LING 110 || Summer 2011, Class #3

McFetridge, Chapter 4: The alphabet

We'll do three things here today:

1. look at the history of the alphabet

2. examine some spelling rules

3. see how the first two affect one of the themes of this course: **REPRESENTATION**

Types of orthography

PICTOGRAPHS represented things by means of pictures. What you see is what you get. Two obvious problems:

1. inferior drawing means your message may not get through
2. how do you represent intangible things like ideas, concepts …?

One solution to problem #2 is to use existing symbols as **homophones** for other concepts.

— this is the **rebus** (Latin "thing") principle, e.g., "h8" for "hate" or "b4" for "before". Another typical example is using an image of the sun to mean "son".

— this is an important step: we're moving toward associating symbols with sound.

Ideographs take this further by using arbitrary symbols to represent concepts both abstract and concrete.

— this advances expressive power and allows for a system understandable by speakers of different languages. We see this with Chinese languages which may otherwise differ in their oral usage.

— one problem, however, is the vast number of symbols required

Another possible writing system is the **syllabary**: each symbol represents — you guessed it! — a syllable.

— Japanese has two such systems, hiragana for native words and katakana for foreign borrowings.

— this only works because Japanese has highly constrained syllable structure rules and thus a limited number of syllables.

— a syllabary for English would be unwieldy

The Korean **Hangul** system is a marvel of efficiency in which syllables are formed from vowel and consonant symbols, the latter representing simplified outlines of the tongue and the parts of the mouth used to pronounce them.

Anyway, in an ideal world an alphabet would have one symbol for each sound. But …

1. Spelling lags changes in spoken language, i.e., orthography doesn't keep up to date. Consider the "k" in "knee" or "know". That sound used to be pronounced, but obviously isn't anymore.

2. We don't want to record sounds that are not **distinctive**. This means sounds that don't contribute to meaning differences, e.g., the little puff of air (**aspiration**) that accompanies voiceless stops at the beginning of words, e.g., "pit").

3. Nor do we necessarily want to record dialectal differences.

4. We do, however, want to retain relationships among words, e.g., sure we don't pronounce the "t" in "nation" like we do in "native", but we want to keep the relationship between the two words intact.

There are other problems with alphabets and particularly with the words that arise from them. For example, what IS a word? Is a combination like "office party" one word or two? Why isn't it spelled the same as, say, "airplane"?

History

Many contemporary alphabets descend from the Greek alphabet (and of course many don't, too). The Roman and Cyrillic alphabets owe a lot to the Greek alphabet.

> — Thai, Hindi, and other Southeast Asian language alphabets come from Devanagari, originally used for Sanskrit texts
>
> — it's even thought that Hangul was influenced by the Roman alphabet.

Why alphabets? Economics!

To track trade agreements, clay figures were put in clay pots that were then sealed. An inscription would be put on the side of the pot.

> These early writings (3rd millennium BCE) were called **cuneiform** from the Latin "wedge".

Of course, if you have the writing on the outside of the clay pot, why do you need anything inside of it?

Cuneiform writing is **pictographic** and over time these symbols became more abstract. You are probably familiar with the Egyptian pictographic system called **hieroglyphics**, literally "sacred writing".

> Hieroglyphics became more abstract over time and the symbols were used to represent both concepts or things as well as sounds — an important development.
>
> — there evolved a set of symbols called **determinatives** that would indicate whether a character was to be interpreted as a concept or a sound.
>
> — the rebus principle was important here, too.

So by ~1600 BCE we have two prominent writing systems.

At around this time Semite groups borrowed the consonants, named them, and put them in a particular order.

— this naming and ordering increased literacy and enhanced education

You might think that an alphabet with only consonants is not a model of clarity or ease of use, but actually it is!

— to this day we find that Arabic and Hebrew have roots built commonly on three consonants to which vowels are added to indicate, say, tense, part of speech, plurality.

e.g., in Arabic the triconsonantal root "ktb" pertains to writing. From this comes katab "he wrote", kitaùb "book", yektub "he is writing" and so forth.

