FACT AND FICTION IN
HISTORICAL LINGUISTICS

John Charles Smith,
University of Oxford
(St Catherine’s College
& Research Centre for Romance Linguistics)

COURSE NOTES

Note:
Bibliographical references throughout are to the section on ‘References and Bibliography’
at the end of these notes.
0. Introduction and suggestions for general reading

The general issues involved in language change are discussed from a rather basic viewpoint in Aitchison (2000), and, more thoroughly, in a dated but still worthwhile book, by Bynon (1977). Useful notions can also be gleaned from Hock (1991), Lehmann (1992), and Jeffers & Lehiste (1979). Three good fairly recent introductions to historical linguistics are McMahon (1994), Trask (2007), and Campbell (2013). Also interesting is Crowley & Bowern (2010) — written by an Austronesianist and an Australianist, it eschews the Eurocentric approach of most introductory works. A more advanced theoretical introduction is Ringe & Eska (2013). Lucid discussion of sources and methods and the problems associated with them can be found in the following more advanced works: Lass (1997), especially chapter 2: ‘Written records: evidence and argument’, pp. 44-103; Labov (1994), especially chapter 1: 'The use of the present to explain the past’, pp. 9-27; and Fleischman (2000). Good specialized dictionaries of historical linguistics are Trask (2000) and Campbell & Mixco (2007), and relevant terminology can also be found in the standard general dictionaries of linguistics, such as Matthews (2014) and Crystal (2008), and encyclopaedias of linguistics, such as Frawley (2003), Crystal (2010) and the monumental 14-volume compilation of Brown (2005).

1. Theoretical Preliminaries

1.1. Two ‘axes’

**synchrony** (the study of) a phenomenon at one point in time

‘vs.’?

**diachrony** (the study of) a phenomenon through time

A topical problem... In recent years, diachronic linguistics has returned to prominence, after almost a century of relative neglect (largely the result of Saussure’s understandable but over-zealous distinction between synchrony and diachrony in the *Cours de linguistique générale*, published posthumously in 1916).\(^1\) This trend has appeared in various guises. Work on grammaticalization\(^2\) and subjectification\(^3\) has led to the suggestion that grammars should be viewed principally as dynamic systems; sociolinguistics has stressed the link between variation and change and has canvassed ‘the use of the present to explain the past’;\(^4\) attention has been drawn to possible analogies between language change and biological evolution;\(^5\) claims about affiliation have led to major re-evaluations of the aims and techniques of linguistic reconstruction;\(^6\) and the application of contemporary formalisms to language change has not only elucidated the diachronic processes involved, but has also enlightened us about the formalisms

---

\(^1\) See Lass (1997:12): ‘[T]he synchronic vs. diachronic division is not a sharp cut but a shifty no-man’s-land, and the action there is often quite as interesting as what goes on behind the clearly defined lines on both sides of the divide.’

\(^2\) The literature on grammaticalization is vast. Convenient introductions are: Hopper & Traugott (2003) and Lehmann (2002), available on line at: [http://www.christianlehmann.eu/publ/ASSidUE09.pdf](http://www.christianlehmann.eu/publ/ASSidUE09.pdf). We return to this issue below.

\(^3\) See, for instance, Traugott (1982; 1989; 2003); Sweetser (1990). Compare, as an example of subjectification, the semantic evolution of the English word *still*, from a spatial item (‘stable in space’: *still water, a still day*), through metaphorical (‘space for time’) interpretation as a temporal item (‘stable in time’: *He’s still there*), to a modal ‘concessive cancellative’ (‘stable in the attitude of the speaker’: *Even if it rains, we can still go for a walk*). Compare also the development of English *whilst*, Italian *mentre*, Spanish *mientras*, French *alors que* (temporal to adversative), English *since* (temporal to causal), etc.

\(^4\) William Labov has used this phrase in many of his publications.

\(^5\) See especially Croft (2000); McMahon & McMahon (2012).

\(^6\) See, for instance, Ringe, Warnow & Taylor (2002).
themselves. In this, as so often, linguistics is simply in step with the age; many disciplines, from physics to economics, have also come to treat phenomena previously considered static as dynamic processes. It’s all very well to say, as Saussure did, that we can’t study language change properly until we’ve first defined language; but what if change were an essential part of the definition of language...? We can now appreciate anew the perceptive ‘pre-Saussurian’ observation that Hugo Schuchardt (1842-1927) made 130 years ago: ‘Jedes Stadium der Sprache ist ein Uebergangsstadium’ [‘Every stage of language is a transitional stage’] (see Schuchardt 1885).

1.2. Various areas of study

phonetics (the study of) speech sounds
phonology (the study of) linguistic sound-systems
morphology (the study of) grammatical exponence, or ‘minimal meaningful units’
syntax (the study of) relationships between words (or between morphemes)
lexis (the study of) vocabulary
semantics (the study of) linguistic meaning
pragmatics (the study of) the relationship between language and the ‘real world’ etc.

1.3. Three levels of enquiry

observation what changes occur?
description how do they occur?
explanation why do they occur?

---

7 Compare the following claim by Kiparsky (2008): ‘If language change is constrained by grammatical structure, then synchronic assumptions have diachronic consequences. Theories of grammar can then in principle contribute to explaining properties of change, or conversely be falsified by historical evidence.’

8 Compare Lass (1997:14): ‘Synchronic structure and historical origin (of course) interact; but the state of the language as a whole is not fully describable in purely synchronic or functional terms. Portions of apparently ‘synchronic’ states are relics of the historical processes that brought them into being, evolutionary scars on the present-day body. The two dimensions are complementary, but in the end history probably has more to say about synchrony than the other way round’.

9 In fact, Chambers & Trudgill (1998), looking at dialect differences, distinguish between pronunciation (neither systematic nor systemic), phonetics (systematic, but not systemic), and phonology (both systematic and systemic). Compare: pronunciation, English schedule [ʃedjuəl] vs. [skedjuəl], but shed [ʃed] vs. *[sked], etc. (just affects one word, so unsystematic); phonetics Spanish la semana [la semana] vs. [la hemana], English better [beta] vs. [beʔə] (systematic, but does not impinge on lexical distinctions, therefore not systemic); phonology Spanish arrollo ([aroʎo] or [arojo]) vs. arroyo (only ever [arojo]), or cierra ([θjera] or [sjera]) vs. sierra (only ever [sjera]), English putt [pʌt] or (Northern) [pɔt] vs. put (only ever [pɔt]) (systematic and does impinge on lexical distinctions, so also systemic).
1.4. Variation as the seed of change; change as the consequence of variation

Languages are not monoliths

How languages don’t change (despite the impression given by many books on the history of French and other languages!) — the ‘metachronic fallacy’:

\[ A \rightarrow B \]

In other words, people don’t suddenly wake up one morning to find that a language has changed overnight! Malcolm Bradbury’s fictional Stalinist state of Slaka, in which this sort of thing happens, is, of course, a parody...

