

What is regular, and what is an exception?

(1) What we have seen so far:

- Phonological irregularity is a property of words (and not solely a property of morphemes)
 - Some form of lexical listing is required
- Need a mechanism that allows exceptional words to surface, but still retains crucial $\mathcal{M} \gg \mathcal{F}$ rankings to enforce regular pattern on rare and novel words
 - Gradient faithfulness can do this, in a way: $\mathcal{F}_{\text{high}} \gg \mathcal{M} \gg \mathcal{F}_{\text{low}}$
 - Predicts that high frequency words may follow any pattern (exceptional or not)¹
 - At some threshold of low frequency, words are unable to retain irregularity, and follow the regular pattern (whatever that may be)
- A puzzle: what is the regular pattern?
 - A logical expectation: if most words in the language follow a pattern, it should be learned and extended to new words
 - The surprising effect from English: trisyllabic shortening not as productive as one might imagine, given its consistency with existing words of the relevant type
 - Why are learners failing to learn TSS, in spite of apparently abundant evidence?

(2) Goal:

- Show why TSS generalization may be largely inaccessible to learners
- Preview: it is linked non-coincidentally to the fact that the suffixes that condition it are themselves non-productive (level 1)
- Learning goes beyond pattern matching; learners try to determine what the target grammar looks like, and whether they need to modify their current grammar to achieve the target

A brief history of TSS

(3) A brief history of trisyllabic shortening (Lahiri & Fikkert 1999)

- At one time, a relatively productive process in English (both inherited & borrowed words)
 - Long vowel \rightarrow short / ___ $\sigma \sigma$
- Caused alternations within inflectional paradigms, derived forms, and compounds

clōver	clāvere	‘clover’ (nom. sg. ~ pl.)
hēafod	hēafodu	‘head’ (nom. sg. ~ pl.)
- Some frozen relics of TSS in native derived forms and compounds

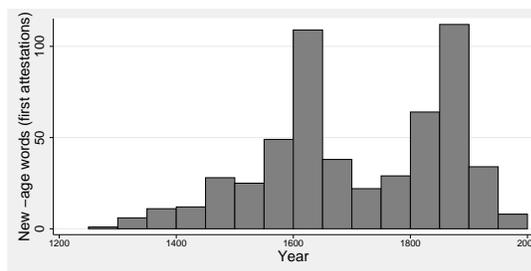
s[au]th (< sūþ)	s[ʌ]thern (< sūðerne)
h[ou]ly	h[a]liday
wh[ai]te	Wh[ɪ]taker (‘white field’)
dear (< dēor)	darling (< dēorlingas) (w/additional change caused by [r])
- Alternations were subsequently leveled in:
 - Inflected words (*clavere* \Rightarrow *clōvere*)
 - Words derived with certain suffixes (roughly, native, or level 2)
 - E.g., older *estern*~*estrin* \Rightarrow newer *eastern*

a1300 Cursor M. 11388 A prophet of **estrin**land, hight Balaam, crafti and bald.

¹Actually, once learnability is considered, we see that if all high frequency words followed the “exceptional” pattern, then the learner may not have any basis for learning the lower ranking. Thus, we may actually predict that this configuration can arise only in case the exceptional pattern is “truly exceptional”; more work is needed to explore this issue.

- (Gradually, later) leveled in a handful of borrowed (Level 1) formations, such as *obesity*
 - Another example: *mundanity*
 - OED says: UK [mændænitɪ] ~ [mændemɪtɪ], US [mændemɪti]
 - In other words, “exceptions” like *obesity* are in a sense part of a long-term trend in English; the challenge is to understand why the alternation is left intact in a certain set of forms
- (4) Why didn't borrowed forms regularize?
- One possibility: the alternation was so robust among this set of words that it was retained, even after it was lost elsewhere
 - Probably not the right explanation: the alternation was robust in the entire language, not just among this subset of words
 - Another possibility: borrowed affixes are not productive, and Level 1 formations must be listed as exceptions anyway
 - In other words, Level 1 formations behave as if they are monomorphemic, and don't participate in paradigms of their (apparent) bases
- (5) The link between productivity and leveling: three borrowed affixes
- *-ity*: relatively unproductive, TSS largely intact
 - *-((a)c)y*: somewhat more productive, many exceptions to TSS
 - *-age*: quite productive (esp. in Early Modern English), only a few TSS forms (*lineage*), while most resist (*brokerage*, *foliage*, etc.)

