1. Historical linguistics: The history of English

1.1. Proto-Indoeuropean (roughly 3500-2500 BC)

1.1.1. Proto-Indoeuropean and linguistic reconstruction

<table>
<thead>
<tr>
<th>Sanskrit</th>
<th>Greek</th>
<th>Latin</th>
<th>Gothic</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>pater</td>
<td>pater</td>
<td>pater</td>
<td>fadar</td>
<td>father</td>
</tr>
<tr>
<td>palaam</td>
<td>poda</td>
<td>fota</td>
<td>foot</td>
<td>*ped- *foot</td>
</tr>
<tr>
<td>b'tatar</td>
<td>g'tater</td>
<td>frater</td>
<td>brother</td>
<td>brother</td>
</tr>
<tr>
<td>f'aramu</td>
<td>p'ero</td>
<td>fero</td>
<td>bear</td>
<td>*b'ero - 'carry'</td>
</tr>
<tr>
<td>jirah</td>
<td>kwim</td>
<td>kwim</td>
<td>quick</td>
<td>*qweoi - 'live'</td>
</tr>
<tr>
<td>sanah</td>
<td>sinex</td>
<td>sinex</td>
<td>senile</td>
<td>*sen - 'old'</td>
</tr>
<tr>
<td>vrath</td>
<td>/wi/</td>
<td>wair</td>
<td>were(wolf)</td>
<td>*wi-ro- 'man'</td>
</tr>
<tr>
<td>ini</td>
<td>tres</td>
<td>tres</td>
<td>dini</td>
<td>*trew - 'three'</td>
</tr>
<tr>
<td>dasa</td>
<td>deka</td>
<td>decem</td>
<td>tahan</td>
<td>*dakm - 'ten'</td>
</tr>
<tr>
<td>sata</td>
<td>/kentum/</td>
<td>hund(rath)</td>
<td>hundred</td>
<td>*dkm-tom (&lt;*dekm) '100'</td>
</tr>
</tbody>
</table>

Table from www.utexas.edu/depts/classics/documents/PIE.html

- From the table, note that the similarities and systematic differences. Why do they exist?
  - Not due to coincidence, since hundreds of other words show the same pattern.
  - Not due to language contact, as there was no (or not enough) contact with India.
  - Not due to linguistic universals, as there are no such similarities in most other languages in the world (say Mandarin, Maori, Mohawk, Malagasy...).
  - Best answer: genetic relation: the languages in the table all came from the same ancestor language.

- This ancestor language is called (Proto-)Indoeuropean (proto=reconstructed), PIE.
- By comparing similarities and systematic differences between the languages whose origin is investigated, historical linguists can reconstruct words in the original language, e.g. last column in the table. Asterisks (*) indicate unattested, reconstructed words.
- Reconstruction is assisted by knowledge of normal patterns of linguistic change for which direct evidence is available (e.g. development of Romance languages from Latin).

What we know about the people who spoke PIE:

- Not known whether PIE was spoken by a single ethnic group.
- Inferences about PIE speakers, based on vocabulary common to all/most Indoeuropean languages, and hence likely to have existed in PIE:
  - Possible original location: near Caspian & Black Seas, South Russian steppes.
  - They had a word for snow: *oneigh*—Latin nix, Greek niphos, Gothic snaiws, Gaelic sneachta, ‘snow’, so coldish climate.
  - No (known) word for seal/sea/ocean (though words for lake & rowing) and forest.
  - Words for bee, birch, elm, ash, oak, apple, cherry; bee, bear, beaver, eagle.
  - Original location is also deduced from subsequent spread of IE languages.
  - Bronze age technology (gold, silver, copper known, iron wasn’t yet known)
  - They rode horses & had domesticated sheep, cattle. Cattle a sign of wealth (cf. fee/German Vieh ‘cattle’, Latin pecunia ‘money’/pecus ‘cattle’).
  - Agriculture: cultivated cereals *gre-no*—(grain, corn), also grinding of corn mela-(cf. mill, meal); they also seem to have had ploughs and yokes.
  - Weaving with wool.
  - Wheels and wagons (wheel < k(e)-k(lo) > k(e)l ‘go around’)

