

Grammaphonology: A new theory of English spelling

Des Ryan

This study of standard and non-standard English spelling patterns makes steps towards developing a theory of spelling which can explain both within the one model. The study explores the formal differences between standard English spelling and 'constructed homophony', a sub-branch of non-standard spelling, and some existing theories of standard spelling are updated to cover both. In the latter section, the goal is to provide the outline of a visual model which can be used to predict all possible well-formed English spellings of all possible well-formed English words, even allowing for the interchanging of phonograms, morphograms and syllabograms.

Keywords: *Orthography, English spelling theory, standard, non-standard, constructed homophony, grammaphonology, sound-spelling relationships.*

*"I don't give a damn for a man that
can only spell a word one way."*

Mark Twain

1. Introduction

The aim of this paper is two-fold. The first aim is to develop the theory of orthography so that it will be possible to study the form of both standard and non-standard spelling within the one theoretical model. The second aim is to make steps towards predicting possible 'well-formed' English spellings, a kind of orthographic correlate to phonologically well-formed English words (e.g. Giegerich 1992). This means that instead of having a set of existing English spellings that just includes all standard spellings and whatever unquantifiable number of 'non-standard' spellings might happen to exist, the visual algorithm introduced here will make steps towards predicting all of the well-formed spelling possibilities of any possible string of phonemes. The aim is not to predict the likelihood of a spelling occurring, but simply that it could exist, by extrapolating from the graphotactic constraints of English spelling. These intelligible spellings should then be easily translatable back into speech by competent readers.

The study begins by looking at a particular kind of non-standard spelling that Ryan (2010) terms *constructed homophony*. We will need to look firstly at the phenomenon of homophony in English spelling, and its creative development as constructed homophony, something that frequently occurs in the spelling of names (*ibid*; Jacobson 1966; Praninskas 1968; Carney 1994; Anderson 2003) and in computer-mediated communication (Sebba 2007; Shortis 2007; Crystal 2008). We will then look at some formal differences between constructed homophones and standard spellings, and then an effort will be made to update some existing theories of English spelling (Albrow 1972; Carney 1994) so that both standard spelling and constructed homophony can be studied within the one theoretical framework, mapping phonemes onto spelling units in an orthographically shallow manner.

At this point, an interlude will be required to expand upon the existing terminology used to describe this larger field of study. The term *grammaphoneme* will be introduced to describe the set of all orthographic variants that can be used to spell a particular phoneme,

and with it will come associated terms, including the term that gives its name to this essay, *grammaphonology*.

The final section will introduce the grammaphoneme in some detail. Its function is to explain synchronically the existing standard spellings *and* to predict possible well-formed spellings. It is not the intention here to provide an exhaustive account of all of the spelling units used for every English phoneme – that work remains in progress. The model is tiered in a way that allows for the massive variation in how spellings can be constructed. Towards the base of the model are the default spellings for each phoneme, and above these are the lesser used variants that have arisen through the language's complex orthographic history. Above these layers the model allows for overlapping among grammaphonemes, so that it is possible to account for the use of spelling symbols that relate to phonological tiers higher than the phoneme. For example the spelling <C U L8R> (rather than <see you later>), uses <C> and <U> as morphograms, <L> as a phonogram, <R> as a syllabogram, and <8> as 'rhymogram'. This means that it will be possible to provide a map which accounts for the use of the same written symbols occurring at different phonological levels – as phonemes, rhymes, syllables and morphemes. It will also allow for a translation of spellings across different phonological levels.

2. Constructed homophony

2.1 *Non-standard spellings and constructed homophony*

The term *non-standard* spelling is a term which covers a massive range of spelling practices, and this would appear to be growing in recent decades. Accordingly, it is not possible, nor desirable, to try to pigeon-hole all non-standard spelling into one single, simple definition. Instead, it is much more desirable to view it as a range of orthographic practices, so that we can identify subcategories, examine what their internal properties are, and investigate what constraints there may be upon them.

Shortis (2007) examines the language of text and internet messaging, uses the term *TXT language* to describe it as a whole. He sees its spelling as a continuation of traditional *vernacular* spelling, a view echoed by Crystal (2008) in his study of text-message spelling. This latter term, 'vernacular' spelling, may well be a better term than 'non-standard' spelling, but it remains a blanket term which fails to make any distinctions between the different kinds of non-standard spellings.

One such subcategory is 'constructed homophony' (Ryan 2010). A constructed homophone is one which shares its phonological correspondences with an existing word or phrase, but which has been given a different spelling. Its benefit is that the phonological connection allows the newly spelt form to retain semantic links with an 'underlying' standard form, but its 'surface form' gives it a distinct visual identity. Such spellings occur frequently among names (*ibid*; Praninskas 1968; Jacobson 1966), and these include *Def Leppard*, *Mortal Kombat* and *Xtra Vision*. Constructed homophones are a creative development of the accidental phenomenon of homophony, whereby words are identical phonologically but not orthographically, as with *meet*, *meat* and *mete*.

2.2 Homophony and homography

Homophony, along with its complement, homography, is caused by the polyvalence of English spelling. English spelling is not a phonemic spelling system, mapping spelling units onto phonemes and back in a one-to-one fashion. It is not even a predictable many-to-one system like German or French (Rollings 2003). Instead, it is many-to-many system, whereby individual spelling units can correspond to more than one phoneme, and each phoneme can be represented by more than one spelling unit. For example, the spelling unit <th> can represent three different phonemes, as heard in *thigh*, *thy*, and *thyme*. Conversely, /t/ can be written in numerous ways, as demonstrated by *terror*, *totter*, *pterosaur* and *Theresa*. This variation is not as haphazard as one might be led to believe, and there are frequently strong connections between the spelling variants. <t> and <tt>, as Carney observes (1994: 114 *et seq*), are largely in complementary distribution, and this disqualifies spellings such as <ttoter> from standard English.

Such complexity has arisen from two main sources. Firstly, sound changes have occurred since the spelling slowly became standardized several hundred years ago. Phonemic mergers have resulted in homophones such as *meat* and *meet*, or *write* and *right*. Phonemic splits have caused some homography too, giving us *live* and *abuse*, and so has stress shifting (*produce*, *convert*). The other source of complexity has been the continual influx of words from many different languages, combined with a general tendency to maintain the spelling of borrowed morphemes wherever they are deemed compatible. This tendency has led Albrow (1972) and Carney (1994) to analyse English spelling as a set of subsystems. (We shall see this in detail in section 4.) As Carney observes ‘different spelling rules apply to different sectors of vocabulary’ (*ibid*: 97). This results in *buffet* having two wildly different pronunciations, and oppositions such as *mangetout* versus *mange* and *tout*. Examples of subsystemic homophony include *fill* and *Phyl*, or *shoot* and *chute*.

These complexities are redeemed by some major advantages, however. The main advantage of polyvalent spelling units in English is that they can systematically represent morphological alternations, something which Chomsky and Halle (1968: 184n) consider to be an optimal way of representing pairs such as *sane* ~ *sanity* and *serene* ~ *serenity*. For anything that is not systematic, polyvalent spelling units can cause ambiguity for the reader, as with *live* or *read*. This ambiguity can also be played with. Sebba (2007:30) gives a good example: the spellings *Kosov@* and *Latin@* both avoid the contentious choice of either <a> or <o> by representing both simultaneously.

In the other direction, homophones provide a useful way of making a visible distinction between different words that happen to sound alike. There are a substantial number of examples, but nothing like the number that *could* exist. It is only when we look at creative spellings that we see how many ways words can be re-spelt, especially when the language’s graphotactic restrictions are relaxed, and the main constraint is just comprehensibility.

For example, the word *tough* is examined in depth in Section 6 and the visual algorithm introduced there hypothesizes a large number of possible re-spellings. These include <tuff>, <ptough>, <ptuff> and <ttuff>. The first of these four seems the most likely, and indeed it occurs in *Tuff Gong Records*, in the band name *Tuff*, and in the geological term *tuff*. But the others all *could* occur, and indeed <ttuff> features in *Ttuff Technologies*. At the time of writing, there does not exist the spellings <ptuff> or <ptough>, although they appear to be possible, if not probable.¹

These examples demonstrate the way in which the graphotactic constraints of English can be creatively manipulated in order to construct homophones with distinct orthographic identities. In the case of *Ttuff*, initial geminate <tt> happens not to occur in English, but extrapolations from words such as *llama* and *Ffoulkes* would lead us to believe that initial <tt> corresponds incontrovertibly to /t/. Similarly, initial <pt> also corresponds predictably and systematically to /t/, as exemplified above by *pterosaur*. By contrast, adding an arbitrary consonant at the beginning, as in <gtough>, <ftuff>, <rtuf>, would not appear to be a possibility because these spellings have no connection to the standard system of correspondences.

2.3 Constructed homophones in names

As noted above, constructed homophones frequently occur among names. Ryan (2010) provides many examples of band names formed in this way, including *OutKast*, *The Beatles* and *INXS*. These are homophonous respectively with the standard spelling of the words/phrases <outcast>, <the beetles> and <in excess>. Constructed homophony in trade names is examined by Jacobson (1966) and Praninskas (1968), (although without the use of a similar term), and Carney discusses a similar kind of homophony among surnames. Praninskas talks about the ‘graphemic manipulations’ that occur in her corpus, and these go in two directions. The general trend is for the regularization of existing spellings, resulting in names such as *Kleen Brite*, but frequently regular spellings are de-regularized, resulting in names such as *Lads of Knot* and *Septic Aide*. Ryan (2010) also spots similar patterns among band names: so while *Filosofy* regularizes the word <philosophy>, the bands *Phuture* and *Pharcyde* reverse this.

2.4 Capitalisation

Before we analyse spellings as complex as this, we shall see that standard English already contains some very simple forms of constructed homophony. The most basic form is the use of capital letters to distinguish proper nouns from *common words* (Anderson 2007). This is a normal and productive part of standard spelling, but the effect is visual, and has no phonological representation. It is thus constructed homophony. Anderson (2007: 191-3) discusses the examples of *The Channel* and *The New Town* (in Edinburgh), but the list of such examples is effectively infinite. This is also how we distinguish the *Catholic* religion from a *catholic* reader.

