

LING 425: HISTORICAL LINGUISTICS

Prof. Charles Boberg

Fall Semester 2016

Tuesday, 6 September 2016: Introduction

Historical linguistics is defined as “the study of the history and development of languages”. It is interdisciplinary, and often draws from the fields of archaeology and anthropology. Change is inherent to language; thus, any theory of language that does not account for change is incomplete. The goal of historical linguistics is to identify and explain the principles that govern language change, and use those principles to constrain theories about language change. Applications: identifying related language families and proposing hypothetical proto languages; determining which language changes are likely to occur in the future; identifying the social mechanisms that drive language change.

Related Terms Linguistic fields can be separated into one of two subcategories, **diachronic** and **synchronic**. The latter examines a language at a given point in time, whereas the former examines the differences in a language over time. **Philology** (Greek, “love of words”), is defined as “the branch of knowledge that deals with the structure, historical development, and relationships of a language or languages”. **Etymology** is the study of word origins.

Common Problems It is often difficult to examine language change due to the paucity of written records for most languages. Also, written records fail to illustrate many aspects of a language, such as pronunciation. The extent of our understanding of a language’s pronunciation largely hinges on whether the writing system is phonetic or not; for instance, Chinese is not phonetic and English and French are only partially. Nearly all writing systems are phonemic, and only indicate contrasts. However, written texts are valuable for observing certain other aspects of language, such as syntax and lexicon.

Additionally, pervasive literacy is a modern phenomenon. Old texts almost certainly do not reflect the full range of language, but only the type of language used by an educated upper class in certain formal situations.

Attitudes Language change, although inevitable, is often met with resistance. Older language forms are often regarded as the “better” or “pure” forms, whereas newer language is considered degenerate and corrupt. This is true regardless of time period or geographical location. Relatively few celebrate language change, with the exception of slang that is seen to define the cultural ethos of a group or generation.

Even now, language is morphing. Consider the drop in the use of “whom”. Or the substitution of “was like” for “said”. Many changes are ephemeral, but many others stick and are passed down to future generations. Many people consider their language as stable, when this is never the case.

Thursday, 7 September 2016: Introduction (con’t.)

Many linguistic changes are documented in the *Journal of Modern Linguistics*. Examples from recent issues include “Netflix and chill”, “sharewashing”, and “Thanks, Obama”. (Most of these words will most likely not last very long.) These changes in lexicon force lexicologists to habitually add and trim words from dictionaries.

Evolution of English

Old English Exemplified by *Beowulf*, circa 750 A.D.

- Overall, totally incomprehensible to speakers of 21c English. However, certain words (life, world) are relatively unchanged.
- Uses alliteration rather than modern rhyme scheme.
- Contained rich declensions (variation in noun forms), similar to German. For instance, it distinguished between 3 genders (neutral, female, male) and many cases (nominative, accusative, genitive, dative), and plurality). Modern English only distinguishes between singular and plural forms.
- Also had more conjugations, whereas modern English only preserves one unique form (-s on the present plural singular). Some archaic forms from OE preserved in ME: consider difference in “ride” “he rode” vs. “he has ridden”. This distinction is not present in new verbs e.g. there is no difference between “she googled” and “she had googled”.

Middle English Exemplified by *Canterbury Tales*, 1387 A.D.

- Disappearance of Anglo-Saxon graphemes. Familiar rhyming scheme (couplets). Much more recognizable.
- Change in language due to Norman conquest (example of an extralinguistic event with linguistic consequences): Vikings, many whom had spent time in France and absorbed French language/culture came to England; as a result, many French vocabulary words were borrowed into English. This was a lasting change; an estimated one third to one half of modern English vocabulary is derived from French.
- Not shown in text: pre-Great Vowel Shift, so although spelling is same as modern English, pronunciation was quite different.
- Many nouns feature a word final vowel, and declensions (though even these are relatively simplified when compared to Old English).
- Middle English was unstandardized, and had many alternate spellings.

Early Modern English exemplified by *King James Bible*, 1611 A.D.

- Totally understandable by speakers of modern English. Still commonly in use today.
- Considered a “high point” of English prose.
- Also reflective of a non linguistic event (Protestant reformation). Emphasis on having a direct path with god, which led to translation of bible into local languages so as to be accessible to the plebians.
- Given the formal nature of the subject matter, it most likely exemplifies language that is old fashioned for its time. (For instance, the ye/your vs. thee/thou distinction is sometimes ignored in Shakespeare’s works, indicating that its use was already somewhat dated.)
- Post great vowel shift; disappearance of superfluous e’s

Tuesday, 13 September 2016: Borrowing

World War Two Many Canadian/American soldiers went to Britain and were exposed to various dialects, resulting in some standardization between the Englishes. After WW2, US rose to ‘superpower’ status, and US troops were stationed everywhere, spreading American English worldwide. This helped push English to the status of global dominance.

Sound Change Changes in language can occur in sound, grammar, morphology, or lexicon. Sound change is the best studied aspect of language change. Sound change can be phonological (if it affects sounds that were previously contrastive, and became non contrastive or vice versa) or phonetic, if it does not involve contrasting sounds. The **Neo-Grammarian hypothesis** postulated that sound change is systematic and operates without exceptions. Furthermore, it stated that the unit of change was the phoneme, not the lexeme.

If sound change is regular, then we can make sense of the systematic correspondences that we identify between languages to make hypotheses about their history and development. The assumption of regularity allows loanwords to be identified as exceptions to sound change. There are two methods of identifying loanwords in a language. First, linguists can identify source of loanword in another language; secondly, it should also be obvious that the word did not undergo the changes in the host language before it was borrowed. Or, it is possible that the law did not apply in that case, and the Neo-Grammarian hypothesis of regular sound change is wrong.

Hypotheses of sound change are testable: theories about sounds that should have changed over times can be tested with new data from various time periods in the language.

In old changes, it's hard to know whether apparent regularity operated throughout the process of the change or was a result of a gradual diffusion of a change through the available vocabulary. For instance, if p became b, we don't know if it happened one word at a time or all at once; we do not know the mechanism by which regularity resulted.

Sound changes are notated using a series of abridged notations, eliminating the need for verbose descriptions.

$$X > Y$$

X becomes Y.

The textbook distinguishes **conditioned** and **unconditioned** changes. The above change is unconditioned. Conditioned changes are when the change only occurs in a specific environment. Constraints should not be phonetic, not morphological or syntactic (according to Neo-Grammarian theories). In Neo-Grammarian view, phonetics/phonology operates on a different level. Changes certainly should not be social: i.e. this change is only used in *x* social context.

$$X > Y/a_b$$

X becomes Y when it is located between *a* and *b*. This is an example of a conditional change.

$$C > [-vcd]/_ \#$$

C is devoiced when word-final.

$$V[-low] > [+nas]/_ [+nas] \#$$

Vowels become nasalized when word final.

Shorthands	
V	Vowels
E	Elsewhere
0	Nothing
.	Syllable Boundary

Physiological Cause of Sound Change The vast majority of changes have to do with the articulation of a word. Usually, they change the word by making it easier to pronounce. Word evolution is a clash between the need to maintain enough clear distinction among different sounds to communicate effectively and the need to expend as little effort as possible in articulation. Effort involves opening/closing mouth, moving the tongue, lips, larynx, and other articulators. All of these actions are subject to **erosion**, or simplification. (Erosion is due to speaker laziness, or more charitably, **efficiency**).

In connected speech, particularly rapid speech, erosion will most likely be maximized. Sounds tend to become like one another; large consonant clusters are simplified; words closer to the ends of words drop off.

Written language, beyond providing a graphic record, may have changed the language itself. Written and spoken language can evolve independently; typically, written language is much more resistant to change. In formal linguistic study, we have a tendency to study spoken, rather than written language, as it is seen to be a mere reflection of the original phenomenon. Certain forms of communication, like texting, blur the lines between spoken and written communication. Example: in written English, regular verbs (e.g. worked, walked) end in *-ed*, yet we pronounce them with a *t* sound.

Sequence of Sound Changes In a **feeding sequence**, one change indirectly gives rise to another (the first rule creates inputs for the other rule). This can happen synchronically, in a derivation. In a **bleeding relationship**, one rule removes the input for another rule.

It's often the case that a sound change occurs, but later stops applying. We know this because words that enter the language after the change do not have the change applied. Consider the word *visa*. Many Latin words with that word sequence, such as *advise*, underwent a change in the pronunciation of the vowel during the Great Vowel Shift. However, the new word does not reflect this change, because it entered the lexicon after the change occurred.