The Phoenicians (literally "purple people"), who were great traders, took the named and ordered symbols on their travels.

The Phoenicians made it to Greece. The Greeks liked this alphabet, but Greek was an Indo-European language and didn't form words in the same way as the Phoenicians.

— the Greeks took the consonant symbols that they didn't need and made them stand for vowel sounds.

Consider the first symbol "alef" (Greek "alpha"). Originally, it was a picture of an ox, but over time became a stylized symbol with only three strokes. By the time the Phoenicians got hold of it, the symbol was on its side.

— the Greeks rotated it a little bit more and, presto, "A"! (In Phoenician, the symbol stood for a pharyngeal sound not used in Greek, so the symbol was used as a Greek vowel instead.)

— the symbol "beth" (Greek "beta") was a stylized house. Pretty short trip from that to "B".

Noteworthy was the development of kappa or "k". Originally this was a stylized rendering of the hand and was written as 𓎛. The current orientation came about through boustrephedon writing, literally, "as the ox ploughs".

> Once a left-to-right writing style was settled on, the "k" was fixed in its present orientation.

The Greek alphabet is shown on p.82, Table IV.2. Note that "zeta" is the sixth letter, not the last. How did that happen?

Note, too, that the Bible's Book of Revelation uses the Greek alphabet as a metaphor for God in more than one spot: "I am Alpha and Omega, the beginning and the end …".

> Alpha and Omega are, of course, the first and last letters of the Greek alphabet. This usage also reveals the language in which this section of the Bible was written.

The Roman alphabet came to us through the Etruscans who had borrowed it from the Greeks.

— many interesting bits: "koppa" (ultimately the letter "q") was pronounced with lip rounding. Hence you always find "q" with "u".

— the Etruscans did not have a character for [g]. Latin needed a character for that sound, but did not need [z]. So the character "z" was reshaped as "g" (see p.83).

— the letter "c" came about by turning "gamma" Γ by 45°.

— particularly interesting is the evolution of "x". So common was the [ks] sound combination that a single character was devised for it.

In the following slides, we see why it's important to rely on sounds and not orthography when trying to understand word formation.

Representation + Reality

The Latin root √ duc means “leader”; √urb means “city”

nominative case	dux	urbs
accusative case	ducem	urbem
genitive case	ducis	urbis

Analysis

nominative	urbs	urb + s
accusative	urbem	urb + em
genitive	urbis	urb + is

Paradigm

nominative	ROOT + s
accusative	ROOT + em
genitive	ROOT + is

Analysis

nominative	dux	?
accusative	ducem	duc + em
genitive	ducis	duc + is

Representation

nominative	duks (dux)	urbs
accusative	dukem (ducem)	urbem
genitive	dukis (ducis)	urbis

The sound [z] came back because it was needed by English (to distinguish it from its voiceless counterpart [s]).

Both the Romance languages and English needed new characters to represent new sounds:

— the Romance languages created or reused "j" and "v"

— English created "y" and "w"

"Y"

This letter had various duties. It could represent both [y] and [i]. Note the spellings "tyre" and "tyger".

— sometimes it would even represent [ð] as in "ye olde bookstore"

We see the "y" / "i" alternation in what are called **spelling rules**.

cost	costly	costliness
lone	lonely	loneliness
world	worldly	worldliness
cry	cries	crying
die	dies	dying
fry	fries	frying

The alternations between "v" and "w" on the one hand and "j" and "y" on the other are seen clearly in cognate relationships. (Remember that "j" and "v" are Latinate and "y" and "w" are English).

revolve	walk
vibrate	whip
value	wield
January	year
jugular	yoke

More about spelling rules

Languages constantly change. We all know that. So why not just add new characters when needed? There are various (weak) excuses. So what some languages do is add diacritics like the cedilla as in 'ç' or umlaut marking as in 'ü'.

