

Seminar 1

Introduction

Contents

A	Historical linguistics, diachronic linguistics and language change	1
B	Questions to ask	4
C	Languages, grammars and acquisition	5
D	Review	8
E	Further reading	9

A Historical linguistics, diachronic linguistics and language change

§1 This is an introductory course in **historical linguistics**. What is historical linguistics? Here it is useful to make a general distinction between synchrony and diachrony (de Saussure, 1916):

- **synchronic linguistics:** the study of the state of a language at a given point of time (“frozen in time”, as it were)
- **diachronic linguistics:** the study of language over (long periods of) time, the study of language change (or lack thereof!)

Now, historical linguistics comprises both

- the synchronic study of *past* language states (e.g. the study of Old English syntax, as it was spoken and written down in 900 CE), and
- diachronic linguistics (e.g. the study of how OE syntax turned into Modern English syntax)

This course will focus mostly on diachronic linguistics and language change.

§2 That languages change is a *fact*. The reasons and causes of language change are not fully understood, but change can be seen everywhere, and there are theories about why this should be the case (which we will look into in more detail later). Some examples of linguistic change follow.

§3 Here are four ways of saying one and the same thing:

- (1) and ne gelæde þu us on costnunge, ac alys us of yfle

- (2) and lede vs not in to temptacioun, but delyuere vs fro yuel
- (3) and lead vs not into temptation, but deliuer vs from euill
- (4) and do not bring us into temptation, but deliver us from the evil one

This is Matthew 6:13 in the Bible,¹ and the versions are as follows:

1. Wessex Gospels (Old English, OE), circa 990
2. Wycliffe's Bible (Middle English, ME), circa 1382–1395
3. King James Version (Early Modern English, EModE), 1604–1611
4. Christian Standard Bible (Modern English, ModE), 2017

So, each version of the verse is “in English”, but over a timespan of a thousand years. What has changed when we proceed from the Old English (Anglo-Saxon) version to the Modern English version?

1. **Vocabulary.** Perhaps the most evident change concerns the lexicon, i.e. vocabulary. The OE version employs some words which now look totally foreign (*ne*, *gelæde*, *þu*, *costnunge*, *ac*, *alys*). Some of these turn out to be more familiar on closer inspection: so *gelæde* is actually the same word as ModE *lead*, and *þu* is the same as the Early Modern English *thou* that you see in Shakespeare (but that has since been lost to *you*). However, words like *costnunge*, *ac* and *alys* have been lost in those meanings (however, *costnunge* is related to *cost*, and *alys*, infinitive *alysan*, is related to *release*). The word *ne* is a negation word which was lost during the Middle English period in a very interesting process.
 - (a) So where did the new words come from, that replaced the old ones? *Temptation* and *deliver* were both borrowed from French (and ultimately from Latin) – there was extensive contact between English and (Norman) French in the time period between our OE and ME examples.
 - (b) On the other hand, roughly half of the words are the same (in fact, if this wasn't the case, we couldn't recognize the passage as English in the first place). So: *and*, *us*, *on*, *of* and *yfle* continue to be English words (though ModE uses *into* and *from* instead of *on* and *of* in the above example). Important point: *and*, *us*, *on* and *of* are all **function words**, i.e. they do not have a lexical meaning but rather serve to indicate relationships between words, roughly speaking. Function words are much less likely to be removed by processes of language change, though there are some exceptions (e.g. the English pronoun *they* was borrowed from Scandinavian, replacing the original Anglo-Saxon pronoun *hie*).
2. **Syntax.** The syntax has also changed dramatically. Let's look at the OE verse in more detail:

- (5) and ne gelæde þu us on costnunge, ac alys us of yfle
and not lead you us to temptation, but deliver us from evil

¹I've picked a biblical example because Bible translations represent one of the very few genres of writing which have been practiced for a long enough time for strict comparisons to be possible.

It has NEG-V-(S)-O order (*ne gelæde þu us*). By ME, this has turned into V-O-NEG:

- (6) and lede vs not in to temptacioun, but delyuere vs fro yuel
and lead us not into temptation, but deliver us from evil

Finally, Modern English overthrows this order once more by adopting what is known as **periphrastic do**:

- (7) and do not bring us into temptation, but deliver us from the evil one

Here the order is *do*-NEG-V-O. So we have gone from NEG-V-O (without *do*) through V-O-NEG to NEG-V-O with *do*. This is quite puzzling: why would any language want to do this?

