more like inflection and affixed derivation than compounds because any conflicts are usually resolved within the orthographic word. However, socio-economic is often spelt with a hyphen, so as to help in parsing its two morphemes. In this case, it is spelt more like a compound word than affixed derivative. Thus the spelling of neo-classical compounds may be subject to its own peculiar set of patterns.
8e. Blend words

Blend words involve the mixing of two truncated English words. An example of a blend is Oxbridge, blending Oxford and Cambridge. Another is Tanzania, the country formed from the merger of Tanganyika and Zanzibar. IdP-reduction occurs in blend words where there may the concatenation of two reduced CPUs, as opposed to the addition of chopped up elements. Thus, we do not see *

<Coxfridge> or *

<Tanzanyizibar>.

Changes are generally not made on PhM grounds (unless somehow necessary), hence bootylicious is not respelt as <bootylishus>. Such a spelling would diminish the IdP connection with the parent form delicious. Indeed, the CPU <licious> recurs in several combining forms, including sodalicous, popalicous and groovalicious (OED 259095). The distinction between blend words and neo-classical compounds may be one based on etymological grounds more than morphological grounds.

There can be families of blend words. For example, university societies often have names that end in <soc>, as in DramSoc, LitSoc, FilmSoc, Sci-Fi Soc, etc. It is important that these blend words are not spelt with <sock> or <sok>, as this would break the IdP connection with the fuller form of the CPU. The process of IdP-reduction can be rather attritional and it may be difficult to recover the original units. The blend word <Brexit> only maintains two letters of the parent form <Britain> and one clearly needs context to identify the parent words. In this, blends become very similar to abbreviations.

In other cases, IdP-reduction can result in homophonic forms, as in <psy>, <sci> and <sy>. The first of these occurs in psy-trance, psybient and The Psy Press, clearly relating to psychedelic or psychology. The second occurs in Sci-Fi, while the latter occurs in Syfy, a TV channel that rebranded itself using a constructed homophone in order to distance itself from its science fiction roots. CPU-reduction generally occurs on the edges of the ‘parent’ form without changes to the middle. Hence there can be no *

<Pcy> or *

<si>, etc. because the visual connection would be lost.

It is possible for blends to contain some CPU-internal changes. For example, the words mockumentary and rockumentary, respectively blend mock and rock with documentary. The latter word is compromised orthographically, but not phonologically, because a <k> is added inside the reduced CPU <ocumentary>. There is the option of spelling them <ocumentary> and <ocumentary> but this would obscure the visual connection to the words <mock> and <rock> without any phonological gain. There therefore seems to be a certain level of compromise
required between the need for IdP and PhM, albeit the patterns may be very
different from other kinds of English word-formation patterns. The manner in
which sound and meaning are represented in blend words can be a lot more
complicated than other kinds of word formation (see Arndt-Lappe and Plag 2013).

Blend words may represent the limits of phonological word-formation but
it is possible to include much more creativity in written words, because letters are
more easily manoeuvred than sounds. This kind of wordplay occurs in James
Joyce’s novel *Finnegans Wake*. For example, the spelling *<funferal>*>, also spelt
*<funferall>*>, suggests a word homophonous with the phrase ‘fun for all’
(1939: 21:15; 210:10; 186:16). However, the spelling also looks rather
like *funeral* and indeed the word is description of an Irish wake, an event which is
often a sombre jamboree.

Such creative spelling is often used in newspaper headlines, especially
tabloids, in order to render as much relevant, new information in as little space as
possible. These formations rely on meanings being formed through both sound
and sight, but the visual connections between spellings may be more complicated
than the simple IdP discussed throughout this study. In the example of *<funferal>*,
an *<f>* has been inserted inside the CPU *<funeral>*; and yet we may still *see* that
unit. Yet breaking the normal connections between spelling and meaning may
tempt the reader to look further, unearthing *<fun>* and *<feral>*; and these in turn
may add further shades of meaning to the ‘word’ *<funferal>*. Such spellings
destroy the illusion that there is a simple one-to-one mapping between spelling
and meaning, argues Attridge (1988), and once that illusion is broken in the
reader’s mind, the potential for creative interpretation becomes almost limitless.
The downside is that it makes fluent reading very difficult (a major problem for
longer texts) but the upside is that more information can be packed into smaller
spaces, and seemingly disparate ideas can be humorously juxtaposed. This is why
the phenomenon is so common in tabloid newspapers. Regardless of the effect on
the reading process, these words remain rooted in the basic idea that meaning is
represented visually in writing and, even if CPUs are chopped and separated (as in
*Coxfridge* above), it may still be possible to recover the related words and
morphemes, albeit with greater difficulty than is the case in simpler words.
Abbreviations come in several forms. An initialism, such as FBI, abbreviated from Federal Bureau of Investigation, is formed by taking the first letter of each (important) word. An acronym is a specific kind of initialism whereby the letters form a readable word. Hence laser is an initialism formed from Light Amplification by Stimulated Emission of Radiation. However, there are other kinds of abbreviations which may be formed by cobbled together the most important letters, in order to help the reader reconstruct the original form. And there are also some heavily restricted abbreviations, such as airport codes and file extensions, which must be spelt within very strict limits. While these written forms do not form new English words per se (unlike laser), they can tell us a lot about the relationship between spellings and words. Abbreviations have received very little attention in the literature on English spelling, although they have not been exempted from studies of word formation (e.g. Cannon 1989). In this section, I shall argue that abbreviations are spellings par excellence because they reveal to us the very nature of how words are represented in writing.

Carney (1994: 105) omitted ‘most abbreviations’ from his corpus, including examples such as i.e. and viz. He is not clear about what he omits, only what he includes, namely clippings (e.g. veg, meths, discussed below), and ‘abbreviations which would read as the full orthographic form’. An example is cwt., meaning ‘hundredweight’, an imperial unit of measure. In this case, <c> is used as a morphogram, meaning the Latin word centum, ‘hundred’, and <wt> is a consonantal spelling of the world ‘weight’. This kind of mixed spelling is redolent of non-alphabetic writing systems (see Section 2a), but it is rare in English, except in abbreviations, where we can see evidence of several different kinds of writing system: alphabetic, consonantal, syllabic, morphographic, and various combinations thereof.

Like other kinds of English spelling, abbreviations are subject to IdP. A simple initialism such as a.s.a.p., short for ‘as soon as possible’, simply uses the first letter of each word. This is IdP-reduction. It could be argued that this abbreviation also expresses PhM, since <a> ≡ /æ/, <s> ≡ /s/, etc. Yet in a.k.a., short for ‘also known as’, the <k> of known is represented, despite the fact that the phonological word begins with /n/. This suggests that the retention of the first letter provides an easier path to reconstructing the full spelling than the marking of the first phoneme. A notable exception to this, familiar to linguists, is MFM, the Manchester Phonology Meeting. Here the <F> is chosen for the sound of the word, not the spelling. Other kinds of abbreviations may also be subject to IdP.
For example, the days of the week are commonly written as Mon, Tue, Wed, Thu, Fri, Sat, Sun, preserving the orthographic identities of the full forms. A more phonetically transparent rendering might give *<mun>, *<chu>, *<wen>, although this would be subject to accent variation. Some speakers might prefer <too> and/or <wed>. These abbreviations are slightly different from clippings because they are not designed to be decoded as full words. When one sees <Tue>, one reads Tuesday, not the monosyllabic form suggested by <Tue>. One thing that is common in abbreviation, observes Cannon (1988) is that they tend to have exactly three letters, a condition that is very different from clippings, for example.

Such restrictions may be obligatory. For example, file extensions used to be restricted to just three symbols. Examples include (.doc) for Microsoft Word document, (.txt) for a ‘text file’ and (.xls) for a Microsoft Excel file. These examples, despite the stringent constraints, display spelling patterns which are common among the world’s writing systems. A (.doc) file is a clipping of document, and a simple example of IdP-reduction. A (.txt) file is an example of consonantal spelling but it is also an example of a new kind of IdP which we can call IdP-compression, as the CPU is compressed from <text> to <txt>. The letters and their order are maintained, but not all of them. The example of cwt (above) also contains IdP-compression, as <weight> is compressed to <wt>. The example of (.xls) is rather different. <xl> is a constructed homophone of excel, with two nyrophones. The <s>, however, is a simple initialism of spreadsheet.