2.5 Padding

A little more elaborate is the phenomenon of *padding* in surnames, discussed by Carney (1994: 454-8). Here, letters are either added or changed to make names visibly distinct from common words in the language. The three most common processes are gemination, as in *Lett*, *Ladd* and *Foxx*; the addition of an empty <e>, seen in *Cooke*, *Browne* and *Rowe*; and the changing of <i> to <y>, as with *Whyte*. All three processes occur in *Thynne* and *Wynne*. Of course many of these spellings have not been constructed intentionally, but are simply relics of a time when spelling was less regulated.² For our purposes, the historical issue is not a concern. What is noteworthy is that a) there exist such homophones, and b) there appears to be some very strong constraints upon what is permissible.

2.6 Licensed and unlicensed spelling

A great range of orthographic variation is described by Sebba (2007: *passim*), in his in-depth discussion of the relationships between spelling and society. Sebba considers non-standard spellings to be *unlicensed* variants (2007: 30). This is contrasted against *licensed* variation, which in English includes examples such *judgement* and *judgment*. Unlicensed variation, he argues, ‘allows for the original meaning to be conveyed, along with additional social meaning which derives from defying the convention’ (*ibid*). We need not concern ourselves here with this ‘social meaning’, but it is frequently a shared phonological correspondence that allows for the ‘original meaning to be conveyed’. Such cases are constructed homophones. Spellings such as <Kris>, <woz> and <luv> can be used instead of <Chris>, <was> and <love>. The latter two lexical items are restricted to what Sebba calls *unregulated* orthographic spaces (*ibid*: 41-8), although *Kris* is not because it is a name, and thus licensed. This last point is very important for the development of constructed homophones. The orthographic practices of unregulated spaces become licensed in names, so they can be constructed under significantly fewer constraints.

2.7 Studying constructed homophony

While the paucity of research into the issue means that the constraints have not been adequately delimited, Ryan (2010) opens the debate on how to examine the processes of formation at work in constructed homophones. It is to such matters we shall now turn. It will be necessary first to look at what makes constructed homophones formally distinct from standard spelling, so that a justification can be provided for how their form and formation ought to be studied. With that in mind, it will be necessary to examine and build upon the existing theories of English spelling so that it becomes possible to study both standard spellings and constructed homophones within the same theoretical framework.

2.8 Puns and blends: not necessarily constructed homophony

Before we do this, we should be clear on what constructed homophony does not include (and this will also help to justify the term’s existence). It might at first appear that all non-standard spellings are constructed homophones and have only one underlying phonological form, but many puns are different. Ryan (2010) observes that the band name *E-Motion* can be mapped back to either <emotion> or <E. Motion>, a pun on the effects of the drug known as ‘E’. This kind of pun has two phonological mappings, and is therefore not constructed homophony, but instead constructed homography. The same goes for the abovementioned *Kosov@* and *Latin@*. As for blend words such as *ginormous* or *bootylicious*, these are not homophones of anything, and merit their own examination. Despite all that, there do exist rare blends such as *The Beatles* which are also constructed homophones.

While it is hoped that the theory developed here can be built upon in time to include such formations, one must learn to walk before one can run, and simple constructed homophones are the domain of this study.

3. Some formal differences between constructed homophones and standard spellings

In this section we shall see that there are some very important formal differences between constructed homophones and standard spellings, differences that go beyond any social or cultural distinctions or preferences. Carney (1994: 18) emphasizes the fact that the standard English writing system is not simply a way of matching letters to phonemes. Instead it is a deep orthography which tends to retain morphemic connections through spelling, both through morphophonemic alternations (see Chomsky and Halle 1968; Venezky 1970) and also by maintaining the etymology of morphemes when new or borrowed words are brought into the language (Carney 1994). By contrast, constructed homophony ignores the morphological and etymological that is kept in a standard spelling (Ryan 2010). Indeed it has to ignore those factors, because respelling a word requires the speller to destroy that information.

3.1 *Forming standard spellings*

Standard English spelling is not a failed phonemic system but instead a relatively consistent morphophonemic system. Alternations such as *divine* ~ *divinity* and *sole* ~ *solitude* maintain their orthographic links, despite their phonological alternations. This makes it convenient for spotting morphemic links between words such as *electric*, *electricity* and *electrician*. Furthermore new words are frequently formed by combining existing morphemes and retaining their spelling as far as possible.

Using Plag's categories of word-formation (Plag, 2003: 12-3), we can apply this principle to compound words as well as affixed and derived words. This has the added bonus of preserving etymological information, because the origins of the morphemes remain transparent. For example, *television* is clearly formed from the Greek {tele} and the Latin {vision} (although words formed using morphemes that derive from different languages are not the norm). Sometimes the spellings have to be clipped to make for a more euphonious end product, although the ordering of the remaining letters is preserved. So the Greek morpheme {philos}, which appears in *philosophy* and its derived forms, also appears in clipped form in *Anglophile* and *Hibernophile*. Of course the surface phonetic form of the morpheme varies, but the morphological and etymological links are preserved in the spelling.

Exceptions to this pattern of word formation include blends, truncations, clippings, initialisms and acronyms, but these are all outside the scope of this study. Truncations and clippings are part of a category that may form a transition group between standard spellings and constructed homophones. This group would also include anglicized spellings of foreign words. This latter group are constructed spellings *par excellence*. The imported word must first be adapted into English phonotactics and then have its spelling regularized into its basic patterns. Irish words such as *galore* or *The Pogues* have undergone this formation. A truncation such as *telly* undergoes a similar fate: the phonological string /teli/ has been derived from *television* and its spelling adapted to concur with more 'basic' English spelling conventions.

3.2 *Forming constructed homophones*

Constructed homophones undergo this latter patterning. Their phonology is pre-determined, and this must be mapped back into a spelling. Doing this frequently compromises morphological and etymological information. So while the phonetic forms preserve links

between *philosophy* and the band name *Filosofy*, the connections to *Anglophile* and *sophist* are lost or, at best, heavily obscured. In their linguistic form, constructed homophones will thus retain fewer historical connections than standard spellings.

As a result of their formation, constructed homophones are orthographically shallower than standard spellings, and this provides some formal linguistic motivation for separating the two. If we are going to examine the two together, as in this study, it seems reasonable to study them for the direct correspondences between phonology and spelling. This is then a good basis from which we can attempt to study the form and formation of constructed homophones, and to see if they are indeed always an extrapolation from the existing graphotactics of standard spelling. It also means that we can avoid invoking social or cultural criteria to distinguish these two.

This is not to deny the benefits of studying constructed homophony or any other kind of non-standard spelling for the morphological information it can and does retain. For example, the trade name *Specsavers* retains a connection to *spectacles* that would be obscured by the spelling <Speksavers>, or worse again, <Specksavers>. Such connections presumably provide a constraint on the formation of blend words, although that remains beyond this study.

3.3 Proper nouns and morphology

At a deeper level, it is also important to observe that many non-standard spellings (of any kind) may not be subject to the same word-formation processes as common words, either because they are proper nouns, or frequently just nonce spellings used in ephemeral computer-mediated communication. In his study of the grammar of names, Anderson (2003) observes that while names can become common words and develop new morphology, as with *Orwellian* and *sadism*, that is not their initial function. Ryan builds on this to note that the creators of names need not concern themselves with the morphological advantages of choosing <c> instead of <k> in *electric*, because the derived forms are unlikely to ever exist: there will be no *elektrikity* or *elektrikian*.

3.4 Patterns of spelling in constructed homophones

When spellings cease to undergo morphological alternations, the results are orthographically complex. The most obvious result is that marked, unusual or phonetically non-transparent spellings become regularized. Default and unambiguous spellings are used, and this initially makes for a more phonemic spelling system. Ryan (2010) analyses some spellings of *Electric*, and observes that the most frequent substitution is the use of non-ambiguous <k> for <c>: *The Elektrik Zoo*, *The Elektrik Euphoria*, *Elektra Records*, *Dr Lektroluv*. Of course there is a major difference here between the use of <k> for the second rather than the first <c>. The first <c> has no morphophonemic duties, and always corresponds to /k/, so it is a direct substitution. The second <c>, however, re-occurs in *electricity* and *electrician*, and is a valid spelling of /s/ and /ʃ/ respectively. In proper nouns, these morphological connections can be abandoned, and <k> can assume the role of <c>, but only for /e`lektrik/.

Ryan also provides the examples of *Electrique Boutique* and *Electrix*, which demonstrate further complications. The latter case exemplifies the fact that single graphemes are frequently used in non-standard spellings to correspond to more than one phoneme. In this case <x> replaces /ks/, but frequently we see graphemes being used to correspond to

whole syllables or morphemes. The names *U2* and *INXS* demonstrate this. <U> and <2> correspond to <you> and either <too> or <two> (but presumably not <to>). *INXS* combines the forms, leaving <in> as <IN> but using <XS> for <excess>. In the case of *Electrique Boutique*, there are two issues at stake.¹ Here we see a more marked spelling variant <que> replacing <c>. Ryan (2010) postulates that ‘marked’ variants are used when the spelling is already regular, and this tallies with both Praninskas and Carney who both observe de-regularization of regular spellings among their respective corpora.

3.5 *The linguistic features of constructed homophony*

We can thus summarise some features of the spelling of constructed homophones, albeit tentatively.

- (1) Morphological and etymological information is deprioritized, so that:
- (2) Spellings can be regularized, but:
- (3) Regular spellings can adopt marked features, and
- (4) Graphemes are frequently redeployed to correspond with more than one phoneme, moving constructed homophones even further away from being phonemic spellings.

To summarise, three things are clear. First, studying constructed homophones requires us to focus on the direct correspondences between spelling units and phonemes. Second, there is no benefit in developing a theory of spelling which neglects marked or less regular spelling variants, especially since many of them feature in constructed homophony. Third, we need to provide a model that can account for the use of graphemes at higher phonological levels, to see how syllabograms and morphograms can be intermingled with the default phonograms of standard spelling.

