Standardization in Early English Orthography

Over thirty years ago Fred Brengelman pointed out that since at least 1909 and George Krapp's *Modern English: Its Growth and Present Use,* it was widely assumed that English printers played the major role in the standardization of English spelling.[1] Brengelman demonstrated convincingly that the role of the printers was at best minimal and that much more important was the work done in the late 16th and 17th centuries by early English orthoepists and spelling reformers – people like Richard Mulcaster, John Cheke, Thomas Smith, John Hart, William Bullokar, Alexander Gil, and Richard Hodges.[2] Brengelman's argument is completely convincing, but it concentrates on developments rather late in the history of English orthography – developments that were external to the system itself and basically top-down. It necessarily ignores the extent to which much standardization occurred naturally and internally during the 11th through 16th centuries. This early standardization was not a top-down process, but rather bottom-up, arising from the communication acts of individual spellers and their readers – many small actions by many agents. In what follows I argue that English orthography is an evolving system, and that this evolution produced a degree of standardization upon which the 16th and 17th century orthoepists could base their work, work that not only further rationalized and standardized our orthography, as Brengelman has shown, but also marked the beginning of the essentially top-down system that we have today.

English Orthography as an Evolving Complex System. English orthography is not just an evolving system; it is an evolving complex system – adaptive, self-regulating, and self-organizing. Complex systems consist of diverse entities that interdepend and interact, and that adapt to changing environments, allowing the system to evolve through even quite radical environmental changes. Another important quality of complex systems is emergence – that is, their ability to produce new and unpredictable features that can order and complexify the system, sometimes locally, sometimes globally – as, for instance, in the use of syllable-structure to indicate long and short vowel sounds, which emerged in English orthography in the early 13th century.[3]

Complexity theorists argue that emergence is an expectable – though unpredictable and unpredetermined – occurrence brought on by a combination of a system's natural tendency towards order and its having achieved a sufficiently high level of complexity.[4]

Orthographic Codes. In the orthographic system the interdependent and interacting

entities are the human agents – that is, the spellers and their readers. Understanding their interdependence and interaction requires some discussion of orthographic code: The orthographic code is the system of abstract categories, distinctions, and relationships that structure the orthography; it is the set of rules of the orthographic game. Orthographic code resides in two realms of reality Karl Popper has called Worlds 2 and 3. In Popper's scheme human reality consists of three separate but equally real worlds: World 1 is the physical world – the outer world of physical objects and the events and processes involving those physical objects. World 2 is the psychological world – the inner world of psychological states and responses, of feelings, volitions, perceptions, the individual's knowledge, the world of the human mind. World 3 is the world of abstract ideas, the world of cultural values, principles, laws. It is the symbolic world, a product of the human mind, where reside such things as mathematical laws, logical principles, esthetic standards. It is not a Platonic ideal, but rather an evolving work in progress.

It is tempting to say that the orthographic code is of World 3, but that would be only partially correct: To be sure, part of orthographic code does exist in World 3, as the enduring, though evolving and highly abstract set of tactical, procedural, and correspondence rules and patterns that inform English spelling. But there are also the particular versions of the code in the minds of the individual spellers and readers of English – that is, in Popper's World 2. Each of these individual codes reflects the amount of the enduring code that the individual has internalized from his or her particular experience with written English. I will speak of the enduring general code of World 3 as *code3* and of the particular codes held by individual spellers and readers as *codes2.*

It is also tempting to think of the many codes2 as versions of the single code3, but to do so would be to get things twisted around phenomenologically: It is better to think of code3 as an evolutionary global property that emerges as the individual codes2 engage one another in acts of spelling and reading. As a global property, code3 then feeds back into the subsequent engagement of codes2 so as to regulate and reorganize the codes2, bringing them more into line with the emerging code3 and with each other. This circular interaction is a large part of the orthographic system's self-regulation and self-organization, its evolution to more pattern and standardization.

Early English Spellings and Emerging Attractors. The following discussion is based on the list of early variant spellings given in the 2nd edition of the *Oxford English*

Dictionary.[5] It's important to remember that despite the *OED* 's notorious concern for thoroughness and accuracy, we are dealing here with a sample gathered by an army of mostly volunteers. So it is a sample in which there is considerable potential for human error. Yet even with these caveats, the evidence provided by the *OED* 's list of early variant spellings is useful and compelling.

