Bad Writing

developing a complex orthography:
the tale of writing in Zevy

Kenan Kigunda · digital Language Construction Conference · March 7, 2021

A quick history of Zevy writing

the symbols of zevy writing

early zevy writing: one-to-one sound correspondence

"a boring shoe in a brown box"

brown	box	in	boring	shoe

t e i s u b e e t m e d a k a v e r v o d

early zevy writing: one-to-one sound correspondence

"a boring shoe in a brown box"

brown	box	in	boring	shoe

t e i s u b e e t m e d a k a v e r v o d

t e i z u

d a k k a

A scribe in a different dialect might write the same words differently (just as they speak them differently)

early zevy writing: one-to-one sound correspondence

"a boring shoe in a brown box"

brown	box	in	boring	shoe

t e i s u b e e t m e d a k a v e r v o d

t e i z u

d a k k a

A scribe in a different dialect might write the same words differently (just as they speak them differently)

"in the square under the circle"

circle	under	square	in

k ou s uu n b o b u m e

modern zevy writing: complex sound correspondence

fossilized affixes   ligatures

sound changes

$\widehat{ts}$
t ei s un be t$^{\theta}$ d a k a v e r v o d^{ð}
əɪ   z

spatial abbreviations   me

"a boring shoe in a brown box"

homophone disambiguation

əʊ
k u u b o bu
unstressed vowel deletion   me

"in the square under the circle"

Deep vs shallow orthographies
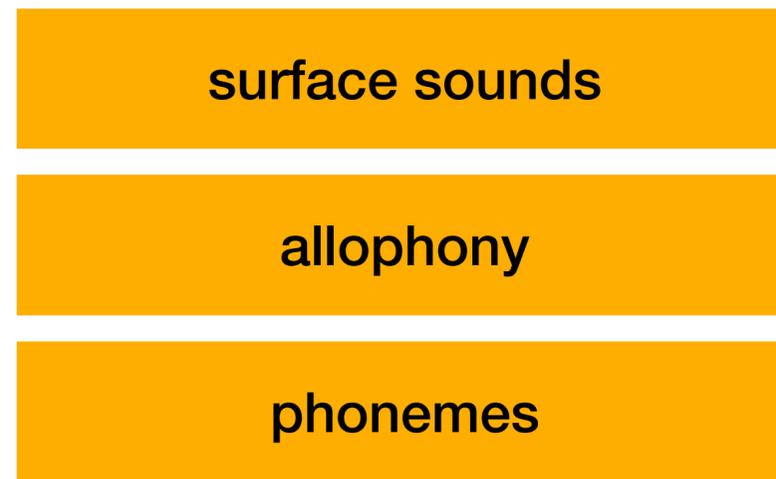
surface sounds
allophony
phonemes

Shallow orthographies
tend to represent
sounds more directly

e.g. Swahili

Deep vs shallow orthographies

surface sounds

allophony

phonemes ◀━

word-by-word memorization

morphophonolgy

sound changes

source languages

e.g. English, French

Deep orthographies tend to pull from some combination of these
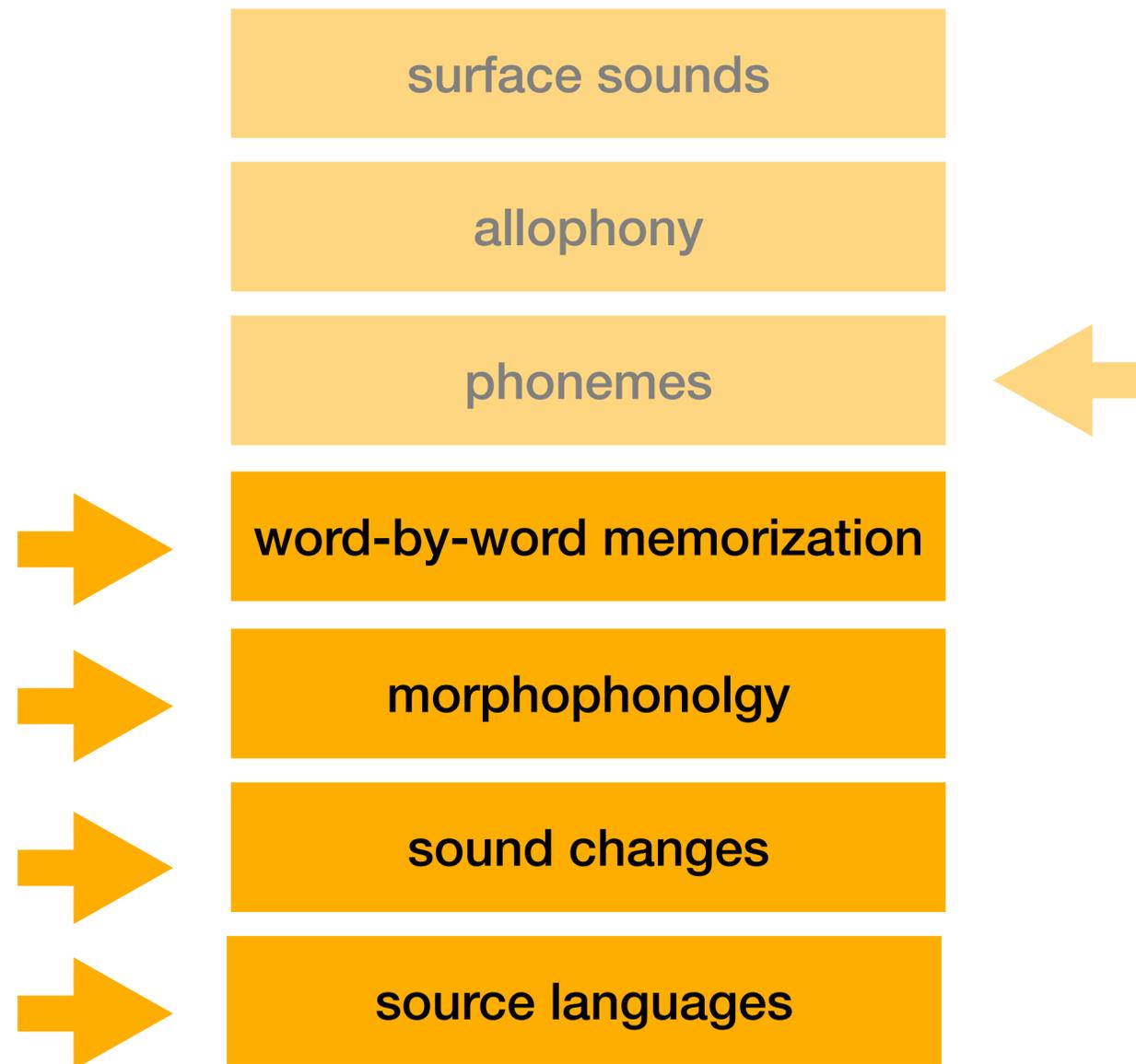
Shallow orthographies tend to represent sounds more directly