3. **Spelling.** The way words are written down has also changed (e.g. *us* vs. *vs* or *temptacioun* vs. *temptation*). This is not linguistically interesting *per se*, unless it tells us something about e.g. sound change, etymology or morphology. The spelling of *evil* in the OE verse is a good example: it is *yfle*, with a final *e*. In most manuscripts from the period, you find *yfel* or *yfl*, without final *e*. However, when the word follows a preposition such as *of*, the final *e* appears. This is, in fact, evidence of a **case system** in OE, not unlike the case system of German (in fact, historically speaking it is one and the same system, since both English and German descend from a common ancestor language). The preposition *of* requires the following noun to be in the dative case, which in OE for a word like *yf(e)l* meant adding a final *e* (cp. German *das Böse* but *von dem Bösen*).
4. **Sounds.** The way words are pronounced also changes over time (and also varies tremendously across geographical regions, even across short distances – just compare the English spoken in Liverpool and Manchester!). But how can we know how a language which is no longer spoken by anyone, and of whose speakers we have no recordings, sounded like? This is a really tricky question, and one that we will return to. For now, however, let's focus on the word *temptation* in the above examples. Notice that in ModE, the last vowel in this word is the schwa /ə/. Now, it is very unlikely that when ME scribes wrote down the word *temptacioun*, they would have used a sequence of *three* letters <iou> to encode just *one* sound. So we can hypothesize that in ME, the final vowel in *temptacioun* was something more complicated than a schwa, and thus that sound change has occurred. This is also supported by the fact that Modern French has the full vowel /ɔ̃/ with the glide /j/: *temptation* /tãtasjɔ̃/. And similarly in Spanish: *temptación* /tenta'θjɔn/. It is likely, then, that the reduced vowel /ə/ for the last syllable in *temptation* in Modern English is a relatively recent English innovation and wasn't part of the Middle English pronunciation of this word.



The funny letters here, such as ə and ɔ̃, belong to the International Phonetic

Alphabet (IPA) and are used to transcribe speech sounds. You will probably have seen them before but, if you are like me, may have forgotten (some of) them. 😊 We will later take a more detailed look at them, because good command of a phonetic alphabet is essential if we want to describe historical sound changes.

A related point: when I enclose something in slashes //, it is a **phoneme** (a speech sound that is used to distinguish words from each other in a language), or a sequence of phonemes. When something is enclosed in square brackets [], I mean the actual realization of this sound in some context (also known as a **phone**), or a sequence of such. When something is in angle brackets <>, what is meant is the written form (a **grapheme** or a sequence of them). To take a (pseudo)random example, in Modern English the phoneme corresponding to the letter <o> in the word *for* is /ɔ/, but in rapid speech (when the word is not stressed) it is actually realized as the schwa, [ə]. I'll return to all of this in more detail later, so don't worry if this sounds like gibberish to you now.

B Questions to ask

§4 In this course, and in historical linguistics more generally, we are interested in finding answers to the following kinds of questions:

1. What was some past state of a language like? (E.g.: did Old English have an indefinite article like ModE *a/an*?)
2. How did a language change from state S_1 into state S_2 ? (How did Latin turn into French?)
3. Why did a language change from state S_1 into state S_2 ? (Once we understand how French evolved from Latin, we can ask what the motivating factors were.)
4. Why do different languages take different historical paths of change? (Why did the Latin word *tentatio* /tenta:tiɔ:/ (as far as we know) turn into *tentación* /tenta'θjɔn/ in Spanish, but *tentation* /tãtasjõ/ in French and *temptation* /tɛmp'teɪʃən/ in English?)
5. Why does a certain change happen in one language but not in another? (Why did Estonian lose word-final vowels but the closely related Finnish didn't, or hasn't so far?)
6. How can languages be classified into families? (Why is it correct to say that French and Spanish are more closely related than French and English?)
7. How much faith can we have in historical reconstructions and classifications? (It is generally assumed that the words meaning 'two' in Indo-European languages (e.g. English *two*, French *deux*, Spanish *dos*, Italian *due*, Albanian *dy*, and so on) descended from a Proto-Indo-European numeral *dwo*. But there is no direct evidence for this; the word has to be reconstructed instead. How much faith can we have in it?)

8. Are there universal tendencies of language change? Are some kinds of change more common than others? (E.g. many languages have developed a definite article (like English *the*) from a demonstrative pronoun (like English *that*), but no language seems to have developed a demonstrative pronoun from a definite article. What does this mean?)
9. Can language change be modelled mathematically or with computer simulations?
10. What is the relationship between language change and language acquisition?