There are two basic dimensions of sound change: similarity and strength. **Strength** is difference: the more features that must be changed between phonemes, the more difficult the pronunciation. Thus, we should eliminate phonemic contrasts between sounds to facilitate pronunciation. This is known as **lenition**. The strength of consonants can be measured in hierarchies: similarity to vowels (least similar to vowels are the strongest). These are voiceless stops. (No voicing or continuacy: maximally different from consonants) Then voiceless continuant (-cont, + voice), voiceless fricatives (+ cont, -voicing). Voiced fricatives are most similar to vowels, then approximates. The **Weakening Hierarchy** describes the relative strength of sounds: generally, they have a tendency to become weaker rather than stronger (lenition rather than fortition). **GET A PICTURE**

$$[p > b > v > w]$$

(assimilation to adj. vowels)

$$[g > k > x > h]$$

(assimilation to adjacent silence: removal of phonation, occlusion)

Sound change: unconditional changes affect all sounds, regardless of environment. Example: Great vowel shift. Consider tornado vs. tacos: *taco* entered the lexicon after the shift, so it's not pronounced like *tornado*.

Tuesday, 20 September 2016: Borrowing

Assignment will be due on Thursday; it should only take half an hour.

Phonetic Processes **Lenition** is much more common than **fortition**. Whole segment processes = involve not features (like voicing, but entire segments).

Aphaeresis (initial loss); syncope (medial) apocope (final) ; note that we say *loss* rather than deletion. (Deletion refers to phonological processes and implies it can be recovered).

Epenthesis: Addition. Prothesis: initial position, anaptyxis (medial V), excresecece (medial C); paragoge (final).

Aphoresis of "silent k" in English: knife, knee, know, etc.

Syncopé of Latin /d/ in Spanish

Apocope of final -an in ModE verbs: OE findan, cuman turn in to *find*, *come*

Apocope of final syllables (and silent e) in French: Eg. Latin *amare* to French *aimer*

Epenthesis is the addition of sounds.

Assimilation and lenition tend to be lexically regular, phonetically gradual; whole segment processes tend to be lexically specific and phonetically abrupt.

In **metathesis**, two sounds change positions. It occurs abruptly, as it cannot have medial stages.

Other Changes:

Vowel shifts are more difficult to qualify. (Andre Martinet). *L'économie des changements phonétiques*.

Phonological space: symmetry, fields of dispersion, margins of security.

Vowels can undergo changes in quantity (shortening/lengthening) or quality (raising, lowering, backing, fronting, centralization, diphthongization, monophthongization).

Shifts of one vowel can lead to **chain shifts** (GEVS, NCS) or **mergers**. Mergers can initiate pull shifts or inhibit push shifts. **pull and push shift**

First Germanic Consonant Shift (Grimm's Law)

Look in the book

PIE to PGmc

$*/p, t, k/*$ to $*/f, th, h/*$ (kept place, changed manner . Went from noncont. stops to cont. fric)

$*/b, d, g/*$ to $*/p, t, k/*$ (devoicing)

$*/bh, dh, gh/$ to $*/b, d, g/$ (dropping of final h)

Thus, English *father*, *foot* v. French *père*, *pied* due to Grimm's Law.

Grimm's Law has some systematic exceptions. If the stop is *NOTE*: Asterisk indicates that it's a theoretical reconstruction rather than a proven form.

Vowels tend to arrange themselves symmetrically in the available space: this is the governing principle behind vowel shifts. (Analogous to people in an elevator who like to spread themselves out as much as possible.)

Consonants can also undergo coordinated sets of shifts.

i u

e o

a

Prothesis: addition rather than loss.

Thursday, 29 September 2016: Analogy

Previously discussed: Borrowing also affects other aspects of languages, yet phonology and other levels of grammar are often borrowed as well. Core vocabulary tends to be immune from borrowing, whereas more complex words are open to borrowing, and are more likely to change internally within the language.

English can also gain new phonemes as a result of lexical borrowings. v and f are two examples of words that used to be the same phoneme; they're now contrastive. /oy/ also from French loans. Has also influenced phonotactic constraints (English rules are relatively lax. -sh used to only appear before r, not l; this changed with the importation of Yiddish words like *shpiel* or *shlep*, shmuch).

Morphological Influence Combining forms (-able, -ation, -ist, -ity, mini-, mega-) which originally only occurred in borrowed forms, are now attached to other native forms of the words. Despite German and English's close relationship, a large amount of vocabulary differs due to **calques**, or loan translations, which English has mostly avoided and German does systematically. Borrowing in German is relatively rare. English *skyscraper* is an example of calquing, to German *Wolkenkratzer* and French *gratte-ciel*. American term was adapted as the same types of buildings that originated in the US spread to other parts of the world. Architectural phenomenon has been dispersed, but term has not (rather, it has been calqued to various native languages).

Phonological Adaptation As a word comes in lexical transfer, speakers from recipient language have to figure out how to transfer. In the case of bilinguals, native pronunciation may be kept more. (For instance, in Montreal.) In another situation, there may be no general, daily contact between languages. Or, people may see a written form of a word from another language (shares same writing system vs. different writing system). Each of these cases may be characterized by phenomena of adaptations. We (Americans) cannot read Arabic, Cyrillic, Hindi, Chinese etc., so borrowings are based on oral forms or transliteration. Often, people match foreign sounds to closest candidate among a set of native phonemes. If it's an oral input, it'll be on the basis of phonetics; if it's a spelled form, it may have more to do with other factors, such as spelling conventions of each language.

When donor language is of high prestige, there is an advantage to keeping a borrowed form as close to its original pronunciation as possible. However, there is a difference between *merLOT* and *merLOH*: comfortable path between overnativization (pompous) and undernativization (ignorant). French - pervasive apocope, word final stress language; English as a Germanic language is a initial stress language. French stress is generally preserved in North American English, not British English (garage, ballet, frontier).

jaʒ in English mapped to a of Trap /æ/, Calm /ah/, or face /ey/. Foreign *a* in English: potato, chianti, tobacco, lava, llama, mafia, nachos, nirvana, pasta, plaza, spa. (Example of nativization: all words assigned to one of the native sounds. However, note demographic shift in younger Canadians toward an American pronunciation. Imitation may be a stepping stone on the way to nativization.) /ey/ can no longer be assigned; American pronunciation tends to maximize the CALM assignments, whereas Ontarians tend to maximize TRAP assignments. Younger Canadians are shifting towards the American assignments of words.

Examples of Constraints on nativization: Phoneme Inventory, Phonotactics Hawaiian has no *t* phoneme. So it replaced it with a *k*. In Japanese, English words pose problems because of their large consonant clusters, and are therefore significantly altered.

Analogical Change Sound change is conditioned by purely phonetic factors. Yet, many changes have little to do with phonetics, such as borrowing. In Ch4, we examine changes to phonetic structure that are conditioned by non phonetic factors. Also called morphological change. Two kinds: **leveling** and **analogy**. Generally not regular, as they whole segment processes. They can be systematic, but they often leave a residue of unaffected form. They are subject to factors like word frequency and social influence. Seldom as regular as sound changes.

Sturtevant's Paradox states that sound change is regular in its operation but introduces irregularities (in morph. paradigms). leveling and analogy are irregular in their operation but restore regularity to paradigms.

Sound changes doesn't care about morphological paradigms, and as a result, some members of a morphological paradigm change and others don't, leading to **allomorphy**. Speakers may consciously rid a language of a needless alteration, thus restoring regularity.

Sound change introduces irregularity because it operates regardless of the status of words as parts of morph. paradigms, analogy restores the regularity of paradigms by sometimes reversing, sometimes extending sound change. Sound change introduces irregularity; analogy restores it.

Leveling: giving up in making distinction of a paradigm. **Analogy**: change in a morph. pattern based on a more common pattern (semantically or functionally analogous).

Reminder: Midterm exam next Thursday.

Tuesday, 4 October 2016: Analogy

Reminder: **Midterm** exam next Thursday in class. Topics include: Neogrammarian hypothesis, sound change, conditions, assimilation, lenition, various phonetic processes, segmentation processes and examples, Grimm's Law, Great English Vowel Shift, principles of lexical transfer and sound change, analogical change. Problems include: lexical transfer and sound changes to languages.

Overall, morphological change is less regular than sound change. There are two types. One is **leveling**, which reduces the number of allomorphs a form has; it makes paradigms more uniform. Leveling need not affect the entire paradigm. **Suppletion** is the deletion of parts of the morphological paradigm. Different

dialects show different types of leveling. For instance, some dialects of English standardize verbs by deleting the third singular /s/.

The **four part analogy** can be represented with $a : b :: c : x$, in which less common patterns are replaced by more common, productive patterns.

cow : kine → cow : cows

There is a distinction in English between strong verbs (which become past tense by the alteration of an internal vowel) and weak verbs, which become past with an affix, such as -d (for example, *eat* is a strong verb because its past tense is *ate*, whereas *control* is a weak verb because it takes a suffix in its past form, *controlled*). In English, there has been a drift from strong to weak. This is an example of leveling.

Another instance of simplification/leveling: in Old English, the -s suffix was only one way that nouns could be pluralized. Some words were also made plural with the -a and -um suffixes. NOTE: This leveling is not universal; many other languages (such as Russian and Lithuanian) retained complex plural forms and cases that were leveled in English.