‘But on the third day, when he rises, and goes down to the lobby, to collect, on his way to breakfast, the red-masted newspaper, he senses that something has changed. Then, over the slow breakfast, he sees what the change is, a perfectly small one: P’ryuu Populatuuu has become P’ryii Populatiti again.’


How languages do change:

\[ A \rightarrow A/B \rightarrow B \]

How languages sometimes ‘stay the same’:\(^{10}\)

\[ A \rightarrow A/B \rightarrow A \]

1.5. Two stages of change: actuation and propagation

A change emerges, giving rise to variation (‘actuation’). Subsequently, this change may spread through the language (‘propagation’).\(^ {11}\)

An analogy from evolutionary biology: Mutations arise, selection favours certain mutations, which thrive at the expense of others.

The analogy is not exact, as linguistic ‘mutations’ are in general less random than biological ones (they can be motivated — see below) and linguistic ‘selection’ generally has a greater element of randomness (although it may to some extent be socially determined — again, see below). However, it is thought-provoking.


A high-level discussion of the analogies between linguistic change and biological change can be found in Croft (2000). Croft argues, *inter alia*, that plant evolution, which allows cross-fertilization and hybridization (see §1.7.2 below), may be a more relevant analogy for language change than animal evolution.

---

\(^ {10}\) On this type of ‘failed change’, see Postma (2010).

\(^ {11}\) On this crucial point, see Weinreich, Herzog & Labov (1968).
A typology of language variation

**diachronic** variation  
variation through time (what historical linguistics is all about)

**diatopic** variation  
variation according to place; geographical variation

**diastratic** variation  
variation according to social class

**diagenic** variation  
variation according to sex/gender

**diaphasic** variation  
variation according to register or style

**diamesic** variation  
variation according to medium (written vs. spoken, etc.)

1.5.1. The propagation of change as a social phenomenon

Sociological factors  
creation of identity (generally with respect to a ‘reference group’)
move towards élite norms (‘overt prestige’)
move away from élite norms (‘covert prestige’)
conscious change (‘change from above’)
unconscious change (‘change from below’)

These factors may work on an **intralinguistic** level (i.e., dialect contact) or an **interlinguistic** level (i.e., language contact).

1.5.2. The uniformitarian hypothesis


‘The U[niformitarian] P[rinciple] holds that we can constrain our hypotheses about the structure and history of languages of the past only by reference to what we know of contemporary language structures, linguistic behaviour and changes in progress, since the recoverable information about any language or speech community of the past is always far more limited than what we can know about languages whose native speakers we can still observe; and, further, that we can extrapolate into prehistory (and across gaps in the historical record) only on the basis of what we know from the study of contemporary languages and the actually documented past. Positing for any time in the past any structure or development inconsistent with what is known from modern work on living languages is unacceptable, and positing for prehistory any type of long-term development that we do not observe in documented history is likewise unacceptable, unless it can be demonstrated that there has been some relevant change in the conditions of language acquisition or use between the past time in question and later periods which can be observed or have been documented.’

1.6. Internal and external change

1.6.1. Internal change

**physiological factors**  
production (articulation) i.e., speaker
perception ([acoustics], audition) i.e., hearer

**psychological factors**  
ease of processing
sense of pattern (‘regularity’ — but in fact more complicated!)

---

12 In practice, the distinction between written and spoken language is often more diaphasic than diamesic.

13 Classic studies of the use of linguistic variables to express or create social identity include Labov (1963), Milroy (1987), and Eckert (1989; 2000). On overt and covert prestige, see Trudgill (1983a), chapter 4
FURTHER READING: The idea that speakers have a role in language change is not hard to grasp. On the role of the hearer in language change, see Ohala (1989; 1992; 1993).

1.6.1.1. What is regularity? Synchronic regularity vs. diachronic regularity

Sturtevant’s paradox (Edgar H. Sturtevant, 1875-1952), first articulated in Sturtevant (1947):

‘Phonetic laws are regular but produce irregularities.
Analogic creation is irregular but produces regularity.’

Diachronic regularity often produces synchronic irregularity, whilst diachronic irregularity often produces synchronic regularity.

The dominant view of sound change in the nineteenth century (echoed in the quotation from Sturtevant) was that it was exceptionless — that is, a given change took place whenever its structural description was met, regardless of the particular word involved. This position was associated with the Neogrammarians (Junggrammatiker), a group of German scholars, of whom the most famous are probably Karl Brugmann (1849-1919), Berthold Delbrück (1842-1922), Hermann Osthoff (1847-1909), and Hermann Paul (1846-1921). However, Hugo Schuchardt (1842-1927) disagreed, and, following pioneering work on linguistic geography (especially the work on the Atlas linguistique de la France by Gilliéron & Edmont) at the turn of the twentieth century, a rival hypothesis grew up, according to which sound change took place according to the model of lexical diffusion — in other words, a sound change took place in some words before others. The ‘slogan’ of the lexical diffusion camp was: ‘Chaque mot a son histoire’ [‘Every word has its [own] history’].

A problem for historical linguists with access to only scant data is that the end result of a sound change which applies ‘across the board’ at the same time and the end result of a sound change which is lexically diffuse but which does eventually propagate right through the lexicon are identical, and the data may not permit the two processes to be distinguished.

Throughout much of the twentieth century, the lexical diffusion hypothesis was rather airily dismissed. In fact, there is some evidence that at least some sound changes are lexically diffuse. Examples of such changes are not hard to find. Extremely frequent items (such as Spanish Vuestra Merced > Usted, French Mon Sieur > Monsieur) may undergo irregular sound change ([bwestra merθede] > [usteθ], [mɔsjɔr] > [maʃθ]) (maʃθ]; although ‘one off’ changes are not generally regarded as examples of lexical diffusion, they do indicate that sound change may affect individual words (this type of development may be especially frequent when the item in question becomes grammaticalized: compare Usted, Monsieur, and English do not > don’t, is > ’s, has > ’s, isn’t it? > innit?, etc.).


Consider now the following example of regular changes operating on regular forms to produce irregularity. Regular and transparent masculine/feminine adjective pairs in Latin (here given in the accusative case) undergo regular sound change to yield opaque outcomes in French.
As an example of **irregular** change leading to **regularity**, we might take the past tense of certain verbs in English. *Strove* has become *strived* for many speakers; *throve* has become *thrived* for most speakers. This change makes the past tense of these verbs more regular — *strive, strived* and *thrive, thrived*, just like *arrive, arrived*. But it’s not a regular, ‘across the board’ change — there’s no sign of *drove* becoming *drived*, for instance (possibly because of its high frequency; see below).