These abbreviations, which may or may not be defined as ‘spellings’ have a lot in common with the writing seen in text messages (Crystal 2008; Shortis 2016), internet messages (Silva 2011), and increasingly elsewhere. Silva argues that the abbreviated messages seen in internet messages (in both English and Portuguese) have resulted in users inventing an ad hoc mixed writing system showing features of non-alphabetic scripts. She writes:

Although net users do not realise it, they recreate features mainly from morphographic systems and consonantal scripts and, consequently, they seem to be searching for a more economical way of conveying meaning through writing.

Silva (2011: 143)

Silva argues that the aim in instant messaging is to abbreviate words while maintaining an adequate visual connection to the full spelling. Consonants tend to be retained a lot more than vowels and syllable onsets in particular, so that common Portuguese word like tudo, meaning ‘all’, is reduced to <td>. This suggests that consonant letters are more important in readers’ word recognition.
Silva (2011) argues that representing syllables is particularly important in Portuguese writing but in English the most important unit is the word. This may be due to English’s more complex syllable structure or because it is a stress-timed language, making it more important to represent stressed syllables than unstressed.

In this world of compressed spelling, it becomes clear that the purpose of writing is to represent words (and hence meaning) in an expedient fashion, and that the isomorphism between spelling and sound is simply an illusion. That is to say, an alphabetic writing system may provide a sizeable amount of phonetic information about a word, but it does not provide all of it. Abbreviations are designed to provide just enough information to help the reader to reconstruct the fuller form of the spelling and, from there, the actual word itself. IdP, in its various manifestations, provides a direct orthographic path between the abbreviation and the full spelling, whereas PhM would introduce an unwelcome phonetic intermediary. The bigger take-away from this, however, is that the relationship between abbreviations and spellings is similar to the relationship between spellings and words. Knowing that an abbreviation is just a mnemonic for a full spelling, it is not much of a jump to consider a spelling to be a mnemonic for a word, an idea expounded by Lass and Laing (2008–13). This brings us back to Sproat’s (1996) observation that the primary purpose of writing is not phonetic transcription but the representation of words and morphemes. Representing phonological information is simply a useful way of helping readers reconstruct a word from its written form, but it is not necessary. Individual writing systems tend to prioritise one kind of information over another and yet, as we have seen, all of them include information about both the sound and the meaning of words, an observation that applies as much to abbreviations as it does to normal English spelling as it does to Chinese writing, and so on.
8g. Clippings

Clippings, such as *phone* and *telly*, are words formed by truncating a full word. These words have great potential for theories of spelling because what shortens well in sound does not always shorten well in spelling, so the solutions to correspondence conflicts can tell us a lot about the workings of the writing system, and how certain principles may be prioritised over others. Despite this, clippings have been largely neglected in studies of English spelling, perhaps because they are informal variants of existing lexemes (a ‘telly’ is the same thing as a ‘television’) and perhaps because their spellings are not always faithful to the parent form. Why, for example, is the spelling <telly> preferred over <tele>?

8g.1 Plag’s analysis of clippings

Plag’s (2003: Ch 5.2) discussion of clippings begins with an examination of clipped names, which are particularly common, and then extends to clippings of lexical words. He observes a great deal of systematicity in the interaction of clippings with prosodic categories. Examples of clipped names include *Mike* from *Michael* and *Trish* from *Patricia*. Examples (*ibid*: 154) of lexical clippings are the following:

1. ad – advertisement
   condo – condominium
   demo – demonstration
   disco – discotheque
   lab – laboratory
   photo – photograph
   porn – pornography
   prof – professor

2. fax – telefax
   phone – telephone

The same kind of prosodic rules apply to both clipped names and clipped lexical items, allowing us to study them together. Where the beginning of the word is retained, as in (1), they are called *back clippings*. Often it is just the first syllable, as in *ad*, *lab* and *prof*, but it can be a whole foot, as in *demo* and *disco*. Where the end of the word is retained, as in (2), they are called *fore clippings*, as in *phone* and *fax*. Other examples include *blog* from *weblog* and *ludes* from *Quaaludes*. Plag gives no lexical examples of middle clippings although he had *Trish* –
Patricia among his nominal examples. Other examples of middle clippings include *flu* from *influenza*, and *fridge* from *refrigerator*.

8g.2 The spelling of clippings

Most of Plag’s examples are only of interest here in that they reinforce existing patterns of spelling. This is most obviously the case where there is a simple linear mapping from spelling to sound and back. *<ad>* and *<lab>* are undoubtedly the most basic possible spellings for those syllables. The same could be said of *<pop>* and *<fan>*, taken from *popular* and *fanatic*. All of these spellings involve IdP-reduction and none of them affects PhM. The spellings *<phone>* and *<photo>* are a fraction more complex since they might have been simplified to *<fone>* or *<foto>*. Instead, the *<ph>* is retained and the clipped form shows that IdP can be an important factor in the spelling of clippings. All of the examples above obey IdP, except for the middle clippings, *Trish* and *fridge*.

The fact that these words are middle clippings is not important, from the viewpoint of spelling. Instead, it is the fact that the reduction occurs after a palatalized consonant. Johnson (23/1/12) observes similar issues in clipped forms of *pleasure*, *usual*, *casual*, *legend*, *tragic*. In each of these spellings, the recognition of the palatalized consonant is reliant on the presence of succeeding unstressed syllables. The following clipped spellings do not indicate palatalization clearly: *<pleas>*, *<us>*, *<cas>*, *<leg>*, *<trag>*. Thus there is a demand for these IdP-reductions to be amended in order to obey PhM. And this time we cannot re-define IdP to help explain the data. No amount of theoretical manoeuvring can consider *<plezh>*, *<yoozh>*, *<cazh>*, *<ledge>* and *<tradge>* to be examples of IdP. Thus we must conclude that the spelling of clippings is different to the spelling of other kinds of words. The question is how different.

If we assume that PhM is the highest-ranked constraint, then this explains why *<Trish>* ousts *<Tric>* and *<fridge>* ousts *<frig>*. But it does not explain exactly why *<frydge>* and *<frijj>* are excluded. Thus there must be some mechanism in the analysis for explaining why one kind of spelling pattern, e.g. §Basic, is preferred over another. But we must also explain why *<veg>* and *<Reg>*, short for *vegetables* and *Reginald*, do not change to *<vedge>* and *<Redge>*. One possible factor is DSTNCT. Since *<frig>* is already an existing word (OED 74675), then that spelling could be re-used as a clipping for the much newer word *refrigerator*. DSTNCT may also play a role in the clipped form of *microphone*. The IdP-reduction draft spelling is *<mic>* and this appears to be a bad phonographic match for /maɪk/. Thus we might expect *<mike>* to be preferred, except that this too may be blocked by the name *Mike* (although it is
capitalised; see also, Wells: 10/2/12). And yet <mic> is the more common spelling (although not in the inflected forms, as we saw in Section 4c.1.4). The problem again arises due to decoding patterns that rely on words being polysyllabic. In <mi.cro>, the g-syllable <mi.> is open and can thus be read as a long vowel. In <mic>, however, it appears to be closed. A similar problem arises for <Raph>, short for Raphael, where <a> ≡ /eɪ/ in a closed g-syllable.

A full analysis of clippings would need to assess the relative merits of IdP, PhM and DSTNCT, and to analyse the degree to which readers will tolerate correspondence mismatches (e.g. <mic> ≡ /maɪk/) in order to preserve IdP. Working in the reverse direction, it is worth asking to what extent spellings can be changed to match sound, at the expense of maintaining connections between the clipped and the parent form, (e.g. <yoozh> for usual, a spelling which uses a correspondence which is not established, and may be difficult for many readers). Other factors may play a role. The frequency of the clipping’s occurrence, the register of language in which the word is used, and the written domain in which it occurs. Clippings occur in informal language and such a setting may allow for a rather different set of spelling constraints.
8h. Conclusion

This chapter makes two major claims. One is that etymological spelling is a special case of IdP, and therefore conceptually linked to morphemic spelling, and the other is that different kinds of word formation category may be subject to different kinds of spelling patterns. It is hoped that enough evidence has been presented here to show that each group merits further analysis.

Two important kinds of etymological spellings stand out. Adoptions without adaptation, such as gateau and macchiato, occur when the identity of the borrowed word’s spelling is preserved whole; and adoptions with adaptation, such as philosophy, larval and possible, where a word has been borrowed from another language (often a combination of French and Latin) but its spelling has been adapted somewhat to fit in with English writing system. Members of the first group can contribute new correspondences to the writing system, or they may strengthen existing, minor correspondences (e.g. <cch> ≡ /k/, <eau> ≡ /əʊ/). These words are often subject to spelling pronunciations, due to the difficulty of integrating foreign borrowings into English phonotactics (e.g. the syllable /kjə:/).