e.g. Swahili

Deep vs shallow orthographies
we have both in english! sort of…
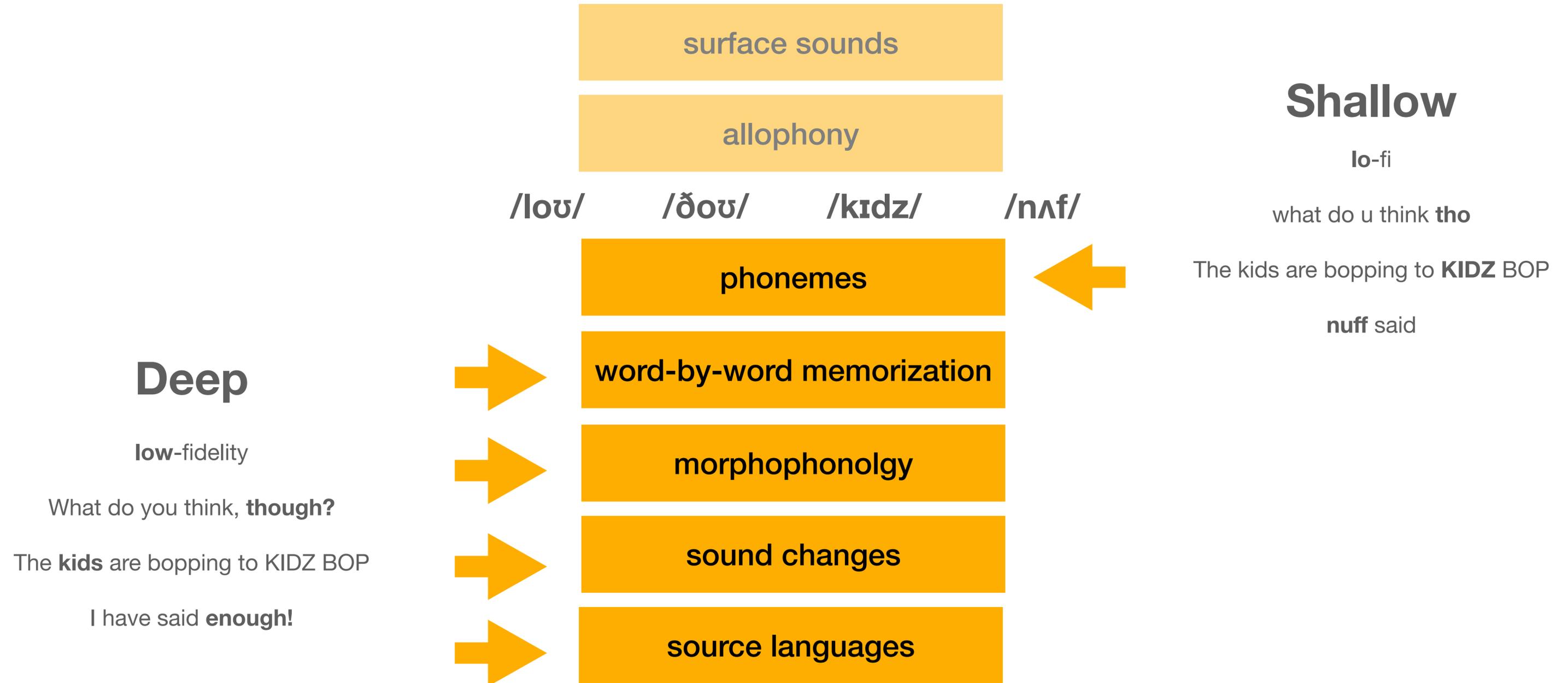
| surface sounds |
| allophony |

/loʊ/ /ðoʊ/ /kɪdz/ /nʌf/

phonemes

word-by-word memorization

morphophonolgy

sound changes

source languages

Shallow

lo-fi

what do u think **tho**

The kids are bopping to **KIDZ** BOP

nuff said

Deep

low-fidelity

What do you think, **though?**

The **kids** are bopping to KIDZ BOP

I have said **enough!**

Three principles of deep orthographies

1. Don't get it write the first time

2. Standardize. Standardise. Standardize.

3. Watch words drift as you refuse to let go of the passed

Principle one:
Don't get it write the first time

Don't get it write the first time

- Not everything that's spoken needs to be written down
- Humans are great at noise correction!
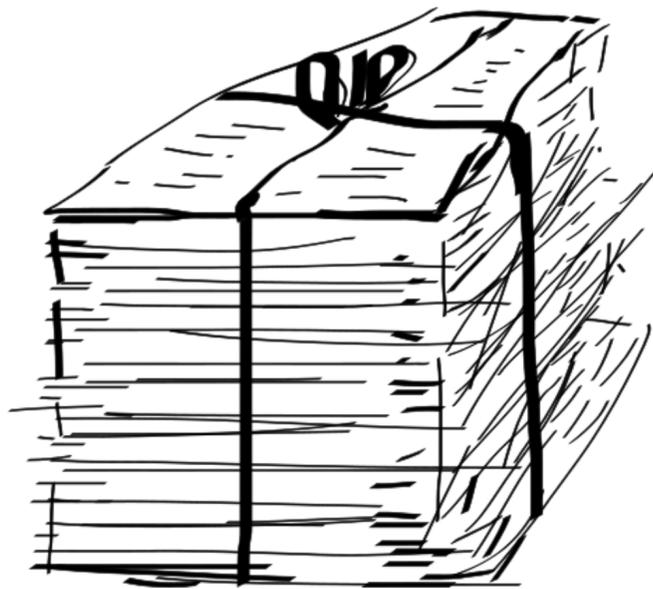 - Even in speech, listeners can "hear" sounds that were lost

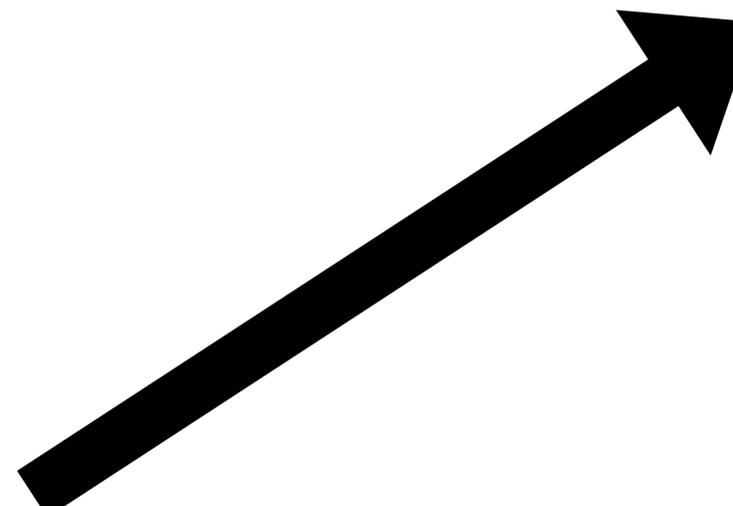it's all thanks to me, the phoneme restoration ghost

- Writing has the further advantage that the reader can reread!

Don't get it write the first time

- Writers are lazy!
- *Especially* when writing by hand
  - Jotting down quick notes?
  - Transcribing long texts?

Work less, relax more 😎

Yikes! Take shortcuts!