In English, some verbs are in between strong and weak. For instance, in the verb *light*, *lit* and *lighted* are both acceptable past tense forms. Likewise, both *sneaked* and *snuck* are considered acceptable past tenses of the verb *sneak*. These are modern issues, and proof that leveling is still taking place in Modern English.

Another type of analogical change is **metanalysis**, in which word boundaries can shift. For instance, the OE 'mine uncle', in Shakespeare's time, became 'my nuncle'. *Nuncle* is still used today in some nonstandard English dialects.

In **contamination**, words that frequently occur in conjunction change to become more like one another. *Male* and *female*, taken from the French *male* and *femelle*, were originally *masle* and *femele*. However, given the frequency with which these words occur together, the forms became more similar.

In **hyper-correction**, a speaker (often motivated by sociolinguistic factors) will overcompensate for a particular idiosyncrasy in speech. For example, British speakers often drop *rs* at the end of words, but not when they are in the middle of the sentence (Thus, a British speaker would say 'I like the pastor without an *r*, but not 'the pastor is nice'. Hypercorrection occurs when they insert the *r* in words that don't have it in the first place, as in 'I like to eat pasta[r]'.

Thursday, 6 October 2016: Midterm

Midterm exam in class.

Tuesday, 11 October 2016: Midterm Postmortem; The Comparative Method

Questions.

1. *What is the Neogrammarian hypothesis and why important to historical linguists.*

Sound change is regular and operates without exceptions in specified contexts

Comparative reconstruction depends on the assumption of regularity. Can also discuss scientific method introduction/ testing hypothesis, can also discuss implication of loanwords.

2. *Give an example from any language of assimilation and lenition.*

Many choices.

3. *What is the most important different between phonetic and phonological change? Name a specific type of change that exemplifies each cat. and explain resp.*

4. *Explain how the concept of phonological space can help us understand vowel shift using Great Vowel Shift as an example.*

Vowel systems tend to retain **margins of security** among vowel phonemes so that when one vowel moves, others move too in order to maintain phonological contrast.

5. *What are the 2 main forces that cause lexical transfer (borrowing) among languages? Give two examples.*
Necessity and prestige. The latter came mostly from European languages.

6. *Explain one way in which lexical transfer can have effects beyond the lexicon of the recipient language, giving examples to illustrate your answer.*

Many choices. Phonemic inventory (introduce new phonemes); can affect phonotactics by introducing new sound sequences; can create new combining forms in the morphology (telegraphy)

7. *State Sturtevant's paradox about the relationship between sound change and analogy.*

Sound change is regular in operation but creates irregularity in morphological paradigms; whereas analogy is irregular in operation but creates morph. regularity.

8. *What is Grimm's Law? State the language or language families it applied to, the sounds it affected, and how it affected them.*

See textbook.

9. *Word relationship: realize when languages are genetically related; or if they were in contact later on. Realize direction of borrowing.*

10. *Recognize various phonetic processes, including:*

prothesis; anaptyxis; fronting; affrication; apocope; metathesis; syncope; degemination; palatalization.
NOTE: will be on the final.

11. *Understand the sonority hierarchy.*

Chapter Five: Systematic Correspondences, The Comparative Method Now that we understand the changes in languages, we can try to reverse them to go back to the earlier state of a language. We do this by making comparisons between words that we think are genetically related. Examine **cognates**, or words in different language that come from the same source and usually mean something similar and have similar phonological shapes. Look for regular systematic relationships or **systematic correspondences** that relate a sound in one language to a sound in another.

From two languages with systematic correspondences, we can attempt to reconstruct some sound from which the two forms developed in a hypothetical ancestral language. We must make sure that two languages are truly genetic (which is quite difficult). Two language sharing vocabulary could merely be indicative of lexical transfer, not a common ancestor. For instance, English and Japanese share many words, but are obviously not genetically related.

In French, consider *chambre*, *chanter*, and *camion*. Only the last word, *camion*, has a hard *k*. It is misleading and should not be considered when analyzing genetic relationships, because it is a loanword.

Consider words that mean the same thing, but have totally different shapes. For instance:

German : English → Himmel : sky

Clearly not related at all...(trivial). Words that look similar but mean different things are also out of contention. These are **accidental resemblances**. They can be much more problematic when they *seem* to satisfy both semantic and phonological criteria for genetic relations.

German : tag : (day)

English : tag

Although they are phonologically similar, their semantic meaning is not similar at all. In almost two languages, there are enough accidental resemblances for one to think there are a genetic relationship. However, these sound correspondences are not systematic.

We must look for general semantic similarity rather than exact correspondence, as semantic meaning tends to shift over time. For example:

German : tier : (animal)

English : deer

There is not an exact semantic relationship between the two, although they are cognates.

German : Bein : (leg)

English : bone

Or:

German : lust : (happiness/joy)

English : lust

English has restricted *lust* to a very specific domain!

In some cases, words not only change their meaning; they become obsolete and are replaced. Not as much of a problem in European languages, in which there is a long written history from which we can draw earlier forms, but in a language in which we only have access to contemporary daily vocabulary, there may be many words that are no longer accessible due to lexical replacement. E.g. In English, the German *magen* has been completely replaced by *stomach*, a French borrowing.

Examples of systematic correspondences:

ENGLISH	GERMAN
tongue	Zunge
tide	Zeit
tin	Zinn
token	Zeichen

Note that correspondences are systematic. Semantic meaning is not always preserved, but the forms between the two languages are always at least vaguely related.

Thursday: more of the comparative method. We will be examining the assignment in the textbook on Polynesian for the third homework.

Thursday, 13 October 2016: The Comparative Method (con't.)

Discussed last class: systematic correspondences that underlie the historical method of historical reconstruction. Usually, in establishing historical correspondences, we examine the earliest reliable data that is available.

Example of a Systematic Correspondence Latin *f* was ‘debucklized’ to *h*, and *h* was finally lost altogether. If we didn’t know that Latin *h* was ancestral, we would have to explore multiple explanations for synchronic diversity. Either there was *f* that became *h* in Spanish, or an original *h* that became *f* in Italian and French. This latter hypothesis is less likely because it requires changes in two languages rather than two. It is also unlikely because fortition is far less likely than lenition. Or, there may have been a third sound that became *h* in Spanish and *f* in French and Italian (yet, this violates the principle of Occam’s razor.) Phonetic arguments strongly indicate that it was far more likely that *f* became *h*. In this case, actual historical evidence corroborates our theory.

Similar conjugations in discrete languages can also indicate a common ancestor. For instance, see verb clauses in OE, Old High German, Gothic, Old Norse. Or in the endings of Romance verbs (Italian, Spanish, French): most end with -er or -re, and are also found in Latin (confirming genetic relationship). Every -er verb in Spanish, there's an -ar verb in Latin.

Bottom Line Systematic correspondences can be determined by corresponding morphology, phonetics, or phonology.

Historical Reconstructions Very difficult; complex; take a lot of practice. Our examples will be made up or notoriously simple.

Reconstruct unattested earlier forms, and in so doing so establish relationships between modern language forms and formulate a theoretical proto-language. This proto-language is, of course, theoretical and abstract. (It is nearly impossible to do with evidence from only one language, **internal reconstruction**.) The causes of changes become obscure over time. For instance, the alternation that caused the difference between *foot* and *feet* is now obsolete. (The final vowel that caused the alteration is no longer present.)

Reconstructions should be natural/realistic in terms of how languages evolve, drawing on knowledge presented in Ch 1-4. Favor hypothesis that are simpler and conform to our sense of what is more likely and what is less likely. For instance, assimilation is more likely than dissimilation; lenition is more common than fortition. Secondly, should appeal to natural and regular changes (**phonetic naturalness**). Third, reconstructions should be maximally simple.

Steps of Reconstructions 1: Group together putative **cognates**, or words in both languages that are semantically and phonetically similar.

2. Establish sets of corresponding sounds. To complicate this, there is often more than one correspondence for a given sound.

3. Reconstruct a proto-segment, and theorize the changes that would have led to alternate forms in different languages. These are abbreviations, or shorthands, and do not have to be precise. Initially, assume that every correspondence could be separate proto-segment. But then, see if any of the sets of sound correspondences are in complimentary distribution, which suggests that they collapse into a single proto-form. Must keep contrasting sounds separate.

Certain distributions of complimentary distributions are much more likely than others. For instance, the complimentary distribution of ch, s, θ is likely. However, the complimentary distribution of m, h, t, and b is unlikely.

Reviewed example data sent in email; constructed theoretical proto form. Next homework due Tuesday.

Tuesday, 18 October 2016: Relations Between Languages

Reminder: Language change is constant and unavoidable. Certain factors can accelerate change. Change starts in a smaller group of innovators and then (sometimes) diffuses among the rest of the society. Some changes stop before they reach the level of complete diffusion.