A complex system regulates and organizes itself via the pull of emergent attractors, states to which the system naturally tends. Acts of spelling and reading are communicative transactions between spellers and readers, which achieve varying degrees of success. Successful transactions generate and strengthen emergent attractors, which in orthography are simply privileged patterns. A spelling system evolves via the pull of many small interrelated attractors, which are molded by qualities favored by any evolving system: economy, reliability, pattern, ruliness, efficacy, endurance. Although specific attractors cannot be predicted, they can be identified and rationalized after they have emerged. Thus we can look back to earlier spellings and watch as later rules and patterns emerge, molded by the regulating and organizing force of the patterns that attracted spellers and readers to them.

For instance, Table 1 displays the various spellings recorded for the native English word *din*. The numbered columns represent centuries in which the variant spellings occurred throughout the four periods of English — Old (OE), Middle (ME), Early Modern (EMnE), and Modern (MnE), stretching in this table from the 11th through the 19th centuries. The paucity of forms recorded from Old English is due mainly to the fact that so few early manuscripts survived. Assuming that code3 emerges from the engagement of codes2 in the acts of spelling and reading, Table 1 gives us a simplified and schematized example of nine centuries of that engagement, in which we can watch four different attractors emerge and grow more dominant over, and thus standardize, subsequent codes2, as writers tend to choose the spellings they stipulate and readers come to expect them:

Table 1. Early Spellings of the Native Word *din*

Spelling	OE		ME			EMnE		MnE	
	11	12	13	14	15	16	17	18	19
dyne	●	●	●	●	●				
dyn	●	●	●	●	●	●	●		
dynne					●	●			
dynn							●		
dune			●						
dine			●	●	●	●	●		
dinne					●	●	●		
dinn							●		
din			●	●	●	●	●	●	●
dene			●						
deone			●						
deane							●		

Table 1 reveals the typical variety early on with gradual convergence to a smaller number of spellings, usually, as in the case of *din,* just one. It shows the early system state with several weak attractors being replaced by a state with one set of attractors that are very strong. It also illustrates the two fundamental questions involved here: First, how are we to agree to pronounce this word? And second, how are we to agree to spell that pronunciation? In the case of *din* although the initial and final consonant sounds pose no problems in pronunciation, there is some variation in the spelling of [n]. And there is considerable variation among the vowel spellings, indicating considerable variation in the vowel pronunciation. The Old English spelling with <y> was pronounced with the rounded short high front vowel [ü]. The <y> spelling persisted into the 17th century, although by then surely the vowel had unrounded to something more like modern [ɪ]. The rare variant with <eo> suggests some dialectal diphthongization. The 13th century *dene* and surely the 17th century *deane* suggest lowering to something much like modern [i] or perhaps [ɛ] or perhaps even [e]. However, the <i> spellings and the later <y> spellings indicate that the pronunciation with [ɪ] came very early to dominate.

The first attractor is the sound-to-spelling correspondence stipulating that word-initial [d] is regularly spelled <d>, a correspondence that, as the table shows, has controlled the spelling of this word from the very beginning. This correspondence is motivated by the larger systemic demand for economy: There is no reason for a spelling of [d] here more complicated than the simplest, <d> – thus, [d] = <d>.

The second attractor is the sound-to-spelling correspondence stipulating that in native words when stressed [ɪ] occurs in word-medial position, it is regularly spelled <i>. Thus, after the first question – How shall we pronounce this vowel? – has settled on [ɪ], the second question – How shall we spell it? – is settled by the emergent [ɪ] = <i> attractor.

The third attractor is the tactical pattern in which most word-final single consonant sounds that follow a short vowel are regularly spelled with a single consonant letter– or putting it more specifically, word-final [n] following a short vowel is regularly spelled <n>. The only common modern native words with word-final <nn> are *wynn* "a rune" and *inn*. *Wynn* has the more regular variant *wen*. *Inn* has its second <n> due to the Short Word Rule, which limits two-letter spellings to function words.[6]

The fourth attractor is closely related to the third: Over the centuries most Old English inflections reduced and converged to a neutral schwa sound, which came to be spelled <e>. By the end of Middle English that final <e> was no longer pronounced, and due to the systemic demand for economy, there was a growing tendency to eliminate it from the spelling unless it was filling some diacritic function, such as marking a long preceding vowel or a soft <c> or <g> or a voiced [ð] or insulating an otherwise word-final <u>, <v>, <s> or <z>.[7] So it gradually disappeared in the spelling of *din*, where it serves no such diacritical function.

