Why Is English Spelling So Wierd Weird?

James Harbeck
English is a Caesar’s body

It has had a lot of... *input*... from many different sources.
Old English

Hwæt?
† Old English †

✧ Hwæt! We Gardena in geardagum, þeodcyninga, þrym gefrunon, hu ða æþelingas ellen fremedon.

✧ (Yo! We spear-Danes in yore-days have heard fame of folk-kings, how those noblemen did great deeds.)
Time and tide

Old English: 625~1066
Middle English: 1066~1476
Early Modern English: 1476~1776
Modern English: 1776~last week
† Time and tide †

✦ Letters we stopped using: þ, ð, æ, ð, ʒ

✦ Sounds we stopped making:

✦ /y, Y, x, χ, ç/ spelled y, h, g, ʒ

✦ true length distinctions in vowels and double consonants

✦ Onset sound combinations we stopped using:
hl, hr, hn, cn, gn, wr

✦ Inflectional endings that got worn off: nearly all of them, with mostly just some e and s and es left
† Time and tide †

❖ Sounds we added:
❖ f/v distinction
❖ s/z distinction
❖ n/ŋ distinction
❖ “zh”
❖ Letters we added:
❖ q
❖ z
❖ i/j distinction
❖ v/u distinction
❖ w for p
French

1066 and all that
† French †

✦ Large percentage of modern English vocabulary

✦ *Large percentage of modern English vocabulary*

✦ Much from Norman French

✦ Brought scribes with them

✦ But the words didn’t all come at the same time – we have more recent borrowings as well

✦ How French the word looks depends on how recently we “borrowed” it
Less changed:

- ballet, corps, macaque, cigarette, rendezvous, chateau, petit, etiquette, avant-garde

More changed:

- peasant, money, leaven, castle, petty, ticket, vanguard
In Norman French loans, *ie* usually stood for *e*; while *ei* usually stood for *ɛ*:

- As we will see soon, these became *i* and *e*; with the *ɛ* > *e*; finally becoming *ɛi* or sometimes *i*:

- French gave us:
  - *g* before *e/i* = “*j*” and “*zh*”
  - *c* before *e/i* = “*s*”
  - *c* ≠ “*ch*”, *g* ≠ “*y*”
  - “*z*” and “*v*” as distinct sounds
† French †

✧ Not just words from French; French respellings of existing English words:

✧ hwæt → what
✧ ecg → edge
✧ sinder, is → cinder, ice
✧ mus, mys → mouse, mice
✧ cwen → queen
✧ cyning → king
✧ feond → fiend
Medieval calligraphy!

Lovely, but have you tried reading it?
† Scribes †

- Got paid by the letter
- No standard reference, so they went by what sounded right and by apparent precedent and analogy
- Respelled words for ease of reading and under influence of French
- Worked in large part for lawyers and law courts; established the “Chancery standard”
And then...

I, aye, eye!
The great vowel shift

Like someone just... swirled the long vowels
† The great vowel shift †


していました。

† Mostly completed by the time of Shakespeare

† It’s why we think:

♀ “long a” is not /a:/ but /eɪ/: bat ~ bate
♀ “long e” is not /e:/ but /i/: bet ~ beet
♀ “long i” is not /i:/ but /aɪ/: bit ~ bite
♀ oo is not /o:/ but /u/ or /ʊ/: boot (and how about bout)
### Vowel changes

Data from *Problems in the Origins and Development of the English Language, 5th ed.*, by John Algeo and Carmen Acevedo Butcher