However, the waters appear to be muddied by the fact that, in many varieties of American English, *dived*, the regular past tense of the verb *dive*, has been replaced by the irregular from *dove*. How can we account for a variety in which *strive* becomes *strived*, yet *dived* becomes *dove*? Is this simply random or chaotic?

Probably not. The arguments here lead us to distinguish between **global** and **local** environments. Globally, most English verbs are regular, forming their past tense in -ed. The shift from *strove* to *strive* is therefore regularization in a global sense. But, if we take monosyllabic English verbs indicating manner of motion with an /aɪ/ in their stem, we find that, fairly systematically, they form their past tense by Ablaut or apophony — i.e., by changing their vowel, in this case to /əʊ/: compare drive, drove; ride, rode; write, wrote; rise, rose; stride, strode, etc. Seen within this local environment, *dive* becoming *dove* is a regularization.

### 1.6.1.2. Frequency and language change

**Frequency** is also an important issue here. In general, very frequent items are both more susceptible to irregular sound change and less likely to undergo analogical regularization. In other words, as a rule of thumb, the more frequent an item, the more likely it is to exhibit irregularity (although some very frequent items are none the less regular).

It is important when discussing frequency to distinguish between **types** and **tokens**. Types are individual items, tokens are occurrences of those items. You might like to think about how many words there are in the sentence *The cat sat on the mat* (there are two correct answers!). We sometimes get different answers according to which class of items we are dealing with. For instance, the statement that ‘most English verbs are regular’ is almost a truism (that’s what ‘regular’ means) — but only as a statement about types (i.e., most of the verbs listed in the *Oxford English Dictionary*, etc. form their past tense and past participle in -ed). If we look at tokens, the situation changes, and the statement is probably false, because, in an extended sample (say, a million words) of English text, most of the verb forms are going to be of common verbs like *be, have, go, make, do*, etc., all of which are irregular.

### 1.6.1.3. Inferential reasoning and language change

At this point, it is worth considering how mechanisms of language change relate to the mechanisms of inference familiar from philosophy.
**Deduction** takes a case, applies a law to it, and reaches a conclusion (or result). If the law is valid, then the conclusion is of necessity valid. Example:

- **Case:** Daphne is a cat
- **Law:** All cats like milk
- **Conclusion:** Daphne likes milk

**Induction** infers laws by generalization from observed phenomena. Example:

- **Observation 1:** Daphne is a cat
- **Observation 2:** Daphne likes milk
- **Conclusion = Law:** (Therefore) All cats like milk

Induction is our natural way of forming hypotheses about the world; it is how human knowledge advances. Of course, the hypotheses are not certain and can be falsified by further observation — but they are normally plausible.

Both deduction and induction (often jointly) play an important role in language change. If we take the relevant locus of change to be the child’s acquisition of the language, then the child is constantly formulating hypotheses (induction) and then applying them (deduction). Clearly, the inferential reasoning involved is not conscious in these cases; but it equally clearly exists.

For instance, it is clear how a combination of induction and deduction can account for analogical change of the type described above. Just as a child observes Daphne the cat drinking milk with gusto and concludes that cats in general like milk, so it hears a verb form *wanted*, and realizes that this is the past tense of the verb *want*. By induction:

- **Observation 1:** Want is a verb
- **Observation 2:** Want forms its past tense by adding *-ed* at the end
- **Conclusion = Law:** (Therefore) All verbs from their past tense by adding *-ed* at the end

and thence by deduction

- **Case:** Fight is a verb
- **Law:** All verbs form their past tense by adding *-ed* at the end
- **Conclusion:** (Therefore) The past tense of *fight* is *fought*

In this case, subsequent evidence modifies the hypothesis, and the child comes to learn the form *fought*. But in certain circumstances (for instance, when the irregular form is infrequent: see the section on analogy above), it may be that the child’s hypothesis is not overridden, resulting in an analogical extension of the regular form and a change in the language — hence, for instance, *throve* > *thrived*.

There is a third logical possibility, **abduction**, which, unlike deduction and induction is logical only in the sense that it exists — it does not result in a conclusion that can be sustained in logic. Consider the following

- **Case:** Daphne likes milk
- **Law:** All cats like milk
- **Conclusion:** (Therefore) Daphne is a cat

Clearly, this conclusion does not hold — Daphne may indeed be a cat, but she could just as well be a woman, a mongoose, or a hippopotamus.
Despite being logically flawed, abduction does seem to play a significant role in language acquisition, and hence in language change. A relevant example concerns the ‘prepositional object’ of Spanish (and some other Romance varieties, such as Neapolitan and Sardinian; a similar phenomenon, but involving dative vs. accusative case, is also found in many north Indian languages). Contrast Spanish *Vi la pared / Vi a la pared ‘I see the wall’ with Vi a la mujer / *Vi la mujer ‘I see the woman’. As a rule, definite human objects in Spanish require the preposition a, even when they are direct objects. This state of affairs is not found in Latin — how did it arise?

The dative is the rarest case in Latin (about 5% of tokens) — almost its only function is to encode participants who are affected by an action: beneficiaries or ‘maleficiaries’. The Latin case system does not survive into Spanish, and the functions of the dative are taken over by the preposition a (< AD). (Note that we must distinguish this dative use of a from the lative [sic] use, encoding motion to or towards: in Spanish they behave differently with respect to pronominalization: compare Di un libro a Santiago [person] > Le di un libro with Voy a Santiago [place] > *Le voy.) The overwhelming majority of participants affected by an action are definite and human — a trawl through texts will confirm this, as will a moment’s reflection (one generally gives presents, sings songs, reads letters, etc. to an identifiable human being rather than an inanimate object). Consider, therefore, the following abductive reasoning:

Case: This object is definite and human
Law: Dative objects are definite and human
Conclusion: (Therefore) This object is a dative object

A child reasoning along these lines will make all definite human objects, regardless of their function in the sentence (direct or indirect object) dative. In other words, the semantic and pragmatic characteristics of the object come to determine its grammatical form through a process of abduction.

On abduction in language change, see especially Andersen (1973), and, for a rejoinder, Deutscher (2002).

