

## REPLIES

### The redundancy of self-organization as an explanation of English spelling

GEOFFREY SAMPSON

*Sussex University*

Kristian Berg and Mark Aronoff (2017; henceforth B&A) use historical data on the spellings of various English suffixes to argue for the role of ‘self-organization’ as a process by which system emerges spontaneously in an unregulated linguistic domain. They argue that choices of spelling for given phonological suffixes gradually settled down in such a way that spellings became indicators of part of speech. This is an attractive idea, but B&A’s data fail to support it, because they do not adequately consider the availability of other explanations. Data cannot constitute evidence for a novel theory, if they are already convincingly accounted for independently of that theory. This, I believe, applies to most if not all spellings cited by B&A. Many are explainable in terms of phonological detail, or of the nature of education during the relevant period; some have more specialized explanations.

B&A discuss four cases: the suffixes *-ous*, *-ic*, *-al*, and *-y*, together (in each of the four cases) with suffixes that are phonologically similar but spelled differently. To keep the length of this commentary within bounds, I do not deal exhaustively with every example word cited by B&A. I begin by showing that for their first case, *-ous*, all of their data are convincingly accounted for by explanations independent of ‘self-organization’. Then from B&A’s three other cases I pick out examples that introduce considerations different in kind from those already discussed.

I do not claim to be able to ‘explain away’ every last one of B&A’s data. But that is to be expected, considering that I do not have (and perhaps no one has) an ideally complete knowledge of the evolution of English spelling and English pronunciation (particularly the pronunciation of vowels in unstressed syllables). However, the proportion of B&A’s examples not accounted for by any of the principles I discuss is not large enough, I believe, to leave the case for self-organization persuasive.

B&A introduce the *-ous* case by displaying a table (their Table 2, p. 45) that lists alternative graphic patterns for words that they regard as sharing a common phonological suffix,<sup>1</sup> which they transcribe as /is/.

(1) <ous>	as in	hazardous, nervous
<us>		bonus, genius
<is>		glottis, tennis
<ess>		hostess, princess
<ice>		office, service

B&A note (p. 41) that the CELEX database used for their research makes a phonetic distinction between [əs] as in *nervous* and [ɪs] as in *office*, but they see that as not relevant because ‘[t]his distinction is not justified phonetically in many varieties of English’; they suggest, citing Flemming and Johnson (2007), that in American English the distinction only occurs in the special case of pairs like *Rosa’s*, *roses*. However, what matters for B&A’s argument is not the phonology of any present-day English dialect, but that of the language during the period when spelling conventions were still evol-

<sup>1</sup> Not all of the final syllables discussed by B&A are ‘suffixes’ in the etymological sense, as they know (their p. 40). In *office*, <fic> is from the root of *facere* ‘make, do’.

ing. (For *-ous* and all but one of the other phonological suffixes discussed by B&A, the endpoint of the period they consider is 1710.) One cannot assume that this resembled present-day American English in the relevant respect, considering that in standard British Received Pronunciation (RP) the five graphic patterns listed above correspond to at least three different phonological patterns. A. C. Gimson (1967) gives *hazardous*, *nervous*, *bonus*, and *genius* with /əs/; *glottis*, *tennis*, *office*, and *service* with /is/; *princess* with /es/ (thus not wholly unstressed); and *hostess* with /is/ but /es/ as a 'less frequent' variant. Failure to distinguish the reduced vowels of *tennis* and *office* from those of *hazardous* and *nervous* is unusual in Britain, as demonstrated by the fact that it is a striking feature of the speech of Tony Blair, prime minister from 1997 to 2007, which otherwise conforms to RP norms.<sup>2</sup>

B&A's decision to ignore the contrast between [əs] and [is] would be justified if American English reflected the original position in this case and RP usage were innovative. For some of the obvious phonological differences between the two dialects, that is so, but here the truth seems to be the reverse. Roger Lass (1999:135) says that it is hard to give a coherent account of the history of English reduced vowels, but that '[a]s early as the fourteenth century the incipient standard had at least two reduction vowels (still so in RP and many other varieties): a higher and fronter one identified with short /i/, and a lower one, perhaps [ə]'. This suggests that it is inappropriate for B&A to see their five graphic patterns as having evolved as contrasting renderings of a single phoneme sequence, even if that is what they now are for American speakers.