Don't get it write the first time

- Writers are lazy, less is more
- This creates a tension between the writer and the reader that may be familiar:

Ease of articulation
Speech: easy to say
Writing: easy to transcribe

Ease of perception
Speech: easy to hear
Writing: easy to read

Don't get it write the first time
Examples in natural languages

- Abjads!

Can we skip the vowels? I bet the reader will do just fine!

Sure! Works well enough for me

Don't get it write the first time
How this played out in early Zevy

Don't get it write the first time
How this played out in early Zevy

1. Spaces? What spaces?

eneu keki

"special pair"

ene ukeke

"famous coat"

Don't get it write the first time
How this played out in early Zevy

1. Spaces? What spaces?

2. Ligatures galore, even across words! Because… no spaces, remember?

eneu keki

"special pair"

ene ukeke

"famous coat"

all the vowel ligatures!

Many of these only occur across word* boundaries, rather than inside a word, but all are common because ligatures can span words

*Technically speaking, morpheme

all the vowel ligatures!

	a	e	i	u	o
a	aa	ae	ai	au	ao
e	ea	ee	ei	eu	eo
i	ia	ie	ii	iu	io
u	ua	ue	ui	uu	uo
o	oa	oe	oi	ou	oo

only these two remain unjoined!

vowel ligatures!

For most ligatures, the position of the second character shows whether it comes first or second

vowel ligatures!

For most ligatures, the position of the second character shows whether it comes first or second

But:

The tail of the "u" character always comes first

au

ua

uo

oa

ou

vowel ligatures!

For most ligatures, the position of the second character shows whether it comes first or second

But:

The tail of the "o" character always comes last

vowel ligatures!

For most ligatures, the position of the second character shows whether it comes first or second

But:

The tail of the "u" character always comes first

Taken together, these rules create three pairs of sounds that are each spelled identically

The tail of the "o" character always comes last

Don't get it write the first time
How this played out in early Zevy

1. Spaces? What spaces?

2. Ligatures galore, even across words! Because… no spaces, remember?

3. Let's keep the tails consistent, not the sound order

atu at
"instantly go"

atau
"your son"

Don't get it write the first time
How this played out in early Zevy

1. Spaces? What spaces?

2. Ligatures galore, even across words! Because… no spaces, remember?

3. Let's keep the tails consistent, not the sound order

atu at
"instantly go"

↓

atuat, a tuat, at uat, atu at, atua t
ataut, a taut, at aut, atau t

atau
"your son"

↓

atua, a tua, at ua, atu a
atau, a tau, at au, ata u

Don't get it write the first time
How this played out in early Zevy

1. Spaces? What spaces?

2. Ligatures galore, even across words! Because… no spaces, remember?

3. Let's keep the tails consistent, not the sound order

4. While we're at it, let's not write long vowels when we don't have to

eneu keki
"special pair"

ene ukeke
"famous coat"

Don't get it write the first time
How this played out in early Zevy

e n eu k e k i

e n eu k e k e

All of these are long vowels, but only word-internal are marked
This is because vowels are always long at the end of a word

t e i s u b ee t m e d a k a v e r v o d

Don't get it write the first time
How this played out in early Zevy

1. Spaces? What spaces?
2. Ligatures galore, even across words! Because… no spaces, remember?
3. Let's keep the tails consistent, not the sound order
4. While we're at it, let's not write long vowels when we don't have to
5. Look, we can save time by abbreviating postpositions!

kou suun bobu me
"in the square under the circle"
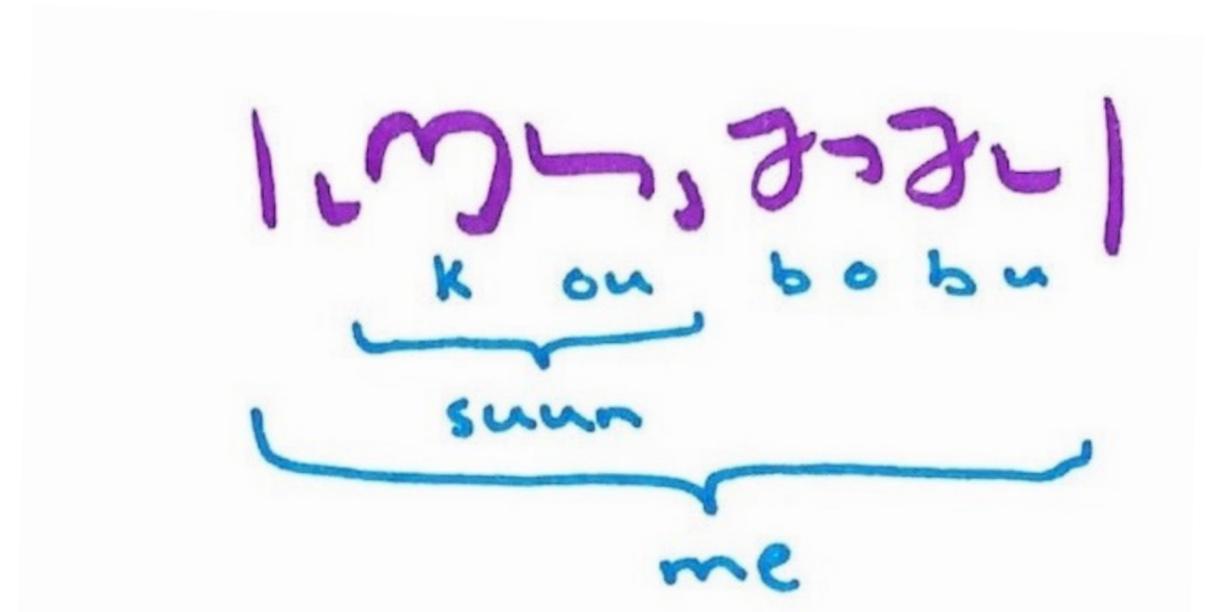
spatial abbreviations!

⸜(ー)⸝
- suun

under -
below -
bottom of -

(ー‿)
- muun

on -
above -
top of -

)ー(
- teen

to -
towards -
go to -
-going
-bound

}ー⊦
- deeden

from -
come from -

|ー|
- me

in -
inside -
part of -

⸂ー⸃
- ni

with -

spatial abbreviations!

Problem

Solution

Knowing how much space you need for the opening marker *ahead of time* is hard

Only write the first opening marker; all other markers have only the closing marker

spatial abbreviations, modified

- suun

under -
below -
bottom of -

- muun

on -
above -
top of -

- teen

to -
towards -
go to -
-going
-bound

- deeden

from -
come from -

- me

in -
inside -
part of -

- ni

with -

spatial abbreviations, modified

- suun

under -
below -
bottom of -

- muun

on -
above -
top of -

- teen

to -
towards -
go to -
-going
-bound

- deeden

from -
come from -

- me

in -
inside -
part of -

- ni

with -

uh oh

spatial abbreviations, modified

- suun

under -
below -
bottom of -

- muun

on -
above -
top of -

- teen

to -
towards -
go to -
-going
-bound

- deeden

from -
come from -

- me

in -
inside -
part of -

- ni

with -

phew

Don't get it write the first time
How this played out in early Zevy

1. Spaces? What spaces?
2. Ligatures galore, even across words! Because… no spaces, remember?
3. Let's keep the tails consistent, not the sound order
4. While we're at it, let's not write long vowels when we don't have to
5. Look, we can save time by abbreviating spatial relations!