1.6.2. External change — language contact

Some definitions... A substratum (or substrate) language is one that is already spoken in an area when speakers of another language arrive and become dominant — socially, culturally, politically, and linguistically — but which none the less influences this new language. A superstratum (or superstrate) language is one that is spoken by new arrivals in an area who, however dominant socially and politically, ultimately fail to dominate linguistically, but still influence the language which is already spoken in the area. An adstratum (or adstrate) language is one that influences another ‘from the side’, as it were, without any mass geographical displacement of speakers. Schematically:

```
SUPERSTRATUM
  ↓
  language ‘L’ ← ADSTRATUM
  ↑
SUBSTRATUM
```
Substratum hypotheses come with a health warning! Let us take Romance as an example. The major problem is that the languages invoked as influences were in most cases displaced by Latin and often had no written form; we therefore have little evidence of them, and cannot assess the claims with any confidence. Sometimes, the only ‘knowledge’ we have of a substratum language is its alleged influence on Latin! Beware of this type of circular argument.

FURTHER READING: A sober rehearsal of the problems associated with substratum hypotheses in Romance is Wanner (1977).

Even superstratum and adstratum languages may present this type of difficulty (although normally to a lesser degree; but see the comments on Frankish below). The fact that we often have modern languages that are ‘related to’ (for this metaphor, see §1.7.1) the earlier ones doesn’t always help us to disentangle the problems: Italian is ‘related to’ Latin, but is very different from it!

A particularly interesting language from this point of view is French. The major substratum influence invoked for French is Celtic, specifically the Gaulish language which is known to have been spoken in Gaul before (and for some time after) the arrival of the Romans. Our knowledge of this language is limited to several hundred largely fragmentary inscriptions. It is related, at some distance, to Welsh and Irish, but also has some affinities with the Italic languages, of which Latin is one. The major superstratum influence is Germanic, specifically the Western Germanic Frankish language spoken by the Germanic tribe which occupied the northern part of what is now France. We have even less direct knowledge of this language, although it is assumed to be the ancestor of modern Dutch. Other parts of Gaul were occupied by tribes speaking Eastern Germanic languages — the Visigoths (albeit briefly) in the south and the Burgundians in the east. It has been argued that the tripartite linguistic division of Gaul/France into French (langue d’oïl), Occitan (langue d’oc) and Franco-Provençal (map on facing page) can be attributed to these different superstratum influences — see George Jochnowitz (1973), who also points out that the linguistic boundaries correlate with non-linguistic differences such as traditional legal systems, patterns of crop rotation, and the pitch of roofs on houses. The major adstratum influences in the history of French have been Latin (as the language of the church, law, and scholarship), Italian (through cultural and court influence in the sixteenth century), and, more recently, Arabic (as spoken in France’s North African colonies) and English.

Different types of ‘colonization’ and ‘invasion’:

The colonizers send an administrative élite, but don’t really populate the new territory
   (the Romans in Gaul? — compare the British in India)

The colonizers populate the new territory
   (the Romans in Dacia (now Romania)? — compare the British in Australia)

The colonizers settle in the new territory, but adopt its language
   (the Franks in Gaul — compare the Normans in England)

Note that this neat map, with its well-behaved boundaries between ‘languages’, ‘dialects’, and ‘dialect areas’, represents a massive simplification of the true picture. In fact, there is a multitude of transitional zones (some of which are represented by the hatching on the map). Isoglosses (lines which separate areas with different linguistic characteristics) are at best an idealization. See Chambers & Trudgill (1998), chapter 7: ‘Boundaries’ and chapter 8: ‘Transitions’.
Why, by the way, did the Franks, who were the socially dominant group and had their own language, adopt Latin? Latin was the language of an established culture and civilization, which could certainly be seen as more advanced than that of the Franks; but this sort of argument is ultimately subjective. The Franks had laws and literature, too. There is, however, one striking objective cultural difference between the two languages — unlike Frankish, Latin was written. Compare Toon’s comments on Old English (Toon 1983:2):

‘A preliterate people had conquered a new land and had established there a new social order. [...] Members of a culture as pervasively literate as our own need to be reminded of the magnitude of that change. In oral societies, culture and history reside exclusively in memory; social and political structures depend entirely upon present and immediate balances of power. Spoken agreements are valid only as long as those who spoke them choose to remember them; material possessions are limited to what an individual can grasp and defend. Literacy makes possible a past independent of memory. Written documents outlive convenience and document ownership; bookkeeping makes possible more complex administrative structures; a king can make a permanent ally by granting a privilege for all time in a written charter.’

However, this can’t be the whole story, as many Germanic languages that came into contact with Latin (such as the Old English that Toon is talking about) evolved their own written form as a result. There must have been other cultural and/or demographic factors at work, as well. For instance, it’s possible that — like the Normans in England — the Franks, despite their power and influence, just weren’t numerous enough to impose their language in the long term.

1.6.2.1. L1 and L2 in language contact — borrowing and imposition

We may distinguish two different types of language change through contact, according to whether the initiators of the change are first- or second-language speakers (L1 vs. L2). One of these is borrowing, which is defined as transfer due to recipient-language agentivity. In this case, speakers of a language adopt elements of a foreign language into their own. This is perhaps most characteristic of adstratum, but is also found to an extent with superstratum (the Latin-speakers of Gaul, for instance, will have adopted Frankish words into their speech). Contrasting with this mechanism is imposition, defined as transfer due to source-language agentivity. Here, speakers of a language attempt to speak a foreign language; but, as L2 speakers, they are unable to master the new language fully, and so inevitably transfer elements of their native speech into it. This process is characteristic of both substratum and superstratum — in terms of the history of French, of Celts learning Latin and of Franks learning Latin.

FURTHER READING: Although similar ideas are to be found in earlier work on language contact (see, for instance, Thomason & Kaufman 1988), the theory of contact outlined above is especially associated with two scholars — Frans van Coetsem and Donald Winford. See especially van Coetsem (1988; 2000) and Winford (2003; 2005; 2010).

1.7. Two simple models of language change — trees and waves

1.7.1. The Stammbaumtheorie

All of the above should make us slightly suspicious of the traditional view of language change, embodied in the metaphor of the ‘family tree’ (the Stammbaumtheorie of August Schleicher (1821-1868); see Schleicher 1853, 1863, 1865), which is probably still the dominant view — at least amongst non-linguists — of the historical relationships amongst languages). In this model, Spanish, French, Italian, etc. are all ‘related to’ the other Romance languages, of which they are ‘sisters’ or ‘cousins’, and all of them are ‘descended’ from Latin (note the extensions of the ‘family’ metaphor).
1.7.1.1. An arguably more accurate ‘family tree’

Latin fragmented into many different dialects, essentially a continuum; subsequently a handful of these dialects were ‘selected’ to become ‘standard’ languages.

(In fact, even this representation is not strictly accurate, because the ‘standard’ language is often based on a dialect which has been koinéized through contact — see §1.9 below.)