Dates of first attestation in the OED:



1639 WHATELEY Prototypes I. xix. (1640) 222 *These gracelesse young men think the threats but words of sport, counterfeit words which have no truth nor substance in them, but were very mockery and scoffage.*

1673 Phil. Trans. VIII. 5194 *In their [the snow flakes'] continual motion and waftage to and fro touching upon each other.*

- (6) The scenario that this suggests
- Overall increasing pressure to level within paradigms; trisyllabic shortening still active outside paradigms, but eliminated within paradigms
 - Forms that are derived with unproductive morphology don't act like members of paradigms
 - Thus, unproductively derived forms are simply lexicalized, and are exempt from leveling
- (7) Getting our theory of exceptions to capture this intuition
- We need to understand why words like *serenity* gain independence from their bases, even though they bear a clear relation (semantic, morphological, and phonological)
 - The reason for this is by now obvious:
 - If the word derived by unproductive morphology, then it could not be re-created on-line, and must be stored
 - Storing a word involves remembering both its morphological and phonological form
 - Thus, morphological irregularity can lead to phonological irregularity²

²Burzio (2002) pursues the opposite line: morphological irregularity is correlated with phonological regularity. More work is needed to understand the relation between these claims.

- (8) So how do learners decide whether or not a word needs to be lexicalized?
- Zuraw (2000): involves reasoning about relative likelihood that speaker could have produced the form on-line using their grammar, vs. likelihood that it was simply a listed form
 - Every time a word is heard, learner asks “What is the likelihood that my interlocutor synthesized that form with their grammar?”
 - If it seems likely that the speaker used their grammar, but the learner’s grammar doesn’t predict it, then some reranking is needed
 - If it seems likely that the speaker used a listed (lexicalized) form, then the learner needs to list the form, too

- (9) Some possibilities that the learner might consider, on hearing [səɾɛnɪti]

- The speaker’s input was /səɾɛnɪti/
- The speaker’s input was /səɾi:nɪti/
- The speaker’s input was /səɾi:n + ɪti/

A difficult inference: what is the probability that the speaker intended /səɾi:n + NOMINALIZ./, given that I just heard [səɾɛnɪti]?

- (10) Reducing hard questions to easier questions:

- It’s hard to know what the speaker really intended (i.e., what the input was), but it’s easier to know what you yourself would do in the same situation
- So, could at least check which of the options is most likely under your own current grammar

Example:

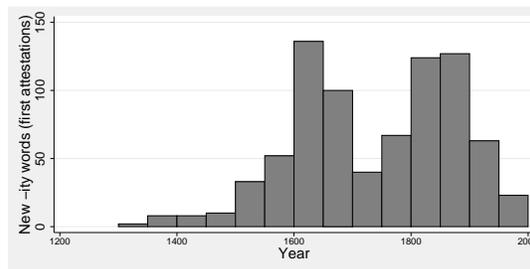
- What is the probability that the speaker used input /səɾɛnɪti/?
 - That depends on: (1) what is the likelihood that there is a listed form /səɾɛnɪti/, and if so, (2) what is the likelihood that it would be pronounced [səɾɛnɪti]?
- What is the probability that the speaker used input /səɾi:nɪti/?
 - That depends on: (1) what is the likelihood that there is a listed form /səɾi:nɪti/, and if so, (2) what is the likelihood that it would be pronounced [səɾɛnɪti]?
- What is the probability that the speaker used input /səɾi:nɪti/?
 - That depends on: (1) what is the likelihood that speaker combined morphemes /səɾi:n/ and /-ɪti/, and if they did, (2) what is the likelihood that it would be pronounced [səɾɛnɪti]?