1.1.2. Indo-European language families

PIE split into distinct dialects/languages/families due to migration, language contact, conquest, etc. Ten main families: Tocharian (extinct languages in Western China), Indo-Iranian (Sanskrit, Hindi, Urdu, Persian, Pashto...), Armenian, Anatolian (extinct languages in Turkey, Syria, incl. Hitite), Albanian, Greek, Balto-Slavic, Italic (Latin, Romance Iges), Celtic, Germanic (Gothic, English, German, Danish...).

1.2. Proto-Germanic/Common Germanic (very roughly 2000 BC - 250 BC)

- Proto-Germanic speakers: originally IE nomads who settled in an area in Nth Germany and Southern Scandinavia, around 2000 BC, give or take several centuries.
- References to them by Roman authors after about 200 BC.
- Very little common Germanic is recorded: a few words written down by Roman writers in 1st century BC & a couple of artefacts assumed to reflect this stage of language.

1.2.1. Proto-Germanic grammar

- Initial stress: PIE did not have stress on any particular syllable, but Gmc. put stress on the root, i.e. mostly first syllable (also in compounds). This eventually caused the erosion of inflectional suffixes.
- The effects of Grimm’s Law (=The (First) Germanic Sound Shift):
  - A) Voiceless unaspirated plosives became fricatives
    | PIE | Germanic | Examples |
    |-----|----------|----------|
    | p   | f        | pedis/<foot, pecus/<Vieh, per/for, poly/greek/viel/German, piscis/<fish, Off |
    | t   | o        | tonitus/<thunder, tenuis/<thin, tres/<three |
    | k   | x/h      | canis/<hound, sequo <follow/saun/horn <see, cornu/<horn |
  - B) PIE voiced unaspirated plosives lose their voicing
C) PIE aspirated stops end up as unaspirated (they became voiced fricatives first, which is ignored here):

<table>
<thead>
<tr>
<th>PIE</th>
<th>Germanic</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>bʰ</td>
<td>b̥arami/strēt/bear, frater/brother</td>
<td>(PIE b&gt;- Latin f)</td>
</tr>
<tr>
<td>dʰ</td>
<td>d̥ace/foris/door, vidua/widow</td>
<td>(PIE d&gt;- Latin f/d)</td>
</tr>
<tr>
<td>gʰ</td>
<td>hostis/gardien, homo/gymnəmb</td>
<td>(PIE g&gt;- Latin h)</td>
</tr>
</tbody>
</table>

- When did this happen? Clue: shift in *hemp* from Greek *kάναβις*. The Germanic people learned about hemp from the Greeks, who themselves didn’t know about it till about 500 B.C. Hence the sound shift was still in effect some time after 500 B.C. The sound shift was no longer in action by the time the Germanic people had contact with the Romans (1st century B.C.), since Latin borrowings don’t undergo the shift (pepper<piper, street<via strata, peach<persica, pound<pando, tile<regula).

1.3. Old English (450-1100)

1.3.1. British Isles before Anglo-Saxon Invasions

- 6000-4000 BC: Non-Indoeuropean people in British Isles (Stonehenge 2000 BC)
- From about 700 BC: Celtic settlements
- 55 BC-410 AD: Roman presence in British Isles (abandoned 410 AD).

1.3.2. Anglo-Saxon Invasions

- From 449 AD: hordes of Germanic speakers (Angles, Saxons, Jutes, Frisians, Franks) from what is now Northern Germany/Southern Denmark move into Britain in waves.
- By late 6th cent, Anglo-Saxons dominated British Isles & marginalised the Celts.
- The (differing!) Germanic varieties they spoke are collectively called Old English or Anglo-Saxon.