1.7.2. The Wellentheorie

Contrast the Wellentheorie (‘wave theory’) pioneered by Johannes Schmidt (1843-1901) (see Schmidt 1872) and taken up by Hugo Schuchardt (1842-1927) (whom we’ve already met in §1.1), partly as a result of his interest in pidgins and creoles, in which he was well ahead of his time. In this view, changes are ‘waves’ or ‘ripples’, which propagate through a language from different points, at different rates, and to different extents, yielding a highly differentiated pattern over time. By extension, a given language can be seen as the intersection of different influences, or successive ‘waves’.
1.7.3. Are they both right?

Recent work by Labov (2007) seeks to reconcile these two approaches to change in an interesting way, by claiming that the ‘family tree’ model accounts for linguistic changes which are the result of child language acquisition, whilst the ‘wave’ model must be invoked to explain change resulting from imperfect learning by adults. Once again, the L1/L2 distinction turns out to be crucial.

1.7.3.1. Transmission and diffusion

Transmission, involving first-language (L1) learners (by definition, children), is monotonic — i.e., by and large, each generation continues to move the language in the same direction. The ‘ratchet principle’ of Lieberson (2000), regarding monotonicity in fashion, is relevant here. Through transmission, speakers acquire both structures and systems, even though the systems may undergo modification from one generation to the next. That change proceeds in a particular direction is largely the result of social cohesion. Groups that are socially separated (characteristically through distance) will tend to develop differently, and their speech will grow apart — hence the splits typical of the ‘family tree’. This view of the differentiation of speech varieties through the separation of communities had already been articulated by Paul (1880). It bears obvious resemblances to the arguments for speciation through separation put forward by evolutionary biologists. Transmission represents continuity in change.

Diffusion, involving second-language (L2) learners (who may be adults or children beyond the ‘critical period’ for L1 acquisition), is not necessarily monotonic. Through diffusion, speakers acquire structures, but, crucially, fail to acquire the systems which go with them and determine their distribution. The result is the piecemeal influence of one language or dialect on another — hence the rather random intersections which characterize the ‘wave model’. Diffusion represents discontinuity in change.

1.7.3.1.1. Transmission (internal change; first-language acquisition)

Transmission, as we have just seen, crucially involves native-language acquisition as the locus of change. For this reason, it has been particularly investigated within the generative tradition, with its notions of the innateness hypothesis, the ‘language acquisition device’ (‘LAD’), and parameter setting. See (especially) two works by Lightfoot (1991; 1995). If we accept that variability is the seed of change, then a particularly important issue is the acquisition of variability, through variable rules and/or competing grammars at the level of the individual (Anthony Kroch calls the latter phenomenon ‘diglossia’, an unfortunate choice of term, as the word had already been used in a totally different sense by Ferguson 1959 and Fishman 1967) — see Kroch, (1989; 2001).

1.7.3.1.2. Diffusion (external change; language contact)

Diffusion, as we have seen, represents discontinuity in change. An extreme example of this discontinuity can be found in creole languages, which, in one sense have no transmission element. Creoles, viewed as nativized pidgins, are languages which, in the words of Muysken & Smith (1995), have ‘come into existence at a point in time that can be established fairly precisely — we can, in a sense, see them coming into being. If we ask: ‘What was the language that the present-day Spanish or French speech community spoke two thousand years ago?’, the answer is clearly: ‘Latin’. There has been diffusion of other elements into that original Latin, which has also been altered by transmission — but there is an unbroken thread of transmission over thousands of years. The same is not true of creoles. If we ask: ‘What was the language that the
present-day Haitian creole speech community spoke two thousand years ago?’, the answer is: ‘There wasn’t one’. Haitian, like other creoles, came into existence when a generation which was exposed to a pidgin (nobody’s native language, and arguably not a language at all) acquired it as a native language, and, in so doing, vastly expanded it, possibly through the ‘bioprogram’ — for this concept, see especially Bickerton (1991). Creoles, in this sense, are pure contact languages, representing absolute diffusion and total discontinuity. Of course, once they exist as native languages, they undergo transmission, just like all other languages.

1.8. ‘Central’ vs. ‘peripheral’ languages — ripples in the pond?

The Italian linguist Matteo Bartoli suggested many years ago that language change took place differentially, according to whether the variety involved was central or peripheral to the linguistic area involved (in the case of the Romance languages, the Roman Empire). His idea was that language change spread from the centre of a linguistic area towards the periphery rather in the way that ripples spread in a pond when you throw a stone into it. (The similarities with the Wellentheorie, above, should be obvious.) According to this theory, central areas should be more innovating and peripheral areas more conservative. This works well for some features in a variety of languages, including Romance; unfortunately, there are also many examples of the peripheral area innovating and the central area remaining conservative. So, if this sort of distinction does exist, it probably isn’t for Bartoli’s reasons.

Further Reading: Bartoli’s main statements of his theory are Bartoli (1925; 1929; 1933; 1945). Hall (1946) is a fairly comprehensive hatchet-job on Bartoli — it’s also useful if you don’t read Italian, because it contains a summary of Bartoli’s main ideas in English. Hall’s paper provoked a rather pained reaction in the form of Bonfante (1947) — see especially pp. 368-374. Two subsequent commentaries are Mańczak (1965; 1976), which support Bartoli’s findings, but reject his explanation of them.

1.9. Simplification and complexification

A lot of ink has been spilt on trying to answer the question: Do languages grow more simple or more complex? Hand in hand with this debate, some scholars have argued that languages evolve in the direction of greater ‘efficiency’. However, we don’t really know what an efficient language looks like, and, if we were to try to define the concept, we’d probably begin by saying that an efficient language was one that stood still — change in itself is communicatively inefficient. Needless to say, that doesn’t help us! There’s another problem, too — what’s efficient for the speaker isn’t what’s efficient for the hearer. For the speaker, the ideal language would be some form of telepathy — expend no physical energy at all and still have your interlocutor understand every nuance of what you want to say. What hearers would find helpful, on the other hand, is repetition and redundancy (you’ll observe that I’ve used two words there), so that if they miss something first time round, they can find it elsewhere in the message. In other words, what speakers ideally want is incompatible with what hearers ideally want — and, just to make matters worse, we are all both speakers and hearers, so this battle is constantly being fought out in the mind of every one of us.

However, the debate on simplification vs. complexification has recently been placed on a firmer footing by Trudgill (2010; 2011), who begins by pointing out a paradox. Sociolinguists claim that language contact leads to simplification (often through the process known as koinization or dialect levelling — see, for instance, Siegel 1985), whereas linguistic typologists see contact as a cause of complexification. Who is right?
Trudgill suggests the following typology:

**Isolation (little or no contact):** These languages are conservative and retain existing complexity. Icelandic is a good example. The mainland Scandinavian languages (Danish, Norwegian, Swedish), which derive from the same ancestor, but which have had much greater contact with other languages (especially Low German during the Hanseatic period), have lost the nominal case system, ‘quirky’ subject constructions, and a host of other complexities. Many (but not all) of their dialects have also reduced the number of genders from three to two.