Physical barriers (mountain ranges, seas) prevent the spread of changes in one community from spreading to another community. Often, undiffused changes on both sides of the communication barrier lead to the divergence of the communities and the development of two separate dialects, that may become mutually unintelligible over time. If two languages are **related**, they are derived from a common source, or **proto language**. The changes between them are largely due to the different effects of language changes on each community. When two languages share a large basic vocabulary and have similarities in morphology, phonology, syntax. etc, we assume that these similarities could not have happened by chance, so the languages must be related. Must look for similarities at more than one level of structure: the more abstract the structure or more basic the vocabulary, the less likely it was the result of borrowing.

1786: Discovery of Indo-European Roots The discovery of language relatedness is one of the greatest achievements of linguistics. This began with **Sir William Jones**, who had been living in India as a British colonial. He was exposed to many languages, including Sanskrit (the historical language of Northern India, and ancestor of Hindi). He noted some parallels about what he knew about Latin and Greek and Sanskrit. He noticed an overlap of basic vocabulary that seemed very unlikely to emerge all on its own. He hypothesized that all three sprung from a proto language that no longer exists. These discoveries would lead to the discovery of the Indo-European language family as well as the application of the same method and thinking to figure out connections between different language families worldwide.

In the 19th century, many others Frang Popp, Rasmus Rask and Yakov Grimm began a formal study of German languages. Reassertion of Neogrammarian hypothesis, these laws led to the comparative method of linguistic reconstruction, which allowed majority of languages in India, Europe, Iran, to have a common ancestral language. This later led to the discovery of many other language families (Sino-Tibetan, Algonquian, Austronesian...). More similar languages form a **branch** of a language family. (For instance, Scandinavian languages are more related to each other than they are to English, although they're all Germanic.)

There are multiple ways to represent language families. The **Tree Model** associated with August Schleier. Easier to read/draw, most straightforward representation. Language "branches" are represented as literal branches, in which branches closer to the root represent changes that occurred earlier on.

Another model, the **Wave Model** is associated with Johannes Schmidt. The trees are arranged in a way to represent isoglosses to represent common innovations. Innovations are represented as circles around the languages or language groups affected. This model, unlike the Tree Model, can represent a change that spread from language X to language Y.

Shared innovations are the means of establishing subfamilies. If two families share an innovation that isn't a normal process (such as assimilation), languages must have still been in close contact at the time of the innovation, whereas languages without the innovation must have been in sparse contact. Friesian is the language. most closely related to English, which is unsurprising because of the close contact between the two communities.

Bible is a great source for direct comparison (got to spread the word of Jesus).

1,000+ language families exist. Not necessary to memorize all, but should be familiar with some of the major ones. Check email list of language families. (Hinted that you should know a family or two from different geographical areas, and say where they're spoken, provide examples of those families.) In this course, we will focus on Indo-European, primarily because it is the most well studied. Many family classifications are controversial or unclear. You should also know some well known examples of **language isolates**. Logically, we know that they must have had some sort of relationship with other languages, but we have not been able to trace such a relationship. Some examples include: Basque, Japanese, Sumerian, Zuni (Southwestern US). Effort to categorize languages has been overall successful, but there are still a number that we haven't been able to classify.

Thursday, 20 October 2016: Reconstruction Review, Models of Linguistic Change

Homework #3 Review Polynesian Languages Reconstruction.

First step: construct a table of sound correspondences with a hypothetical proto-phoneme for each correspondence and its surface in each language. You must assign a unique correspondence for each symbol: don't collapse until you've demonstrated complementary distribution.

In the sample homework provided by Prof. Boberg, at the initial stage of determining sound correspondences, examine initial and medial consonants separately. (However, this is not always necessary.) Determining which sounds are in the protophonemes depends not only on commonality (which ones are found in the most languages) but the likeliness of the changes that would result in the change, loss, or addition of a phoneme. For instance, the glottal stop occurs only in one language, but we hypothesized that it was in the proto language because loss is much more likely than epenthesis.

NOTE: reconstruction problem (guided) will be on the final.

Polynesian Languages Most successful revival of a Polynesian language is Maori in New Zealand. Hawaiian in Hawaii is much more difficult: people say *aloha*, but its use is greatly diminished.

Chapter Seven: Models of Linguistic Change Review definitions of tree and wave model (explained above). Tree model: A problem is that tree models obscure complex series of shared innovations that may cut across branches of a tree. Tree diagrams fail to show the innovations that language share even after they branch off. Tree model explicitly denies this type of contact. P

Palatalization of *k* in French: Dialect geographers show that different isoglosses for this change have a different distribution. (Opposite to the prediction of the Neo-Grammarians.) There are places where some retain the Latin *k* and others don't, in a way that cannot be predicted by phonetic environments. This type of change not unique to French; also found in Dutch and other languages.

To account for these changes, the **wave theory** was proposed. In the wave model, changes were seen as diffusing from a focal point, until they reached a limit due to geographic, linguistic, or political factors. Waves can cross dialect or even language boundaries. Wave models can demonstrate a number of isoglosses and provide a more nuanced picture of the relationship between two languages; however, they are often less straightforward than tree models. A larger bundle of isoglosses indicates an earlier split, whereas languages with fewer divisions represent splits that occur later. The wave model can be mapped to the tree model fairly easily (automatically with software).

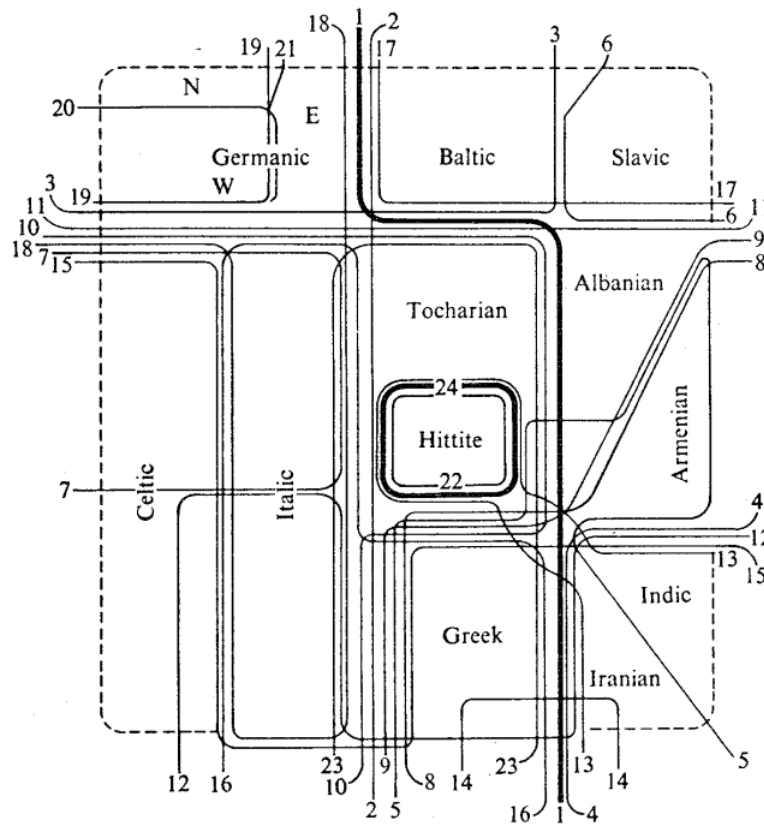


Figure 1: The Wave Model of Indo-European Languages

Given documentation of surviving cognates and similar forms, a computer can objectively create a tree model from the various linguistic data.

Dialectology Usually studied with maps that divide distributions with complimentary features. Compare geographical regions and linguistic trends. Examined by Neo-Grammarians to prove language change, ended

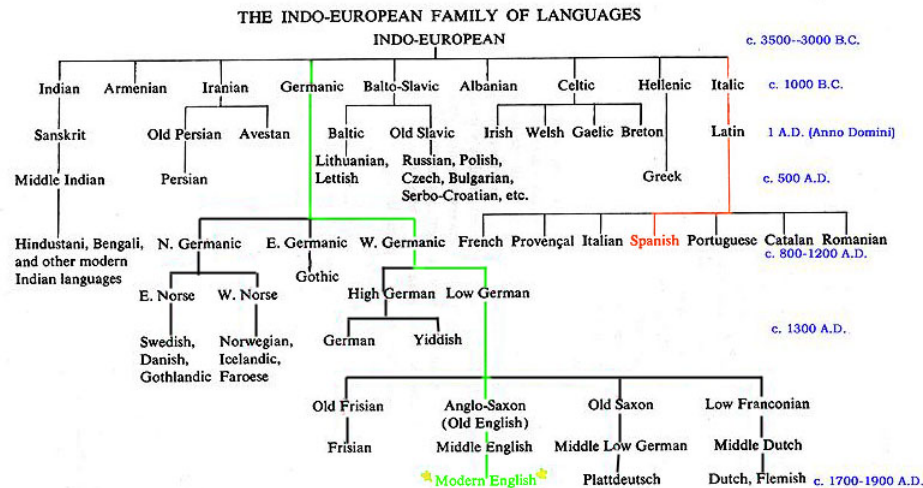


Figure 2: The Tree Model of Indo-European Languages

up somewhat disproving it.