1.3.3. The era of the Viking invasions (787ff)

- 787: Scandinavian (=Viking, Norse, Danish, Norwegian) invasions. Continued for nearly 200 years, culminating in a period where England was ruled by Danes in early 11th cent.
- Linguistic aspects of Scandinavian invasions:
  - Old Norse and Old English were mutually comprehensible, but inflectional forms differed, and resulting compromise eroded inflection (standard assumption, at least).
  - 1000-odd words borrowed into OE in late OE period:
    1. anger, bag, both, cull, die, egg, flat, get, husband, knife, leg, low, sister, steak, take, until, want, window, wrong
    2. Pronouns: they/then/their
    3. Words starting with /sk/ were Norse: sky, skin, skill, skull. Doublets (often with semantic differentiation): shirt/shirk, shriek/screeh, ship/skipper, shatter/scatter.
    4. Other doubles: bathe/bask, church/kirk, whole/hale, ditch/dike
    5. Borrowing yields near-synonyms: heaven/sky, carven/cart, craft/skill, hide/skin, sick/ill
    6. Place names: -by (Derby, Rugby), -thorp (Linthorpe, Althorpe)

1.3.4. Inflectional morphology

- OE had a fairly rich inflection system compared to modern Eng., but less rich than PIE. Also inflection on adjectives.

<table>
<thead>
<tr>
<th>Masc.</th>
<th>Neuter</th>
<th>Feminine</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strong</td>
<td>Weak</td>
<td>Strong</td>
</tr>
<tr>
<td>Singular</td>
<td>Nom./Acc.</td>
<td>genitive</td>
</tr>
<tr>
<td>Plural</td>
<td>nom./acc.</td>
<td>genitive</td>
</tr>
</tbody>
</table>

1.3.5. Syntax

- Verb-final order possible in subordinate clauses; verb-second in main clauses (i.e. one constituent, not necessarily a subject, before the inflected verb).

1.3.6. Text (The Lord’s Prayer)

Father our thou that art in heavens
Si fin nuna gehalod
be thy name hallowed
to become fin rice
come thy kingdom
gewurfe bin willa
be-done thy will
on eordan swa swa on heofonum.
unre gedaglwamlican hlaf syle xode todag
and forgis us ure gyllus
and forgive us our sins
swa swa we forgysa urum gylendum
and we may have sins against us
ne gedel us on costunge
and not let us into temptation
ac alys us of yfel lufel
1.4. Middle English (1100 – 1500)

1.4.1. The Normans in England (1066ff)

- **Normans**: people in Normandy; descendants of Norse invaders; spoke Norman French.
- **Norman Conquest**: Normans under William the Conqueror took over England in 1066.

1.4.2. Linguistic aspects of the Norman Conquest

- *(Norman) French replaced English in upper classes, but the vast majority spoke English.*
- England bilingual. Norman words imported into the English vocabulary *en massep.*

12. **Government**: crown, parliament, state, tax, baron, count, duke, prince; **Law**: accuse, attorney, court, crime, judge, prison, punish; **Church**: abbey, clergy, parish, prayer, religion, saint; **Fashion**: apparel, costume, dress, fashion; **Culture**: art, beauty, chant, colour, music, paint, poem; **War**: armour, battle, war

Why do you think the following pairs of words differ semantically in the way they do?

13. **house** (*Germanic*) / **manor** (*Norman*)
14. **pig** (*Germanic*) / **pork** (*Norman*)

1.4.3. Decline of Norman French and resurgence of English

- **Decline** of French in England: 1362: English replaces French in parliament, courts;
- 1399: Henry IV: king again native English speaker;
- 1423: parliament records no longer kept in French.
- Reasons for decline: (a) 1204: Normans lose Normandy to the French, severing connections to Normandy. (b) Norman French considered ‘bad’ due to rise of Parisian French as standard in 13th c. (c) The Normans had always been a minority in England anyway. (d) Hundred Years’ War (1337-1453) against France: French was the enemy’s language. (e) Black death (1349) kills 30% of people → labour shortage → surviving Anglo-Saxons get better pay/status → English-speaking middle class.
- In the 15th cent. London dialect became a new standard, since London was a trading, cultural and political centre. Geoffrey Chaucer (1340 –1400) also used London English.
- **Middle English period defined**: Usually dated from either 1100 or 1150, since by then the linguistic effects of the Norman Conquest were starting to take hold. End of period: 1500 or so (due to standardising effects of printing press).