The examples of B&A's Table 2 also demonstrate the relevance of literate speakers' knowledge of etymology. B&A quote Aronoff (1985:28) as calling written language 'the objectification of spoken language', as if those who were influential in developing English orthography thought of it as a device for transcribing the sounds of speech and nothing more. But, as I have argued elsewhere (Sampson 2017:Ch. 8), it is normal for orthographies to be based mainly or wholly on pronunciation when they are young but to give other considerations an increasing role as they mature. This was unquestionably so for English. During the period when English spelling remained fluid, the minority of the population who had more than very elementary education received it in schools, many of which were called 'grammar schools', reflecting the fact that from the Middle Ages onward the only or chief subject taught in all of these schools was Latin, with special attention to its grammar. When the famous educator Samuel Butler became a headmaster in 1798, his decision to make Latin grammar a less dominant aspect of the curriculum was a large innovation (Barnard 1961:72–73). Even when I applied for undergraduate admission to Cambridge University in 1962, Latin was the one individual subject of which I was required to demonstrate knowledge, though it was not the subject I aimed to study there. In this context it was inevitable that people who made decisions about how to spell the many English words that derive from Latin considered not only pronunciation but also etymology. To quote Dr. Johnson, recognized by B&A (p. 37) as one of the individuals influential in fixing English spelling: 'The great orthographical contest has long subsisted between etymology and pronunciation' (Johnson 1747:9). In some cases the spellings that became conventional gave etymology extreme priority over pronunciation, as with the <b> of *debt*, *doubt*.

<sup>2</sup> For what it is worth, as a British speaker one generation younger than Gimson my usage agrees with his transcriptions of all these words, except that what was for Gimson the less frequent variant of *hostess* is the only version I recognize.

From the etymological point of view, it would not have been tempting to spell *office* and *service* with the same vowel as *hazardous* and *nervous* even if the syllables had sounded alike. The former words have <i> in Latin (*officina*, *servitium*); the <ous> suffix reflects the Latin suffix -os-, as in *nervos-us* from *nerv-us* or *glorios-us* from *glori-a*. (*Hazardous* does not derive from Latin, but once the Latin suffix had been naturalized as English <ous> it was available to be attached to other nonnative nominal roots—*hazard* was a loan from French which is thought to have come to that language from Arabic. I do not know why the suffix is spelled <ous> rather than <os> in English; perhaps this was a compromise between etymology and the English pronunciation with schwa.)

Thus both English phonological history and the nature of English education make B&A's theory about 'self-organization' redundant as an explanation of the difference between the <ous> and <ice> spelling patterns, and phonology probably accounts for the <ess> pattern.

There are also explanations independent of self-organization for the <us> and <is> spellings in B&A's Table 2. In the first place, it is clear that the words *bonus*, *genius*, and *glottis* were originally used in English as nonnaturalized Latin words (in the case of *glottis*, ultimately from Greek), because they include Latin inflectional endings whereas naturalized loans from Latin conventionally omit these, as elements having no role in English grammar. In order to make the self-organization theory redundant in these cases, it might be enough simply to point out that the words were nonnaturalized loans. But it will be more convincing to give specific reasons why these words should have been left unadapted, and this can be done. According to the *Oxford English Dictionary* (*OED*), *genius* was originally used in English as the proper name of an allegorical figure, and many quotations as late as the nineteenth century spell it with capital G and/or pluralize it as a Latin word, *genii*.<sup>3</sup> Non-English proper names are not normally assimilated to English spelling conventions. *Glottis* has always been a learned term belonging to a profession, medicine, that made heavy use of Latin to a late date; hence it was naturally resistant to respelling in an anglicized form even if speakers pronounced its second syllable like that of *nervous*. As for *bonus*, this word includes not just an inflectional ending but an incorrect one, since the sense would require the neuter, *bonum*. The *OED* describes *bonus* as an 'ignorant or jocular' usage which probably originated as Stock Exchange slang, and its earliest quotation is from 1773, by which time orthographic fluidity had largely been eliminated. The suggestion seems to be that brash young traders, with little awareness of or interest in the normal philological conventions for naturalizing Latin loans, simply used what they remembered from their schooldays as the dictionary headword corresponding to English 'good'.