C Languages, grammars and acquisition

§5 Any science needs to try and define its object of study as precisely as possible — so it is with historical linguistics. So it seems like the first question we need to ask is: what is a language? This turns out to be a surprisingly thorny question. Here's some food for thought.

§6 First, let's think about the problem purely synchronically. Take Modern English, and take a look in any modern dictionary, and you'll find that the word *first* is to be pronounced /fɜ:(r)st/ (the r-sound is included if the variety is **rhotic**, e.g. most American accents, Irish and Scottish English, and so on). From this, it looks like we can make the following statement:

(8) English has the vowel /ɜ/.

(The same vowel appears in words such as *bird*, *nurse*, *blurb*, *kerb*, *third*, *courtesy*, and so on — it's not just *first*.) This seems like a very simple statement and, indeed, linguistics textbooks and research articles are filled with similar statements. On closer inspection, they turn out more problematic than you might at first think.

The problem is that saying “Language X has Y” is ambiguous. Compare it with this biological statement:

(9) The domestic cat (*Felis silvestris*) has fur.

(Here “the domestic cat” is a collective noun — it doesn't refer to any particular cat, but to the whole species). Now, it seems that (9) must mean one of the following two things:

(10) Every individual that is a cat (i.e. belongs to the species *F. silvestris*) has fur.

(11) Some individuals that are cats (i.e. belong to the species *F. silvestris*) have fur.

The problem here is that, on the one hand, (10) is false, and on the other, (11) doesn't say what we intended to say.

(10) is false since (a) the breed of cat known as Sphynx does not have fur, and (b) in principle you can shave the fur off of a cat.² So it is not true that every cat has fur.

²This is just a thought experiment — I don't endorse mistreatment of animals!

On the other hand, (11) is not what we really mean when we say that the domestic cat has fur. The statement is too weak. We want to say something like: “Ordinarily, unless certain extraordinary things happen (such as a cruel shaving, or the cat happening to be a Sphinx instead of a conventional cat), cats, on the whole, have fur.”

§7 Now, the linguistic case is exactly analogous. We can understand (8) in one of two ways:

- (12) Every speaker of English has the vowel /ɜ/.
- (13) Some speakers of English have the vowel /ɜ/.

Statement (12) is false:

1. Liverpool English (also known as Scouse) does not have the vowel /ɜ/. In this dialect, the vowel has merged with the vowel /ɛ/, and *first* is pronounced /fɛ:st/.
2. Ghanaian English (spoken in Ghana) never had the vowel /ɜ/ in the first place. When English began to be spoken in Ghana (due to colonization), the locals never adopted the /ɜ/ vowel but instead used the vowel /ɛ/ which was part of the vowel inventory of most of the local African languages.³
3. And these are just two examples — they could be multiplied. For example, in Nigerian English the vowel in question becomes /a/ and *first* is pronounced /fast/.

On the other hand (13) does not have the right force, since we were looking for a generalization.

§8 The problem can only be overcome when we realize that

1. a language only exists as some sort of conglomeration of several individual languages spoken by individual people;
2. a language usually has a number of varieties, which differ from each other in some ways;
3. within these varieties, however, speakers behave relatively uniformly.

Point 1 ought to be self-evident (if all the speakers of a language die, the language itself surely dies). Point 2 is challenging, since it is often difficult to decide where the boundaries between different varieties or dialects go. Point 3 is equally challenging: sometimes it is difficult to decide whether a particular speaker speaks a particular variety.

§9 All this philosophizing brings us to a very important point: the distinction between E-language and I-language. **E-language**, as Chomsky (1986) defined it, is language in the external sense, i.e. external to the mind: English, French, Finnish, and so on. **I-language** (I as in internal) refers to the mental representations of a

³Look for speeches of Kofi Annan, former Secretary General of the UN and a native of Ghana, on Youtube to hear this in action!

speaker. We can summarize the relationship between E-language and I-language as follows:

- (14) An E-language is a population of utterances drawn from a number of I-languages.

I-languages are also called **mental grammars** or just **grammars** (warning! – this is *not* the same as a grammar of English/French/Finnish that you can check out from the library).

(14) has the following important consequence: an E-language can only change if there is change in the I-language. For English word order to change, for example, a change must happen in the mental representations of speakers of English. The usual mechanism for this to happen is when a generation of children somehow “mislearn” the grammar of their parents (Figure 1). This is sometimes called the “Z-model of language change”.