“Rhenish Fan”: an isogloss bundle that goes across northern Germany. (Rhine river between Germany and France). Fans out to a number of different isoglosses for different words. What was a regular change in Germany was irregular in the surrounding rural areas: some forms were shifted and others were not (unsystematic).

Tuesday, 25 October 2016: Dialectology, Sociology

Rhenish Fan (con’t) : Dialectology in models of linguistic change; wave vs. tree models. In lexical forms, you wouldn’t necessarily expect isoglosses to all occur in the same place. When we see major glosses of all isoglosses, they indicate major dialectal divisions. Irregular change, steady along a river, “fans out”. Could be a result of many languages and contact in the region. River can be a barrier as well as an important channel of communication (only channel in pre-modern era). This is called a **transition zone**. Often the difference between two dialects is not clear-cut (except when communities are separated by a clear geographical barrier). This led for dialectologists to claim (contrary to Neo-Grammarian thought) that the unit of the change was the word, not the phoneme.

Dialectology Typically studied with dialect maps, which link dialects and their ling. features spoken to topographical features. Delimit isoglosses. Bundles of isoglosses with similar locations denote regional differences.

A riff arose between Dialectology and Neo-Grammarians when Neo-Grammarians found that the regular sound change they expected to find were not so regular. Modern theoretical linguistics is the descendant of historical linguistics; preference for generalization and theoretical development. On the other hand, dialectology prefers empirical methods and focused on detail.

Often, there was merit in both views. Changes were regular in some places, but changes were often unsystematic in the transition zone. (The Rhenish fan is an excellent example of this).

Another example can be seen in the variety of r versus r-less varieties of English. Dialect maps allow us to track a change from its place of origin to its later diffusion. Instance of gradual diffusion on a generational basis. Some dialectologists don’t believe in dialects as an empirical reality. If we take the conclusion of all of these communities and transition zones. We only find ‘pure’ dialects in an area well away from the transition zone, but in between those areas is a ‘mishmash’, may better be represented as varying probabilities of different features rather than discrete dialects. We can only characterize communities as



Figure 3: Rhenish fan

having the probability of using certain forms/structure. To assume that those frequencies can be extrapolated to a “dialect” is an abstraction. To explain all irregularities and something that’s NOT sound change is impossible. Both approaches have merit. Some changes are highly regular and some are not: both theories are relevant.

Sociolinguistics Only very briefly discussed in textbook. To recap: until 1960s, most researchers thought it best to ignore synchronic variation, as it was assumed that it was only ‘white noise’ that got in the way of making generalizations and theories.

Chomsky: linguistics should be the study of the totally normal speaker in a homogeneous speech community. **Competence**, not **performance** should be studied. (*Langue* more important than *parole*, to use de Saussure’s terminology). Many considered language change that was so slow and gradual as to be unobservable in real time. This was a major impediment to historical linguistics.

The belief that variation was unprincipled (something to be sidestepped) was based on the assumption that the change was random and unpredictable. This view was challenged in the 1950s (John Fisher) and in the 1960s by William Labov (Martha’s Vineyard and NYC department stores). Variation which appeared to be unsystematic was actually dependent on many social factors (like age, and social class); was actually an instance of **orderly homogeneity**. Predictable variations could account for the seemingly random variations between forms. (In NYC, high class people are less likely to use formal -ing form, this probability becomes gradually less likely as style becomes more formal. Pronunciation also depends on the formality of a particular situation.) Entire community is involved in a common, but variable, element.

This view of language was very different than Chomsky’s view: it’s clear that the ‘ideal speaker in a homogeneous community’ simply does not exist. Brings the need for lots and lots of data, as grammatical judgements cannot depend on one speaker’s impression.

Observer’s Paradox Best way to observe language is to observe it directly, yet observation causes the listener to speak in a different style (usually more formally) than they would have otherwise. The natural state of language occurs outside the laboratory. This can be merited by either collecting data in such a way that people aren’t aware they’re being studied. Or, you can distract people from their speech by asking to talk about something that’s more emotionally engaging (for instance, by asking them to talk about a scary, funny, or happy memory).

Labov found a significant difference between young and old speech styles. He extrapolated that synchronic variation was often the manifestation of diachronic change. As a change is occurring, you will have a time period in which some use the old form and some use the new form. The belief that language change was unobservable stemmed from the belief that synchronous variation was of no significance.

Labov: while not all variation indicates change, all change involves variation (for instance, some variation can remain stable for many generations without the overall population shifting from one form to another).

Social diffusion similar to geographic distribution: spread from small group of innovators out to rest of community. Sometimes, these changes can cross language barriers as well as social ones.

Apparent Time Hypothesis Merger: between aspirated and aspirated w phoneme (English). Distinction is gone in virtually all major and most dialects of English. Mapped, a higher and higher frequency uses identical forms. Our view of the 1920s depends on an assumption (the 'apparent time' hypothesis: people maintain their grammar around the same time that they acquire it). Language acquisition shuts down at puberty, so grammatical patterns tend to remain fixed, suggesting that adult grammars are fixed (so, people in their 80s would reflect the language patterns they learned when they were around 10). By looking at a succession of these groups, we can get a picture of language change over time).

Age Grading Hypothesis However, this may not be the case: Age Grading Hypothesis suggests that people may switch their language with time. People may change their language over time as they enter different phases in life (for instance, a worker may adopt a new style of speaking to fit in in his workplace; a parent may swear less to make a good impression for his children; a parent may learn new change from his teenage son).

To reconcile the two: phonology and syntax may be more resistant to change than other areas, like morphology. For instance, adults can easily acquire new vocabulary when they learn about a new concept.

S curve pattern is common in language change: they start off slowly over a long period of time, then pick up very quickly before leveling off again near the top.

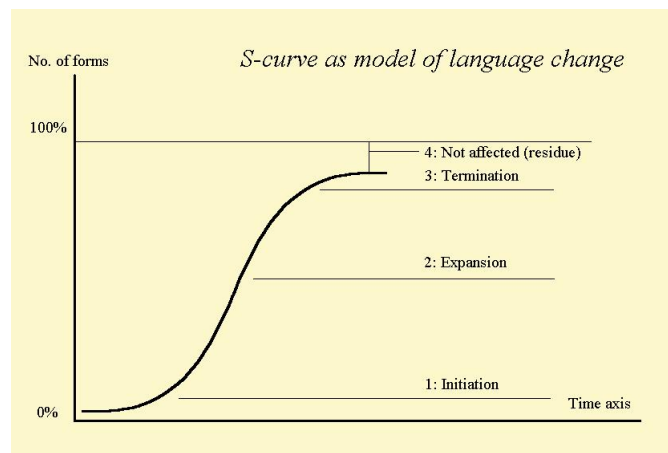


Figure 4: S-Curve common of language change

Thursday, 27 October 2016: Title

To do: Finish discussion of models of linguistic change, looking at models of lexical diffusion and the challenges it poses to the Neo-Grammarian view of regularity.

Emphasis of sociolinguistics in study of language change. Language change must involve language variation (contrary to Chomsky's idea of the ideal speech community).

Social class and sex play a role in language change. In 95% of changes, women are up to a generation ahead of men; they are usually the ones to initiate and adopt language change. Most changes also come from the higher social classes and diffuse down to the proles.

Attitudes In the comparison of Latin and the Romance languages, many think that the loss of distinction, crumbling of phonological rules, ease of articulation over fine distinction represent laziness and the decline of civilization.

Study of language use in Ottawa and Toronto suggested that it was the middle class who used the most language changes. Presumed that they were the group that had the most to gain by using prestigious forms (either from a national or local level). People on the top can only go down, people on the bottom have disengaged from society. Middle class women are the first to adopt new ling. changes, but they are also the first to drop them once they go out of fashion and are perceived negatively. However, not all changes go out of fashion.

Successful changes are associated with prestige, either local or global. These are often opposing values. Distinction between **overt prestige** and **covert prestige**, which is associated with counterculture and a rejection of mainstream values.

Orderly Heterogeneity See example from previous class, discussing use of -ing and -in forms among different classes of New Yorkers.

Example of Language Change Merging of vowels in *cot* and *caught*. In 11 years, merger went from being a minority to being nearly categorical. Once a distinction has been lost, it is impossible to restore, unless it is directly related to spelling.

Language change in Cincinnati: in this instance of a merger (caught vs. cot), we can see that there's a general decrease in distinction with lower age population, but in the middle of the population, there's a large variation in respect to whether they merge their vowels. Cannot see systematic pattern if you only examine individuals: you must examine a large set of speakers to determine an overall trend (a trend barely starts to emerge with twenty speakers, and becomes much clearer with a larger sample size).

Labov: Journal of Modern Linguistics Addresses central issue in HL: is the unit of change the phoneme (Neo-Grammarian view) or the word (lexical diffusion). Bloomfield: supporting phoneme as unit of change (Grimm's Law). Phonetic change in Chinese languages indicate an alternate theory of change, lexical diffusion: changes operate on words, only affecting a few words at a time. Under this theory, regularity is the result of the diffusion of a change through an entire lexical set. Similar conclusions have been drawn from European languages (Dutch, French: some words show palatization of k from Latin).