In evolutionary terms, then, the <din> spelling was selected from among the competing variant spellings because it best fit the niche being defined by the four relevant components of code3, which over the centuries emerge as increasingly strong attractors. These four attractors come to dominate all future codes2 and the performance based on them. The spelling <din> has won the evolutionary competition and has become the standard. The history of English spelling, especially of native words, shows this same pattern again and again, though, of course, the sets of attractors differ from word to word.

Native and Latinate Strands. There are two quite different strands running through

the evolution of English orthography – the native strand from Germanic and Old English and the Latinate strand from Latin and the Romance languages, especially French. These two strands are driven by different expectations and priorities, molded by different attractors, productive of different kinds of rules and patterns. In English the spelling of Latinate words, especially those adopted directly from Latin, underwent considerably less change than did the spelling of native words. Other than the clipping of endings, the spelling of Latinate words tended to stay close to their spelling in Latin, French, or Spanish. Thus, the most privileged pattern, or attractor, in the spelling of Latinate words was the spelling of the Latinate source. Words adopted directly from Latin in particular tended to be spelled the way they had been spelled earlier rather than the way they had come to be pronounced, for Latin was spelled the same across Europe, though its pronunciation in, say, Italy, France, Germany, Ireland, and England would have varied considerably.

One type of interaction between the native and the Latinate strands can be illustrated by the native noun *boil* "pus-filled inflammation" and the 13th century French adaption *boil* "to heat". Table 2 lists the early spellings of the native noun *boil*:

Table 2. Early Spellings of the Native Noun *boil*

Spelling	OE		ME			EMnE		MnE	
	11	12	13	14	15	16	17	18	19
bȳl	●								
bile			●	●	●	●	●	●	●
bele					●	●			
biel				●					
byil				●					
bule				●	●				
byle				●	●	●	●		
beel					●				
byelle						●			
boyle						●	●		
boile							●		
boil							●	●	●

In Table 2 the Old English vowel was the long rounded high front vowel, [ȳ]. The spellings suggest that it unrounded early on (except, perhaps, in the areas of the <bule> spelling in the 14th and 15th centuries). It tended to remain high front and tense [i], or to have lowered somewhat toward [ɛ]. However, in the 16th century the <oy> and <oi> spellings suggest a convergence to the Latinate verb *boil* , adopted in the 13th century from Old French *boillir* – perhaps due to a perceived connection via the senses of heat and inflammation.

The evidence here is complicated by the English variations involved: For instance, in certain British dialects long [aɪ] is pronounced [ɔɪ], and notice the persistence of the *bile* spelling at least through the 19th century. In his treatment of the native noun *boil*, Samuel Johnson says "See *bile,*" and at *bile* he says, "This is generally spelt *boil*; but, I think less properly" – which, from an etymological point of view would be accurate.[8] Later, John Walker agrees, not even listing the native noun *boil*, and defining *bile* as "a sore angry swelling. Improperly *boil*."[9] Noah Webster, on the other hand, in his treatment at the native noun *bile* "an inflamed tumor" says, "See *boil*, the correct orthography" – which from an evolutionary point of view is accurate.[10] The *OED, Shorter Oxford,* the *Dictionary of American English,* and *Webster's 2nd* and *3rd International* all list *bile* as either a dialectal or obsolete variant of *boil*.

As with *din* in Table 1, with the native noun *boil* in Table 2 there are four components of code3 tending to attract the variants to certain correspondences:

First, the initial [b] is spelled <b>, consistent with a very powerful and early established sound-to-spelling correspondence similar to that for initial [d] in *din*: [b] = <b>.

Second, Though the native noun *boil* did not originally have the [ɔɪ] vowel sound, it evolved into a homonym of the Latinate verb *boil*, which did. By the 16th century the medial vowel comes to be pronounced [ɔɪ] and spelled <oi>, echoing and converging to the Latinate verb *boil* and consistent with the now dominant correspondence in English that stipulates that [ɔɪ] is regularly spelled <oy> only in word-final position and is regularly spelled <oi> everywhere else.

Third, after a vowel digraph [l] is regularly spelled <l> rather than <ll>.[11]

The fourth attractor is again the tendency to omit silent final <e>'s that are serving no diacritical function.