=

Lots of complexity built up very early on!

Three principles of deep orthographies

1. Don't get it write the first time ✅

2. Standardize. Standardise. Standardize.

3. Watch words drift as you refuse to let go of the passed

Principle two:

Standardize.
Standardise.
Standardize.

Standardise

- Individual scribes -> Printing press = Standards easier to enforce
- But… by whom?
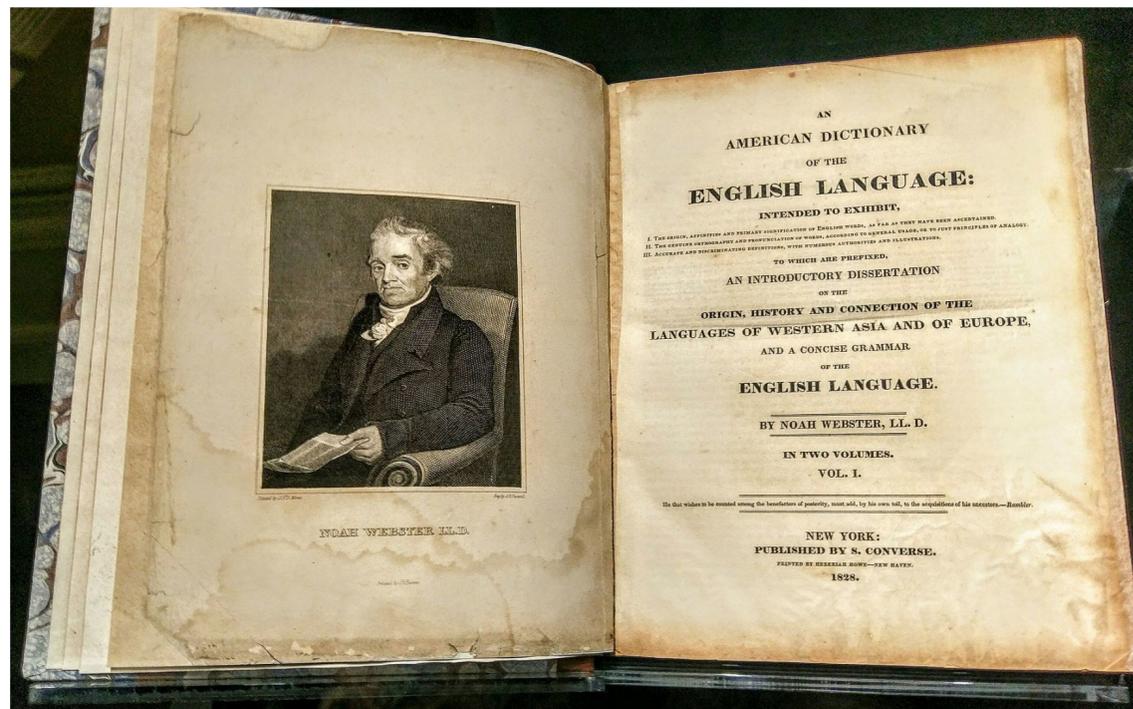
Cullen328, CC BY-SA 3.0 <https://creativecommons.org/licenses/by-sa/3.0>, via Wikimedia Commons

By Samuel Johnson - Eighteenth Century Collections Online, Public Domain, https://commons.wikimedia.org/w/index.php?curid=3334877
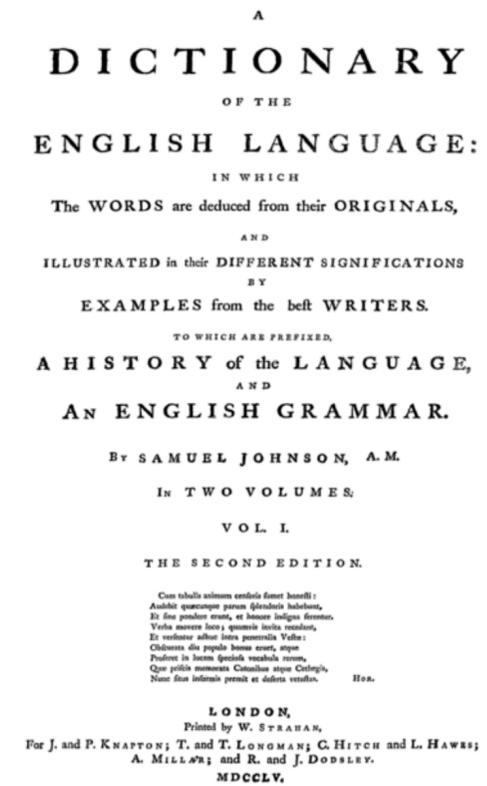
Standardize
US vs UK (or both)

- *-or* (directly from Latin) vs *-our* (through Old French, now *-eur* in Modern French)

It's an hono**o(u)r** to make your acquaintance on this col**o(u)r**ful day, but I'm afraid I don't know the hon**or**ific for one of your level of glam**our**

Standardise
US vs UK (or both)

- *-re* (from French, Latin, or Greek) vs *-er* (matching the pronunciation)

The met**er** measured the met**re**s it would take to get to the theat**er** to perform theat**re**

Standardize
US vs UK (or both)

- *-ize* (directly from Greek) vs *-ise* (through French)

I real**ize** that you must anal**yse** the way you want to... improv**ize**?

Standardise
How this played out in early Zevy

the cultural center of the Zevy linguosphere

VS

the administrative capital of the Zevy political sphere

↓

the **Bemi** dialect
"Literary"

↓

the **Kuuvi** dialect
"Capital"

MATCH AT 5 O'CLOCK! GET YOUR TICKETS NOW

Standardise
How this played out in early Zevy

the **Bemi** dialect
"Literary"

VS

the **Kuuvi** dialect
"Capital"

STARRING IN

The Great Debacle of /oi/ and /ou/

Standardize

The Great Debacle of /oi/ and /ou/

starting context

- The diphthong /ui/ existed in Old Zevy, but merged with /oi/ by Middle Zevy

- Synchronous derivations and inflections that would create /ui/ were also spoken as /oi/ instead

- Before the printing press, scribes varied in how they would render this in writing according to personal preference

Standardize
The Great Debacle of /oi/ and /ou/

the **Bemi** dialect
"Literary"

VS

the **Kuuvi** dialect
"Capital"

- The Matrama publishing house set the direction when they made a decision to always respect etymology, using **oi** and **ui**

Standardize
The Great Debacle of /oi/ and /ou/

the **Bemi** dialect
"Literary"

- The Matrama publishing house set the direction when they made a decision to always respect etymology, using **oi** and **ui**

VS

the **Kuuvi** dialect
"Capital"

- Early iterations of the government drafting style was shifting in the opposite direction, using **oi** consistently

Standardize
The Great Debacle of /oi/ and /ou/

the **Bemi** dialect
"Literary"

- The Matrama publishing house set the direction when they made a decision to always respect etymology, using **oi** and **ui**

vs

the **Kuuvi** dialect
"Capital"

- Early iterations of the government drafting style was shifting in the opposite direction, using **oi** consistently