§10 A crucial point about the Z-model is that *there is no direct mode of linguistic transmission from generation to generation*. Every new generation of speakers must learn the grammar (the I-language) from scratch, as it were. It is not transmitted in the genetic material, for example (if this was the case, then it would be impossible for a child of Chinese immigrants to the USA to learn to speak English as his or her first language). It is also indirect in another sense, which is that language learning is an extremely difficult task. Parents, for example, don’t explicitly instruct their children that “the proper word order in our language is VO” – rather, the child must discover this on his/her own.

Except that sometimes the child doesn’t, and postulates OV order instead – and this is where language change can happen. If, *for any reason*, generation $t + 1$ of speakers internalizes a different mental grammar than generation t , language change has occurred. This will then be represented in the linguistic output (the E-language) of generation $t + 1$.

§11 Change need not happen in initial language acquisition (at an early age). Sometimes, when people move to a different geographical area as adults for example, their linguistic production may change. (Thus a speaker of British English, having

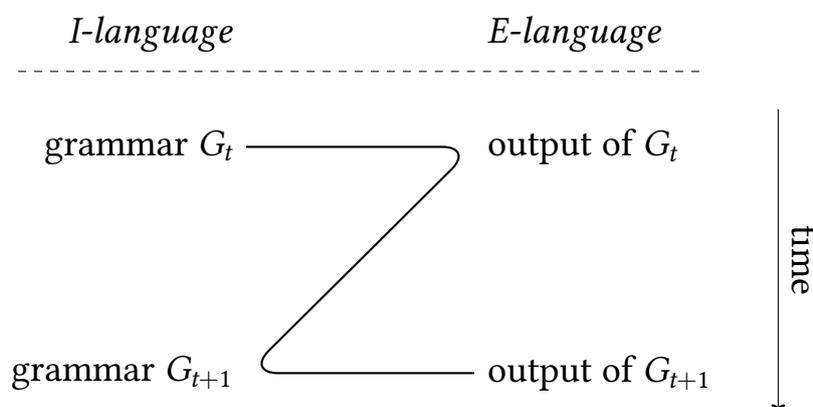


Figure 1: The “Z-model” of language change.

lived for a long time in America, may start approximating the vowel qualities of American English.)

§12 To summarize: language change is, in a sense, an epiphenomenon: it is just evidence that mental grammars of speakers have changed, and it is these mental grammars that we need to try and study. This is extremely difficult, of course. It is already extremely difficult in synchronic linguistics, and doubly so in historical linguistics, when you try to recover the grammars of people who no longer live and whose language only remains in a scattering of written documents. It can be done, however.

§13 In this endeavour, it is crucial, however, that contemporary synchronic linguistics and diachronic linguistics are in constant connection with each other. To understand how and why languages change or changed, we need to have solid theories about what languages (I-languages) are, synchronically speaking.

On the other hand, some synchronic facts can only be explained by diachronic means. Here's one example, after Lightfoot (1999): in Modern English, adjectives usually precede the noun they modify, thus *small book*, *tall woman*, and so on. A small set of adjectives, however, are postnominal. *Ablaze* is one of these:

(15) He photographed the buildings ablaze.

(16) *He photographed the ablaze buildings.⁴

This is puzzling from a synchronic perspective, but becomes understandable from a diachronic point of view, when we realize that *ablaze* in fact derives from a preposition phrase in Middle English, *on blase* ("on fire"). Since preposition phrases must go after the noun in English, the adjective *ablaze* does so, too, as a historical relic. Hence, here we have a historical explanation of a synchronic fact.

D Review

§14 After this seminar, you should be able to explain what the following terms mean:

synchronic linguistics	E-language	Z-model
diachronic linguistics	I-language	
function word	mental grammar	

E Further reading

⁴The asterisk (*) preceding the sentence indicates that the sentence is ungrammatical, i.e. not accepted by native speakers of English.

§15 The best introductory textbook on historical linguistics, in my opinion, is Hale (2007). On a more advanced level, Lass (1997) is indispensable. Classic choices for introductory reading would also include Campbell (1998) and McMahon (1994).

References

- Campbell, L. (1998). *Historical linguistics: an introduction*. Cambridge, MA: MIT Press.
- Chomsky, N. (1986). *Knowledge of language: its nature, origin, and use*. New York, NY: Praeger.
- de Saussure, F. (1916). *Cours de linguistique générale*. Paris: Payot.
- Hale, M. (2007). *Historical linguistics: theory and method*. Malden, MA: Blackwell.
- Lass, R. (1997). *Historical linguistics and language change*. Cambridge: Cambridge University Press.
- Lightfoot, D. (1999). *The development of language: acquisition, change, and evolution*. Malden, MA: Blackwell.
- McMahon, A. M. S. (1994). *Understanding language change*. Cambridge: Cambridge University Press.