(This is a form of Bayesian inference; we won’t go into the formalism, but the intuition should be clear even without it)

- (11) Ruling out the morphologically complex analysis /səɾi:n + ɪti/:

- The pronunciation [səɾɛnɪti] is fully compatible with this hypothesis
- However, it would be relatively unlikely that the speaker would have created the word by productive combination of /səɾi:n + ɪti/
- Why is this unlikely? Certainly there are plenty of *-ity* words; why don’t they make it seem like a plausible formation?
 - Profile of new *-ity* words entering language

On the face of it, chronology of neologisms is quite similar to *-age*:



- However, comparing the first 20 new words to enter after 1500 for each suffix reveals a substantial difference:

shewage, tyage, saveage, marinage, schoolage, cellarage, bailliage, stirrage, winage, metage, advowsonage, farmage, endowage, lettage, mastage, tribulage, orphanage, pollage, brewage, shootage	mundanity, facility, callidity, cardinality, consubstantiality, absurdity, aquosity, calidity, cecity, improsperity, facundity, concinnity, equability, frugality, implacability, carnosity, miserity, imperiality, morosity, muchity
--	---

- Novel *-age* forms are sort of ordinary (and typically on free-standing roots—including many native ones); novel *-ity* forms are “aureate” (learned borrowings/creations)
- Furthermore, *-ity* nouns tended to be in competition with zero-derived deadjectival nouns (the more productive pattern at that time)
 - 1610 Histrio-mastix II. 264 *Our heavenly poesie, That sacred off-spring from the braine of Jove, Thus to be mangled with prophane **absurds**.*
 - 1615 CHAPMAN Odyss. XXII. 585 *That both on my head pour'd **depraves** unjust, And on my mother's, scandalling the court.*
 - 1628 FELTHAM Resolves I. lii. Wks. (1677) 84 *The power of the Gospel, in crying down the **vains** of men.*
 - 1760-72 H. BROOKE Fool of Qual. (1859) I. 220 *No more than ye can see the gloom of last winter in the smiling **serene** of a summer's evening.*
 - 1778 WOLCOTT (P. Pindar) Ep. to Reviewers ix, *I never question'd your **profound** of head.*
- The suffix *-age*, on the other hand, was deverbal/denominal,
 - *saveage, tyage*, etc. = act of X (zero derivation from verb tends to mean “result of X”)
 - 1546-7 in Leland Collect. IV. 320 *Take Bow and Shaft in Hand, learn **Shewtage** to frame.*
 - 1545 R. ASCHAM Toxoph. II. 107 *For in a rayne and at no marke, a man may shote a faire **shoote**.*
 - *swannage* = money you pay for the right to own swans
 - 1610 W. FOLKINGHAM Art of Survey III. iv. 70 *Wrecks, **Swannage**, Warrenage, Commonage, Piscage.*

(12) Putting this together:

- If you heard a novel *-ity* form in 1600, you would probably not have thought “I’ve never heard that, but I probably would have said it the same way”
- Novel *-age* forms, on the other hand, might have been quite plausibly productive
- Thus, probability of compositional interpretation would have been less for *-ity* form

(13) What this means for TSS:

- Particular set of suffixed forms (now Level 1 affixes) were exempt from leveling; TSS alternations remained in them
- However, the same factors that make them immune from leveling also mean that they are less informative about the alternation itself
 - The best data for an alternation is two inflectionally related forms; or, at least, two forms that are completely unambiguously derived from the same root
- Thus, although the alternation is statistically strong (applies in many or most of the cases where it “should” apply), these words may not inform phonological learning

(14) Moral:

- Determining whether a pattern is learned (=grammatically encoded) and extended to new/rare words requires a theory of learning
- It is not simply a matter of counting forms and determining what the dominant pattern is
 - Which forms to count? Which patterns count as competition?
- Cases like English TSS provide an important key to the solution
 - Robust patterns in the lexicon are not applied as productively as one might expect
 - Sometimes taken as evidence that learners are unable to construct grammars elaborate enough to capture them; this cannot be right (plenty of evidence that speakers notice lots of generalizations and patterns)
 - Key is to understand why some patterns are “inaccessible” to the learner