Now that we can see some formal differences between standard spellings and constructed homophones, it is time to look at the existing theories of English spelling so that we can build upon them to include standard spelling and constructed homophony into the one theoretical framework.

4. Developing the existing theories of English spelling

4.1 *Subsystems of English*

English has had a long history of maintaining morphological spellings among a great number of its lexical borrowings, and this has allowed for layers of regularity within the system. These layers have been considered as sets of overlapping and incomplete subsystems which

¹ There is presumably a stress-shift in the spelling *Electrique Boutique*, but the model developed in this essay cannot deal with such complications.

correlate strongly with the languages that have provided the bulk of the language's vocabulary. This includes its Germanic roots, Norman French, Latin, Greek, Italian and Modern French. The first linguist to analyse English spelling as a set of subsystems was Albrow (1972), who was trying to adapt Firth's polysystemic theory of phonology (*ibid*: 13). Albrow broke down English spelling into individual correspondences between spellings and phonemes, and split them into three subsystems of correspondences. Carney (1994) approaches them lexically, arguing that since spellings are maintained throughout borrowed morphemes, entire morphemes are spelt according to one subsystem or another. We shall look at the work of both scholars here, and then synthesize their work so that it can be updated to explain both standard spelling and constructed homophony within the one theoretical framework.

4.2 Carney's analysis

Carney's (1994) monumental *A survey of English spelling* provides much of the theoretical and statistical basis for the theory developed in this work. Carney analyses in great detail the relationships from both spellings to phonemes and vice versa, adding a lot of useful theory as he justifies his rationale. Carney firstly examines each RP phoneme individually and provides exhaustive statistical accounts of their realisations as spelling units. His initial distinction is between 'conditioned' and 'competing' spelling variants (Carney, 1994: 16). Geminate consonants are chosen as examples of conditioned variants, so that <p> and <pp> are considered to be in complementary distribution (except in exceptional cases, and Carney duly notes these for each consonant phoneme). Competing variants come in different guises, including 'default' and 'minority' variants. For example, <f> is the default spelling for /f/, while the <gh> of *laugh* and *tough* are minority variants. However, no rigorous definitions for these terms are provided, and the choices are frequently based solely upon statistical frequency, and do not take into account the consistency of the spelling-sound correspondences. Just because a correspondence is rare, that does not make it irregular. All words beginning with <pt> correspond to /t/, but not all words beginning with <ch> have the same correspondence.

4.3 Carney's subsystems

Carney (1994: 96-103) subdivides English spelling along the lines of entire morphemes and their etymology, rather than spelling units and their phonological correspondences. His reasoning is consistent with the observation made above that standard spellings retain the spellings of their constituent morphemes. Words such as *philosophy* and *pneumonia* have §GREEK correspondences throughout their morphemes (his notation), even though those morphemes have been clipped. Putting it another way, Carney says that 'different spelling rules apply to different sectors of vocabulary' (*ibid*: 97). This disqualifies inconsistent spellings such as <filosophy> or <pnewmonia>, since they blend §GREEK and §BASIC correspondences. He also observes that words from different subsystems often have different English word-formation patterns. For example, the native word *shoot* alternates with *shot*, whereas the French borrowing *chute* becomes *chuted*.

Carney's subsystemic categories are §BASIC, §GREEK, §LATINATE, (modern) §FRENCH, §ITALIAN and §EXOTIC, and his use of <§> to mean *subsystem* will be retained here. A final category, §ROMANCE, refers to late-mediaeval loanwords which do not necessarily correlate with either §FRENCH or §LATINATE spelling conventions.

Carney makes no effort to be exhaustive in his analysis of subsystems; instead he introduces his categories tentatively to aid his explanations of English spelling, and hopes that the breakdown might be developed further in the future, which is what we shall do after we look at Albrow's work.

4.4 Albrow's subsystems

Albrow has three subsystems, numbered 1, 2 and 3; the divisions are crude, although, to be fair, just a preliminary outline. Albrow aims to be the first word on the matter, rather than the last, and Carney (1994: 96-104) is right to criticize many of the subdivisions. Albrow's system 1 refers to units that are mostly the default correspondences for each phoneme. System 2 includes the commonly occurring correspondences that come via Latin, French, and Greek, as well as the more marked native features (such as <igh> for /aɪ/). System 3 mops up all the remaining correspondences, §EXOTIC or otherwise. This system includes, for example, the §FRENCH <ch> of *chauffeur* and the §GREEK <pt> of *pterodactyl*.

Albrow's categories are muddled and his choices are not always explained explicitly. It seems that frequency of occurrence might be his criterion for allocating spelling units to his subsystems, rather than breaking it down into groups that share degrees of predictability. For example, in system 3, he makes no allowance for the fact that the strange and rare occurrence of word-initial <pt> is at least consistent and predictable. This is unlike, for example, the <ch> - /ʃ/ of *chauffeur* or *machine*. In such words we can only predict the spelling if we know it is §FRENCH, otherwise confusion will occur between §FRENCH *chauffeur*, §BASIC *chowder* and §GREEK *chaos*.

4.5 Synthesizing the analyses: the issues.

The formal differences between standard spelling and constructed homophony actually mirror the divide between Carney's and Albrow's analyses. Standard spelling is formed by blending entire morphemes, and retaining their individual spellings. Constructed homophones are formed by interchanging individual spelling units, with little or no regard for morphological or etymological information. For example, <f> and <ph> can be interchanged (Ryan 2010), allowing for the formation of the band names *Phuture* and *Filosofy*. As it turns out Carney actually disqualifies this latter example as a possible standard English spelling (Carney, 1994: 96) although he recognizes that such combinations are possible in 'commercial jargon' such as *washeteria*. In this subsection I want to look at some of the concerns that will arise in a synthesis of Carney and Albrow's observations.

The analysis provided later on will be based on direct correspondences between spelling and phonology, although unlike Albrow there will not just be a simple list of correspondences. We could look at /f/ and observe that there are the following correspondences:

- (5) <f> - graft
- (6) <ff> - gruff
- (7) <ph> - graph
- (8) <gh> - rough

(9) <ffe> - giraffe

Of course not all of these correspondences carry equal weight. Some are unambiguous, and predictable from the spelling (<f>, <ph>),³ some are conditioned by context (<ff>),⁴ and some are restricted to certain positions; <gh> and <ffe> only occur syllable- or word- finally (e.g. *rough*, *laughter*, *Jefferies*, *Cuffe*). Furthermore, how should we deal with the /f/ in <4> and <5>? We could ignore such complications, and concern ourselves solely with phonograms, but morphograms such as these are far too important in non-standard spellings to be discounted.

The next issue is that many spelling-sound correspondences are conditioned by their neighbouring graphemes. A good example is <c>. When it is followed by <e>, <i>, <y>, <æ> or <œ>, <c> corresponds to /s/.⁵ Otherwise it is usually /k/. So while *Electric* can be respelt <elektrik>, *kind* can't become *<cind>. We must have some way of accounting for the strings of spelling units (such as *war*) which cannot be broken down into their individual components (<w>, <a> and <r>). Finally there are issues concerning phonetically conditioned correspondences such as <ti> ~ /ʃ/.

A proper theory of spelling must take all these complications into account, and the goal of the final section of this essay is to provide a simple and elegant visual model which not only describes the existing patterns of standard spelling but also predicts possible spellings. This model will be layered in various ways:

- (10) In accordance with the subsystems.
- (11) In accordance with the phonological level of correspondence: phonemic, syllabic or morphemic etc.
- (12) In accordance with any orthographic conditioning that the spelling unit may undergo.

4.6 Redefining the terms

Much of the terminology introduced by Carney (1994: 96-104) will be retained here, but with some important amendments. The labels §FRENCH, §GREEK, §ITALIAN, §LATINATE, §ROMANCE and §EXOTIC will be retained where appropriate, and the categories §DEFAULT, §NATIVE and §NONCE will be introduced. The term §BASIC will be avoided, along with Carney's use of the term 'minority' variant. The section that follows will begin by explaining the problems with this term, and that will help us to understand the need for the three new terms.

4.6.1 §NONCE spelling units

Carney never actually describes what the term 'minority' variant means. The most likely criterion for entry to this category seems to be frequency of occurrence, rather than any formal characteristics of the spelling units. So, for example, initial <pt> ~ /t/ is considered to be 'minority' and so is <th> ~ /t/, as in *thyme*. While neither correspondence is phonetically transparent, the difference is that initial <pt> ~ /t/ is predictable, whereas <th> ~ /t/ is not.

Notice the difference in the respellings of *tab* as <ptab> or <thab>. *Ptab* is a well-formed and intelligible homophone of *tab*. *Thab* clearly isn't.

The category §NONCE will thus be introduced to include those correspondences which are neither phonetically transparent nor predictable. Other examples include <gh> ~ /f/, and those awkward consonant clusters with empty letters (*debt*, *salmon*) or inert ones (*sign*, *handkerchief*).⁶ This category will also include vast numbers of vowel correspondences. For /i/ alone, Carney (1994: 137) lists five different 'minority' correspondences, seen in *English*, *breeches*, *sieve*, *women* and *busy*.

4.6.2 *Breaking up §BASIC*

Carney's §BASIC category seems to include everything that could be described as a default correspondence or a native one. He notes that words such as *map* and *dot* have no observable markers and are thus §BASIC by default (*ibid*: 101). Quite clearly <r> is the default spelling for /r/, yet Carney leaves <wr> in the §BASIC category. This is a valid position for standard spellings because <wr> does not occur in foreign-based subsystems. But for our purposes, it is clear that <r> and <wr> will have different roles in constructed homophony, especially since <wr> only occurs word-initially (or syllable initially if we include *Lawrence*, and contrast it with *Larry*). §BASIC will thus be split into §NATIVE and §DEFAULT.

4.6.3 §NATIVE

§NATIVE spelling units are mostly residualisms from earlier English which have survived for various reasons but are not §DEFAULT. There are not many consonantal correspondences in this subsystem: Initial <kn> and <wr> are included here, despite their restricted distributions. Many vowels fall into this category, including the <ea> of *meat* and the <oa> of *moat*. The <igh> of *sigh* and *night* is also in this category. It too is a predictable correspondence, despite its history of vilification, and it must be distinguished from the §NATIVE <eigh> of *eight*, *neighbour*, *inveigh* and so on.