<table>
<thead>
<tr>
<th>OE sound</th>
<th>OE</th>
<th>ME sound</th>
<th>ME</th>
<th>ModE sound</th>
<th>ModE</th>
</tr>
</thead>
<tbody>
<tr>
<td>æː, æːə, ə</td>
<td>clæne, dæl, beam, stelan</td>
<td>ɛː</td>
<td>clene, deel, beem, stelen</td>
<td>iː</td>
<td>clean, deal, beam, steal</td>
</tr>
<tr>
<td>eːː æː, eːə</td>
<td>ges, slæp, seoðan</td>
<td>eː</td>
<td>gees, sleep, sethen</td>
<td>iː</td>
<td>geese, sleep, seethe</td>
</tr>
<tr>
<td>iː, yː</td>
<td>hrim, pif, hydan, mys</td>
<td>iː</td>
<td>rim, wyf, hiden, mys</td>
<td>ar</td>
<td>rime, wife, hide, mice</td>
</tr>
<tr>
<td>OE sound</td>
<td>OE</td>
<td>ME sound</td>
<td>ME</td>
<td>ModE sound</td>
<td>ModE</td>
</tr>
<tr>
<td>----------</td>
<td>---------</td>
<td>----------</td>
<td>----------</td>
<td>------------</td>
<td>-----------</td>
</tr>
<tr>
<td>a:, ɔ</td>
<td>ham, bat, þa, þrote</td>
<td>ɔ:</td>
<td>hoom, boot, wo, throte</td>
<td>ɔ: (ɔʊ, ʌ)</td>
<td>home, boat, woe, throat</td>
</tr>
<tr>
<td>o:</td>
<td>mona, god, flod</td>
<td>o:</td>
<td>mone, good, flood</td>
<td>u:, ʊ, ʌ</td>
<td>moon, good, flood</td>
</tr>
<tr>
<td>u</td>
<td>full, bucca, sunne, cuman</td>
<td>u</td>
<td>ful, bukke, sonne, comen</td>
<td>u, ʌ</td>
<td>full, buck, sun, come</td>
</tr>
<tr>
<td>u:</td>
<td>mus, bru</td>
<td>u:</td>
<td>mous, brow</td>
<td>aw</td>
<td>mouse, brow</td>
</tr>
</tbody>
</table>
Remember those consonants we lost?

Guess what they did to spelling as we passed through Middle English
<table>
<thead>
<tr>
<th>OE sound</th>
<th>OE</th>
<th>ME sound</th>
<th>ME</th>
<th>ModE sound</th>
<th>ModE</th>
</tr>
</thead>
<tbody>
<tr>
<td>æj, ej, æəx</td>
<td>hægl, dæg, segl, peg, eahta</td>
<td>æi</td>
<td>hail, day, seil, wey, eighte</td>
<td>ej</td>
<td>hail, day, sail, way, eight</td>
</tr>
<tr>
<td>æːw, æəw</td>
<td>læpede, feape</td>
<td>æu</td>
<td>lewed, fewe</td>
<td>ju, rw, u:</td>
<td>lewd, few</td>
</tr>
<tr>
<td>iːw, eːə</td>
<td>nipe, Tipesdæg, cneop</td>
<td>iu</td>
<td>newe, Tuesday, knew</td>
<td>ju, rw, u:</td>
<td>new, Tuesday, knew</td>
</tr>
<tr>
<td>aːw, aːɣ, aːx, oːw, cːɣ, oc</td>
<td>snap, agan, dah, gropan, boga, dohtor, brohte</td>
<td>ɔu</td>
<td>snow, Owen, dough, grown, bowe, doughter, broughte</td>
<td>oː (ou, ʌu), ɔ, ə</td>
<td>snow, owe, dough, grow, bow, daughter, brought</td>
</tr>
</tbody>
</table>
The printing press

This is why 1476 is treated as the end of Middle English
Caxton wanted to print a lot of books that would sell so he could make money: poetry, historical & religious prose.

European typesetters came with it, with their ideas:

- gost → ghost
- noticed that silent e’s seemed to be in near-random distribution, and treated them accordingly to help lines fit

Proofing, when it was done, was done by apprentices.

Later printers (e.g., Pynson) hewed more to the Chancery standard.

Pronunciation moved on; print… not so quickly.
Classical influences

Because the Greeks and Romans were just so goshdarn *noble*
† Latin via French †

- creation
- principle, disciple
- leopard
- machine
- actually most words from French because French is descended from Latin
† Latin direct †

✧ pronounce it according to English standards
  ✧ vacuum, ratio, via, excel, aqueous

✧ pronounce it according to some version of Latin standards
  ✧ via, brachial, paternal, aquatic
Greek via Latin

Still look like Greek:
- physic, Terpsichore, psyche, chaos, elephant, phoenix, rhythm, xylophone

Digested all the way through French and/or Old English:
- butter (βούτυρον), chair (καθέδρα), bishop (ἐπίσκοπος), priest (πρεσβύτερος), blame (βλάσφημος), surgeon (χειρουργός), jealous (ζῆλος)
Greek direct