English avoids this by using vowel combinations, consonant combinations, and the silent 'e' to indicate pronunciation.

For example, when the [ʃ] sound entered the language, the symbol "sh" was created to represent it. This kind of "double symbol" is called a **digraph**.

Other kinds of spelling rules can mimic phonological rules of a language. Thus the silent 'e' of English mimics a sound change in Old English, i.e., "car" vs "care".

First we need to understand a little about syllable structure:

i) a syllable can begin with a consonant or a vowel, e.g., "bat", "at"

ii) a syllable can end with a consonant of a vowel, e.g., "set", "see"

iii) The sequence VCCV is divided VC | CV

iv) The sequence VCV is divided as V | CV

Two Old English rules

i) A closed syllable ends in a consonant; all closed syllables are short

ii) An open syllable ends in a vowel; all open syllables are long

So in addition to representing a long vowel with two symbols as in "food" or "feed", we can represent a syllable as open, e.g., "pin" vs "pine".

— how does this work? Well, if we add a silent 'e', we have a sequence VCV. By our earlier rule, this divides as V | CV. That means that the first V is an open syllable and hence long.

— if you don't add a silent 'e', the syllable is closed and therefore short.

We can see all of this at work in "stopping" vs. "groping"

Doubling the character "p" makes the preceding syllable appear closed so that the vowel is "short", viz., "stop | ping"

— In "groping", the syllable appears to be open so no 'e' is required, viz., "gro | ping"

A complication

Once 'e' is used for reasons other than representing a single sound, it can be press-ganged into use for other purposes …

— for example, in words like "lettuce" and "menace", the 'e' is used to signal how to pronounce the letter 'c'. This is nothing to do with the length of the previous vowel.

Rule conflict

Note: invoke, but invocation; revoke, but revocation ...

Why do we need two characters , 'c' and 'k', for the same sound?

— because 'e' at the end of the verb is needed for the vowel length rule, but not for the pronunciation of 'c'. So we use 'e': to get the pronunciation of the vowel right. Then, using 'k' in the verb avoids the French spelling rule that would have made us pronounce 'c' as [s].

Spelling reform

Pretty much a dead issue in Canada these days although there are small groups of diehards that continue to push for it.

— major attempts at reform in Germany led to only modest changes

— pending changes in Brazilian Portuguese are designed to reduce the number of accent marks

Remember, spelling helps to retain relationships among words.

These morphological relationships are perfectly clear in some cases:

govern	government
commit	commitment
content	contentment

In other cases we can see a relationship but it takes some analysis to get at its precise nature:

divide	division
conclude	conclusion
deride	derision

In yet other cases, relatedness is being lost and the words end up with meanings that obscure their common roots…

food	fodder
house	husband
home	hamlet
holy	Hallowe'en
out	utmost
goose	gosling

Further spelling reform would likely lead to this kind of thing happening throughout the entire lexicon. Although perhaps that's only of historical interest to most.

Typography

Not strictly related to the concerns of this course at the moment, but relevant to the way the written word is presented. And this may impact words in unknown ways. Consider, for example, the tiny font size on the screens of newer cell phones.

Which of the following do you find easier to read?

The quick brown fox jumped over the lazy dog.

The quick brown fox jumped over the lazy dog.

The first is 44-point Times; the second is 36 point Verdana.

Times is a serif font; Verdana is sans serif.

Sans serif fonts are claimed to be appropriate only for short passages. Yet you find these fonts all over the place. Determining the "best" fonts for web pages is a field with many self-appointed experts.

— the issue of readability is terra incognita

Consider that at one time there was no space left between words and readersseemedtogetalongOK.

And what are we to make of new research last Fall (2010) published in the prestigious journal **Cognition** that demonstrated better learning in participants who had to read material in grotty, hard-to-read fonts than those who read the same material in conventional fonts?

BTW, I do these overheads in Hypatia Sans Pro at 36 points. Hypatia is an Adobe font that has lots of "air" around each letter. It was not designed for use as a small point font.