1.4.4. Inflectional morphology

- ME inflectional system much simpler than that of OE. Examples:

<table>
<thead>
<tr>
<th>Sing.</th>
<th>Pl.</th>
</tr>
</thead>
<tbody>
<tr>
<td>sg 1</td>
<td>-th/-es</td>
</tr>
<tr>
<td>sg 2</td>
<td>-(e)st</td>
</tr>
<tr>
<td>sg 1</td>
<td>-eth/-e(n)</td>
</tr>
<tr>
<td>sg 2</td>
<td>-(e) / -dest</td>
</tr>
<tr>
<td>past</td>
<td>-eth/ -eth</td>
</tr>
<tr>
<td>pl</td>
<td>-eth/-e(n)</td>
</tr>
</tbody>
</table>

- Nearly all nouns ended up in one inflectional class (=OE strong masc).
- Shift from grammatical gender to natural gender by 13th century.

1.4.5. Syntax

- Word order became fixed SVO, even in subordinate clauses. It had to be fixed, because case morphology wasn’t able to guarantee right thematic interpretation for NPs.
- Prepositions used in lieu of case morphology.
- Development of progressive construction: *be on working* > *be aworking* > *be working*. The suffix was a nominalisation suffix in OE and into ME. Progressive could have a passive meaning until the 19th century: *the road is building*. English starts to develop condition that all sentences have a subject, so subjectless impersonals like the following become rare:

15. **And happyd so, they coomen in a toun**

1.4.6. Middle English Texts

Geoffrey Chaucer (1340-1400): end of Troilus and Criseyde:

And for ther is so gret diversite [great diversity]
In English and in writyng of oure tonge, [tongue – i.e. language]
So prey I God that non myswrite thee, [miswrite – i.e. copy it wrongly]
Ne thee mysymetre for defauyte of tonge; [mismetre – i.e. get the rhythm wrong]
And red whereso thow be, or elles songe, [red = read; elles = else; songe = sung]
That thow be understonde, God I beseche! [understonde = understood]
1.5. Early Modern English (1500-1700)

1.5.1. General observations

- 1476: William Caxton brings printing to England, enabling massive transfer of language in a fixed form (i.e. without scribes transferring texts into their own dialects), helping standardisation (since books had to be in a widely understood type of English). Caxton used spelling from round London, and some of these conventions are still used.
- Renaissance (16th c.): renewed interest in classical languages, much borrowing from Latin.
- The first attempts at linguistic purism in English, as well as spelling & other school books.

1.5.2. Morphology and syntax

- 2nd pers. sg. pronouns/verb inflection (thou/you) replaced by plural forms ye/you in 17th c. (except in religious usage and some dialects). From the 13th c., plural forms used for polite address. Eventually, it was considered rude not to apply it. E.g.:
  (16) SIR WALTER RALEIGH: I do not hear yet that you have spoken one word against me; here is no treason of mine done. If my Lord Cobham be a traitor, what is that to me? SIR EDWARD COKER: All that he did was by thy instigation, thou viper; for I thou thee, thou traitor.

- Nouns: No dative or accusative marking on nouns survived. In 16th century the genitive –s ceased to be a inflection of nouns, and became a clitic added to whole NP.

- Adjectives: all inflection lost save comparative/superlative.