4.6.4 §DEFAULT

Discussing simple spellings such as *map* and *dot*, Carney (1994: 101) says that they are '§BASIC by default'. What this analysis means implicitly is that there is a large number of letters in the alphabet which have the same correspondence, no matter which (roman) alphabetic language they have come from. <m> ~ /m/, whether it be from Latin, French or Old English. Greek transliterations often fall into this pattern too. This set of letters provides much of what Scragg (1974) calls the 'ancient phonemic core' of English spelling, a core which has been there ever since Irish monks applied their spelling system to English. These correspondences, including among them ~ /b/, <d> ~ /d/, <f> ~ /f/ etc., provide the bulk of the §DEFAULT set of correspondences.

Despite this core group, it is not a simple task to enumerate all of the default correspondences, and it would be useless to pretend that there is exactly one default spelling for each phoneme and vice versa. Praninskas (1968) provides a rough outline, and Carney attempts to provide a default spelling variant for each phoneme, but the problem he faces is that he fails to define exactly what he is looking for, especially since he limits himself to standard spelling. Ryan (2010) also attempts the task, but argues that without adequate evidence from non-standard spelling it is impossible to provide a definitive list of productive, predictable, bi-directional default correspondences for every phoneme.

Constructed homophony displays a general tendency towards spelling regularization (Ryan, 2010), so default variants are frequently used here. We have seen already the spelling of *was* as <woz> and *love* as <luv>. What such variants have in common is that they provide an unambiguous path from spelling to sound. Default correspondences are not necessarily the most common variant in standard spelling. For example <s> corresponds to /z/ more frequently than <z> (Carney, 1994: 238), but in constructed homophones, <z> would appear to be the unambiguous default.

/i:/ provides a more complex picture. Is the ‘split digraph’ of *mete* the default correspondence, on analogy with *mate*, *mite*, *mote*, and *mute*? Or is it the ‘simple digraph’ of *meet*? Clearly it is not the §NATIVE <ea> of *meat*, nor the §FRENCH <i...e> of *machine*. Carney labours over this question (1994: 155-64), and still does not yield a satisfactory answer because he confines himself to standard spelling. Had he examined constructed homophony he would have had some evidence to indicate which unit people *choose* to use when they can make such choices.

Ryan’s (2010) found that <ee> was by far the most productive way of re-spelling /i:/. Examples include *Dreem Teem*, *Lovefreakz*, *Mis-Teeq*, *The Bee Gees*, *Skee-lo* and *Skeewiff*. Only one occurrence of the split digraph was found among the data,⁷ in the band *Breze*, and the standard spelling of this word already has <ee>. The only use of <ea> was in *The Beatles*, and notice that this spelling adds a pun on the word *beat*. The problem with <e...e> is that, like <ea>, it is ambiguous. Compare *theme* with *creme* and *meat* with *measure*

Ryan also notes that among vowels there is often a different default correspondence word-medially than word-finally. A simple example is that /oi/ is <oi> word-medially by default but <oy> word finally. Less obviously, <a..e> is the default for /ei/ word medially, but <ay> is the default word finally, itself a variant of the word medial <ai>. Compare *way*, *waist* and *waste*. The same can be said for all unchecked vowels.

§DEFAULT spellings must therefore be predictable correspondences in both directions, and they should also be productive among constructed homophones. Of course the lack of research in this area means that there is a lack of data, but that should not stop us making efforts towards defining an important orthographic concept.

4.6.5 §EXOTIC

The only remaining category is §EXOTIC. This category is a catch-all for any remaining predictable correspondences which do not derive from the usual sources. It includes the <kk> of *pukka* or the <dh> of *dhal*.

4.7 Summary

The subsystems used in this study will be §DEFAULT, §GEMINATE, §NATIVE, §LATINATE, §GREEK, §FRENCH, §ITALIAN, §ROMANCE, §EXOTIC and §NONCE. Default spellings must be unambiguous, going from spelling to sound. Geminate are orthographically conditioned in standard spelling, but can be creatively used as padding where it does not mark a preceding vowel as short. §NONCE spellings are neither phonetically transparent nor predictable, and all the other subsystems have to be predictable from the spelling, even when they are not phonetically transparent (e.g. <ph>, <pt>).

5. Terminology

5.1 *The need for terminology*

The study of the relationship between written symbols and their phonological correspondences has been beset by a number of concerns. Firstly, is writing secondary to speech, and should it only be studied for its ability to record spoken language? Is writing a system unto itself which merits linguistic analysis independently of speech?

There is not space here to discuss this debate in any detail, although it should be very clear from the discussion of constructed homophony above that written symbols (and sets of symbols) can be creatively redeployed without affecting the resulting phonological correspondences. In that sense writing is a system unto itself, independent of phonology, and we need some terminology to describe it. We need distinct terms to describe the abstract symbols, and also the sets of such symbols with phonological correspondences.

5.2 *The grapheme*

In her essay *On the use of the term grapheme*, Henderson (1985: 146) observes that there is a need to refer to:

- (13) A letter or cluster of letters that can usefully be regarded as corresponding with a phoneme.
- (14) The minimal distinctive unit of a writing system.
- (15) An abstract letter identity in all its physical realisations.

For (14) the contentious term *grapheme* shall be used for the remainder of the essay. Henderson discusses the existence of two very different and incompatible definitions of this term, and Kohrt (1986) also identifies the same split in the literature, calling them the ‘analogical view’ of the grapheme and the ‘referential view’. In brief, the analogical view considers the grapheme to be the minimal distinctive unit of a writing system, whereas the referential view considers meaningful spelling units to be the units of phonological reference. In the analogical view each symbol <t>, <c> and <h>, is a distinct grapheme, whereas in the referential view, <t>, <th> and <tch> can all be graphemes of English. In the analogical view, graphemes are defined by their visual form, not their phonological function, so that <k> and <q> are considered distinct graphemes, despite the fact that both almost always refer to /k/.

An inventory of graphemes for English would thus include the twenty-six letters of the alphabet, the ten Arabic numerals, and assorted morphograms such as <&> and <@>, as well as various punctuation marks and the space.⁸ These latter symbols, while marginal in standard spelling, feature heavily in non-standard spelling, and must be included in any theory of spelling that includes both.

5.3 Spelling units

We can now revert to Henderson's first terminological challenge. If the term 'grapheme' is to be used to describe the form of the minimal distinctive units in the writing system, how should we describe the function of those units, or clusters of units?

As it happens, the issues are greater than a simple bi-directional mapping of spelling units onto phonemes and back. When we take constructed homophony into account, we will see that the inter-relationships of those spelling units become key to our understanding of how spelling-phoneme correspondences operate.

5.4 Four sets of units

In a fully phonemic system of spelling there are only two sets of units. There is a set of phonemes and an isomorphic set of spelling units. Ideally each spelling unit would be represented by a single symbol, although in a quasi-phonemic system there can be digraphs and trigraphs. In a many-to-many system such as English, there are in fact *four* sets of units, something which is not, to the best of my knowledge, made explicit in any of the relevant literature. Once again, there are the phonemes and the spelling units; but there are also the *sets* of spelling units which correspond to each phoneme, and the sets of phonemes which correspond to each spelling unit.

It is these sets which make homophony and homography possible. Recall that morphological alternations in English are expressed by using polyvalent spelling units, while different spelling units with the same phonological correspondence can be creatively interchanged in order to construct homophones. It is now time to provide some terminology to account for all of these facts, rather than continuing with the circumlocutions.

5.5 Grammaphonology, and related terms

With all this in mind, I want to expand upon Henderson's terminological challenges, and where possible, provide inter-related terms that can help in our description of the relevant issues.

- (16) There is some need for referring to a grapheme or cluster of graphemes, whether it be a monograph, digraph or trigraph, which can be usefully regarded as a unit corresponding to any phoneme, regardless of how many phonemes it can correspond to.

This unit of spelling will be henceforth referred to as a **grammaphone**. This term covers the dual nature of a spelling unit as being both a visual form and a phonological correspondence. In English all the letters are grammaphones, as are the geminates, digraphs, trigraphs and ligatures: so <t>, <tt>, <th> and <tch> are all grammaphones, as is initial <pt>.

- (17) There is some need for referring to each *set* of such units that can usefully be regarded as corresponding with each phoneme.

A set of grammaphones that are used to realise a particular phoneme in writing will be referred to as a **grammaphoneme**. Accordingly, each of the grammaphones in a set will be called an *allogrammaphone* or **allogram**, for short.

For example, the grammaphoneme /f/ is a set that contains the allograms <f>, <ff>, <ffe>, <ph> and <gh>. We shall see later how morphograms such as <4> and <5> interrelate with these allograms.

Note that since a grammaphoneme is just a set of ways to write a phoneme it does not need its own bracketing convention.

Since English is polyvalent in both directions, grammaphones are not tied to any particular grammaphoneme (just as phones are not tied to phonemes). So <th> is an allogram of /θ/, /ð/ and /t/, (*thigh*, *thy* and *thyme*) while <ch> is an allogram of /tʃ/, /ʃ/ and /k/, as in *Chaucer*, *chauffeur* and *chaos*.

The study of the relationships between phonology and graphemes will be called **grammaphonology**.

- (18) There is some need for referring to each *set* of phonemes that can usefully be regarded as corresponding with each of these spelling units.

No terms will be provided here to describe these sets of phonemes. Apart from their use in morphophonemic alternations, there has not been enough research done on their uses or potential uses, so there is no need to force the issue, especially since they will not feature in the remainder of this study.

6. The Grammaphoneme

Most grammaphonemes contain many allograms, and these need to be subcategorised according to their orthographic behavior. This will be done in accordance with the updated set of subsystems provided in Section 4. We shall begin with /f/ because its allograms cover a wide range of subsystems, so it can work as our template for other grammaphonemes. We shall then look at /t/ which has some allograms with a heavily restricted distribution, and, as an introduction to vowels, we shall look at /ʌ/, because of its simplicity. With an understanding of these three grammaphonemes, we can then attempt to predict the possible homophonous spellings of /tʌf/, a string that occurs in the homophonous English lexemes *tough* and *tuff* (a geological term).