"Phonetically abrupt, lexically gradual" ↯ Lexical Diffusion

If you condition the phonetic conditioning enough, then you may only target a few words, so it would appear to just be lexical conditioning. Test lexical identity vs. phonetic conditioning by testing homonyms, like not and knot. Determined that it was phonetic environment, NOT lexical identity (score 1 for Neo-Grammarians).

laxings/tensings = movements across subsystems = display irregular lexical diffusion

Both views are correct - evidence for both, they are fundamentally different processes

phonetically abrupt shifts of words are lexical diffusion (eg. tone changes in

changes at phonetic feature level (voicing turned on, off, continuant on/off, these changes are lexically abrupt).

Tuesday, 1 November 2016: Internal Reconstruction

(Info in class today also in Ch. 8 of the textbook)

Thursday, 3 November 2016: Nothing

Prof. Boberg on conference; class cancelled today.

Tuesday, 8 November 2016: Semantic and Lexical Change

Note Information in today's class can be found in Chapter 9 of the textbook.

Anwaves: gathering of language variation and change researchers; many presenting apparent time studies theorizing about language change

Much of historical linguistics involves the study of sounds, although we see changes at all levels of structure.

Semantic change refers to change in meaning, whereas **lexical change** refers to change in word choice. Semantic change usually refers to lexical semantics (the meanings of words), rather than the meanings of phrases or sentences, which are usually the focus of semantics classes.

New Word Formation You should know on the exam examples of new words added to the lexicon by various means

Compounding I.e., jumpsuit, fingerboard, keyboard, bedroom. Both show general form of compounding in English. English is a head final language, so the semantically dominant element comes last.

The original sense of a compound can fade over time. For instance, take the word *jumpsuit* originally referred to a garment worn by parachuters. Now, it more commonly refers to an article of women's apparel. Not all compounds directly combine two words. Some are more literal than others, "bedroom" is literally a room with a bed, whereas "ponytail" is completely distinct from an actual ponytail, it merely shares a resemblance.

Formation process is different in different languages. For instance, French doesn't compound how English and German do, one cannot say "coucher chambre" but rather "chambre à coucher".

Derivation Unlike compounding in which both lexical items are free morphemes, in derivation a word is built out of a root morpheme plus derivational affixes. An extreme example of this in English is *antidishestablishmentarianism*, which is composed of a great variety of affixes and a root, "establish".

Some affixes are more productive than others. For instance if blork were a verb, "blorkhood", "blorkitude", "blorkishness", "blorkship", "blorkity", "bloorkdom" could all be nouns, yet some suffixes are much more likely. Some involve Old English morphemes (such as -ship and -hood) whereas others (such as the -ity) suffix are from French.

Convert to a verb? Blorkify, blorker, etc.

Conversion Also known as 'zero derivation'. Both words for changing the POS adherence without adding any affixes. Like going from the noun 'table' to the verb 'table' without changing the word at all. This is relatively easier in English compared to other languages. The most common change is to make verbs out of existing nouns: for instance, tweet, Google, access, impact, text. Even a generation ago, the use of words like access and impact as verbs was innovative and controversial. (For instance, "access" versus "gain access to").

Initialism and Acronym Initialisms are not pronounceable as words, but as strings of letter names, whereas acronyms are pronounced as letter names.

Initialisms include: French's RSVP, America's USA, CIA, FBI, Europe's EU, Canada's CRA. Many of these are strictly regional. For instance MTA (metropolitan transport Authority) exists only in New York. Acronyms include: SCUBA, CEGEP, UQAM, laser. ASAP can be an acronym or initialism, depending on how you pronounce it.

Sometimes, there is a combination of acronym and initialism, for instance, RAMQ (Régie de l'assurance maladie du Québec). The RAM part is pronounced as an acronym, but the Q is pronounced as a letter.

Various Examples of Innovation in English "Outsourcing", a new type of verb.

"Trial" and "Transition" as verbs. The opposite can also occur, in which verbs become nouns. One very new example is "ask" (as in, "that's a big ask"). Note that this innovation is not strictly necessary, as

English already has a word with the same semantic meaning (request). In a similar vein, “fail”, “reboot”, “install” are all instances of this process. Somewhat unclear why these take on, as often there is already a noun form in existence (e.g. failure instead of fail). Another instance: Many politicians use “going forward” instead of “in the future”, fairly ubiquitous. Again: a go-to X. (Go-to has become an ubiquitous noun).

Many of these terms originate in the fashion industry, or in TV shows or films.

Blends Two parts of words are blended. For instance “Agrivision” for “Agricultural exhibition”; also called **portmanteaus**, or amalgamations. Oxbridge is a portmanteau of Oxford and Cambridge. Canora is a blend of Canadian Northern Railway. Franglais, Spanglish, megacenter, miniskirt, minimall, bromance, motel (motor + hotel), skort, jeggings, brunch, blog.

Creation and Loss New roots can arise spontaneously through need, or become obsolete. Some words have no traceable etymological history, such as *boy*, *quiz*, and *pig*.

Clipping When a longer word is shortened, usually but not always at a morpheme boundary.

Clipped	Original
Stereo	stereophonic
zoo	zoological garden
bus	omnibus
gym	gymnasium
movie	moving picture
flu	influenza
mounties	mounted police
resto	restaurant
ado	adolescent
labo	laboratoire
métro	train métropolitaine
ordi	ordinateur
ciné	cinéma

Semantic Change Sometimes can be abrupt, when a word is adopted in a new context. The process of shift usually involves an intermediate process of **polysemy**, in which the word has multiple meanings as the new meaning takes hold and the old version becomes obsolete.

Many words can keep two meanings simultaneously. Classical, swing, punk, wave, goth, grunge, rap, house: all of these words have non musical senses that have developed “Mouse” now is more commonly used to mean the computer part rather than the small fluffy animal.

Shipping words were applied to air travel, whereas carriage words were applied to trains. “Cabin, pilot, captain, steward, port, starboard” were all originally used to refer to ships in the 19th century and before.

Sometimes new meanings force out old meanings (for instance ‘gay’, which used to mean happy, today is almost exclusively used to mean homosexual). “Presently” used to mean the next day in the future, rather than “now”. Meanings can become more wide or more narrow. They can shift from being concrete to being abstract, or literal to figurative, or vice versa.

Mental process of analogy can be used to shift the process from words.

Synechism, Amelioration and peturbation (in which words go up and down or

Taboo, social stigma: In English, sex and bodily fluids are taboo, so every dialect has other inoffensive words to refer to the place we go to the pee. (Loo, restroom, water closet, bathroom, washroom.) Even “toilet” is a euphemism, it refers to the small towel on which you wipe your hands. Euphemisms are rapidly replaced as their euphemistic value wears out. For instance, a succession of terms are used for people with physical problems: people who cannot walk were originally crippled, then disabled, physically challenged, differently abled.

Or, consider how we refer to illegal aliens, illegal immigrants, undocumented immigrants, foreigners, undocumented immigrants.

Black: Negro (originally not loaded term); colored; black; African-American; Person of Color. Hyperbole, litotes

Thursday, 10 November 2016: Morphological and Syntactic Change

Previously Discussed Semantic and lexical change. One area in which general public also takes interest in. Often, conclusions achieve descriptive adequacy, but it's more difficult to make generalizations in which semantic and lexical change operates. These phenomena are not as subject to generalization as phonemic or syntactic changes.

Morphological Change This information can be found in Chapter 10 of the textbook.

Morphologization Can refer to 1. the development of individually independent words into bound morphemes, otherwise known as **grammaticalization**. The other process is the development of originally phonological rules into morphological rules. Phonology and syntax can both affect morphology.

Some linguists have proposed that most bound morphemes originate in free morphemes that habitually appear in certain positions. Eventually it loses its status as an independent lexical item. These morphemes usually undergo **Semantic bleaching** and **phonological reduction**. They shift from having lexical meaning and function to grammatical meaning and function.

A classical example is the adverbial ending -ment in the Romance language, which developed from Latin *mente*, meaning "mind". "Rapidement" in French doesn't mean "of quick mind", but rather quickly. This is an example of grammaticalization and semantic bleaching.

Another example can be found in the -ly suffix, which used to come from German "leiche" body or corpse, but now is regulated to a purely grammatical context.

You also have the morphologization of syntax. "Will" originally meant "want to" (still can be seen in the noun, meaning what you want to happen after you die. In German, "I will go" means I want to go, yet in English "will" has been reduced to a marker indicating future.

Semantic bleaching in English from "going to" (which used to literally mean you were on your way to perform a task), and now has undergone semantic bleaching to just mark something as happening in the future.

Phonological alterations can be morphologized. In pre old english, "mus" and "musi" In pre-OE, "musi" became "mysi", then "mys". In modern English, these two words have become "mice" and "mouse". What used to be purely phonological became a phonological alteration that marks plurals.