**Short-term contact** involving L2 learning by adults: These languages will exhibit simplification, often drastic. The extreme example is pidgin.

**Long-term contact** where territory is shared, and involving child bilingualism: These languages will exhibit complexification, mainly through additive borrowing. The extreme example is the Sprachbund or ‘linguistic area’, such as the Balkan Sprachbund (see Sandfeld 1930), where a variety of unrelated (or, in some cases, only very distantly related) languages — Romanian (relevant for our purposes as a Romance language), Greek, Albanian, Bulgarian, and, to some extent, Turkish — share a number of characteristic features which are not found in much more closely related languages.

Note that, as was the case with the transmission/diffusion distinction discussed in §1.7.3 above (to which Trudgill’s discussion is clearly related), the distinction between L1 and L2 speakers is crucial.

Note, too, that an interesting (and mildly controversial) conclusion to be drawn from Trudgill’s work is that languages are not equally complex. The ‘equal complexity’ hypothesis has been an article of faith for many linguists. But, as Ruhlen (1994:148) notes: ‘all extant human languages are today considered of equal “complexity” by virtually all linguists, despite the fact that there is no recognized way of measuring complexity in language’. Compare also Lewontin (1990:740): ‘We would all agree that a human being is more complex than a soap bubble [...] but we cannot all agree that a dog is more complex than a fish, although fishlike forms preceded doglike forms by 500 million years and were their ancestors. [...] What in fact do we mean with complexity?’

### 1.10 Social structure and language change

It is clear from the foregoing that language change is not independent of the type of society in which the language is spoken.

#### 1.10.1 Social networks and language change

Milroy (1987) investigated the role of social networks in language variation and change, studying a number of communities in Belfast, the capital of Northern Ireland. Speakers will characteristically have interchanges with: i) members of their family; ii) their friends; iii) their neighbours, and iv) their workmates. The greater the overlap between these categories of people (i.e., the denser the network), the higher the network strength score of the individual involved. The broad finding from Milroy’s work is that the greater an individual’s social network strength score, the more likely that person is to exhibit local (i.e., non-standard) variants. This has a number of important consequences. One of these is that the role of variables such as social class, gender/sex, etc. in language variation and change may be epiphenomenal. In some of the areas studied by Milroy, men went out to work, whilst women did not; moreover, when they married, women tended to move to live with their husbands rather than vice versa. Women therefore had
a lower network strength score on two counts (no interchanges with workmates; fewer interchanges with family). In these communities, Milroy observed the ‘expected’ pattern of sociolinguistic variation — men used more local variants and women more standard variants. However, in one area, there was high male unemployment and most women went out to work (in the same place). Women therefore had a higher network strength score than men. Here, the ‘expected’ pattern was not observed — rather the reverse. Women used more local non-standard variants than men. It appears therefore that linguistic behaviour, at least in the communities studied by Milroy, is not directly correlated with variables such as social class, gender/sex, etc., but rather with network strength score.

1.10.2 Societal typology and language change

Andersen (1988), Røyneland (2004), and Kerswill (2015), building on the work of Milroy and others, distinguish two typological parameters of societal variation: a sociocommunicational or demographic parameter (open vs. closed) and an attitudinal or ideological parameter (endocentric vs. exocentric).

Closed societies are essentially societies in which the majority of the population have high network strength scores (à la Milroy).

Open societies are essentially societies in which the majority of the population have low network strength scores (à la Milroy).

Endocentric societies set great store by solidarity within the community and are suspicious of or hostile to outside influence.

Exocentric societies accept or even welcome outside influence and regard solidarity within the community as secondary.

This yields a four-way typology, as follows.

Closed endocentric societies. Characteristically, these are isolated rural communities, although many metropolitan inner cities now also conform to this definition. Inner London, where a strong sense of community holds amongst people of many different ethnic and geographical origins has led to a high degree of linguistic contact and the emergence of ‘new London English’ or ‘multicultural London English’ (see Cheshire et al. 2011, Kerswill 2015).

Open endocentric societies. These are typically urban centres with both a strong sense of community and strong external contacts. They tend to be sources of innovation, from which new forms diffuse outwards; they are ‘optimal senders’ of language change. Cities such as London and Paris have played, and to some extent continue to play, this role, which is also played at a regional level in Britain by Manchester, Liverpool, and Glasgow.

Closed exocentric societies. These are low-contact communities which none the less have a positive orientation to outside norms. Because of this orientation, they are receptive to change, but, in Kerswill’s phrase (Kerswill 2015), this is essentially ‘change driven by ideology, not change driven by contact’. A good example in the U.K. is inner-city Glasgow.

Open exocentric societies. These are typically very small communities, which are being subsumed into another community’s identity. They are therefore likely to be transitional and unstable. These are often rural communities (compare Marshall 2004 on the village of Huntly in Aberdeenshire in Scotland), but may also be mobile suburban areas (such as outer-city Belfast). They are ‘optimal receivers’ of language change.
What emerges with some force from these recent studies, is that endocentricity and exocentricity (i.e., attitudes) appear to be more significant determinants of language than closed vs. open networks. Note that there had been indications that this may be the case since Labov (1963).

1.11. A problem of focus — the microscope and the telescope

A characteristic feature of the working-class speech of south-east England is the vocalization of ‘dark’ [l] (i.e., /l/ in non-prevocalic position) to [u]. Recently, this (together with some other features of this variety) has been spreading through the speech of young people of all classes in most parts of the country. In other words, it is becoming a marker not of a particular class and/or geographical area, but of a particular age-group. This phenomenon, which originally entered the national consciousness and was remarked upon in the 1980s debate over so-called ‘Estuary English’ (the spread of new demotic speech patterns along the River Thames), may have had its origin in the ‘Punk Rock’ movement. Until the mid-1970s, as shown by Trudgill (1983b), it had been fashionable for British rock musicians to attempt (and often fail) to imitate an American accent. With the success of British rock music, the reverse became fashionable. It was at about this time that ‘Punk Rock’ bands, made up of young working-class south Londoners came to the fore. Many of their speech habits began to propagate through the community as a result of ‘covert prestige’ — very schematically, they came to mark groups of adolescents and young people who wished to distinguish themselves from (or even rebel against) the generation of their parents. Subsequently, they came to mark an identity as a young person in a broader sense.