Adoptions with adaptation are common in affixed derivation, because the spelling of English suffixes varies considerably from the Latinate equivalent. The spelling of words containing the {al} suffix is relatively straightforward as these words are all modelled on the Latin ending <alis>. Thus we see the usual amendments, including <y>-to-<i> swapping (denial), CLD (referral), <e>-dropping (modal), and <a>-dropping (vaginal), a new kind of dovetailing. The latter example has two English pronunciations, both decodable from its spelling, thus providing further evidence for the argument, presented in Chapter 5, that variant pronunciations which are sanctioned by the spelling may gain traction over time.

The spelling of words containing the {-ble} suffix is more complex because there are two etymological forces, Latin and French, and it is their morphophonology which leads to the otherwise inexplicable difference in spelling between possible and probable. The two languages often provide different models for English spelling, resulting in variation such as <contractible> and <contractable>. In all cases, however, the Latin and/or French spellings have a major impact on the English spelling. As a result, etymology interfaces strongly with the usual needs of representing English morphology and phonology within the confines of the writing system. Thus, the spelling of affixed derivation is similar to the spelling of inflections, but significantly more complex. And yet the two groups share many
features, most notably the tendency for orthographic conflicts to be resolved within the spelling of the word.

This contrasts significantly with the spelling of compound words where conflicts are solved between the spelling of the input words, using a space or a hyphen. This is why still life requires a space, unlike pondlife. The spelling patterns of compound words may interface strongly with syntactic categories, as recent literature has begun to show, but not so much with morphophonology. Hence compound words have not been of great interest in this study. The spelling of neo-classical compounds may involve patterns similar to those seen in affixed derivation and also in compounds. Once again, we see conflicts being resolved within the spelling. Dovetailing explains why sociology is not spelt *
<socioology>, and mineralogy is not *
<mineralogy>. However, socio-economics may be spelt with a hyphen. The same goes for the word neo-classical itself.

The spelling of clippings may require a rather different analysis. Some clippings adhere to IdP (e.g. veg, mic) but others require changes (fridge, Trish). It seems that PhM and DSTNCT can play a very important role. The <sh> of <Trish> is introduced to represent /ʃ/, while <fridge> is distinct from the existing word <frig>, even though the latter spelling may have otherwise sufficed.

Blend words seem to preserve the input CPUs (e.g. bootylicious), despite major orthographic attrition (Brexit) which may obscure the parent forms. The examples presented here are not plentiful enough to make even the most tentative conclusions, except that the spelling of these words may be a lot more complicated than the spelling of other words. This reflects the complexity that underscores their original formation. Adding to the mix is the fact that it is also possible to have blend words which only function when seen (e.g. funferal), and such spellings can combine words in ways that can never be done in speech alone.

Abbreviations reveal more about the nature of writing than any other group of spellings, even though they may not be considered as real spellings themselves. The purpose of an abbreviation is to index the full spelling and IdP is very important in that task. Hence aka represents the initial <k> of known, not the initial /n/. However, some abbreviations may involve drastic reductions of the parent form, using a combination of writing-system strategies. For example, cwt uses a Latin initialism/morphogram <c>, mixed with consonantal English writing <wt> in order to represent the English compound word hundredweight. Such examples show that it is not always necessary for written forms to represent phonological forms faithfully. This brings us back to Sproat’s (1996) point, discussed in Section 2a, that the primary purpose of writing is to represent words
and morphemes. Just as an abbreviation is a mnemonic for a spelling, so too a spelling is a mnemonic for a word. The fact that the spelling may represent the pronunciation of the word with a high degree of accuracy is only its auxiliary function, and that explains why spellings can vary in their degree of phonography, both within English (compare <8>, <eight> and <ten>) and across the writing systems of the world.

***

***
9. Conclusion

This study attempts to explain how new English spellings are formed and methods from previous studies of English spelling are combined with methods from studies of English word-formation. Instead of studying the writing system as an amorphous whole, the aim has been to subcategorise it according to different word-formation categories. The focus has primarily been on the spelling of polymorphemic and derived words, most notably regular inflections and, to a lesser extent, affixed derivatives. The latter chapters look at non-affixed derivatives such as clippings and abbreviations. Evidence is also drawn from names and non-standard spellings. These categories allow us to examine productive patterns of spelling, and to see how complex units of spelling are combined. Such insights are not possible in studies which focus on the correspondences between letters and phonemes in the more common English words (notably Venezky 1970; Carney 1994).

9a. Summary of findings

In Chapter 2, six questions were put forward and we can now answer them in full.

1) To what extent does English spelling represent morphemes consistently?

2) How does this interact with the need to represent the phonological form of a word?

Maintaining orthographic connections between related words is perhaps the most defining aspect of English spelling, and it applies to almost all kinds of formation categories, to varying degrees. In earlier theories, these connections could only be explained through the concept of morphemic spelling (Venezky 1970), or lexical spelling (C. Chomsky 1970). Such theories failed to explain why pairs such as \(<\text{profane}, \text{profanity}\>\) are not completely morphemic, as the spelling of one of the words contains an \(<\text{e}\>\) and the other doesn’t. Instead, the system was considered to be morphemic ‘as much as possible’ (Venezky 1970: 120) or ‘to a large extent’ (Carney 1994: 18). Berg et al. (2014) argue convincingly that the spelling of affixes, in this case \(<\text{ity}\>\), are less variable than the spelling of base forms, in this case \(<\text{profane}\>\). However, Berg et al. do not adequately explain why the spelling is not simply *\(<\text{profaneity}\>\).

A partial solution to this problem was to introduce a much more flexible concept known as the Complex Pleremic Unit (CPU), whereby a string of letters can be associated with a set of semantically related words, with the advantage that
the string can be changed at the margins. Thus <profan> and <profane> are instances of the same CPU. In the following set of words, we see the same recurring element with minor alterations: <social>, <sociability>, <society>, <sociolinguistic>, <associate>, <soccer> (OED 183733), <FilmSoc> and <DramSoc>, etc. All of these words contain <soc> and it is not permissible to spell them with <sock> or <sosh>, despite the pronunciation, because those spellings would suggest unrelated words.

It was therefore hypothesised that the default English spelling pattern is to concatenate the spelling of the input forms, and this theoretical act was formalised as the Identity Preservation principle (IdP). The simplest kind of IdP is morphemic spelling, and it is best seen in the spelling of compound words, where no changes are made to the spelling of the input forms. Where conflicts occur, they are either ignored (e.g. uphill, hothead), or they are resolved with a space or hyphen. This accounts for why <pondlife> is one orthographic word but <still life> is two, although the study these patterns was beyond the scope of this study.

The focus instead was on word-categories where orthographic conflicts have to be solved within the spelling, allowing for multiple CPUs to be integrated into a single complex spelling. This required us to recognise that the writing system has a number of amendment strategies, such as dovetailing and CLD (consonant-letter doubling), used to maintain relations between spelling and phonology. In this way, it is possible to obey IdP, as much as possible, but also adhere to the Phonographic Matching principle (PhM).

**Inflections**

Regular inflections provide the perfect testing ground for examining English spelling formation because they are highly productive, they pattern identically across the spelling subsystems, and they do not affect the prosodic shape of the base form. Thus we can test the spelling patterns without any interference from phonological changes (such as stress-shifting, for example). Preserving IdP, or keeping the spelling of the morphemes constant, is both the ideal and the default pattern, hence jumping, going, singing and shoeing. In many cases, however, amendments are required to maintain readable relations between spelling and sound, and the most common involve CLD (consonant-letter doubling) or <e>-dropping. Examples include <hop, hopped, hopping> and <hope, hoped, hoping>. CLD can be considered as an example of IdP-extension because the extra letter is predictable from the spelling of the base (c.f. Berg 2014) while <e>-dropping is a case of IdP-reduction. Berg (2016) argues that CLD can be predicted solely from
the spelling of the base form, but it was argued here that both the orthographic and phonological conditions must be met. Thus \textit{fix} meets the orthographic conditions (ending in \textit{<(C)VC#>}), but not the phonological conditions /VC#/; hence we have \textit{fixed, fixing}. Further evidence comes from words whose base forms contain spelling patterns which deviate from the norm: \textit{<quiz, quizzed, quizzing>, <dé but, dé buted, dé buting>, <ok, okayed, okaying>, <mic, miked, miking>}, etc.