As we shall soon see, the geometric shape of grammaphonemes allows them to overlap, and this will help us to explain allograms with a restricted distribution, as well as strings of allograms that may not always be predictable from their individual parts. To complete this, the word *roughed* will then be modelled. After that, it will be time to introduce more complex issues such as long vowels, and the overlapping nature of grammaphonemes will allow us to model both rhotic and non-rhotic accents simultaneously. Finally we will look at the relationship between <4> and its homophones, *four*, *fore* and (for some speakers) *for*. This will allow us to see how suprasegmental graphemes can interact with canonical allograms, providing us with a direct mapping from alphabetic writing into morphogrammatic writing.

6.1 Modelling /f/

The §DEFAULT allogram of /f/ is <f>, its geminate is <ff> and <ph> is §GREEK. By contrast, §FRENCH <ffe> is restricted to word-final contexts, as in *giraffe* and *pouffe*,⁹ so it will be marked <ffe#> to indicate this. <gh> is a §NONCE allogram, since it is neither predictable nor phonetically transparent.

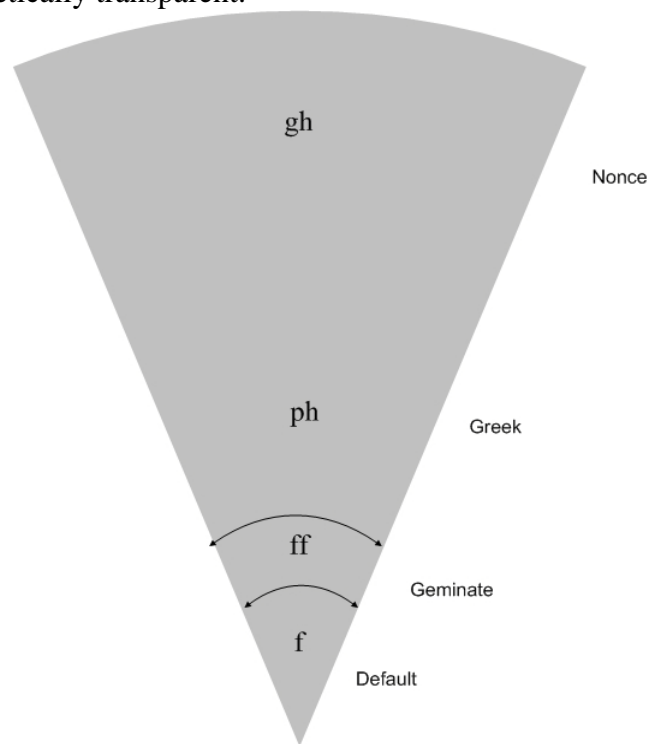


Figure 1 The /f/ grammaphoneme

6.2 Modelling /t/

We can do the same with /t/ and /ʌ/. <t> is the §DEFAULT and <tt> its geminate. <th> is §NONCE, and <pt> is §GREEK, but restricted to word-initial position (*ptarmigan*, *pterodactyl*), and will be marked <#pt>. By contrast, the past-tense morpheme <ed> and <d> only correspond to /t/ in word-final position (*stashed*, *moustached*), so they will be marked <ed#> and <d#>.

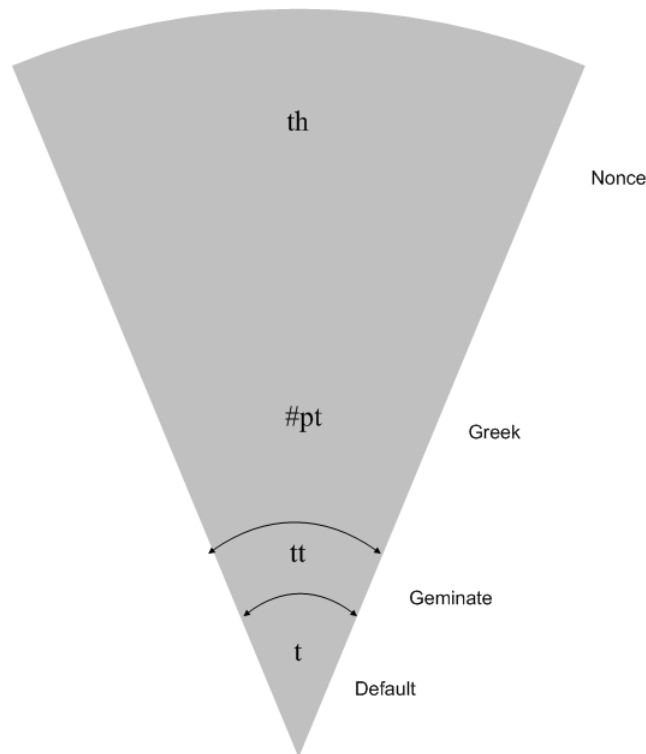


Figure 2 The /t/ grammaphoneme

6.3 Checked vowels: /ʌ/

Given the lack of graphemes available for representing vowel phonemes, they are much harder to model than consonants. They also tend to have a great deal of §NONCE allograms. We shall begin by looking at the relatively simple phoneme /ʌ/. The §DEFAULT spelling is <u>. There are some §NONCE allograms: the <oo> and <ou>, as seen in *blood* and *country* – neither of these is phonetically transparent or predictable. The same might be said of the split digraph <o...e> of *come* and *love*, but <o> on its own presents a problem case which raises some important issues.

According to Carney (1994: 147), /ʌ/ is spelt as <o> in 7% of lexical items in which it occurs, but in 27% of all examples in his corpora. This indicates that it appears in many frequently occurring words, (such as *son* and *money*). This spelling occurs frequently before <v>, <m> or <n> (Scragg, 1974: 44), ‘especially in §Romance words’ (Carney, 1994: 147).

The issue at stake here is the fact that <o> ~ /ʌ/ overlaps with the §DEFAULT correspondence <o> ~ /v/. The question that arises is whether we can predict the pronunciation from the orthographic context? No research tells us the answer to this, so I will tentatively underline such allograms. For now, I will underline it as <o>, and file it under §NATIVE.

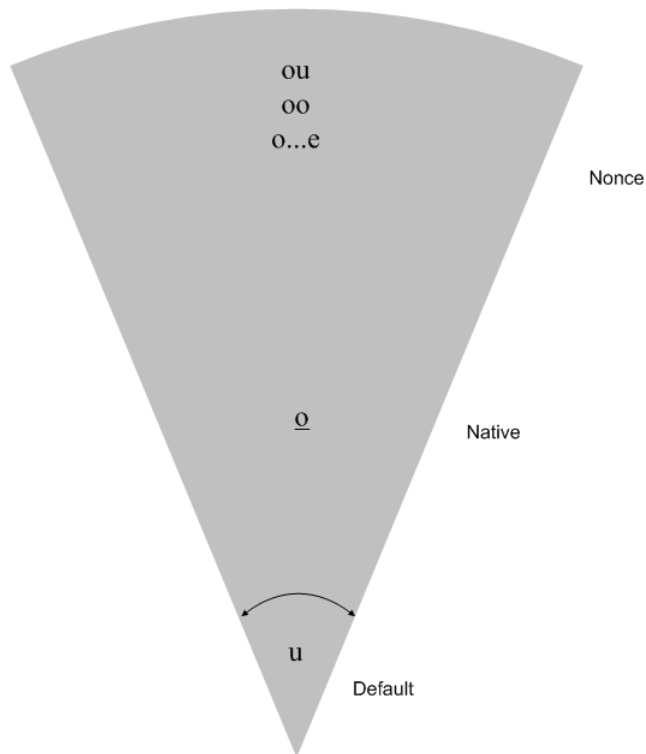


Figure 3 The /ʌ/ grammaphoneme

6.4 Constructed homophones of tough

Earlier in this essay we looked at some of the ways in which *tough* could be respelt. We shall now make this list exhaustive. Explaining the problem cases will introduce us to how grammaphonemes overlap. The §DEFAULT spelling is clearly <tuf>. <tuff> and <tuffe> are also options, as is <tuph>. It is not possible to spell it <tugh>, because §NONCE <gh> ~ /f/ only occurs after <ou> or <au>. We must therefore find a space for the §NONCE sequences <ough>, <augh> ~ /ʌf/, and they fall into the overlapping space caused by juxtaposing the grammaphonemes /f/ and /ʌ/.

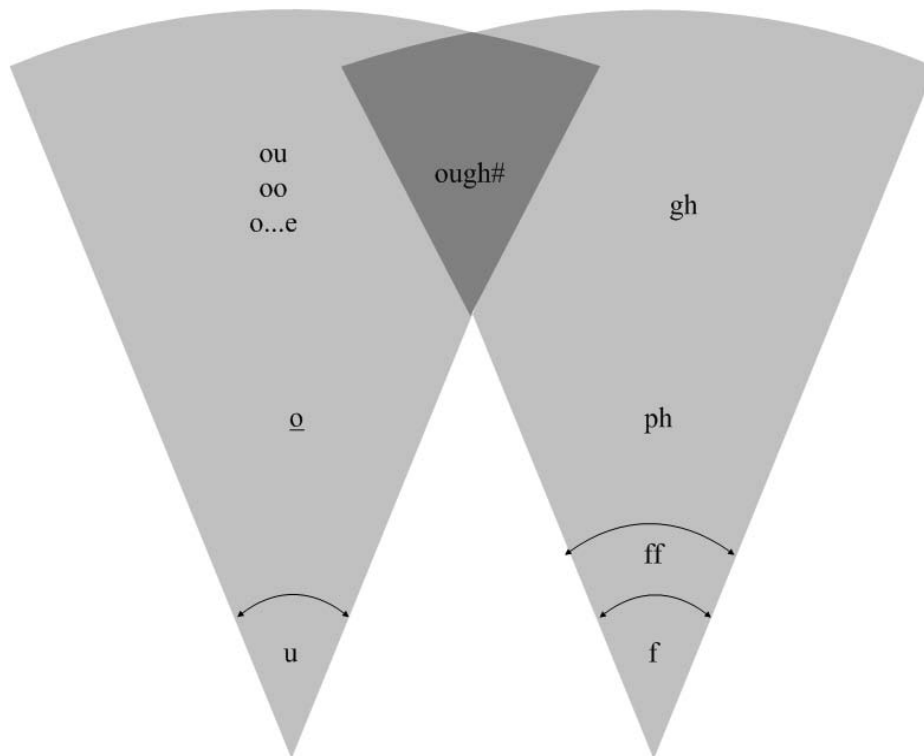


Figure 4 The /Δf/ grammaphonemes overlapping

Retaining the <ough> spelling requires the retention of <t>, so <ptough> seems to work, as do the related <ptuf>, <ptuff>, <ptuffe> and <ptuph>. Since <ou> is §NONCE, it cannot be retained without <gh>, disqualifying *<touf>, *<touff>, *<touffe>, *<touph>, and thus *<ptouf> etc. As for <ed#> and <d#>, these are restricted to word final position, and then only after fortis consonants. Clearly, there can be no <duff> or <eduph> etc. A more detailed visual model of grammaphones will have to account for such constraints, along with stress-imposed constraints which are often not marked in the orthography.