That seems to be that — a specific, circumscribed change, which has taken place in the last thirty or forty years, and which has a specific, almost punctual, social motivation. But if, instead of looking at this change through a microscope, we shift focus and use a telescope, an interestingly different picture emerges. At some stage in the mid- to late seventeenth century — certainly by 1700 — non-prevocalic [r] started to disappear from British English. This was a gradual development; it appears to have started in London and spread slowly through the country. There are pockets of resistance to this day — south-west England, parts of Lancashire, and parts of Northumberland are still rhotic, as, of course, are most of Scotland and Ireland. The change did not affect much of North America, which was separated from England by a vast distance (and much of which would shortly become politically separate, as well), except in a handful of words (cuss < curse, bust < burst, etc.) — thereby indicating that the change was probably lexically diffuse. But the change spread out from London and gradually reached most of England. It appears to have begun as a vocalization of the [r] to schwa; thereafter, in some postvocalic contexts, the schwa fell, leaving us with the situation we have today.

Now, [l] and [r] are members of the same small class of sounds — the liquids. Indeed, in English, they are the only two members of this class. A linguistic historian of the year 3000 is going to tell the following story: there was, in the history of English at the end of the second millennium, a change which we might refer to as ‘liquid vocalization’. First of all, it affects [r], beginning in the late seventeenth century. Only about three hundred years later is [l] affected; but we are clearly dealing with a single, if somewhat drawn out, change.

The microscope offers us a detailed sociolinguistic account. The telescope gives us a structural and systemic account. The unanswered question is: how do we reconcile the two?
2. An Interlude: the name of the language

An apparently uncontentious statement...

‘2000 years ago, people spoke a language called Latin. Nowadays, people speak Spanish, Portuguese, French, Italian, Romanian, etc., which are (largely) mutually unintelligible languages related to Latin.’

This apparently anodyne analysis leads to a non-question: when did Latin become Spanish, French, or Italian? The futility of this question becomes clearer if we consider the following equally ‘uncontentious’ statement:

‘2000 years ago, people spoke a language called Chinese. Nowadays, people speak Mandarin, Hokkien, Hakka, Cantonese, etc., which are (largely) mutually unintelligible dialects of the same language, which is still called Chinese.’

‘When did Latin become Spanish, French, Italian, etc.?’ and ‘Why didn’t Chinese stop being Chinese?’ are diachronic analogues of the famous (non-)question: ‘What is the difference between a language and a dialect?’. The answer is the same: two speech varieties are different languages (synchronously or diachronically) if that is how speakers perceive them (or wish to perceive them). The question is cultural and ideological, rather than linguistic stricto sensu.

This issue is discussed, with specific reference to Romance, by Lloyd (1991) and Janson (1991). Janson quotes the example of a group of southern African language varieties, which were mutually intelligible and had no separate names. However, different groups of their speakers were proselytized by missionaries from Britain, who named the language Setswana, missionaries from France, who called it Sesotho, and missionaries from Germany, who designated it Sepedi. There thus grew up a new sense of linguistic separateness which had not existed previously, even though nothing about the language itself had actually changed. Similar examples abound. Chambers & Trudgill (1998:9-12) point out that the area known as Skåne, now the southernmost region of Sweden, was Danish until 1658, when it was conquered by the Swedes. Until then, the inhabitants had claimed to be speaking a dialect of Danish; a generation later, with their language practically unchanged, they claimed to speak a variety of Swedish. The language had not changed, but social, political, and geographical perceptions of it had (to use the appropriate technical term, it had been heteronomous with — i.e., perceived as a dialect of — Danish, but became heteronomous with Swedish). Equally striking is the example of Afrikaans (also mentioned by Chambers & Trudgill, loc. cit.); until the 1920s, this variety was considered heteronomous with Dutch, but, with the advent of greater Afrikaner nationalism, both Dutch and South Africans came to view it as an autonomous language. A slightly different, but related, phenomenon is discussed by Gerritsen (1999).

Crucial in this context are the Carolingian reforms of c. 800 AD. Charlemagne (c. 742-814) was King of the Franks from 768 until his death and Emperor from 800. He was not a native speaker of Latin/Romance, and surrounded himself with courtiers who were likewise non-native speakers (one of the most famous of whom was Alcuin of York, c. 735-804). By this stage, the spoken vernacular had developed considerably since the Classical Latin period (the first centuries BC and AD), but the written language had remained more or less the same. There seems to have been a feeling that the gulf between the written and spoken language should not be so great, perhaps coupled with a certain nostalgia for the days of great authors such as Caesar, Cicero, and Livy. So, whereas the obvious resolution of the discrepancy might have been to bring writing closer to speech, the opposite path was taken, and it was determined that the language should be pronounced as it was written. It’s as if a new ruling class, who were not native speakers of English, arrived in twenty-first-century Britain, observed the huge differences between the
language’s phonology and its spelling system, and decreed that, since the spelling represented the way English was spoken in a heyday of its literature (which it does — it’s a pretty accurate reflection of the pronunciation that Chaucer would have been used to), it should henceforth be given a literal realization. Although only a small minority of the population could read and write during the Carolingian period, the reforms were far-reaching. Church services, which most people attended, suddenly became incomprehensible because of the new pronunciation; and, whilst the ritual parts of the Mass could be understood precisely because of their ritual nature, sermons fared less well. The Council of Tours, in 813, sought to remedy this problem by stipulating that priests should preach in *rustica romana lingua* — i.e., the way ordinary people spoke, rather than a letter-by-letter pronunciation of the written language. It is obvious that these circumstances created for the first time a situation which favoured a conceptual distinction between ‘Latin’ and ‘the vernacular’ (although it would be several centuries before this vernacular was given a precise name in any Romance-speaking area). An associated issue is that it soon became necessary to write the vernacular down; and, since it wasn’t Latin any more, a new spelling system had to be developed, by trial and error. We shall have cause to return to these points. They are addressed most famously by Wright (1982), and, latterly, from a different perspective, by Banniard (2013).

3. Methodological Preliminaries

All diachronic linguistics and the synchronic study of any language other than in its contemporary state (even, say, something as recent as nineteenth-century English or Spanish) involves using ‘dead data’. This fact constrains us in significant ways.

3.1. Sources denied to us

**Until very recently, speech in general.** We can never be certain, therefore, how a ‘dead’ language was pronounced; this fact, in turn, can vitiate claims about its phonetics and its phonology.

**By definition, native-speaker judgements and intuitions.** This limits us, by depriving us of a natural mode of investigation, and raises the fundamental issue of *attestation* (see below).