The evolution of the native noun during the Early Modern English period parallels that of the Latinate verb, whose English spellings are summarized in Table 3:

Table 3. Early Spellings of the Latinate Verb *boil*

Spelling	ME			EMnE		MnE	
	13	14	15	16	17	18	19
boille	●	●					
boili	●	●					
boile(n)	●						
boyle		●	●	●	●		
boyll			●	●			
boyl				●	●		
boile					●		
boil				●	●	●	●
buyle		●					
?bayl		●					
?bele		●					
bule			●				
bulle			●				
bylle			●				
byle			●				
boll			●	●			
bull (Scot.)				●			

The rare 14th through 16th century spellings recorded in the bottom half of Table 3 were probably caused in part by the fact that the diphthong [ɔɪ] was a late import from French and not native to English, so it varied greatly in pronunciation. But true to the Latinate tendency, the verb *boil*, in spite of all of the variation over the centuries, is attracted finally to a spelling close to that of its French original.

The French Frontshift Preemption. In addition to self-regulation, robust complex systems have the ability to reorganize themselves to adapt to radical change. One of the largest self-reorganizations in our orthographic system was the Gallicization of English starting with the Norman Conquest and extending until the late Middle English period. There was, of course, the importation of hundreds of French words. There was also the French-trained scribes' preference for Gallic spellings, such as <qu> for Old English <cw> or <wh> for Old English <hw>.[12] But there were also less obvious, deeper changes, as in the stress frontshifting, described below, which led to a regular preemption of wider tactical expectations:

In French, words were stressed on the final syllable unless that syllable was a weak <e>, in which case they were stressed on the penult. As French adoptions were integrated into English, the stress shifted forward, especially in nouns and in line with the native English pattern of stressing the first syllable of the base. But before the stress frontshifting occurred, French adoptions would have been pronounced with the first vowel unstressed and either reduced or short. Then, when the stress moved forward in the word, these vowels, now stressed, would be unlikely to lengthen. As a result there are in English hundreds of French adaptions that have short stressed vowels in patterns that would lead us to expect long vowels. This stress frontshifting preempts the more general VCV vs. VCC pattern, which stipulates that a stressed vowel letter followed by a single consonant which is in turn followed by another vowel letter will be orthographically long, as in a words like *demon* and *dame*.

However, in a French adaption like *damage,* although the stress is on the first vowel, which heads a VCV string, that head vowel is short, [æ]. And the spelling is with single <m> rather than the doublet <mm>: It is not *dammage, with the VCC pattern normal for medial short stressed vowels. This preemption of the VCV vs. VCC pattern is perfectly regular in French adaptions. And the competition among variant spellings for some representative words suggests the evolution of this preemption across the centuries.

Table 4 displays aspects of this competition in a small sample of disyllabic French adaptions. There are three steps in the preemption of the VCV vs. VCC pattern: (1) At first the new adoption carries the French stress, on the final syllable rather than the native English stress on the first. This non-native stress pattern is often reflected in spellings that load the final syllable more heavily by either using a consonant doublet or other cluster, or sometimes, it would seem, a vowel digraph. (2) As the stress shifts to

the first syllable, there is a tendency to load the first syllable more heavily in the spelling, usually by the use of consonant doublets or clusters, to reflect the heavy stress and short vowel. (3) In time the preemption is encoded and the loading disappears, with the domination of the variant spellings by a simpler spelling, one that is often exactly that of the original French, in keeping with the powerful attractive force of Latinate words' source spelling.

To speak of three steps does not imply any neat unfolding of the three in a series of historical stages. The three steps are meant to suggest some of the contentions at work among emerging attractors as English spellers spelled such words, interweaving the native and Latinate strands. The second and third columns of Table 4 provide representative spellings with the back and front loading. The numbers indicate the centuries in which the *OED* records spellings. The table is not exhaustive, excluding several spellings without any loading. But the contentions were clearly there as spellers spelled and readers read, as their various codes2 engaged one another and as this preemption gradually emerged in code3 from the interaction:

Table 4. Back and Front Loading of French Adaptions

French Spelling	French Stress with back loading	English Stress with front loading	English Spelling
habit	habitt 16 habette 16-17	habbet 16 habbett(e) 16	habit 15-
limite	lymytt 15 lymit(t)e 15-17 limitt 16	lymmit 16 lymmet 16 limmit 17	limit 16-
pitet, pitez, pité, pitié	petey 14 pitee 14-15 pytye 14-16 pytee 15 pytie 15 pytye 15-16 pitye 15-16 petie 15-17	pitte 14-15 pytte 14-16 pittie 15-17 pittye 15-17 pyttye 16 pitty 16-17	pity 16-
tenant	tenaunt(e) 14-16 ten(e)awnte 15 ten(e)awunt 15	tennaunte 15-16 tennant 16-7 tennent(e) 16-17 tennent 17	tenant 14-

French Spelling	French Stress with back loading	English Stress with front loading	English Spelling
visite	visitte 14 vysitte 14 visytt 15 vysett 15 visitt 15-17 visett 16	vissite 15 vissett 15	visit 15-

In Table 4 the back loading reveals English spellers' attempts to mark the un-English final stress. And the frontloading reveals their attempts to mark the initial stress consistent with the emerging and broader VCV vs. VCC patterns. But eventually the more powerful pull of Latinate words to be adapted with their Latinate spellings won out in the evolutionary competition, resulting in a regular, but complexifying and more local, preemption of the general VCV vs. VCC pattern.

* * * * * * * * *

Bottom-Up and Top-Down Standardization. More than five centuries of early bottom-up self-regulation and reorganization in the English orthographic system led to a level of standardization and regularization that made possible the work of the 16th and 17th century orthoepists. Their work marked the beginning of the end of the dominance of bottom-up regulation. After them came a sharp increase in top-down influence – for instance, via the great dictionaries – Johnson, Walker, Webster, the *OED*. Also important was the spread of literacy and the generations of English teachers, who enforced increasingly rigid notions of orthographic correctness. Thus, the system grew more locked into place, as dictionaries and teachers and the growing mass of printed texts, preserve, among other things, such spelling oddities as *forty, fiery,* and *ache*. Just as it did with the many variant spellings illustrated above, robust bottom-up regulation would eventually have worked to rid the system of such anomalies, urging the more morphologically regular *fourty and *firy, and the more etymologically and phonetically straightforward *ake. But that would seem to be no longer possible. About the only bottom-up regularization possible now would be to pick the more regular spellings from sets of accepted variants – *renig* rather than *renege*; *milage* rather than *mileage,* and at least in American English, *worshiping* rather than *worshipping*, and *traveler* rather than *traveller* – relatively modest standardizations when compared with the accomplishments of the 11th through 16th centuries.

Notes

1. Fred Brengelman. 1980. Orthoepists, Printers, and the Rationalization of English Spelling. *Journal of English and Germanic Philology,* 79. 333.

2. Brenglelman, 332-54.

3. Fernand Mossé. 1952, 1961. *A Handbook of Middle English,* trans. James A. Walker Baltimore: Johns Hopkins University Press. 16-17.

4. A good introduction to complexity theory is M. Mitchell Waldrop. 1992. *Complexity: The Emerging Science at the Edge of Order and Chaos.* New York: Simon and Schuster. Somewhat more technical is Stuart Kauffman. 1995. *At Home in the Universe: The Search for the Laws of Self-organization and Complexity*. New York and Oxford: Oxford University Press. See also Nick C. Ellis and Diane Larsen-Freeman, eds. 2009. *Language as a Complex Adaptive System.* Chichester UK: Wiley-Blackwell. and Diane Larsen-Freeman and Lynne Cameron. 2008. *Complex Systems and Applied Linguistics*. Oxford: Oxford University Press.

5. 2002. Second Edition on CD-ROM, vers. 3.1. Oxford: Oxford University Press.

6. Otto Jespersen. 1909, *A Modern English Grammar on Historical Principles, Part1: Sounds and Spellings*. London: George Allen & Unwin. Reprint 1954. London and Copenhagen: Allen & Unwin and Munksgaard. 70, 149.

7. D. W. Cummings. 1988. *American English Spelling.* Baltimore: Johns Hopkins Universtiy Press. 145-54.

8. 1775*. A Dictionary of the English Language*. Reprint 1983. London: Times Books.

9. 1791, 1837. *A Critical Pronouncing Dictionary and Expositor of the English Language*. Glasgow: Blackie & Son.

10. 1828. *An American Dictionary of the English Language*. Reprint 1970. New York and London: Johnson Reprint Corporation.

11. Cummings, 445-46.

12. Mossé, 10-12.