It seems that correspondences between spelling and sound have become so stabilised over the course of the writing system’s long history that \textit{almost} all inflected spellings can be predicted from the letter-sequences. This makes it very difficult to distinguish between orthographically- and phonologically-motivated spelling patterns. However, total synergy between spelling and sound has not been achieved and variation will occur in cases where the relations have not yet stabilised, as in \textit{<sync, sync(h)ed, sync(h)ing> and <zinc, zinc(k)ed, zinc(k)ing>}. This suggests that writers are very aware of the conflicting needs in English spelling, namely representing the sound and meaning of words as unambiguously as possible.

Finally, it is worth reviewing when \textit{IdP} is used by default, as occurs when changes are either impossible or unhelpful. Words such as \textit{<shoeing> and <eyeing>} need to maintain their \textit{<e>} because dropping it would overly compromise the (irregular) relations between the pronunciation and the spelling of the base form (e.g. \textit{<ey> ≡ /aɪ/ in *<eying>}). A very small number of words maintain \textit{<e>} because it allows for a \textit{DISTINCT} spelling such as \textit{<singeing>} or \textit{<dyeing>}. This principle is not obeyed in cases where the complex spelling results in a homograph of a monomorphemic word (e.g. \textit{moped, evening}).

\textbf{Affixed derivation}

Affixed derivatives show very similar spelling patterns to inflections, but they can be complicated by etymological influences (see Question 3 below). Once again, \textit{IdP} provides the default spelling (e.g. \textit{functional, hexagonal}), but amendments can be made, if necessary, where possible. In the case of \{al\} forms, we once again see CLD (\textit{rebut, rebuttal}) and \textit{<e>-dropping (mode, modal)}, but also new changes such as \textit{<a>-dropping (angina, anginal)}. This suggests that dovetailing is an important design feature of the writing system, an hypothesis which requires further testing across a much wider range of affixes. Shifts in stress are not marked explicitly in spelling but readers can learn to recognise how to decode such words (see below).
Other kinds of word-formation category
The study also looked briefly at non-affixed derivation such as clippings and abbreviations where it was found that IdP and PhM have different roles to play. Clippings frequently adhere to PhM at the expense of IdP (e.g. fridge, Trish) and this is because what shortens well in speech may not shorten well in spelling. The reason is that decoding patterns may span the foot, rather than the syllable. Thus the <c> of <Pa(tri.ci.a)> can be reliably decoded as /ʃ/ in this position, but not in <(Tric)>. Abbreviations, however, may index the original spelling, hence the <k> of a.k.a., rather than a phonologically motivated <n>. In this way, abbreviations function as mnemonics for spellings and this is a reminder that spellings themselves are, essentially, just mnemonics for words (c.f. Lass and Laing 2008–13).

Neo-classical compounds seem to involve two kinds of patterns, some taken from normal compounds, hence the optional hyphenation in <socio-economic>, and from affixed derivation, as can be seen from the dovetailing in <sociology> (not *<socioology>.

Finally, some spellings, including most non-standard spellings and certain names, obey PhM but not IdP, a situation that can give rise to DSTNCT spellings. For example, the name Phuture is an intelligible respelling of the lexical word future, and it retains a familiar sound while also creating a new orthographic identity. DSTNCT is a principle which only occasionally occurs within lexical spelling (e.g. flower and flour) but it seems to be common in the spelling of new words and names. This suggests that there is a strong desire among users to attach one spelling to one meaningful item, but that desire can only be expressed in specific domains of spelling.

3) How can etymological spelling be included in a theory of English spelling formation?

The model developed in this study can easily explain etymological spelling because the focus is on the orthographic relations between related words. Not only does English tend to preserve the spelling of related words within the writing system, it also tends to preserve the spelling of borrowed words. Both morphemic spelling and etymological spelling are therefore examples of IdP. Spellings such as gateau and macchiato involve simple IdP, as they preserve the connection between English and other writing systems. Many affixed derivatives involve complex IdP, as the borrowed spelling may require amendments. For example,
<philosophy> and <fraternity> have been amended slightly, in order to follow English spelling patterns, and thus they differ from French philosophie and fraternité (OED 142505/ 74285). Notably the CPUs recur in related words such as <philosopher>, <philosophical> and <fraternal>.

4) What is the actual process of spelling formation?
5) How can spelling formation be modelled?
6) How do we know if an English spelling is well-formed?

A model of English spelling was proposed where the first step involves forming a draft spelling by applying IdP. That is to say, adding the spelling of the two input forms, as in ?<jumping>, ?<dopeing>, ?<boping>, ?<eyeing>, ?<oweing>, <singeing>, etc. The draft spelling can then be tested against the known phonological form, and amendments be made, if necessary and possible. Amended forms include <doping>, <bopping> and <owing>, while the others adhere to IdP: <jumping> because it is ideal; <singeing> because it is DSTNCT; and <eyeing>, by default, because *<eying> provides a bad phonographic match.

The problem with this analysis is that it relies on readers’ intuitions of what makes a good English spelling. Therefore, two methods were introduced in Chapter 5 for testing how we know if a spelling is a good phonographic match or not (e.g. ?<dopeing> or ?<doping>). The first method involves visually checking correspondences at all levels of the analysis, including phonemes, syllables, sub-syllabic elements, and the foot. Evertz’s (2014) study provides the theoretical model for this innovation. He introduces a graphematic hierarchy, arguing convincingly that supra-segmental information is frequently encoded in English spelling, and this helps to explain patterns such as CLD, mute final <e>, and the relationship between stress and the number of letters in a syllable rhyme. In Section 5d, the graphematic hierarchy was placed underneath the phonological hierarchy, allowing us to visually compare mappings at all levels. This method is particularly good for deciding whether a spelling is a good or bad fit for the phonological form (e.g. ?<doping> or ?<dopeing>). However, it is not useful for deciding whether one bad spelling is better than another bad spelling. For that purpose, Optimality Theory (OT) is necessary (Section 5d.4).

The first thing to recognise is that PhM is a bundle of constraints, unlike IdP, because it involves mappings at all levels of analysis. Following Evertz (2014), it was shown how g-foot structure can be assigned to spellings, in accordance with a number of ranked constraints. Thus we can see that the candidate spelling ?<dopeed> suggests a spondaic structure < (do)(peed) >, while
<doped> suggests the trochee < (do.ped) >. Comparing these two forms against the known phonological form, /daʊpt/, it was possible to determine that <doped> is the less bad spelling, and is thus chosen. This explanation provides concrete evidence for spellings patterns which may be obvious to the experienced reader but which had always relied on the implicit understanding of the reader (i.e. that <doped> is a less bad spelling than <dopeed>).

There are two very useful and closely related theoretical side effects of understanding the relations between spelling and sound at every level of analysis. We can better explain the decoding process and also the path by which spelling pronunciation occurs. If we assign foot structure to a word’s spelling, it becomes much easier to decode the word’s pronunciation. This is because we can see more easily which syllables are stressed and which are unstressed, and thus the act of mapping from a word’s spelling to a word’s pronunciation is reduced to a set of simpler mappings from stressed and unstressed graphematic syllables to stressed and unstressed phonological syllables. Take the words normal and normality. If we know that {al} words always end in either a trochee or dactyl, then we can metrically decode the word as < (nor.mal) >, with the first syllable being stressed by default, and the second syllable having a reduced vowel. Hence, <nor> ≡ /nɔː(r)/ and <mal> ≡ /mæl/. If we know that {ity} words always end in dactylys, then we can decode the latter word as < nor(ma.li.ty) >, with an unstressed <nor> ≡ /nə/, and a stressed <ma> ≡ /mæ/, followed by two unstressed syllables. Therefore the functionality of the writing system relies on readers having a deep knowledge of the language’s stress and word-formation and patterns. This statement echoes Chomsky and Halle’s (1968: 49) remarks that ‘orthography is a system designed for native speakers’.