- An exception! The **ui** ligature could be used to distinguish words which were homophones, such as **zui** "victory" vs **zoi** "sight" (both pronounced /zoi/)

Standardize
The Great Debacle of /oi/ and /ou/

the **Bemi** dialect
"Literary"

VS

the **Kuuvi** dialect
"Capital"

- The Matrama publishing house set the direction when they made a decision to always respect etymology, using **oi** and **ui**

- Early iterations of the government drafting style was shifting in the opposite direction, using **oi** consistently

- An exception! The **ui** ligature could be used to distinguish words which were homophones, such as **zui** "victory" vs **zoi** "sight" (both pronounced /zoi/)

- Similarly, they disambiguated **kou** "space" from **kuu** "circle"

Standardize
The Great Debacle of /oi/ and /ou/

the **Bemi** dialect
"Literary"

VS

the **Kuuvi** dialect
"Capital"

- The Matrama publishing house set the direction when they made a decision to always respect etymology, using **oi** and **ui**

- Since there was no u in the etymology of any of these words, Matrama resisted this change, preferring homonyms

- Early iterations of the government drafting style was shifting in the opposite direction, using **oi** consistently

- An exception! The **ui** ligature could be used to distinguish words which were homophones, such as **zui** "victory" vs **zoi** "sight" (both pronounced /zoi/)

- Similarly, they disambiguated **kou** "space" from **kuu** "circle"

Standardize
The Great Debacle of /oi/ and /ou/

the **Bemi** dialect
"Literary"

VS

the **Kuuvi** dialect
"Capital"

- The Matrama publishing house set the direction when they made a decision to always respect etymology, using **oi** and **ui**

- Since there was no u in the etymology of any of these words, Matrama resisted this change, preferring homonyms

- Moreover, due to a different suffix, Bemi speech already had **koun** "space" vs **kou** "circle"

- Early iterations of the government drafting style was shifting in the opposite direction, using **oi** consistently

- An exception! The **ui** ligature could be used to distinguish words which were homophones, such as **zui** "victory" vs **zoi** "sight" (both pronounced /zoi/)

- Similarly, they disambiguated **kou** "space" from **kuu** "circle"

Standardize
The Great Debacle of /oi/ and /ou/

the **Bemi** dialect
"Literary"

VS

the **Kuuvi** dialect
"Capital"

- The Matrama publishing house set the direction when they made a decision to always respect etymology, using **oi** and **ui**

- Since there was no u in the etymology of any of these words, Matrama resisted this change, preferring homonyms

- Moreover, due to a different suffix, Bemi speech already had **koun** "space" vs **kou** "circle"

- The clout of the publishing houses ultimately ended up winning the battle of etymology

- Early iterations of the government drafting style was shifting in the opposite direction, using **oi** consistently

- An exception! The **ui** ligature could be used to distinguish words which were homophones, such as **zui** "victory" vs **zoi** "sight" (both pronounced /zoi/)

- Similarly, they disambiguated **kou** "space" from **kuu** "circle"

Standardize

The Great Debacle of /oi/ and /ou/

the **Bemi** dialect
"Literary"

vs

the **Kuuvi** dialect
"Capital"

- The Matrama publishing house set the direction when they made a decision to always respect etymology, using **oi** and **ui**

- Since there was no u in the etymology of any of these words, Matrama resisted this change, preferring homonyms

- Moreover, Bemi speech already had **koun** "space" vs **kou** "circle"

- The clout of the publishing houses ultimately ended up winning the battle of etymology

- Early iterations of the government drafting style was shifting in the opposite direction, using **oi** consistently

- An exception! The **ui** ligature could be used to distinguish words which were homophones, such as **zui** "victory" vs **zoi** "sight" (both pronounced /zoi/)

- Similarly, they disambiguated **kou** "space" from **kuu** "circle"

- But through the education system, the capital added and enforced its disambiguating changes

Standardize
The Great Debacle of /oi/ and /ou/

the **Bemi** dialect
"Literary"

🤝

the **Kuuvi** dialect
"Capital"

Standard Written Zevy

- Today, Zevy exists in a situation of light diglossia

- The two dialects remain prominent and distinct in spoken language, and are roughly 90 percent mutually intelligible

- Meanwhile, there is a single written standard used throughout the Zevy linguosphere

Standardize
The Great Debacle of /oi/ and /ou/

the **Bemi** dialect
"Literary"

🤝
Standard Written Zevy

the **Kuuvi** dialect
"Capital"

- "fish" as **nui**, due to an etymological **u**, despite the fact that it had no homophones
  - a remnant of the transitional Literary standard

/noi/

Standardize

The Great Debacle of /oi/ and /ou/

the **Bemi** dialect
"Literary"

🤝

the **Kuuvi** dialect
"Capital"

Standard Written Zevy

- /zoi/ "victory" as **zui**, to disambiguate it from /zoi/ **zoi** "sight", despite the fact that it had no etymological **u**

  - a remnant of the transitional Capital standard

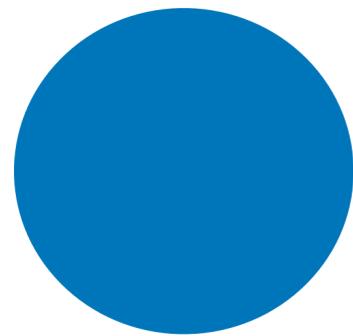/zoi/

Standardize

The Great Debacle of /oi/ and /ou/

the **Bemi** dialect
"Literary"

🤝
Standard Written Zevy

the **Kuuvi** dialect
"Capital"
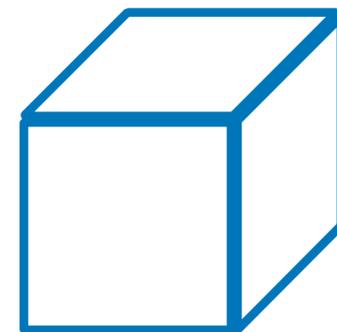
- "circle" as **kuu**, which ironically became redundant as **kou** "space" was eventually supplanted by **koun** in both dialects
  - a remnant of the interaction between the two transitional standards

/kou/

/koun/

Standardize
fossilizing affixes

the **Bemi** dialect
"Literary"

🤝
Standard Written Zevy

the **Kuuvi** dialect
"Capital"

- other words like **teisun** "brown" retained the -un suffix of the Bemi dialect, even though the Kuuvi dialect reduced this to /u/ in speech

/teisun/

/teisu/

Standardize
ganging up on the smaller dialects

the **Bemi** dialect
"Literary"

🤝
Standard Written Zevy

the **Kuuvi** dialect
"Capital"

- long vowels, which existed in both major dialects, were written, but long *consonants*, which existed only in smaller dialects, were not written