- Verbs: (a) loss of all inflectional endings except those we still have today; (b) Northern –s suffix replaces Southern –eth in 16th c. in 3rd pers. sg. present; (c) Many strong verbs became weak; weak class is now default/regular class.

- Auxiliaries:
  - Have: begins to replace be as perfect auxiliary.
  - By late EME, do-support was used as it is now. In early EME, it was not necessary in the structures where we now need it (questions, negation etc.), but it had a now non-essential use which is often assumed to be semantically vacuous. (Hypothesis: it serves to keep the V inside the focus domain (=VP).) Examples from Shakespeare (1564-1616):
    (18) that suggestion whose horrid image doth [ VP unfix my hair] (Macbeth)

- (19) He heard not that (Two Gentlemen of Verona)

- (20) And what says she to my little jewel? (Two Gentlemen of Verona)

1.5.3. The Great Vowel Shift (roughly 1400-1700)

- The Great Vowel Shift (GVS; roughly 1400-1700):
  - All long vowels moved one step higher.
  - The high long vowels which could not move higher became diphthongs.

- Examples:
  (21) [aɪ] → [ai] .. [u] ... [ʊ] [iː] → [iː] ... [oʊ] [aɪ] → [iː] ... [eɪ] ... [uː] [æ] → [æ] ...

- Examples:
  (22) ME 1700 ME 1700

- Two proposals about how GVS happened (various compromise positions exist as well):
  - GVS as a pull chain (drag chain shift): high vowels diphthongised first, then [o, e] moved into the space these vowels had vacated and dragged [eɪ, aɪ] into the former positions of [o, e], and so on. (Possible cause for drag chains: a desire to maximise distinctions between vowels.)
  - GVS as a push chain shift: one or more vowels moved upwards, encroaching on the space of other vowels, forcing them to be pronounced differently (lest communicative efficiency be reduced).

- Everybody agrees that not all vowels shifted at the same time. Part of the disagreement about pull vs. push chain theories revolves around disagreement as to which vowels shifted before which others. See e.g. Lass (1999: 72ff) for detailed discussion.

- There were regional, sociocultural and ideological differences in the rate at which each vowel shifted. For instance, some Northern dialects still have pre-GVS [u:] in words like house.

- English spelling of vowels largely reflects their pronunciation before the GVS, hence differences between Engl. spelling and that in other languages.

- Due to vowel lengthening/shortening processes in late OE/early ME (not mentioned above), we find vowel quantity differences in (originally) morphologically related pairs of words. Since GVS affected long vowels but not short vowels, we find that the differences in quality result in differences in quantity as well.

- (23) chaste/chastity, mania/manic, fable/fabulous, grade/gradual, grain/granular, grateful/gratitude/gratify, navy/navigate, sane/sanity, state/static, pen/penny, serene/serenity, sincere/sincerity

- (24) athlete/athletic, discrete/discretion, gene/genesis, legal/legislate, penal/punitive, pleased/pleasure

- (25) analyse/analytic, child/children, alive/living; apply/applicable, Christ/Christmas

- Auxiliaries:
  - Oblige/obligation, sign/signature

- (26) abound/abundance, house/husband, out/utter, profound/profundity, south/southern

1.5.4. Changes in consonant system (not reflected in orthography)

- Palatal fricative [ç] was dropped in words like night, right (with compensatory lengthening: [niːt] → [niː]) and the velar fricative [x] was either dropped or became [ʃ]: laugh, laughter, draught, enough. The old spelling *gh* was retained.

- Loss of syllable-initial velar plosives before /n/ in 16th c. in 3rd pers. sg. present.

- Examples:
  (21) [ŋ] had been an allophone of [n] (conditioned by velar sounds after it). When /g/ after [ŋ] became [f]:

- (22) Loss of /l/ in certain syllable-final clusters: walk, half, calm.

- (23) Loss of /w/ in certain syllable-final clusters: walk, half, calm.

- Loss of /v/ in certain syllable-final clusters: walk, half, calm.