3.2. Sources we do have at our disposal

**Essentially, written texts — a ‘dead’ language is inevitably a corpus-based language or ‘text language’.** (Even sound recordings of earlier stages of a language, whether wax cylinders or .mp3 files, constitute a finite corpus.) A corpus-based approach to any language can yield useful results. However, some scholars have objected to the use of corpora in linguistic research. They argue that the aim should be to investigate the system that underlies, or lies behind, the data (in other words, to study competence rather than performance), and that a linguistic hypothesis must apply to the potentially infinite set of sentences (etc.) of the language, not just to some subset of utterances (etc.). This view is now rather less prevalent than it used to be. And, anyway, in the case of a ‘dead’ language, the subset (and often a rather random subset at that) is all we’ve got — if it is to take place at all, the study of a ‘dead’ language has to adopt a corpus-based approach.

3.3. The problem of attestation and the issue of ‘negative evidence’

**If a form or structure is not present in our corpus, does this mean it is absent from the language, too; or is its absence accidental?** In the case of a living language, we would go and
check with a native speaker — but, of course, with a ‘dead’ language, we can’t. We can make intelligent guesses — but, methodologically, that is all they’ll ever be... The importance of the ‘uniformitarian hypothesis’ (see §1.5.2 above) may become clearer at this point.

3.4. The unrepresentative nature of the corpus

Most of what has survived is the result of cultural transmission — that is, texts with a perceived literary or historical value. Such texts tend to be formal, conservative, or even archaic. At the other extreme, they may be experimental! Compare George Steiner’s definition (Steiner 1971) of literature as ‘language in a condition of special use’. Ephemera (almost by definition!) tend not to survive, or not to anything like the same extent. (There are, of course, famous exceptions, such as the Pompeian graffiti. Some commercial and legal documents also survive from the Middle Ages.) In addition, there is an element of randomness about the corpus, due sometimes to human factors and sometimes to natural disasters — manuscripts have been lost or destroyed in fires or floods; conversely, the inscriptions in everyday Latin scratched on walls at Pompeii and Herculaneum have come down to us only because they were buried by the eruption of Vesuvius in 79 A.D.

All of the foregoing means that our perception of the language is skewed — imagine a description of contemporary English which was based almost entirely on poems, novels, and (to a lesser degree) plays, together with a handful of wills and mortgage deeds, a shopping-list or two, and the odd spray-paint ‘tagging’ of a railway carriage or motorway bridge, and which took no account at all of the spoken language or of other written registers, such as journalism.

A majority of mediaeval literature is poetry — literary prose is a relative newcomer. This clearly causes problems, because poetic form imposes special constraints (metre, alliteration, assonance, rhyme, etc.). However, paradoxically, these same poetic conventions can sometimes help us in determining how something was pronounced (or even whether or not it was pronounced). For instance, in Latin verse, when two vowels came into hiatus at a word boundary (i.e., when one word ended in a vowel and the next word began with a vowel), this yielded a single (vowel for purposes of scansion. Now, when the first word ended in ‘vowel + <M>’ and the second word began with a vowel, exactly the same thing happened — the sequence ‘vowel + <M> + vowel’ was also counted as a single vowel. Thus, schematizing, both BONA ET VERA AMICA and BONAM ET VERAM AMICAM ‘good and true (female) friend’, in the nominative and accusative, respectively, would both be scanned as BON’ ET VER’ AMICA(M). Poetry (certainly Latin poetry) is usually formal and conservative. The obvious conclusion to be drawn is that final [m] was no longer pronounced in any register; and, indeed, we have evidence from other sources for this development. Likewise, we find that, in the earliest assonating Old French verse, a vowel followed by a nasal consonant may assonate with the same vowel in any other environment. When this ceases to be the case, and we no longer find vowels followed by nasal consonants in assonance with vowels not followed by nasal consonants, we may tentatively assume (despite the fact that the evidence is negative rather than positive — see §3.3 above) that phonemic nasal vowels have emerged, distinct from their oral counterparts.

3.5. The problem of manuscripts and editions

• **Time.** For example, the Serments de Strasbourg, allegedly the earliest text in French (it can be dated with precision to 14 February, 842 A.D.!), comes down to us in a manuscript dating from 150 years later. How authentic is the text by this stage? Does it represent the language of 842 A.D., the language of c. 1000 A.D., or some mish-mash of the two?

• **Place.** If scribes of manuscripts are not from the same area as the author of the text, they will often add features of their own dialect.
• Hybridization. When there are several manuscript versions of a work, of which no two are the same, their competing claims have to be assessed, and editors sometimes produce composite texts, which may well correspond to no single manuscript source.

3.6. The role of editors in falsifying texts

Studer & Waters (1924): i) ‘we have eliminated later Anglo-Normanisms’; ii) ‘it [an emendation] supplies a more homogeneous [...] text’; iii) ‘A number of Picard forms [...] have been standardized’

Foulet & Speer (1979): ‘In order to make their texts intelligible to the reader, editors usually regularize certain orthographical anomalies. Thus, they distinguish between ce ‘this’ and se ‘it’, between ces ‘these’ and ses ‘his, her’, and between ci ‘here’ and si ‘so, and’ [sic — in fact a marker of non-switch-reference (i.e., indicating that the subject of the verb is the same as the subject of the preceding verb)]. When the scribe blurs these distinctions he is set right [my emphasis].’ Cf. Fleischman (2000).

Many nineteenth-century editions ‘standardize’ orthography, and some even ‘regularize’ the case flexion — a particularly insidious form of prescriptivism. It may make the text more homogeneous and easier to read; but it obviously renders it worthless as evidence for the survival or otherwise of the case system.


REFERENCES AND BIBLIOGRAPHY


Bartoli, Matteo. 1929. ‘La norma linguistica dell’area maggiore.’ Rivista di filologia e d’istruzione classica 57, 333-345.

Bartoli, Matteo. 1933. ‘Le norme neolinguistiche e la loro utilità per la storia dei linguaggi e dei costumi.’ Atti della Società italiana per il progresso delle scienze (1933), 157-167.


Kerswill, Paul. 2015. ‘Sociolinguistic typology, dialect formation and dialect levelling in industrial and post-industrial Britain: vernacular speech since 1800.’ MS, University of York.


Kroch, Anthony. 1989. ‘Reflexes of grammar in patterns of language change’, Language Variation and Change 1, 199-244 (downloadable from:


Ladefoged, Peter. 1997. ‘Sounds of the world’s languages.’ MS, University of California Los Angeles.


Lewontin, R. C. 1990. ‘How much did the Brain have to Change for Speech?’ Behavioral and Brain Sciences 13, 740-741.


Rodríguez Molina, Javier. 2010. La gramaticalización de los tiempos compuestos en español antiguo: cinco cambios diacrónicos. Tesis de doctorado, Universidad autónoma de Madrid.


Schleicher, August. 1863. Die Darwinische Theorie und die Sprachwissenschaft: offenes Sendschreiben an Herrn Dr Ernst Hackel. Weimar: Böhlau.