/daka/ ✔️

Bemi & Kuuvi

/dakka/ ❌

everyone else
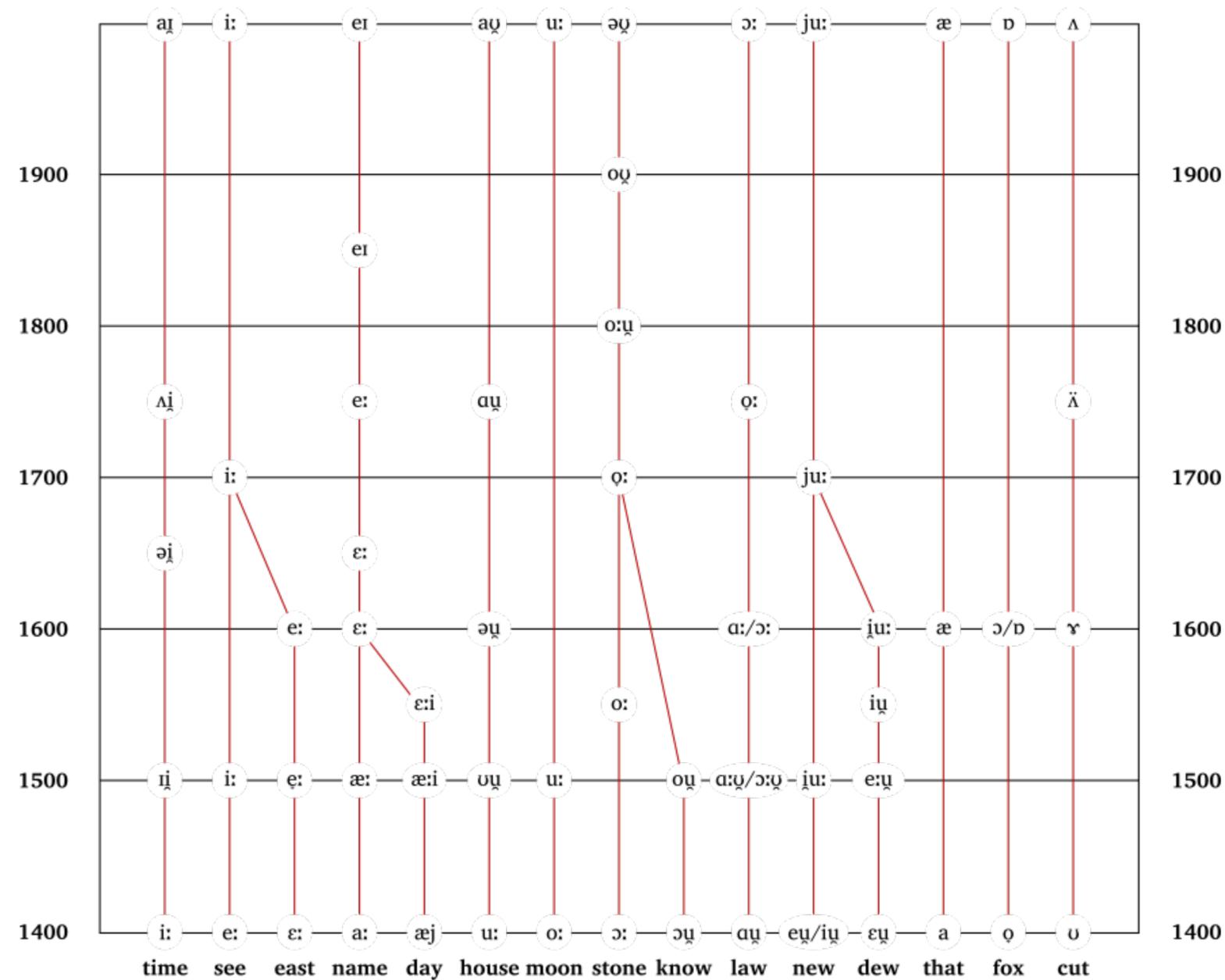
Three principles of deep orthographies

1. Don't get it write the first time ✅

2. Standardize. Standardise. Standardize. ✅

3. Watch words drift as you refuse to let go of the passed

Principle three:

Watch words drift as you refuse to let go of the passed

Watch words drift as you refuse to let go of the passed

English: Great vowel shift

Watch words drift as you refuse to let go of the passed

Zevy: Impact of historically long vowels

Standard Written Zevy	Middle Zevy	intervening period	Modern Zevy *Kuuvi*	Modern Zevy *Bemi*
dit "me"	/dit/	-	/dit/	/dit/
diit "change"	/diit/	/djit/ length to onglide	/zit/ affricatization + eventual fricativization	/d͡zit/ stopped at affricatization
tiri "rope"	/tiri/	-	/tiri/	/tiri/
tiiri "perceive"	/tiiri/	/tjiri/ length to onglide	/t͡siri/ affricatization	/t͡siri/ affricatization

This merger of /d/ with /z/ before historically long vowels is a distinct marker of the Kuuvi dialect

Watch words drift as you refuse to let go of the passed

Zevy: Impact of historically word-final consonants

Standard Written Zevy	Middle Zevy	intervening period	Modern Zevy *Kuuvi*	Modern Zevy *Bemi*
dit "me"	/dit/		/dit/ [dih]	/dit/ [dɪθ]
diit "change"	/diit/		/zit/ [zih]	/d͡zit/ [d͡zɪθ]
dit su "by me"	**dit suun** "under me"		/ditsu/ [diθsə]	/ditsu/ [dɪθsʊ]
diit me "in change"	–		/zitme/ [ziθmɛ]	/d͡zitme/ [d͡zɪθmɛ]

affricatization

Kuuvi generally renders /i/ as [i] in free-morpheme monosyllables, but Bemi does. Both tend towards [ɪ] in pollysyllabic morphemes

The vowel in the postpositions *su* and *mu* merges with the schwa in Kuuvi, but retains its distinct quality in Bemi

The two dialects disagree on the value of /t/ at the end of phonological word…

…but agree on its value at the end of morpheme within a phonological word

Watch words drift as you refuse to let go of the passed

Zevy: More impacts of historically long vowels

|	**Middle Zevy**	**intervening period**	**Modern Zevy** *Kuuvi*	**Modern Zevy** *Bemi*
isiit "seated"	/isiit/	/isjit/ length to onglide	/iɕit/ [jɪɕɪh]	/iʑit/ [jɪʑɪh]
isi "sit"	/isii/	/isji/ length to onglide	/iɕi/ [jɪɕ]	/iʑi/ [jɪʑ]
isit "town"	/isit/	/izit/ intervocalic voicing	/izit/ [jɪzɪh]	/izit/ [jɪzɪh]
si "am" from **disi** "I am" /disii/		irregular; hypothesized to have been hypercorrected by analogy with the similar abbreviation **ti** from **tide**	/zi/ [zɪ]	-
		/disji/ length to onglide	-	/ɕi/ [ɕɪ]