The spelling <ttuff> was also mentioned in Section 2, and morpheme-initial geminates were discussed there as being extrapolations from spellings such as *llama* and *Ffoulkes*. The final homophones we can construct for *tough* are therefore <ttough>, <ttuf>, <ttuff>, <ttuffe> and <ttuph>. Listing them all out, this gives us fourteen options:

- (19) tough, ptough, ttough
 tuf, tuff, tuffe, tuph
 ptuf, ptuff, ptuffe,
 ttuf, ttuff, ttuffe, ttuph.

6.5 Overlapping grammaphonemes

We saw with <ough> and <augh> that the geometric structure of grammaphonemes allows them to overlap, and we can use this feature to place allograms with different restriction. Consonant allograms which can occur in both onset and coda position go into the central section. Onset allograms such as <#pt> go into the right section, and coda allograms such as <ed#> and <ffe#> go into the left section. §NONCE allograms can go into the dummy spaces

at the top of the central section, and from that position their role is negated. <ough#> and <augh#> go to the left, but in its dummy space (above <ffe#>).

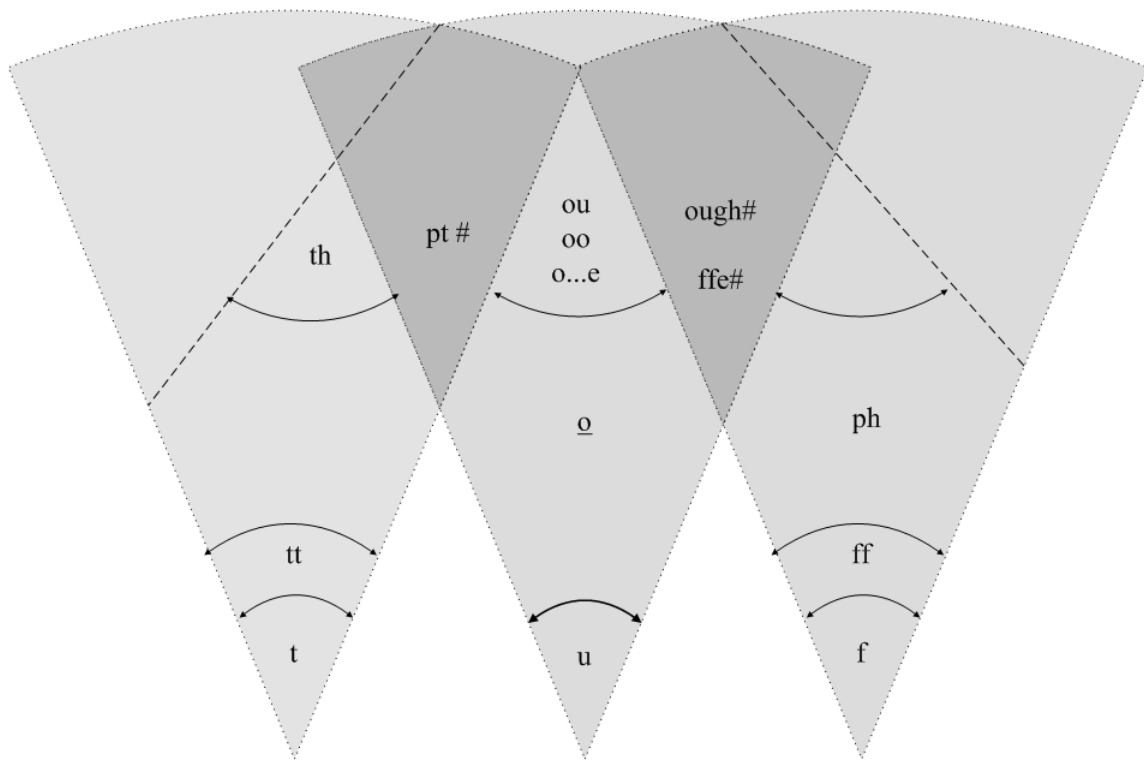


Figure 5 The /tʌf/ grammaphonemes overlapping

We can now extend the model somewhat to show longer words and their potential homophones. Notice firstly that we can add as many grammaphones as we like and they will all intersect.

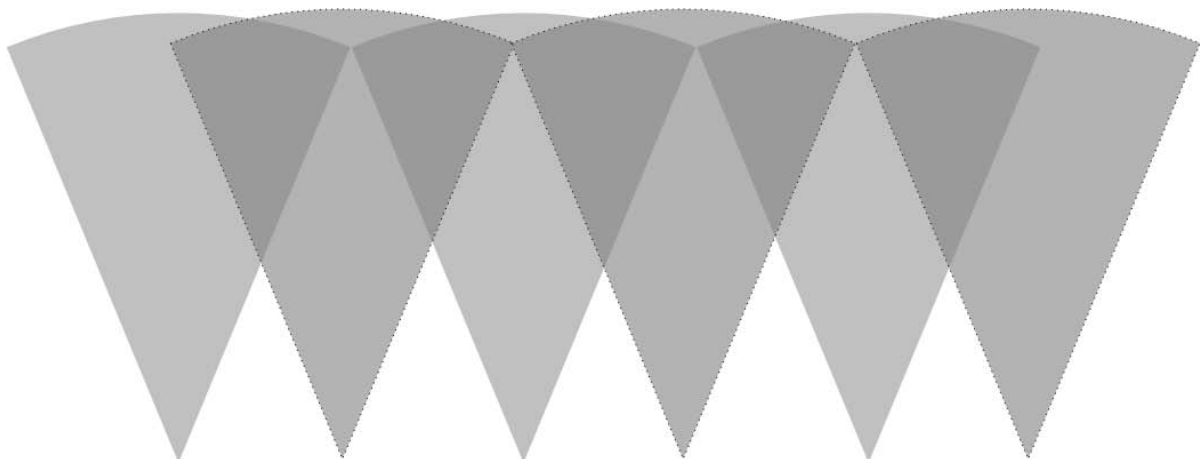


Figure 6 Overlapping grammaphonemes

6.6 Constructing homophones of <roughed>

We shall now model /rʌft/. Firstly though, we see that /r/ has §DEFAULT <r> and geminate <rr>. <rh> is §GREEK, and so is its geminate <rrh> (*gonorrhoea* etc). Restricted to onset position is the §NATIVE <#wr>. Modelling the four grammaphonemes /rʌft/ provides this model

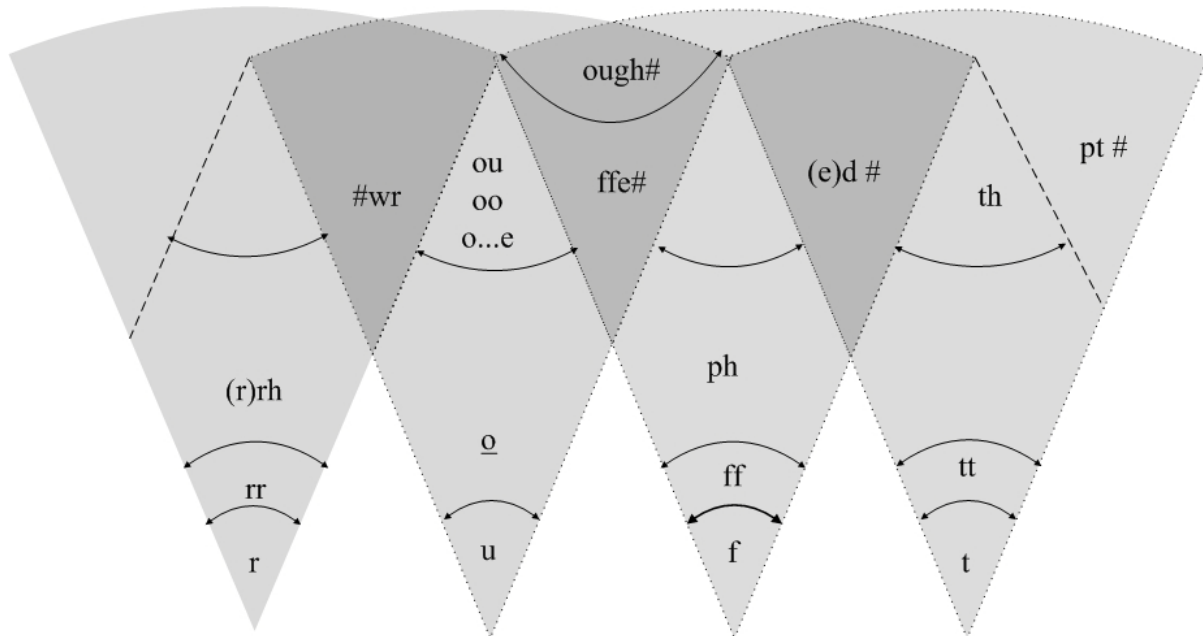


Figure 7 The /rʌft/ grammaphonemes overlapping

Notice now that <ed#> is now an option, whereas, <#pt> is not, and indeed <ed> occurs in *roughed* itself. The number of options is startling:

- (20) Ruft, Rruft, Rhuft, Wruft.
 Rufft, Ruffed, Ruffd,
 Rupht, Rrupht, Rhupht, Wrupht.
 Ruphed, Rruphed, Rhuphed, Wruphed
 Ruphd, Rruphd, Rhuphd, Wruphd

*Rufed is not an option, because the single <f> marks <u..e> as a unit.