Still heavily influenced by Latin conventions; many are modern confections from old Greek roots

- phytochemical
- psilocybin
- helicopter
- photograph
- pterodactyl
- kinetic

Some borrowed from modern Greek

- bouzouki, ouzo
Pretentious nitwits

“People should know our debt to the classics”
† Pretentious nitwits †

- debt, doubt
- people
- isle, island
- ptarmigan
- foreign
- schism
- indict
- ache
- scissors
- falcon
- adventure
- hiccough
Analogy

“That just looks right”
† “That just looks right” †

- cleave vs. cleft
- schooner (from *scooner/skooner*)
- could (borrowed silent *l* from *should* and *would*)
- sovereign (from *soverain*, by analogy with *reign*)
Economy of effort

Our own lazy tongues, and the inertia of published spellings
† Lazy tongues †

vasive

nwl

swain

stcoat

ndkerchief

cupboard

Greenwich, Worcester, etc.
Spelling reformers

Many have tried. Many have failed, caught on a cot in court.
Noah Webster

Who wanted America to have linguistic independence
† Noah Webster †

- color
- center
- analyze
- traveling
- check
- anemia
- magic, music
Rampant theft, part 1

“Stolen words should look like English”
“Make it look like English”

✧ Respelling to suit us
  ✧ gin, catsup, the Hague, pogue, pirogue, Kilkenny

✧ Old ways of transliterating
  ✧ Genghis Khan, Kahlil Gibran, Hiawatha, Chippewa, Ojibway, Calcutta

✧ Shifts in pronunciation
  ✧ colonel, rival
Rampant theft, part 2

“Stolen words should stay the same”
“Leave the spelling as is”

- Same spelling, new pronunciation
  - yogurt, sauna, ski, Paris, maraschino
- Same spelling, original pronunciation (or close)
  - spaghetti, corps, ballet
- Same spelling, some disagreement about how to say it
  - bruschetta, croissant
Rampant theft, part 3

“Stolen words should not look like English”
“Don’t Anglicize them”

- kombucha
- Ojibwe, Anishnaabe
- Beijing
- Pyeongchang
- Xhosa
- Kolkata
“If it looks English it can’t be right”

- Beijing
- Taj Mahal
- bergamot
- claret
- Kahlil Gibran
- Genghis Khan
Teh interwebz

Cuz its teh lulz
O brave new world...

- pwn all teh n00bs for lulz
- CamelCase
- Xperiments & phreakishness
Orthographical CSI

Who did what damage?
ie/ei

- **seize** (OF *seisir*, ModF *saisir*) vs. **siege** (OF *siege*, ModF *siège*) – both borrowed in the 13\(^{th}\) c.

- **weird** (OE *ƿyrd*, respelled in ME under French influence)

- **friend** (OE *freond* with short *e*, respelled many ways but *friend* stuck) vs. **fiend** (OE *feond* with long *e*)

- **grieve** (Fr *grever*, *i* added in 15\(^{th}\) c) vs. **sieve** (OE *sife*, *e* added in 15\(^{th}\) c)
UGH

- **daughter**: OE *dohtor*, ME *doȝter*
- **laughter**: OE *hleahtor*, ME *lauȝter*
- **plough**: OE *ploh*, ME *plouȝ*
- **through**: OE *þurh*, ME *thorȝ*
- **tough**: OE *toh*, ME *towȝ*
- **cough**: OE *cohhian*, ME *couȝe*

All of these have had many spellings in each time
sch

- **school**: Latin *schola*
- **schooner**: 18th c. *scooner/skooner*, respelled by analogy
- **schizophrenia**: Greek roots
- **schedule**: from Fr *cedule*, from L *scedula* but later L *schedula*
- **schism**: earlier *cism/scism*, from L *schisma*, Gr σχίσμα
- **schnitzel**: German
- **bruschetta**: Italian
**ch**

- **macho**: Spanish
- **mocha**: Arabic, *al-Mukhā* (port in Yemen)
- **machine**: French, from Latin *machina*
- **ach**: Scots and/or Dutch
- **loch**: Gaelic
- **yacht**: early modern Dutch *jaght*
silent l

- **calm** (Fr *calme*) vs. **palm** (L *palma*, Anglo-Norman *paume*) vs. **psalm** (Gr ψαλμός, L *psalmus*, OFr *saume*; has had *many* spellings, including *sphalme* and *salme*)