Morphological Typography Three models for languages as follows:

1. Isolating, or analytic language (Chinese)
2. Agglutinating languages (Turkish)
3. Synthetic, or fusion languages (Latin, Old English) One word carries the meaning of several morphemes, that are not as easily separated as in an Agglutinating English. (For instance, it's not as easy to separate the plural and root morphemes of "mice").

It's thought that isolating languages become agglutinating through phonological reduction, which results in globbed morphemes (phonologically a single word), even though there are recognizable morphemes within the fused word. Long strings are then fused together to become fusion languages, as individual morphemes are no longer clearly distinguishable.

Old English was inflectional/synthetic. Latin *Gloria patri* takes five words in English: *Glory to the country*. Is English starting the cycle again, and becoming more isolating? What was previously encoded by affixes and suffixes are now expressed with prepositions. English has moved from a latin type synthetic language to an isolating language: phonological and morphological change has stripped the inflectional system. Yet, this information still has to be conveyed, so semantic meaning is denoted with prepositions.

Syntactic Change Irregularities of parent grammars are simplified by their children grammars. This occurs when the surface structure differs significantly from the deep structure, leading to a reevaluation of the deep structure.

Other processes include **reanalysis and extension**. Changes that begin with reanalysis are then applied in new situations through extension. **Borrowing** is also present but not common; syntactic structures are less likely than lexical items to be borrowed.

Example in English: “I have caught the fish” | “I have the fish caught” The former led to the interpretation of “have” as a past tense agent rather than a possessive, as “have” was originally perceived. Meaning of have has changed from the possessive to a syntactic auxiliary verb.

Word Order English and French are Subject-Verb-Object (SVO) language. PIE was SOV; English and French have therefore made a major transition.

Some remnants of this can be seen in French examples, for instance “je l’aime bien” the object pronoun appears inbetween the subject and the verb.

Tuesday, November 15, 2016

Internal and external factors in explaining language change will add soon

Thursday, November 17, 2016: Labov’s Article

Discuss 2007 Labov article in the *Journal of American Linguistics*.

1. How changes spread in the community in which they begin
2. How those changes spread to other communities

Labov begins the article by explaining the tree and wave models of linguistic change. Problem with tree model: does not indicate changes that are shared between languages after they split from one another. This is a main piece of evidence that linguists have used to propose an alternate model of linguistic relationships (wave model), which explains the “waves” of different innovations, and portray many features that shift between languages, rather than the harsh boundaries between communities as in the wave model.

Often, these are proposed as opposing views. Labov’s purpose is to reconcile these differences. He tries to reconcile the models by finding fundamental differences in what they represent. Tree model involves transmission of grammar. Sociolinguists have learned that data is variable, as data from parents and peers in those contexts they begin to hear variability: a phonetic rule is applied sometimes but not in all contexts. Sometimes, the children acquire grammar that is different from that of the parents. This is called **change from below**, in that it is unconscious, and they’re also from “below” on the social scale, they come from the upper working/upper middle class. This is contrasted with **changes from above**, which are conscious changes. They’re usually consciously imitated because they have social value. For instance: people from a rural town imitate the speech of people from a big city.

Children learn different speech values as variables. They can learn different forms by comparing the speech of people of different ages (e.g., their grandparents, parents, teachers, other children). They acquire not only the variable rule (sometimes people say A and some say B), but they also acquire a sociolinguistic variable and a vector (for instance, as people get younger, the amount of Y increases and X decreases). As members of the speech community, they will take that vector and move it to the next stage. This change is incremental, involving small changes in the extent, scope or specificity of a variance along a trajectory of change inherited from observations. This change can survive geographical change, as evidenced when entire communities of speakers move to a new location.

The wave model accounts best for a different type of change, **diffusion**, not from one generation to another, but from one community to another. In the communities that adopt them, these are changes from above, which usually have some sort of prestige which motivates their adoption. Caused by: conquest, colonialism, etc. Results in dialect continuum (a continuum that exists in real geographic space as well as

in the model), which small incremental differences that gradually add up. Exceptions occur when there is a clear geographical or political boundary between two speech communities, which results in a sharper discontinuity in a continuum. For example, European languages regularized as a result of official political boundaries.

Transition zones are areas in which communities have a mix of features from 2+languages/dialects. Shifting frequencies of competing features.

Transmission vs. Diffusion Labov believes that transmission is the primary source of linguistic diversification. (Children learning a modified version of their parents' grammars.) This assumes that children grow up in a harmonious environment in which there are no nonlinguistic factors that prevent them from learning their parents' grammar. Under this model, we would expect languages once diverged to continue to diverge. Diffusion is a secondary force. To distinguish these types of changes, we cannot find this info in lexical data, as it isn't related to the structural base of a community. Also don't want to examine subtle phonetic shifts; these often diffuse too easily. We must look at word boundaries, systematic relations that drive change shifting.

Past accounts suggest that structural elements are relatively immune to borrowing from other communities. Transmission and diffusion result from different types of language learning. Transmission involves children; diffusion involves adults who adjust their already functioning language model (it is primarily adults who have contact with external communities, not children). It was adults who participated in travel, trade: they are the prime agents in diffusion, which occurs after their native grammar is established. That influence must then be reconciled with native grammar, at a time when they're too old for learning new linguistic patterns. Structural elements are not likely to be diffused because it's hard for adults to learn them (rules and constraints are harder to learn than superficial changes to pronunciation or new lexical items).

Example 1: Norwegian Map of Norway. Isoglosses indicate boundaries. Show that changes are both spreading to other geographical areas and intensifying in the regions in which they began

Example 2: Persian Phonetic shifts in Tehran. Pronunciation shift more common among the lower class in Tehran (where it originated), yet more common in the upper class in other cities in Iran, as those are the people who have more contact with the capital city.

Example 3: NYC English Pronunciation (tense vs. lax) of /a/ in NYC english; not entirely explainable by phonetic environment. Due to function/content word difference (structural constraints). Many complex exceptions that cannot be explained by phonetic constraints. By speakers, the tense and lax versions are considered separate phonemes.

Distinction of splits vs. mergers: Splits are very difficult to acquire (even for adults in their native language); mergers are more easily adopted.

Family tree style. vs. wave style ; superficial features like lexical items or phonetic shifts (isolated elements). Adults can understand isolated elements but not complex systems of constraints.

Tuesday, November 22 2016: Distant Genetic Relationships

Winter Has Come It is snowing.

Final exam Scheduled for the last day of exams, Tuesday Dec. 20 at 2:00.

Last week, we discussed transmission and diffusion (Labov). The information in this class can be found in chapter 14 of the textbook. On Thursday, we will discuss the last articles in the course pack.

This chapter looks at the genetic relationships going back further in time. We examine the possibility of genetic relationships that go back much farther in time than the Indo-European languages (which spread approx. 6,000 years ago). Human language may have had only a single origin in a single place at a single point in evolution, and the hundreds of language families ultimately had a common origin. Scholars disagree as to whether these distant genetic relationships can be demonstrated with comparative linguistic data.

Nostratic

Proto-World Take 2+ established families, like PIE and Uralic that are found in geographically distinct regions, and try to show that they came from a common source. The logical extension of this ideology is the theory that all languages came from the same source. This topic is usually discussed in archaeology and anthropology rather than historical linguistics; it has much more to do with physical human development rather than based on data. The genus *homo* separated from related species of apes about 5 million years ago. 100,000 years ago, *homo sapiens* (modern humans) could be identified. Full behavioral modernity begins around 50,000 years ago; we assume that language appeared around then. Human of this time had changed physiology, including vocal cords similar to those of modern humans.

All of humanity can likely be changed back to one migration out of East Africa. It's possible those migrating humans already had language at that point. Open question: what happened to Neanderthals? Wiped out by migrating *homo sapiens*, or genetically merged with them. If language occurred after initial migration, then there would be several origins of human language.

Ultimate question in linguistics: how far back can the comparative method go? Can we find traces of the original proto-language common to all humans? This question has ramifications about anthropology, development, human culture. Noam Chomsky considers universal grammar to be part of the human language capability. The comparative method has been successful at recovering the immediate ancestors of most of the world's languages; PIE is only one example of this. There are 2 dozen in the "old world", and hundreds in the Americas. Are some of the proto languages ancestors of an even older proto languages? Can we uncover relationships to uncover the pre-proto language?

A few linguists believe that this is possible, and that they have done it. Most mainstream linguists, like Lyle Campbell, disagree that it is not possible. One suggestion is that many Southeast Asian languages go back to **Austic**, and that many others (as in Papua New Guinea) go back to **Indo-Pacific**. Basque and the North Tibetan Sino families all go back to a mega family called **Dene-Caucasian**. Note: the evidence linking these languages is very tenuous; many linguists do not think there is sufficient evidence to link them.