Furthermore, ambiguities in the decoding process can lead to variant pronunciations. The word urinal, for example, has two pronunciations: /juərɪnɔːl/ and /juə’rænəl/, and this variation can easily be explained as two different metrical decodings of the spelling: one as the dactyl < (u.r.i.nal) >, the other as < u(ri.nal) >. Wells (20/01/07)89 observes that the writing system underspecifies the vowel quality in <VCV(C)#> words, and this is especially the case for words with three or more syllables. While he claims that this can result in ‘arbitrary, think it up yourself’ spelling pronunciations, it has been shown here that the possible pronunciations follow highly predictable paths. The word omega has several variant pronunciations, but all of them can be predicted by metrical decoding. Candidate pronunciations can even be analysed within an OT framework to see which ones are more likely, although it may be the case that a statistical analysis might shed further light on the possibilities. The metrical
decoding analysis can also explain the Anglicised pronunciations of borrowed words, such as &lt; (mac.chi)'(a.to) &gt;, where the pronunciation is constructed from the spelling. Until now, spelling pronunciation has always been viewed as a marginal phenomenon in English, but the analysis presented here shows that examples of spelling pronunciation follow predictable paths from spelling to sound.

Even more significant again is the fact that spelling pronunciation involves processes that are almost identical to spelling formation in general. In both cases, a phonological form is constructed from the draft spelling. In spelling pronunciation, that decoded phonological form is retained. In spelling formation, it is compared against the known phonological form, and if these differ, then amendments can be made, if possible.

9b. Significance of the study

Studying the English writing system in terms of how spellings are formed may be a new approach, although it is not an entirely new approach to writing systems per se. DeFrancis’s (1989) re-evaluation of the Chinese writing system was based on his examination of how new characters are formed and, in doing so, he recognised the interlocking needs of representing sound and meaning in each character. It is hoped that this study’s methods can be applied more fully to English and thence to other writing systems. It has also been argued here that writing systems research can benefit from a combined synchronic-diachronic approach, just as in word-formation studies (c.f. Marchand 1969). Just like new words, new spellings is get added to the system, one by one, over time. Furthermore, it is important to remember that all spellings contribute to the writing system, and not just the spellings that are recorded in standard dictionaries. Hence, theories of writing systems must take into account the spelling of clippings, abbreviations and blend words, as well as examining how non-standard spellings (which are themselves constrained by the workings of the core writing system). Finally, the use of spelling to represent non-linguistic sounds (through onomatopoeia) calls into question the purpose of the writing systems to represent language.

The model of spelling formation presented here has some very important implications, most notably in explaining the decoding process, and the importance of recognising foot and syllable structure, as well as simple letter-phoneme correspondences, which themselves are dependent on higher-level structures. The importance of morphological information cannot be downplayed, nor its interconnection with phonological and orthographic information. For this reason, the metrical decoding notation may be of limited use in helping to decode other
writing systems. The spelling-formation model also demonstrates the usefulness of the graphematic hierarchy as a theoretical tool, and it contributes to the growing body of writing systems research which uses Optimality Theory.

9c. Limitations

Base forms
The spelling of base forms was not examined in this study and this is because the model of spelling formation cannot explain how letters are matched to sounds in monomorphemic forms. It can only explain how CPUs are altered in polymorphemic and derived forms, or how borrowed words are amended. No effort was made here to explain, for example, why chair is spelt <chair> and not *<chare>. The spelling of base forms have thus been viewed as inherited input forms, from which new spellings can be constructed.

Data
The absence of a defined data set was both a strength and weakness of the study. The original goal was to get an understanding of the spelling patterns in each group, but it proved too difficult to amass enough examples of clippings, abbreviations, blend words and non-standard spellings, solely by observation. However, keeping these in mind throughout the duration of the study allowed for a perspective on spelling which is not afforded by focussing on standard spelling alone.

Obtaining data for inflections and affixed derivatives was much easier, because they tend to be stored in dictionaries, online and/or off-line. It was also easy to investigate these examples by orthographic and phonological patterns — e.g. all relevant words ending in <k> or /k/, etc. All major studies of English spelling have been structured around phonemic inventories and the alphabet, and this provides some useful structure where there is no clearly defined corpus of words. In future studies, I would recommend gathering a data set for each word-formation category but to supplement this with targeted searches. Studies which have limited themselves to a corpus frequently miss important details. For example, the letter <c> can undergo CLD, as in <spec, specced, speccing>, but instances are so rare that they do not occur in the wordlists of either Thorndike (1941) or Baayen et al. (1995). In this way, we can distinguish between rare and forbidden patterns.
**Known phonological forms**

The known phonological form is a concept which is built in to the model of spelling formation but it has not been comprehensively theorised. It is not clear exactly what phonological form English spellings are “supposed to” represent, although it was argued in Chapter 6 that English spelling functions as a defective pan-lectal phonology of the language, from which modern varieties can be derived somewhat. Each branch can be (incompletely) traced back to a centre which is represented in the spelling, The reason why &lt;profane&gt; and &lt;profanity&gt; can maintain the same string of letters is because the CPU &lt;profan&gt; encodes enough information to allow readers reconstruct the right pronunciation, either /eɪ/ or /æ/, as appropriate. Not all changes are recoverable. The TRAP-BATH split has resulted in the letter &lt;a&gt; representing two different phonemes. This is why RP speakers have /æ/ in the first syllable of *macchiato* but GA speakers have /ɑː/. Such ambiguities are due to underspecification in the writing, rather than a weakness of the notation per se.

**9d. Further Research**

The theory developed herein could be used to provide a much fuller account of English spelling than was possible in Carney’s (1994) segmental analysis. Instead of simple mappings from spelling to sound and back, it is now possible to show how and where foot and syllable structure should be applied in decoding spellings, and in their formation. Such a study would need to account for the strong links between spelling, phonology and morphology. It is not clear, at this point, whether such an account would be directly applicable to the teaching of English spelling, although I certainly recommend that educationalists pay greater attention to the importance of both morphological and supra-segmental information.

Future studies may examine how the English writing system has developed over time, by merging Berg and Aronoff’s (2017) diachronic, statistical methods with the methods developed here. The most obvious candidates for examination are words with variant spellings (e.g. &lt;ageing&gt; and &lt;aging&gt;), and the study might use such cases to examine the changing relative importance of IdP and PhM over time, or between British and American spelling.

The importance of the complex pleremic unit suggests the English spelling ought to be re-categorised among the world’s writing systems. While its base forms may be broadly phonemic, as Venezky (1970) states, we can now see that semantic relations are expressed through the recurrence of letter-strings. Thus new spellings are generally formed out of existing units of spelling, and this creates a
web of interconnections that cannot be explained by viewing the writing system as a series of correspondences between spelling and sound.

It is also hoped that the theory of spelling pronunciation developed here can be used to explain the changing relationships of writing and speech over time. If, as Neuman (2009) claims, the two systems ought to be seen as dialects in contact, then we can now test his hypothesis using more appropriate theoretical tools.

To conclude: English spelling, despite its complexities, is nonetheless systematic. Otherwise, it could be neither read nor written. Creativity is only possible because the system contains so many levels of representation: segmental and supra-segmental phonology; morphology, etymology, and so on. It is hoped that by focusing the analysis on a restricted subset of English spellings (of polymorphemic words), while casting an inclusive net to capture both orthodox and unorthodox spellings, this thesis has uncovered the principal forms of information encoded in English spelling; the principles governing their representation, and mutually constraining interactions; and also, along the way, how this complex system acts as a setting for orthographic creativity that, like all kinds of creative action, stretch the bounds of convention without fully breaking them.

***

***

***
References


Allerton, D. J. (1982 [1967]). Orthography and dialect: How can different regional pronunciations be accommodated in a single orthography? In W. Haas (Ed.), *Standard Language, spoken and written*.


Hickey, R. (2016). Irish English and the English writing system. In V. Cook & D.


containing over 59,000 words in international phonetic transcription (14th ed. / extensively revised and edited by A.C. Gimson. ed.). London: Dent [etc.].


Shortis, T. (2016). *Orthographic practices in SMS text messaging as a case*
signifying diachronic change in linguistic and semiotic resources. (PhD), University of London, Institute of Education.


***

***
Notes

1 To get a sense of the changing terminology used in the field, it is worth observing that Venezky’s (1970) book was called The structure of English orthography. Carney’s (1994) was called A survey of English spelling and Rollings’ (2004) work was The spelling patterns of English, while Cook’s (2004) summary was called The English writing system. My own collection, co-edited with Cook, (2016) is entitled The Routledge Handbook of the English writing system. While the term ‘writing system’ may be slightly helpful within the field (see Section 1c.7) it is a forbidding term and there is surely a strong argument for referring to the topic as just ‘English spelling’.