Bemi: intervocalic /ɕ/ is voiced

Kuuvi: intervocalic /ɕ/ is *not* voiced

Both dialects: intervocalic /s/ is voiced
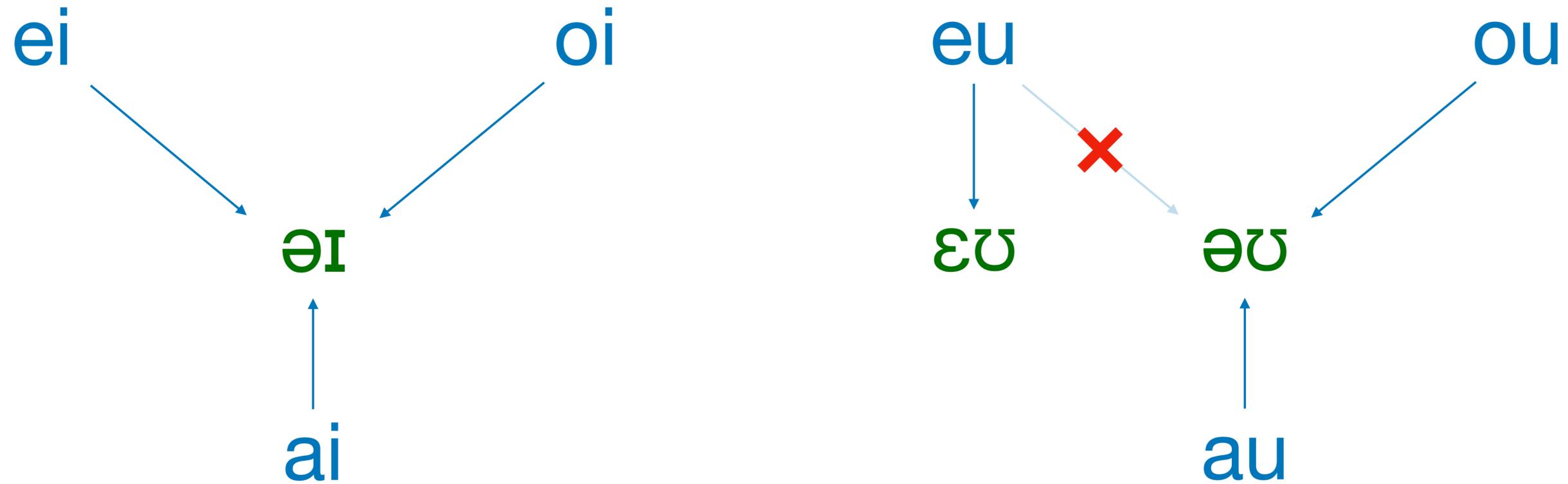
Kuuvi: pronunciation of copula is irregular

Watch words drift as you refuse to let go of the passed
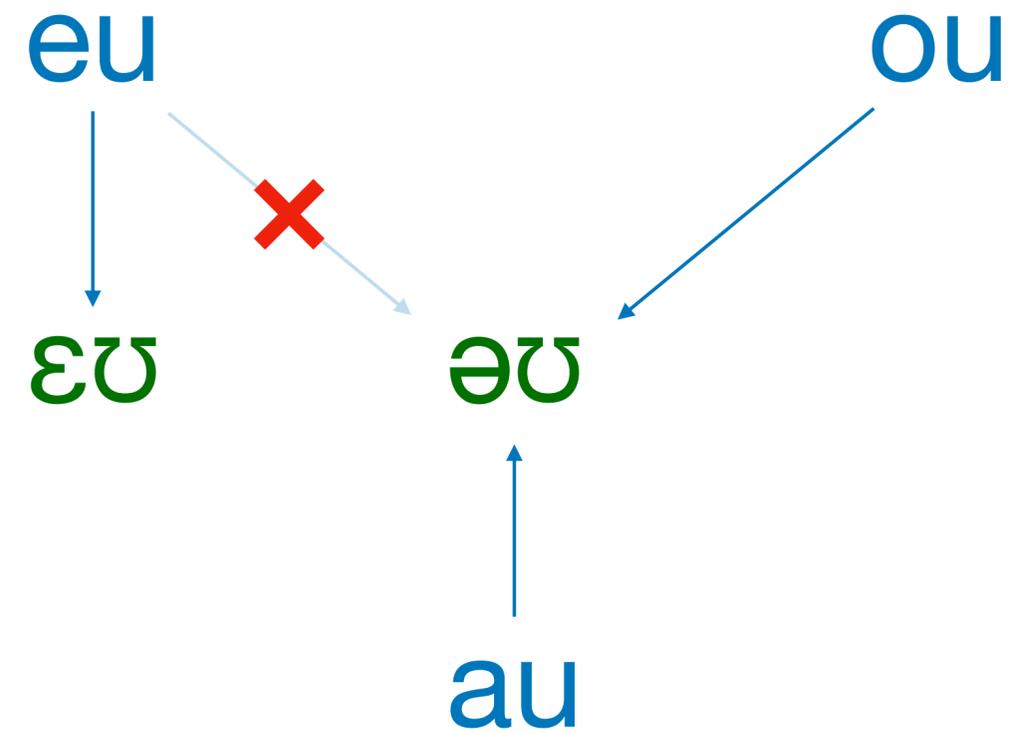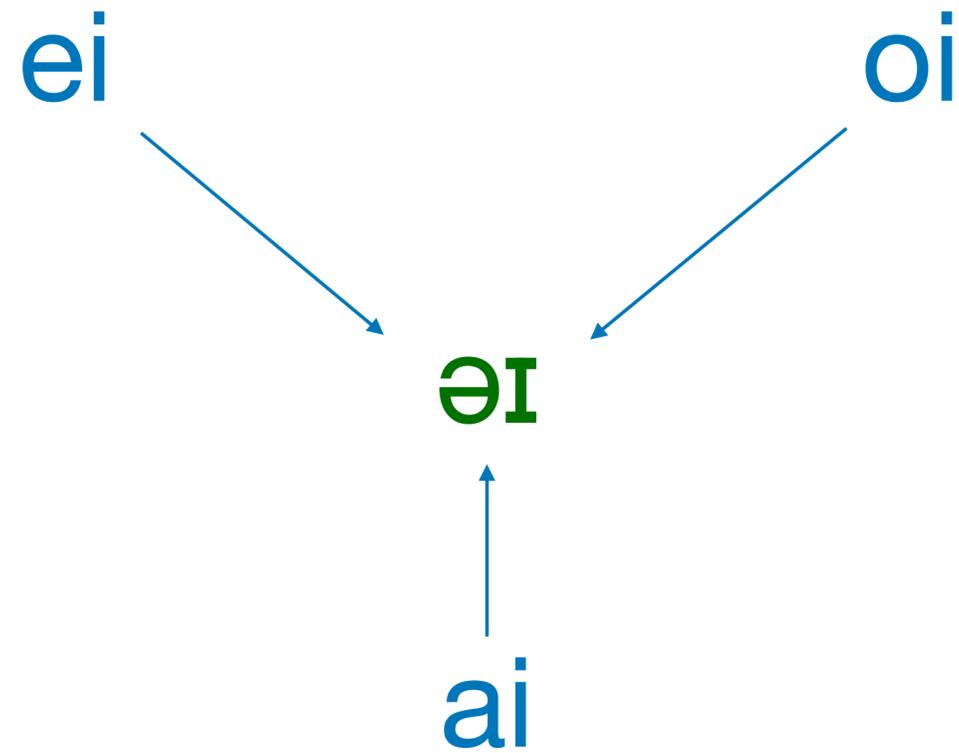
Zevy: Sound mergers