Nostratic First proposed in 1903 by H. Pederson, a Danish linguist. Proposed to include Proto Indo-European, Uralic (Northern Europe), Afro-Asiatic (Semitic languages like Arabic and Hebrew), Altaic (Western Asia), Kartvelian (Caucasian languages) It was later elaborated by two Soviet linguists, A. Dolgoplosky and VM. Illit-Sritich in 1960s USSR. Another well known advocate was ex Soviet V. Shervoski, who later emigrated to the US. Work on Nostratic has followed typical lines. It has a proposed lexicon (700 words; as opposed to 2,000 in PIE), a grammatical system, and phonological changes leading to various proto-languages. About twice the time depth of PIE, about 10-15 thousand years ago. Language offers clues into culture: only clues for wild, not domestic animals, suggesting that they were a hunter gatherer society rather than an agricultural one.

Amerind A proposal more controversial than Nostratic. These correspond to three waves of settlement.

He used **mass comparison** to set compare entire lexical sets at once to identify potential cognate pairs and draw a rough conclusion on the potential for genetic similarity. Very different from meticulous approach of HL that rely on systematic correspondences. One of the three mega families suggested by Joseph Greenberg of Stanford. In 1987, he wrote a classification of Native American languages into three major families, of which Amerind was one. He was a pioneer of **Typology**, a field of identifying language features based on different features.

Native American Languages Most recent to arrive are Eskimos, then Nadene people (Languages in NW Canada as well as SW US, like Navajo and Apache, west) and Amerind, which is all the rest, from Sioux to Quetchua.

African Language Families Greenberg proposed four families of African languages: **Afroasiatic**, **Nilo-Saharan**, **Niger-Congo**, and **Khoisa**. These four classification of languages is now widely accepted, if not

by the author of our textbook.

Criticism of Megafamily Reconstruction Most linguists who are skeptical of such propositions do not disagree that there are pre proto languages, but rather that we can find out anything of use about them due to paucity of available data. Even the input data is hypothetical, meaning that errors will multiply themselves in the construction of pre Proto languages. Languages change so fast in such a short amount of time (consider the change from Old English to Modern English), that trying to reconstruct languages over such a long time frame would be impossible. True cognates should have changed beyond recognition; non cognates could have changed to resemble cognates. If any cognates *did* survive, they would be no more than chance (which accounts for 5-6% of similarities between two unrelated languages).

A second criticism of the construction of Greenberg's megafamilies. The number of cognates claimed for Nostratic does not rise above the ones that would appear by chance, unlike PIE. Amerind construction relies on comparison *en masse* rather than more rigorous methods. Borrowings, too, are not discovered by mass comparison and may falsely be identified as cognates.

Construction of PIE also began with mass comparison, and went through careful analysis of systematic correspondences later. It's possible that Greenberg's work is just the beginning of what may lead to a more rigorous reconstruction of megafamilies later on. Mass comparison may be sloppy, but it's a good start.

Origins of Language **Monogenesis** suggests that language evolved only once in Africa. A less extreme version of this hypothesis is **polygenesis**, which suggests that human language originated at multiple places at different times.

Proto human suggests that all modern languages go back to a single ancestor. This need not imply monogenesis, as other languages could have died out. This is very controversial, and it's unlikely that we'd be able to construct anything meaningful about Proto world. People do not resent the possibility of proto-world, but rather the possibility that we could draw anything meaningful from it. Data going or reconstructed from that far back would essentially be worthless; even basic word forms would have changed beyond recognition.

Linguistic Diversity Renfrew believes that language isolates and centres of great linguistic diversity (like Papua New Guinea) are from the earliest settlements out of East Africa. Indo Europeans are less different from one another because they are the result of a relatively recent dispersion of people across a wide people, resulting in homogeneity from Ireland to India. Areas with more linguistic diversity indicate that settlement occurred earlier.

Quantitative Approach to Reconstruction Some favor a quantitative approach to language reconstruction. The more vocabulary shared between two languages, the closer their relationship. A simple percentage will indicate the closeness of a relationship. As Campbell says, there are many problems with this method: what counts as "shared"? Does something count as a cognate if it exists only on an obscure level of the lexicon? Also, how can you reliably distinguish cognates from borrowings?

More simple than simple lexical comparison is **glotto chronology**. Depends on a simple set of vocabulary (p.449 of textbook), words that are least likely to be replaced by borrowings. We can compare unknown cases to known cases to estimate the connectedness of two languages, assuming a constant rate of lexical replacement. 20% every 1,000 years is the standard rate of replacement. A problem is: how constant is the rate of change? Not very. For instance, Icelandic is very conservative, as it's very geographically isolated. English, on the other hand, is very innovative. It's had frequent and intense contact with other languages. Proposed rates of change cannot account for external circumstances that may hasten or retard the rate of language change. Rates of change differ greatly.

Campbell argues that the only reliable method of comparison is shared innovations. There is a possibility that two languages happen to undergo the same change. It's also difficult to distinguish shared archaisms or shared innovations. While both indicate a genetic relationship, mistaking shared archaisms may make two languages seem like they developed together for longer than they actually did.

Thursday, 24 November, 2016: Distant Genetic Relationships (con't)

Focus of discussion: 1991 Articles in Course Pack

Phillip Broad, *Scientific American* Examines origins of language from hominids to Chomsky's language instincts to the link between genetics and language. The link between language and biological evolution is a new field; previously, linguists had no data to prove or examine their theories. He anticipates the topic of linguistic paleontology.

The main theme he discusses: how far back can we construct a linguistic past? What are the merits of the various scientific processes of reconstruction?

See graphic in article: controversial mashing of many languages into very few families

Lyle Campbell (author of textbook) is cited in the article as an expert in Mayan languages, as is the typologist Joseph Greenberg. Campbell accuses Greenberg of sloppy work that will lead other scholars to waste their time on unscientific work, or "idle speculation". Greenberg states that his methods of mass comparison have the advantage of having a wider scope, enabling specialists to see relationships they would have missed otherwise.

Detectors point out previous errors that have occurred as a result of extrapolating beyond the data. Nostratic is a less radical proposal, closer to the methods supported by more traditional reconstructive linguists. It may have been spoken during the transition from Hunter-gatherer society to agrarian society.

Certain vocabulary words can offer clues into human development. For instance, a society that has a word for "dog" has begun domestication of wild animals.

Anthropologists' "paleoindians" are the Amerind language group. While genetics has no effect on language directly, both languages and genetics change when a population is separated. Thus, differences in genetics are expected to correspond to differences in language, with the exceptions due to interbreeding of previously separated groups, among other factors. Languages can "interbreed" (combine) as human genes do, but it's more common for one language to supplant each other. For instance, the English language is diffusing worldwide in a way that's not connected to people with English ancestry, e.g. black Americans who are native speakers of English.

Chomsky: Are humans born to speak? Article discusses physical changes between humans and chimpanzees. Humans and chimps are very genetically similar, yet they have critical physiological differences with respect to vocal structures. Humans' erect posture caused realignment with head and body. This led to some disadvantages, such as not being able to breathe while we swallow, yet we can produce complex speech.

Robert Wright: Quest for the Mother Tongue (1991) He points out an interesting cultural dimension of issues we've been discussing. This takes place during the collapse of the USSR. It is not coincidental that the debate is between ex-Soviet and European Americans, reflects divide in Soviet and western science divided by iron curtain. After USSR collapse, scientists realized that they had drawn many different conclusions.

USSR's "evil empire" included many ethnically distinct neighbors, including Georgians, Turks, etc. They tried to promote an agenda of the central Russian authority. Idea of shared origin of human language connects with Soviet political interest (advancing the unity of its people), advanced Nostratic hypothesis. Nostratic or even Proto-world's existence would also be a sign of defiance against Hitler, seen as a sign of equality and unity among different ethnic people.

Languages often interbreed through contact, and leave no remains of previous languages, leaving reconstruction extremely difficult. Reconstructing linguistic prehistory is made more difficult by the fact that we have no traces of these interbreedings that must have occurred. This makes it near impossible to prove the existence of megafamilies over a long time span.

Armenian Linguists have proved that Armenian is an Indo-European language. There is proof that Armenian verb is related to Latin "three", despite the fact that the surface forms would not be immediately

recognizable as, say, French *trois* and Spanish *tres*. There is a rule in Armentian: before an r, a t will drop, and no word can start with an r. Because of these rules, the t and r of *tres* were dropped. So, all that has to be explained is the final k sound. This example is used as a cautionary tale as the need for knowledge of sound changes in reconstruction. What would happen to the forms of proto world? All traces of regular change would become totally unrecognizable, making any reconstruction of proto-world totally worthless.

North American Migration Amerind group came first. Later groups (Nadene, Eskimo) were prevented from further migration by other groups. This hypothesis is supported by genetic similarity between the three various groups.

Some believe Nostratic and Amerind are genetically related (go back to before the split). These Eurasians, along with southeast Asians and Africans, make up major ethnic groups.

Scientific Attitudes Lumpers vs. splitters: Those who draw overarching conclusions vs. those to examine narrow data. Risks: if you take risks, you may make a foolish and embarrassing mistake. If you don't take risks, you may miss out on scientific discoveries.

Next time: Linguistic Prehistory Not to be discussed: Ch 15 in Txtbk