2 Sampson puts Korean Hangeul in its own category as a featural system, where the shape of the segmental symbols mimics the point of articulation in the mouth of the corresponding sound. The fact that Korean also manages to represent both segments and syllables at once cannot be captured by Sampson’s arboreal categorisation.

3 See http://www.fizik.com/eu_en/the-brand/

4 Examples of Chinese writing are taken from Sproat (2016).

5 Rather than measuring logography, Neef and Balestra (2011) propose a method for measuring the ‘graphematic transparency’ of alphabetic writing systems, comparing German and Italian. It is thus a measure of how reliably can one predict the phonemic value of any given letter or meaningful letter-cluster.

6 A chair of comparative philology was funded by the wealthy spelling reformer Mont Follick (a former MP), in the hope that the position could be used to help further the reform cause. Since it was deemed inappropriate for a chair to be heavily associated with a particular agenda, the position was first taken up by William Haas in 1966 with the compromise that he would engage in more general research into writing systems. The position, and the department that it is attached to, has precipitated some very important work on writing systems, notably that of Haas himself (1970), the collection of lectures he edited (Haas 1976), Scragg’s history, and, eventually, Edward Carney’s (1994) monumental work, A survey of English spelling. See the obituary of William Haas at: http://www.independent.co.uk/news/obituaries/obituary-professor-william-haas-1271126.html

7 Venezky first worked as a student in Cornell under Charles F. Hockett and then in Stanford where Ruth Weir was his supervisor. Weir’s contribution in particular may be greater than is generally recognised. For more, see Weir (1964) and Weir and Venezky (1965). Weir died in 1965, aged just 39. See the obituary at: http://historicalsociety.stanford.edu/pdfmem/WeirR.pdf

8 Venezky was originally employed by Charles Hockett as an engineer to program the computers, and this was how he developed his lifelong interest in literacy.

9 The bulk of Venezky’s theoretical contribution to the topic is summarised in his 1967 paper, which is shorter and easier to access than his 1970. His 1999 text is aimed at a wider audience and provides no theoretical progress on his earlier work, and often omits the kind of details that are relevant in an academic study.

10 Venezky’s work fitted into two other projects. He was first employed by Hockett to write a computer program to tabulate the 20,000 ‘most common words’ of English, and that work was published in Harry Levin’s (1963) inter-disciplinary study of the reading process. Levin’s book also included Hockett’s (1963 [1961]) study of the spelling of English monosyllables.

11 The list used by Venezky, which appears in Thorndike (1941), is shorter than the list of 30,000 words published by Thorndike and Lorge (1944).

12 This is the reference that Carney gives for the English pronouncing dictionary he used, different from the 1977 dictionary that Wells talks about.
See OED 140671 for several pronunciations of the zoological term, *pereoin*.

<eigh> ≡ /ɪ/ is not an exception to the well-known spelling rule '<i> before <e> except after <c>' since this rule only applies when the sound is /i:/, as in *deceive, receipt, ceiling* (Carney 1994: 68). *Caffeine* is thus an exception (assuming the <ei> corresponds to /i:/). See also §5.3.

Brooks mentions morphemes only a handful of times (Brooks 2015: 57, 65, 71–2, 73, 84, 92, 96, 183, 246, 269, 277, 287, 289, 296, 310, 312, 315, 321, 362, 366, 473), and he discuss not explain its role in any great detail. At first, that may seem like a lot of mentions, but given that the morpheme is a central unit of organisation in how English spellings are formed, it is a relatively low recognition of its role.

Brooks mentions morphemes only a handful of times (Brooks 2015: 57, 65, 71–2, 73, 84, 92, 96, 183, 246, 269, 277, 287, 289, 296, 310, 312, 315, 321, 362, 366, 473), and he discuss not explain its role in any great detail. At first, that may seem like a lot of mentions, but given that the morpheme is a central unit of organisation in how English spellings are formed, it is a relatively low recognition of its role.

Vivian Cook notes that Albrow was part of Michael Halliday’s team in UCL (University College London) and the search for polysystematicity relates to Halliday’s systemic functional grammar (personal communication, February 2015).

While the word *chowder* seems to derive from modern French (OED 32395), it is a word that was borrowed orally and its spelling was constructed segment by segment within the §Basic subsystem. This process is much less common than words like *croissant*, whose §French spelling was borrowed wholesale but whose pronunciation is determined either by mimicking the French (the OED suggest /ˈkwɑːsɒnt/ with a nasal vowel or by adopting a spelling-derived pronunciation /ˈkwɑːsɒnt/).

A huge set of resources for the study of Medieval English spelling is available from http://www.amc.lel.ed.ac.uk/?page_id=19

The primary purpose of morewords.com is as a search tool for word games such as scrabble, anagrams and crosswords and anagrams. The database is the Enable2k North American word list used in well-known word games. It contains 173,528 words – or letter sequences – including inflected forms, something that is very often useful for present purposes. Here is the explanation contained at: http://everything2.com/title/ENABLE+word+list

The ENABLE word list includes over 172,000 words in the English language and describes itself as the most thoroughly researched, and therefore the most authoritative word list and reference available for Scrabble players. It has become a de facto standard, particularly for online word games, and is released into the public domain by its compiler, Alan Beale. Since it is used by many puzzle authors and solvers alike, ENABLE makes the aggressive claim that it is not merely a superior alternative to the Official Scrabble Players Dictionary [OSPD], but that it threatens to supplant and replace it, to squeeze the very life out of it in a process of Darwinian selection.

ENABLE is an acronym for Enhanced North American Benchmark Lexicon, and attempts to combine several other sources, such as Merriam-Webster's Collegiate Dictionary, Tenth Edition and the OSPD in order to claim to be definitive. Much of the work embodied in ENABLE involves resolution of ambiguity, particularly of nouns that admit no sensible plural, grammatical inflections, or lexicographical contradictions, such as definitions that appear in Merriam-Webster that use words that are themselves absent from the dictionary.

The initial version of ENABLE was released in the fall of 1997, and was updated with a "millennium edition" in April 2000. The new edition included words that were added to the source dictionaries after the initial list was released. The 1997 version of the list is now officially retired, replaced by the ENABLE2K version. Of course this indicates the problem inherent with any effort to produce an official word list — that the language is constantly changing.

The ENABLE2K version of the list contains 173,528 words, listed in a plain text format, one word per line, in alphabetical order. They are not limited by length like some other word lists, particularly those targeted for Scrabble. (ENABLE2K for instance contains
ethylenediaminetetraacetates as its longest word). In addition, the ENABLE distribution includes several other files listing the differences between other available word lists. A supplement goes into more detail about the decisions made on each word. Words that have more recently made headline news when included in major printed dictionaries (such as blog) are of course absent. Nevertheless, the ENABLE effort is significant for all kinds of wordplay.

20 Vivian Cook and Jesper Kruse have been of great help with some of the finer points and specifics of RP.

21 The MRC psycholinguistic database can be found at:

22 Potential problems can arise when databases of spellings are intimately connected with linguistic information such as the morphological and phonological composition of words. For the most part, morphological information is stable across dialects but it can be obscured over time. Carney (1994: 1) gives the examples of chubby and shabby which are functionally monomorphemic in the present day, even though they were once transparently polymorphemic (OED 32655 / 177140), just as funny is clearly related to fun. The same can be said of words such as serene or profane which contain fossilized Latin suffixes inherited from French, but which are functionally monomorphemic (OED 176389 / 152024). However, when databases attach phonological information to the spelling of words, one must be very familiar with the accent encoded in the data because one can run the risk of making assumptions about spellings based on modern, somewhat localised pronunciations. For example, an RP-encoded database (such as CELEX, or Carney’s data) would distinguish between two unstressed vowels /ɪ/ and /ə/ (different in Lenin and Lennon) but not between /æ/ and /ɜ/, so that merges and mergers are given as homophones. Hence <es> and <ers> may be cited as spellings of /əz/. Berg and Aronoff (2017) use a corpus that does not make this same distinction between reduced vowels and this leads them to consider the following words to end in /ɪ/: atomic, sputnik, gimmick, eunuch, buttock, etc.


24 !!’s website is chkchkchk.net because <!> is not a valid character in a URL.