Once again §NONCE <ough> - /ʌf/ probably ceases to be intelligible when used in many respellings.

- (21) [?]Roughed, [?]Rhoughed, [?]Wroughed

But

- (22) *Rought, *Rrought, *Rhought

*Wrought is blocked. It is hard to know whether geminate <tt> might also be used instead of <t>, as in [?]Ruphtt and [?]Rufftt etc., but it seems possible. Finally, initial <Rrh> may seem far-fetched, but still within a possible extrapolation of English spelling.

(23) [?]Rrhoughed, [?]Rrhufft, [?]Rrhupht, [?]Rrhuphed, [?]Rrhuphd.

Once again, it remains impossible to tell how intelligible all of these spellings are. For now though, there is no point in ruling them out, simply because we might not think of reasons to use them.¹⁰

6.7 Vowels and rhoticity

Among the vowels, we have so far only looked at /ʌ/. Checked vowels all have a simple grammaphoneme set, because their phonotactics only allow them to occur syllable internally. Free vowels are more complex, and critically, they all display different orthographic patterning morpheme-finally and morpheme-internally. Another concern with vowels involves rhoticity, and we need to be able to devise a model which accounts for both rhotic and non-rhotic accents. To do all this, we shall examine the TAUGHT vowel, firstly in rhotic accents, and then in non-rhotic accents. This will provide a model into which all the other vowels can be structured.

The §DEFAULT allogram for /ɔ:/ morpheme internally is <au>.¹¹ Morpheme-finally it is <aw>. Compare *saw* and *saunter*. Each of these also occurs in the converse position, but less often; here <aw> is a §NATIVE¹² allogram, as in *hawk*, although <au> seems to be distinctly §EXOTIC in final position, given the graphotactic constraint against ending words in <w> (e.g. *Nassau*). §NATIVE <augh> occurs both word finally and word internally (particularly before <t>). Compare *overslaugh* and *slaughter*. Another §NATIVE allogram is <al>, as in *all* and *almost*, *walk* and *talk*, although its predictability in new formations is uncertain, so it will be underlined. A §NONCE allogram is the <ough> of *nought* and *bought*. While neither <augh> or <ough> is transparent, <augh> is predictable, and <ough> (infamously) isn't.

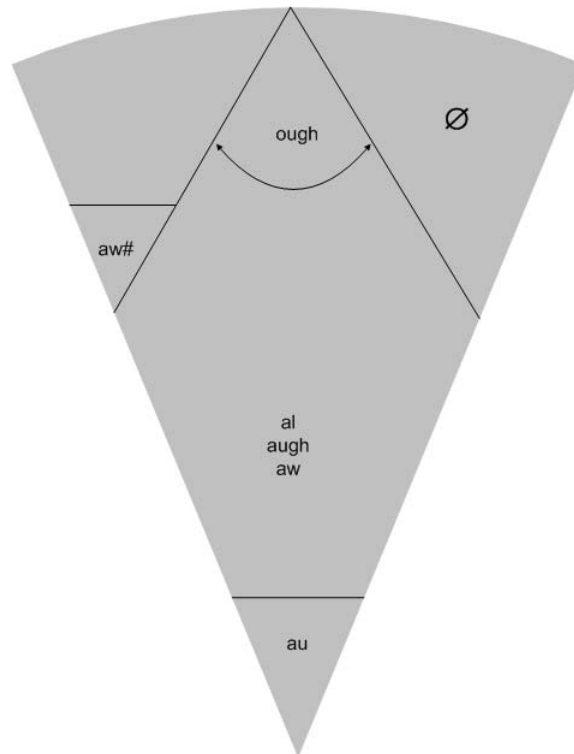


Figure 8 *The /ɔ:/ grammaphoneme*

We see now that the centre and left sections of this grammaphoneme are filled, but the right section is empty. This is because there are no constraints on initial /ɔ:/. This goes for all the free vowels: word-internal spelling constraints are the same as their word-initial constraints. Compare this to <#wr> or <#pt> which only occur morpheme initially (as syllabic onsets).¹³ It is into this empty gap that we can add the spellings of non-rhotic /ɔ:/.

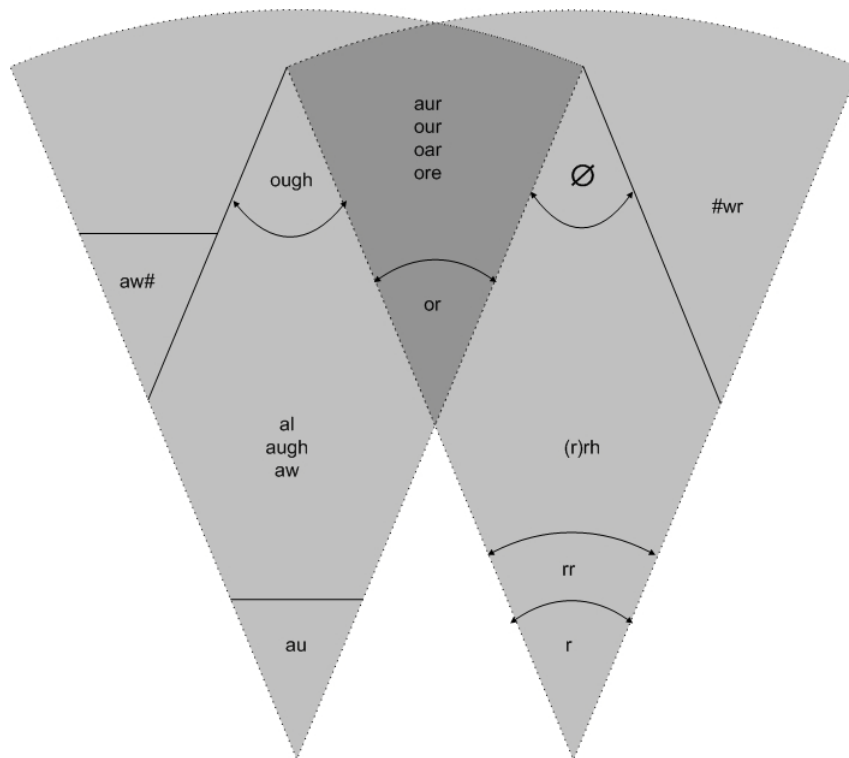


Figure 9 *The /ɔ:r/ grammaphonemes overlapping*

Owing to the great number of mergers which have resulted in the current non-rhotic TAUGHT vowel (Wells, 1982), this set has a massive amount of variation in it. Carney considers the §DEFAULT allogram to be <ore>, although there also exists the <our>, <or>, <oar> and <aur>, giving us *fore*, *four*, *for*, *boar* and *centaur*. *War* is a special case, whose not-irregular pronunciation requires the overlapping of three grammaphonemes to account for it.

Special cases such as this are rare, and <war> acts almost like a morphogram. Indeed it is at exactly this level of analysis where we can now begin to place morphograms. Since we can overlap grammaphonemes, we can map *four* and <4> onto the one model.

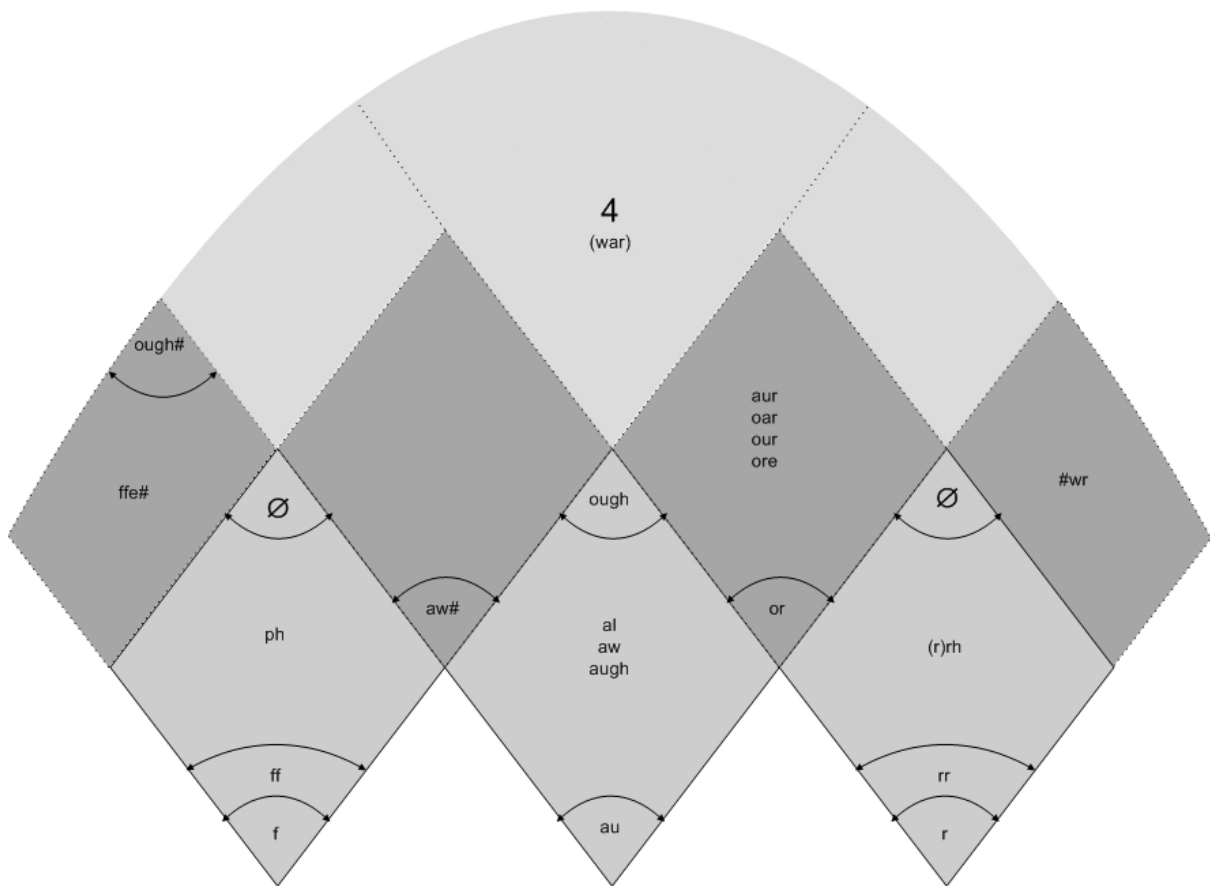


Figure 10 *The /ɔ:ɪ/ grammaphonemes overlapping*

6.8 Syllabograms and morphograms

This final example shows that the tiered model of intersecting grammaphonemes allows for the interspersal of phonograms with morphograms and syllabograms. In the introduction we saw the example of <C U L8R>. We can now see how <C> would go in the overlap between /s/ and /i:/, and <U> in the overlap between /j/ and /u:/ (assuming it is analysed that way). <L> remains as a default allogram, while <8> would go in the overlap between /eɪ/ and /t/. <R> functions as a syllabogram. Notice how in one case <8> can be a morphogram, but in another it is a syllabogram. This is a common feature of non-standard spelling.