- **falcon**: OFr *faucon*, L *falco*; the *l* was readded

- **Ralph**: OE *Radulf* from *rad* ‘counsel’ *wulf* ‘wolf’

- **would** (OE *wolde*; the *u* was added 15<sup>th</sup> c), **should** (OE *sceolde*; the *u* was added 15<sup>th</sup> c), **could** (OE *cuþe*, ME *cude*; the *l* was added ~14<sup>th</sup> c by analogy)
silent b

- **dumb**: OE *dumb*
- **tomb**: Anglo-Norman *toumbe*, OFr *tombe*; also spelled *tome* in ME
- **bomb**: Fr *bombe*; ME also spelled *bome*, then added back the *b*; formerly often said like “bum”
- **comb**: OE *camb*
- **limb**: OE *lim*; the *b* was added 15th c
- **debt**: ME *det*, OFr *dette*, L *debitum*; *b* added back in late ME/early ModE
other silent consonants

- **receipt**: ME receite, OFr recete, from L recepta; p ultimately retained, unlike in *conceit* and *deceit* (but cf. *conception*, *deception*)

- **indict**: ME endite, OFr enditer, ultimately to L indicere and/or indicare with spurious conjectural *indictare*; the c was inserted in the 16th c

- **limn** (OFr lumine, L luminare)

- **damn** (ME dampne, OFr dampner, L damnare), *condemn* (ME condempne, OFr condemnner, L conem(p)nare)
other silent consonants

- **aisle**: ME *eile*, L *ala*; in Fr the *s* indicates vowel quality, but in Eng could just be by analogy

- **isle**: ME *ile*, OF *ile*, L *insula*; *s* added back in during English Renaissance

- **island**: OE *iglond*, ME *yland*, *iland*; *s* added by mistake on model of *isle* in 17th c
wor

- **worse**: OE *pirsa*, ME *wurse*, *worse*
- **word**: OE *peord*, ME *wourd*
- **work**: OE *peorc*, *porc*
- **worth**: OE *peorþ*
- **worn**: was *perede* in OE and until late ME, when it became *woryn* and *worne* by analogy with *sworn*, *torn*
- **wort**: OE *pyrt*, ME *wurte*, *worte*
- **wart**: OE *pearte*
more vowel movements

**one**: OE *an*, *a*, ME *an*, *a*, *on*, *un*, *won*; *aːn* ➔ *oːn* is like *stan* ➔ *stone*, but *oːn* moved further to *uːn* and *un* and then got an on-glide

**two**: OE *tpa*, *tu*, *to*, ME *to*, *two*; *twa*: ➔ *two*: ➔ *twu*: ➔ *tu*:

**sew**: OE *seowan*, *siwian*, ME *sewe*, *sowe*; the vowel shift is unusual, and modern Scots pronunciation of *sew* is like *shoe*

**shoe**: OE *scoh*, *sceo*, ME *scho*, *schoo*; blame it on the great vowel shift, just like *to* and *too* (the *e* indicates length)
blame the GVS

- **bass** (musical): ME bas, long a, respelled as base except in music, where it’s bass by analogy with Italian basso

- **eye**: OE æge, eage, ME eʒe; æj → i: → aɪ

- **river** (Anglo-Norman rivere, OF riviere) vs. **rival** (OF rival): a difference in vowel length before the GVS

- **some** (OE sum, ME soume) vs. **home** (OE ham; a: → o: in southern England but not Scotland; and then the GVS)

- **gouge** (Fr gouge, u: → ow) vs. **gauge** (OF gauge, a: → eɪ)
others

✦ **busy**: OE *bisig*, *bysig*; the *u* spelling based on *y* has won but the pronunciation based on *i* has won, thanks to deliberate efforts by orthoepists in the 16th c

✦ **stranger** (OF *estrangier*) vs. **anger** (Old Norse *angr*, OE anger)

✦ **find** (OE *findan*), **bind** (OE *bindan*), **mind** (OE *mynd*), **wind** (OE *windan*) vs. **wind** (OE *wind*, which would in the normal course of things be said like the verb, but in the 18th the current pronunciation prevailed because of *windmill* and *windy*, which always had short vowels)
james@harbeck.ca

Email me if you want a copy of this