25 <quit> versus <quitted> can be seen here: http://bit.ly/2v3BaGT

26 Carney’s figures do not include inflected forms so the likelihood of that < -ced, -cing> indicating /s/ may differ.

27 The distribution of <axing, axeing> can be seen at: http://bit.ly/2u3bdD0

28 Evidence is again taken from n-Grams. See the discussion at http://languagelog.ldc.upenn.edu/nll/?p=19402

29 The frequencies for <taxiing> and <taxying> can be seen here: http://bit.ly/2v3F16H

30 Another explanation for the rise of <cliché, clichéing> is that the letterform ‘é’ may simply be more available than it used to be in English-based print publications.

31 The only other incidence of <ey> = /aɪ/ is the minority pronunciation of geyser (Carney 1994: 151; Wells 2000).

32 Lidl is a German company but has become prevalent throughout Europe, including Britain and Ireland. https://en.wikipedia.org/wiki/Lidl

33 An example of a non-phonologically motivated linguistic rule would be the relationship between /θ/, /ð/ and grammatical class. Function words tend to have /ð/, as in this, that, these and thy; lexical words tend to have tend to have /θ/, as in thin, thanks, thesis and thigh.

34 Warren Maguire (p.c. 2016) writes that:
The choice of whether to use /hw/ or /ʍ/ is almost an arbitrary decision. It depends on whether you(r analysis) prefers extra ph onomes or extra clusters. You can decide that [ʍ] represents a separate phoneme in some varieties of English, /ʍ/, or you can decide it is the realisation of a phonemic cluster in English, /hw/. There are arguments for and against both: /hw/ kind of patterns with /kw/ and /ɡw/ in English, and if you posit /ʍ/, why not /ç/ in ‘huge’, etc.?

35 Note that I have supplied /eɪ/ here where Giegerich uses a more pan-lectal /e/ for the FACE vowel (c.f. Section 6c.2).

36 The analysis of syllabification is a phonological one, not an account of the phonetic facts. So while atlas might syllabify phonologically as /ˈat.ləs/, in connected speech the phrase at last, will have a [tl] onset in the (stressed) second syllable.

37 The use of angle brackets for <ˈ> is an expedient solution, due to the reading difficulties arising from putting this symbol between inverted commas. The symbol does not indicate that it occurs in spelling (although I do use the symbol to indicate the foot with the primary stress).

38 Several kinds of spelling do not fit the default mappings between vowel letters and phonological vowels, notes Evertz (2014: 75). The following groups have unary phonological vowels, despite appearing otherwise: a) said, dread, cook; b) done, one, come, some; c) camel, honest, lemon. Some spellings appear to have an ‘unnecessary’ final <e>, as in goose, cheese, tense and dense, although Carney (1994: 130) observes that the <e> in the latter two words serves to distinguish these words from the plural forms tens and dens, both with /z/. In other cases, there is a <VCC> spelling but a binary phonological vowel, as in sigh, night, find, cold; ball, palm. Evertz considers scar, fir, for, etc. to be irregular, as they have a <VC> spelling but a binary phonological vowel, and this is a problem encountered by Carney too, but this is because historical /r/ has been lost in the British accents they base their accounts on (the database used by Evertz maps <Vr> spellings onto long vowels). Rollings (2004) has the good sense to categorise the spelling of English vowels according to their orthographic patterns, rather than according to phonological changes that have taken place in a subset of English accents.

39 An example of a heavy g-syllable in an unstressed vowel occurs in words ending in the suffix spelt <less>, as in hopeless and feckless. This seems to be an example of DSTNCT, because the spelling <les> would give us <feckles>, which looks like a plural noun (c.f. freckles, pickles). This problem does not arise for the antonym suffix, spelt <ful>, as in hopeful and rightful, and this makes it possible to drop the unnecessary <l>.

40 Carney (1994: 140) does note four instances of /ɪ ≡ <ei>: forfeit, surfeit, mullein and villein. These four words might be lumped into his loosely defined §Romance category.

41 It may well be the case that the loss of the reduced syllable in {ed} forms is because these forms tend to be followed by another reduced syllable, as in (I)(hoped to) (see)(you) and (I)(hopped in a)(taxi), etc.

42 There are good reasons why moped and refer do not fit the usual pattern. Moped is a blend of motorized pedal cycle, and each stressed syllable indexes the full form of each word. As for refer, it contains a fossilized Latinate prefix, just like reply and apply, and such words tend not to have initial stress (suffer and (pr)ofer seem to be exceptions. Hotel is §French, and thus has final stress. Compare hostel which derives from earlier French and has been integrated into canonical English phonological patterns.

43 In order to have two sets of rules, one for canonical and one for non-canonical spelling, Evertz applies the concept of co-phonology to the English writing system. While I feel that is not needed in this incidence, it would be useful for explaining the spelling subsystems within an OT framework. One set of constraints would apply to §Basic spelling, another to §Latinate, and so on.

44 http://www.phon.ucl.ac.uk/home/wells/blog0701.htm
Unrelated to the presence or absence of /t/, the quality of the vowel in *often* may vary considerably across accents.


http://www.phon.ucl.ac.uk/home/wells/blog0702a.htm

Here are several pronunciations of omega. http://forvo.com/word/omega/#en

http://www.phon.ucl.ac.uk/home/wells/blog0702a.htm

http://www.phon.ucl.ac.uk/home/wells/blog0707b.htm

Many thanks to Jesper Kruse for his valuable comments on an earlier version of this chapter.

The use of the symbol ‘ʌ’ as an English phoneme is on historical rather than phonetic grounds, and its continued use has begun to receive much criticism. Roca (2016) writes:

> We shall adopt the symbol [ä] (centralised [a]) to shun any possible misinterpretations: quite simply, what was once indeed a mid-low unrounded back [ʌ] is now [ä]. The absence of a dedicated IPA symbol for this universally commonest, least marked of all vowels is remarkable, indeed startling. Barry and Trouvain (2008) specifically address the issue. Wells (2014: 2.25, p. 60) acknowledges the [ʌ] anomaly, but still sides with tradition.

Wells notes that Gimson introduced the notation first in his (1962) *Introduction to the pronunciation of English*, but it was the popularised through the *Everyman’s English pronouncing dictionary*, 14th edition, of which Gimson had been the editor since 1964 (Gimson and Jones 1977). For more, see:

http://www.phon.ucl.ac.uk/home/wells/qim-obit.htm

Of course many accents in fact merge these two, especially in America, Scotland and Ulster (see Wells 1982, *passim*).

Wells’s list of words belonging to each lexical set are gathered here.

http://www.yorku.ca/earmstro/courses/phonetics/lexical_sets.pdf

The lists are presented in varying degrees of comprehensiveness inversely proportional to the size and obviousness of membership of the set. PALM, a problematic set, has all of its words represented. KIT has a relatively low fraction.

It has been argued that the *CHOICE* and *PRICE* sets were once merged and the distinction was restored by the spelling, a position that is consistent with the unusually predictable relationship between the spelling(s) <oi, oy> and the phoneme /ɔɪ/. Neuman (2009), following Vachek (1959), favours this argument, but Labov (2001) claims the evidence is inconclusive. What is clear is that it is impossible for a community of speakers to completely reverse a merger, on phonological grounds at least. Individuals may attempt to unlearn a merger but it is as difficult as learning a lexical split, as we saw with Labov’s description of (what Wells calls) the TRAP-BATH split. I myself have no FOOT-STRUT split and it is very difficult to remember which words go into which set, despite it being in my professional interest. It is even harder to reverse a merger than it is to learn a set of words in a new language because, as Labov points out, you have to unlearn your existing knowledge first. And even if individuals do manage to reverse a merger, it is hardly possible for an entire community to do it. Another way of talking about this is to say that merging is a ‘lossy’ process, meaning that information is lost and can’t be recovered. I mention all this to
reinforce how deeply rooted mergers are in one's phonology, so that one can appreciate how arbitrary, and potentially interchangeable, are the different units of spelling which occur for merged vowels.

58 Lists amended from the examples at www.zyvra.org/laffr/hom.htm and corroborated by morewords.com plus the OED (last accessed September 2015).

59 Warren Maguire (p.c. 2016) writes that:

The choice of whether to use /hw/ or /n/ is almost an arbitrary decision. It depends on whether you(r analysis) prefers extra phonemes or extra clusters. You can decide that [n] represents a separate phoneme in some varieties of English, /n/, or you can decide it is the realisation of a phonemic cluster in English, /hw/. There are arguments for and against both: /hw/ kind of patterns with /kw/ and /gw/ in English, and if you posit /n/, why not /ç/ in ‘huge’, etc.?