That leaves *tennis*, which is the only word under B&A's -ous heading whose etymology is quite obscure, and which it was perhaps therefore natural to spell with a 'default' representation for unstressed /ɪs/.

Turning to B&A's -y case (their Table 14, p. 55): B&A suggest that self-organization led to convergence on spellings with <ie> rather than <y> for words that are 'diminutives or hypocorisms'. That might seem plausible for their example *brownie*, but in reality *brownie* is spelled with <ie> not because it is a diminutive but because it was a loan from Scots, which before the political union of 1707 ranked as a separate language with its own spelling conventions; the -y of standard English regularly corresponds to <ie> in Scots: for instance, Macleod and Cairns (1996) give *cairrie*, *cerrie* as Scots for

<sup>3</sup> My *OED* references are to the original edition, published between 1884 and 1928.

*carry*, *yirdie* as Scots for *earthy*, and so forth. (This orthographic difference may reflect a pronunciation difference, but it is not necessary to discuss that here.) The earliest *OED brownie* quotation from a non-Scottish source is from *Jane Eyre*, published in 1847. By then, source-language orthography was normally treated as decisive for spelling loanwords in English even from languages other than Latin. And this principle accounts for some of B&A's other *-y* examples, such as *spaghetti* and *chassis*.

B&A's statement that <trannie> rather than <tranny> is the 'preferred variant' spelling of the slang term for *transvestite* seems straightforwardly mistaken; it does not match my experience, and Google offers me 'about 148,000,000 results' for <tranny> but only 'about 2,390,000' for <trannie>. These statistics are complicated by the fact that <tranny>, but not <trannie>, is also used for *transistor radio* and for (automotive) *transmission*, but that can hardly explain away such a huge difference. There are plenty of cases of <tranny> for *transvestite* in Google's first page of results.

In B&A's *-al* case, of the fourteen words they cite on pp. 49–51 in which /əl/ is spelled <al>, all but one have that spelling because of derivation from Latin.<sup>4</sup> (The exception is *arrival*, from French, and it is easy to surmise that this spelling was chosen by analogy with *rival*, which is from Classical Latin.) Inevitably, words including the Latin adjective-forming suffix *-al* (true of all these Latin-derived words other than *metal*, where <al> is part of the root) tend to be adjectives in English. But another spelling listed by B&A under their *-al* heading, <yl> as in *acetyl*, *ethyl*, *methyl*, *vinyl*, introduces a principle not discussed earlier. These words were coined in the early to mid-nineteenth century. By that period technical terms were disseminated via writing. B&A (p. 50) call <yl> a suffix that 'could be spelled with <al>, but [is] not';<sup>5</sup> the issue facing speakers in practice, however, will not have been 'How should we spell this pronunciation?' but 'How should we pronounce this spelling?'. And speakers have not converged on a single answer to the latter question. Gimson gives the 'chemists' pronunciation' of *ethyl* and *methyl* as /((m)i:θail/ (the only pronunciation I was aware of—they are not words I often encounter), with /((m)eθil/ as an alternative 'commercial and general pronunciation'; he gives *acetyl* only as /æsitil/; he does not list *vinyl*, but in my experience this is the commonest of the four words and is said today with /əl/ not /il/ (though a century ago the *OED* transcribed it with /il/). So these four words cannot be said to share the same phonological suffix, and the only one of them that, phonologically, 'could be spelled with <al>' is *vinyl* in its recent pronunciation.