7. Limitations to the study

The limitations of this study are enormous, but that is only to be expected in a study that attempts to step outside a closed field (standard spelling) and view it as part of a bigger domain of spelling, albeit a demesne that it presides over.

The aim here has not been to give an exhaustive account of all existing spellings, but to leave a model with large, open spaces whose breadth includes all spellings that are constructed by extrapolating from the existing graphotactic constraints. The problem is that

those constraints are not fully known, and can't be known until wider corpora of non-standard (or 'vernacular') spellings are studied, or until more of them actually *exist*.

This study has only examined a handful of phonemes and orthographic combinations. A more in-depth study is required to provide a more exhaustive analysis. We have not looked at polysyllabic words or reduced vowels. Many questions have arisen. To what extent, for example, is <ph> interchangeable with <ff>? When can <o> represent /ʌ/. What role do morphological connections play in recognizing new spellings (as with *Specsavers*)? How are phonetically conditioned spellings such as the <ti> of *nation* to be analysed within the proposed model of the grammaphoneme. All these questions and more need to be asked and answered.

8. Conclusion

This study has developed the existing theories of orthography so that they can explain not just standard spellings but also those creative spellings which are deliberately designed to be their homophones. Approaching spelling from this viewpoint has required an in-depth examination of how standard spelling itself works, and this has allowed us to extrapolate from its patterns to imagine possible constructed homophones which remain within the bounds of intelligibility, if not present-day likelihood.

We have had to analyse spelling solely in terms of its spelling to phoneme correspondences, rather than for its underlying morphological relationships, and doing this has required the introduction of some new terminology (*grammaphoneme* and associated terms), as well as refinements of some existing terms. The subsystems of English spelling have been fine-tuned for this shallow orthographic analysis, paving the way for a breakdown of each phoneme into its §DEFAULT allogram (spelling correspondence), its geminate allogram (where appropriate), and its subsystemic and §NONCE allograms (again, where appropriate).

The end result of this has been the development of a visual model which makes steps towards predicting possible re-spellings of well-formed mono-syllabic words, as well as providing a way of mapping from one kind of writing system (phonogrammatic) to others (syllabic and morphemic) and back again. It is hoped that this model can be used for an analysis of other kinds of spelling. Even though other alphabetic languages do not contain the same amount of inconsistency in their spelling as English, it might be possible to study orthographic code-switching within this model. Furthermore, turning the model upside down might be a useful way of modeling writing systems which use morphograms by default, and phonograms as their exception.

Notes

¹This is according to a Google search, carried out on 22/6/2011. The sequence <ptuff> did occur among some internet user names, but it is impossible to tell whether it was an abbreviation of sorts, and there was no indication that it was ever a homophone of *tough*.

² See also <http://a2dez.com/2011/02/making-a-nayme-for-yourself/> (accessed 22/6/2011).

³ <of> is one exception, <diphthong> another, although the <ph> ~ /p/ pronunciation of the latter may well be receding. Carney analyses *sapphire* as having <p> ~ /Ø/ and <ph> ~ /f/. Rollings (2003) considers <pph> to be the geminate of <ph>. This analysis depends on how we define a geminate. If it is to be a 'conditioned variant' (Carney's term) then it would surely occur more frequently, and *sapphic* would be echoed by <grapphic>. If it is simply another variant then Rollings' analysis would carry more weight.

⁴ Of course <f> is conditioned too, whenever <ff> is *not* called for. Compare *chaffed* and *chafed*.

⁵ *Cedar, cider, cyan, Cæsar* and *cæliac*.

⁶ Carney (1994: 41-2) introduces the terms *empty* and *inert* to distinguish between certain kinds of ‘silent’ letters. Inert letters are silent in one morpheme, but can surface in related morphemes: compare *sign* and *signed* with *signify* and *signature*. Empty letters never surface. Their only possible function is to mark vowel length – compare *salmon* with **samon*.

⁷ It should be noted that Ryan’s study of band-name spelling was done impressionistically, with a corpus that was not exhaustive or carefully delimited. Despite that, the overall patterns are frequently very clear, and merit referencing.

⁸ Providing an inventory of the graphemes of standard English, Venezky (1970: 47) writes: ‘Twenty-six letters of the alphabet <abc...z>, [and] eleven marks of punctuation <, ; : . ? ! – ” () –>. In addition we must include a space, a sort of zero grapheme.’ Herrick (1966, cited in Daniels, 1994: 429n) offers a larger inventory for ‘all Roman alphabets’. This encompasses numerals and some morphograms. He gives the same 26 letters, 10 numerals, 14 punctuation marks (< ‘ ’ / * > are added, <”> is absent), two morphograms (<\$ &>), 2 spaces (word and sentence), and 5 suprasegmentals (initial capital, capitals, small capitals, italics, boldface).

⁹ It occurs syllable-finally in the name *Jefferies*.

¹⁰ These kind of far-fetched spellings sometimes appear in onomatopoeia. For an exceptionally good discussion of the relationship between onomatopoeia and standard spelling, see Attridge, Derek (1988). *Peculiar Language: literature as difference from the Renaissance to James Joyce*. London: Methuen.

¹¹ In sum <au> takes up 25% of all spellings of /ɔ:/, although these seem to occur in lexically infrequent words (Carney, 1994: 182). Given that Carney is analysing the non-rhotic ‘Standard British English’ that figure includes all instances of rhotic /ɔ:r/.

¹² Many German words containing <au> seem to vary between /ɔ/ and /aʊ/ including businesses such as *Audi* and *Braun*.

¹³ In fact <pt> ~ /t/ only occurs word initially because the inert <p> surfaces in *helicopter*.

References

ANDERSON, John 2007. *The grammar of names*. Oxford: Oxford University Press.

ALBROW, Kenneth H. 1972. *The English writing system: notes towards a description*. London: Longman.

CARNEY, Edward. 1994. *A survey of English spelling*. London: Routledge.

CHOMSKY, Noam & HALLE, Morris. 1968. *The Sound Patterns of English*. New York: Harper and Row.

CRYSTAL, David. 2008. *Txtng: The gr8 db8*. Oxford: Oxford University Press.

DANIELS, Peter. T. 1994. Reply to Herrick. In *LACUS Forum*, 1994, vol 21, pp 425-31.

GIEGERICH, Heinz. J. 1992. *English phonology. an introduction*. Cambridge: Cambridge

University Press.

HENDERSON, Leslie. 1985. On the use of the term 'grapheme'. In *Language and Cognitive Processes*, Vol 1, 1985, pp 135-48.

HERRICK, Earl. M. 1966. A linguistic description of Roman alphabets. (M.A. dissertation). In *Hartford Studies in Linguistics*, 1966, vol 19. Hartford: Hartford Seminary Foundation.

JACOBSON, Sven. 1966. *Unorthodox spelling in American trademarks*. Stockholm: Almqvist & Wiksell.

KOVRT, Manfred. 1986. The term 'grapheme' in the history and theory of linguistics. In W. Augst (Ed.), *New trends in graphemics and orthography*. pp 80-96. Berlin: DeGruyter.

PLAG, Ingo. 2003. *Word Formation in English*. Cambridge: Cambridge University Press.

PRANINSKAS, Jean. 1968. *Trade name creation: processes and patterns*. The Hague: Mouton.

ROLLINGS, Andrew. G. 2004. *The spelling patterns of English*. Munich: Lincom Europa.

RYAN, Des. 2010. *Kre-8-iv Spell!nk: Why constructed homophony is key to understanding patterns of orthographic change*. Unpublished MSc Thesis, Edinburgh University.

SCRAGG, Donald. G. 1974. *A history of English spelling*. Manchester: Manchester University Press.

SEBBA, Mark. 2007. *Spelling and society: the cultural politics of orthography around the world*. Cambridge: Cambridge University Press.

SHORTIS, Tim. 2007. Revoicing Txt: spelling, vernacular orthography and 'unregimented writing, in *The Texture of Internet: Netlinguistics*. Ed. S. Posteguillo, M.J. Esteve and M.L. Gea. Cambridge: Cambridge Scholar Press, 2007. pp 1-23.

VENEZKY, Richard. L. 1970. *The structure of English orthography*. The Hague: Mouton.

WELLS, J. 1982. *Accents of English*. Cambridge: Cambridge University Press.

Des Ryan
CLCS, Trinity College,
Dublin University
Dublin 2
Ireland
skajbucket@yahoo.co.uk

<p>In <i>SKASE Journal of Theoretical Linguistics</i> [online]. 2011, vol. 8, no. 2 [cit. 2011-12-16]. Available on web page <http://www.skase.sk/Volumes/JTL19/pdf_doc/01.pdf>. ISSN 1339-782X.</p>