60 The word iamb is actually a nineteenth-century clipping of iambus (OED 90680) and is iambe in French. These may be reasons for the retention of /b/ (for some speakers) beyond simple spelling pronunciation. The word’s use in English comes long after the /mb/ cluster reduced to /m/. Despite this, LPD3 recommends the variant without /b/, which is not listed in the OED.

61 Comb and womb, with their unusual vowel correspondences, pose trouble to both reader and writer.

62 A very small number of words, such as Gaelic-derived ptarmigan, have an unetymological letter-cluster. In other cases, such as tmesis and phthisis, the cluster may be pronounced. In tmesis, an epenthetic schwa is added: /tə mi:sis/; in phthisis the pronunciation varies between /θ(θ)aitiəs/ and /θaɪəsiəs/. Diphthong and monophthong are often pronounced with /θ/ but some speakers have /θθ/. In general, sequences of fricatives are disqualified in English. There are exceptions of course: §Greek sphere, §Italian sforsanza and, in careful pronunciations, of §Basic sixths as /skɪθθ/. Marginal phonotactics are difficult to account for. Bizarrely, English sneeze was once fnese (OED 183157).

63 The spelling <auld> survives in the Scots song Auld Lang Syne and in the Irish expression, the rare aul’ times, where the vowel has fallen in with the MOUTH lexical set.

64 http://phonetic-blog.blogspot.ie/2012/06/happy-again.html

65 It is commonly thought that the ludic re-spelling of fish as <ghoti> was invented by George Bernard Shaw, but it seems that the spelling was around before Shaw’s time, as Zimmer argues on Language Log (23/04/08).

http://linguagelog.ldc.upenn.edu/nll/?p=81

66 Evertz (2014: 161–7) formalises the three-letter rule and claims that there are further constraints beyond simply having three letters in the spelling. For example, the four-letter word free cannot be *<fre>, nor know *<kno>, because a single vowel letter does not carry enough graphematic weight to support a whole phonological rhyme, hence <CCV> spellings are largely disqualified. There must be at least two letters in the graphematic rhyme, so blue is preferred to *<blu> and shoe is not *<sho> (nor indeed *<shu>). <CCY> words do occur as in fry, pry, sly etc., but <Y> tends to be used in place of <i> at word-boundaries across the writing system (c.f. marry, marries). There do exist a handful of foreign borrowings with <CCV> words, as in ski, gnu, chi, but these can be grouped with foreign <CV> words, including pi, or abbreviations, bo, co (Evertz 2014: 164). She, who and fro are permissible because they are grammatical words. Similarly, <VC> is also a ‘restricted structure’. Evertz counts just sixteen of these, and almost of them are either function words (us, of, in, it) or loanwords (Al, id). Hence, /VC/ words are spelt <VCC> or <VCe>, as with ass, odd, axe, etc. <ox> is a rare lexical example from the §Basic subsystem, and <ax> only occurs in American spelling. The OED (14013) claims <ax> ‘is better [than <axe>] on every ground, of etymology, phonology, and analogy’, but they fail to recognise that <axe> has a better length than <ax>. Indeed, Evertz might say that it has two g-verbs. He considers the /V/ words owe, awe, aye and eye to have two g-verbs and these can be analysed (in my notation) as
This formation makes these three /V:/ words permissible graphematic feet, much like \( <(o\\text{we})>\), \(<(a\\text{we})>\), \(<(e\\text{ye})>\) and \(<(a\\text{ye})>\) (Evertz 2014: 163–4). We might also consider \(<(a\\text{xe})>\) to have the same structure (with the dot above the \(<x>\) on account of the underlining). Evertz (2014: 169) thus argues that stressed syllables (and monosyllabic g-feet) have to have a minimal amount of graphematic weight. Hence words must have either a \(<\text{CVC}>\) structure (\(<\text{beg}, \text{hey}\>) or have three letters in the graphematic rhyme (\(<\text{egg, eye}>\)). Not only are \(<\text{VC}>\), \(<\text{CV}>\) and \(<\text{VV}>\) spellings restricted, so too are \(<\text{CCV}>\) and \(<\text{CCCV}>\) spellings (hence \(<\text{blue}>\) and \(<\text{screw}>\), not \(*<\text{blu}>\) or \(*<\text{scr}>\)). The American spelling \(<\text{thru}>\) thus violates the general pattern.

The would-be diphthong in \(<\text{Bucket}=>/bu'keɪ/\) is a reminder that §French words are not pronounced as though one were speaking French, but instead they involve a mapping from the spelling to a conventionalised English approximation of the French pronunciation. Hence \(<\text{duvet}=>/ˈduːvɪt/\). While the \(<t>\) is not \(/t/\), the \(<et>\) corresponds to English \(/eɪ/\) rather than French \(/e/\).

\(^{68}\) oink and joink are onomatopoeic exceptions to the rule disqualifying non-coronals from appendix positions.

\(^{69}\) The reason why I have specified that \(<\text{Aske}>\) does not have a \(\text{historically}\) long vowel is because the long /æ:/ vowel results from the \(<\text{TRAP}-\text{BATH}>\) split, which occurred after English spelling stabilized. This long vowel is not indicated with either a digraph of a final \(<\text{e}>\), as in \(<\text{face}, \text{wait, metre, fleece, price, note, boat, cute, goose}>\). However, an open g-syllable in a phonologically bi-syllabic word spelt \(<\text{CaCV}>\) does often indicate the long vowel /æ:/ rather than /æ/, as in \(<\text{Gaza}, \text{Nazi, gelato}>\).

\(^{70}\) For example, the English rugby players \(<\text{Andy Goode}>, \text{Alex Goode}>\) are rhymed with \(<\text{food}>\). See, for example, this clip of Andy Goode: \(<\text{https://www.youtube.com/watch?v=WmEWY_UqHU8}>\). By contrast, one useful dialectal pronunciation website suggests the \(<\text{FOOT}>\) vowel for other people called \(<\text{Goode}>\) (see \(<\text{http://forvo.com/search/goode/>}\)).

\(^{71}\) Martin Neef (p.c. 2012) has suggested using the term \(<\text{constructed heterography}>\), since it is the spelling which is changed or constructed, not the sound. From the viewpoint of spelling formation and, with it, spelling pronunciation, it seems better to me to emphasise the fact that a homophone is being constructed: the spelling is chosen to ensure a particular known phonological form. The term ‘constructed heterography’ only captures the change in the spelling, rather than the required lack of change in the pronunciation.

\(^{72}\) See \(<\text{https://www.electriksheep.com/>}\)

\(^{73}\) See \(<\text{http://www.elektrikzoo.com/>}\)

\(^{74}\) These are listed here \(<\text{www.zobbel.de/cluk/CLUK_00.HTM}>\) and were accessed frequently between June and August 2010. At the time of writing, the website continues to be updated, although no new names have been added since 2015.

\(^{75}\) Lykke Li on \(<\text{discogs.com https://www.discogs.com/artist/994633–Lykke-Li}>\)

\(^{76}\) The name \(<\text{t.A.T.u.}>\) has had a very unusual history. It began as \(<\text{Тату}>\), a Russian transliteration of the English word \(<\text{tattoo}>\), before being re-transliterated back as \(<\text{tatu}>\). That spelling was already taken by another band, so the full stops and capitals were added to create a distinct spelling. \(<\text{https://www.discogs.com/artist/99356–tATu}>\)


\(^{78}\) \(<\text{http://abbeyrd.best.vwh.net/named.htm}>\) (Retrieved 4/8/10)

\(^{79}\) See a relevant interview with \(!!!\), see: \(<\text{http://www.npr.org/templates/story/story.php?storyId=10236467}>\)
80 See link in previous note.


82 See the link at http://bit.ly/2h6FpvI

83 The chemical *citronellal* has a different [al] suffix (OED 33547 / 4479), denoting its chemical composition, and its formation is thus not subject to the same morphophonological conditioning as *patella(r)*.


86 It should also be noted here that Silva’s PhD thesis (written in Portuguese) also argues her internet-message data provides indirect access to speakers’ internal phonologies.

87 https://www.economist.com/blogs/johnson/2012/01/slang

88 http://phonetic-blog.blogspot.ie/2012/02/

89 http://www.phon.ucl.ac.uk/home/wells/blog0701.htm

   ***

   ***

   ***