On the other hand, I cannot explain the different suffix spellings in, for instance, *crumble*, *channel*, *devil*, *gambol*. To some extent phonological differences may account for them. The RP pronunciation of the second syllables of all these words is identical, but *crumbling* is two syllables while *channeling* is three. *Channel* came from French, which does not reduce vowels in final syllables to schwa, so perhaps an /e/ sound was retained in English for a while (as did happen with *hostess* and *princess*, also from French). According to the *OED*, 'some dialect glossaries' give a pronunciation *gam-BOWling* for *gambolling*, so possibly the <o> of *gambol* was phonologically motivated in the past. But this is an area where some data remain unexplainable (by me at least). I question, though, whether these data could amount to enough to make self-organization a more compelling hypothesis than postulation of now-lost details of early phonology.

<sup>4</sup> In three of these cases, *accidental*, *regimental*, and *spiral*, the words are not attested from the Classical period, but medieval and modern Europeans (writing in Latin or in modern languages) were quite prepared to coin new derived forms using Latin roots and affixes.

<sup>5</sup> Etymologically <yl> is not suffix but root, from Greek *hyl-ē* 'stuff'. (Greek /h/ regularly drops when not word-initial.)

I do not discuss B&A's *-ic* case separately, since doing so would introduce no new principles.

In the foregoing I have appealed to various conventions, such as that of preferring spellings which are better-justified etymologically where pronunciation gives no reason to choose between alternatives; these conventions led (for instance) to the gradual abandonment of spellings like *publick* and *publique* in favor of *public*. It might perhaps be said that such conventions are themselves evidence of self-organization. But B&A are arguing for an idea that goes beyond that: namely, for convergence on a single choice where there are no factors that could point to one choice rather than another. For B&A's data there are such factors. Self-organization might have a role in the evolution of orthographies, but B&A's case is not convincing.

#### REFERENCES

- ARONOFF, MARK. 1985. Orthography and linguistic theory: The syntactic basis of Masoretic Hebrew punctuation. *Language* 61.28–72. DOI: 10.2307/413420.
- BARNARD, H. C. 1961. *A history of English education from 1760*. 2nd edn. London: University of London Press.
- BERG, KRISTIAN, and MARK ARONOFF. 2017. Self-organization in the spelling of English suffixes: The emergence of culture out of anarchy. *Language* 93.37–64. DOI: 10.1353/lan.2017.0000.
- FLEMMING, EDWARD, and STEPHANIE JOHNSON. 2007. *Rosa's roses*: Reduced vowels in American English. *Journal of the International Phonetic Association* 37.83–96. DOI: 10.1017/S0025100306002817.
- GIMSON, A. C. (ed.) 1967. *Everyman's English pronouncing dictionary*. 13th edn of a reference work initially compiled by Daniel Jones. London: Dent.
- JOHNSON, SAMUEL. 1747. *The plan of a dictionary of the English language*. London: Knapton, Longman, et al.
- LASS, ROGER. 1999. Phonology and morphology. *The Cambridge history of the English language, vol. 3: 1476–1776*, ed. by Richard Hogg, 56–186. Cambridge: Cambridge University Press.
- MACLEOD, ISEABAIL, and PAULINE CAIRNS (eds.) 1996. *The Scots school dictionary*. Edinburgh: Chambers.
- SAMPSON, GEOFFREY. 2017. *The linguistics delusion*. Sheffield: Equinox.

[sampson@cantab.net]

[Received 5 April 2017;  
revision invited 14 April 2017;  
revision received 19 April 2017;  
accepted 11 May 2017]