ei ⟶ əɪ ⟵ oi

ai ⟶ əɪ

eu ⟶ ɛʊ ✗ əʊ ⟵ ou

au ⟶ əʊ

Watch words drift as you refuse to let go of the passed

Zevy: Sound mergers

ei   oi   eu   ou

əɪ   ɛʊ   əʊ

ai   au

Kuuvi
"Capital"   ➔   /kəʊvi/

Watch words drift as you refuse to let go of the passed

Zevy: Sound mergers

ei → əɪ ← oi

ai → əɪ

eu → ɛʊ   eu → əʊ ✗   ou → əʊ

au → əʊ

New homophones:
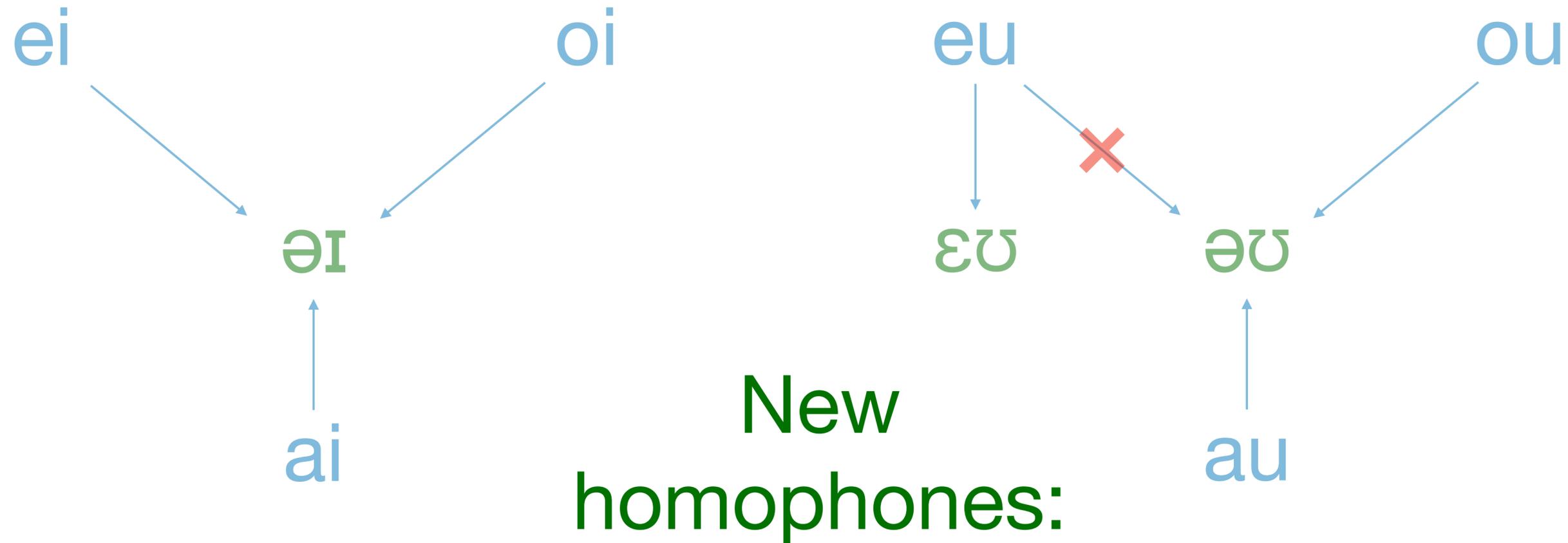
keimi = koimi
"fiction" "problem"

mau = mou
"strike" "puncture"

Three principles of deep orthographies

1. Don't get it write the first time ✅

2. Standardize. Standardise. Standardize. ✅

3. Watch words drift as you refuse to let go of the passed ✅

modern zevy writing: complex sound correspondence

ligatures

sound changes

$\widehat{ts}$ t ei s u b e tθ d a k a v e r v o d $\eth$

əɪ   z

spatial abbreviations

me

"a boring shoe in a brown box"

homophone disambiguation

k u u   b o b u

əʊ

unstressed vowel deletion
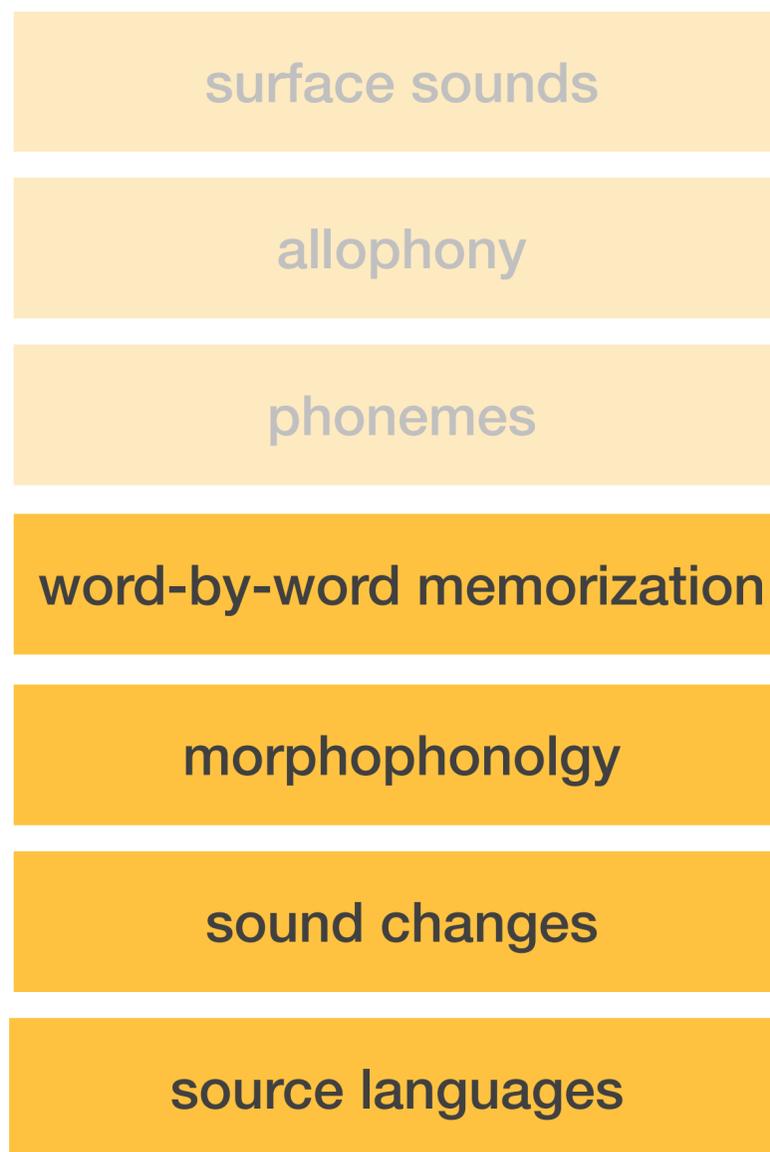
me

"in the square under the circle"

Fin!

Principles of deep orthographies

Don't get it write the first time

Standardize.
Standardise.
Standardise.

Watch words drift as you refuse to let go of the passed

surface sounds
allophony
phonemes

→ word-by-word memorization

→ morphophonolgy

→ sound changes

→ source languages

¿zehi mu?

Questions?

And there's more on Zevy at
connotes.interlocutor.io