- Plosives disappear after nasals at end of syllable: thumb, bomb, numb, wrong. (Retained when syllable boundary intervenes: num-ber, lon-ger, fin-ger, etc.)

- [ŋ] had been an allophone of [n] (conditioned by velar sounds after it). When /ŋ/ after [ŋ] was lost (see last point), the original cause for the allophony disappeared. Result: minimal pairs like single/sing, thing/thing, so [ŋ]-[ŋ] contrast became phonemic.

1.5.5. Texts

A. Beginning of John 8 from William Tyndale’s 1525 translation of the Bible (which had an influence on the standard Authorised Version (=King James Version) of 1603):

<table>
<thead>
<tr>
<th>Rise</th>
<th>Mouth</th>
</tr>
</thead>
<tbody>
<tr>
<td>[iː]</td>
<td>[uː]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Meet</th>
<th>Boot</th>
</tr>
</thead>
<tbody>
<tr>
<td>[eɪ]</td>
<td>[oʊ]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Meat</th>
<th>Stone</th>
</tr>
</thead>
<tbody>
<tr>
<td>[aɪ]</td>
<td>[o]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Name</th>
<th>[a]</th>
</tr>
</thead>
</table>

By late EME,...
Jesus went unto mount olivet, and early in the morning came again into the temple, and all the people came unto him, and he sat down, and taught them. The scribes and pharisees brought unto him a woman taken in adultery, and set her in the midst and said unto him: Master this woman was taken in adultery, even as the deed was a doing. Moses in the law commanded us that such should be stoned: What sayest thou therefore? And this they said to tempt him: that they might have, whereof to accuse him. Jesus stooped down, and with his finger wrote on the ground. As soon as they heard that, they went out one by one the eldest first. And Jesus was left alone, and the woman standing in the midst. When Jesus had lifted up himself again, and saw no man, but the woman, He said unto her: Woman, where are those thine accusers? Hath no man condemned thee? She said: Sir no man, Jesus said: Neither do I condemn thee. Go hence and sin no more.

B. Texts by Shakespeare (1564-1616) available under www.shakespeare-online.com/plays/

4.2.1.1 /æ/ and /ɑ:/

- In the 17th and 18th c., /æ/ shifted to /ɑ:/, inconsistently, depending on the dialect, the phonological environment and even on the word in question.

<table>
<thead>
<tr>
<th>Example</th>
<th>Pronunciation</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. pat, bad, cap</td>
<td>/æ/ [original sound]</td>
</tr>
<tr>
<td>b. path, laugh, grass</td>
<td>/ɑ:/ [before most voiceless fricatives]</td>
</tr>
<tr>
<td>c. dance, demand, chant</td>
<td>/ɑ:/ [before cluster starting with nasal]</td>
</tr>
<tr>
<td>d. part, bar, cart</td>
<td>/ɑ:/ [before /t/ (before it was dropped)]</td>
</tr>
<tr>
<td>e. half, rather, banana</td>
<td>/ɑ:/ [other unsystematic cases]</td>
</tr>
</tbody>
</table>

- The pronunciations above are RP. Examples of variety differences:
  - Midlands, North of England have /æ/ in (b) and (c).
  - American English was isolated from the changes in Britain, so it retains /æ/ everywhere except (d) (because this change occurred earlier).
  - Australia: just like RP, except that some words in class (c) have free variation between /æ/ and /ɑ:/, while others are fixed in a particular way (can’t only with /ɑ:/, romance only with /æ/).

- Since these changes were not fully productive, we end up with minimal pairs like aunt vs. aunt. It is thus not possible to see /æ/ and /ɑ:/ as allophones of the same phone.

1.7. Sources

- Online etymological dictionaries: www.etymonline.com
  www.oed.com;
  www.dwb.uni-trier.de (German, Grimm’s Dictionary)
- Sources on the history of English:

- Sources on language change (=historical linguistics):