DIACHRONEX: Corpus-based exercises for English diachronic linguistics

Katrien Heremans and Hubert Cuyckens[1]
University of Leuven

1   Setting the scene

Over the past ten to fifteen years, interest in diachronic linguistics and (mechanisms of) language change has been mounting. This can be seen, for instance, from the growing appeal (and attendance) of conferences on historical linguistics, or from the recent launching of two new diachronic journals, the *Journal of Historical Linguistics* and *Language Dynamics and Change* (adding to the existing ones *Diachronica* and *Folia Linguistica Historica*). It is no surprise, then, that studies on the history of English, probably the most widely studied and best documented language, have abounded as well. What may have contributed to this renewed interest in (English) diachronic linguistics is the increasing popularity of grammaticalization as a topic of research (possibly because grammaticalization offers a comprehensive perspective on language change allowing linguists to capture phonological, morphosyntactic, semantic, and pragmatic changes under one rubric; cf. Hopper and Traugott 2003). What has undoubtedly given a boost to (English) diachronic linguistics is the wide availability of electronic data, in the form of (tagged and/or parsed) corpora, covering the different historical periods of the English language. These corpora have not only benefited historical research in that they allow targeted searches yielding (sufficiently) large data sets for historical analysis, but they have also promoted a new perspective on language change, namely that language change is gradual and occurs in the context of actual language use (i.e., is usage-based). Indeed, careful diachronic corpus research allows (gradual) changes to be reliably read off from detailed synchronic 'snapshots' of actual language use made at particular intervals in time.

2   *Bringing corpora into the diachronic classroom*

Corpus-based descriptions and explanations of language change have not only found their way into scholarly publications, but they have also informed theoretical courses on (English) diachronic linguistics (both survey courses on the history of the English language as well as specialist courses). While (English) diachronic corpora have observably benefited theory-oriented instruction, they have so far been less established in data-driven or learner-centered diachronic teaching materials or syllabuses (in particular, exercise material), and are thus far less present in the classroom in direct interaction with the learner.[2] And this is somewhat unfortunate, as the availability of (English) diachronic corpora creates new opportunities for the learner as well as for the teacher. Indeed, corpora may offer learners 'hands-on' experience with (mechanisms of) language change and allow them to witness language change in progress; for teachers, corpora are an excellent resource for designing exercises that make language change come alive in the classroom. As such, diachronic corpora allow a data-driven approach that involves observing and interpreting changing patterns of language; they engage the learner in hypothesis building and testing, and may thus have a more lasting learning effect than (more traditional) deductive learning.

Importantly, learners' confrontation with diachronic corpus data best occurs in a stepwise fashion, with edited material offered to them first and less edited material later. In other words, a "guided inductive approach" (Johansson 2009: 42) is preferred, whereby deductive teaching materials and teacher-centered explanation are gradually replaced by learner-centered data analysis and generalization. Diachronic corpus-based teaching materials, then, could be fashioned such that they provide this gradual learning path.

3   *Introducing DIACHRONEX*

Seizing upon the opportunity offered by the wide availability of (English) diachronic corpora, we developed a collection of <u>diachron</u>ic <u>ex</u>ercises (called DIACHRONEX) bringing corpora into the classroom.[3] In particular, DIACHRONEX wishes to provide learners with increasingly data-driven exercises, which – in a step-wise fashion – allow them to discover changes in (patterns of) language use themselves. The exercises are thus meant to offer learners essential hands-on experience with particular phenomena of language change, so that they will be better equipped when – at a later stage – they have to identify and interpret patterns of changing use themselves (e.g., in the context of a final term paper or an MA-thesis).

DIACHRONEX was not only developed to capitalize on, or as a response to, the increasing availability of historical corpora; its aim was also to address some of the limitations of current exercise material in English diachronic linguistics, as it is found in textbooks such as Pollington (1999), Horobin and Smith (2002), Hasenfratz and Jambeck (2005), Singh (2005), Brinton and Arnovick (2011). This material is mainly of a *deductive* nature; its purpose is to allow students to practice/apply the rules and principles they have studied. Further, it tends to be carefully edited, and is often limited to changes within one period (e.g., vowel reduction in *Middle English*; the Great Vowel Shift in *Early Modern English*). As well, it often focuses on phonological and morphological changes, paying less attention to semantic and syntactic changes. Here are some examples:

(1) Phonology (cf. Singh 2005: 136)

> *c* typically represents [tʃ] before front vowels in ME. There are, however, cases in which its value is [s], as in ME *citee* 'city', *ceptre* 'sceptre', *cessyd* 'ceased', 'ended'. What is the reason for such exceptions?

(2) Morphology (cf. Hogg 2002: 66)

> For each of the following verbs identify the strong verb class to which they belong and give their four principal parts: *helpan* 'help'; *brecan* 'break'; *brēotan* 'break'; *lūcan* 'lock'; *līðan* 'travel', *weorðan* 'become'; *slēan* 'slay'.

(3) Morphology (cf. Singh 2005: 101)

> What markings of case, gender and number are evident in the following phrases? In phrases that contain adjectives, what inflectional declension (that is, strong or weak) is used?: *ðām fēorðan dæge*, *sēo sunne*, *ðām dæge*, *ðām ylcan dæge*, *þā twēgen men*, *se mōna*, *ðre nihte*, *sēo*

(4) Morphosyntax (cf. Pollington 1999: 36)

> Translate the following into OE:
> (a) I am teaching this child.
> (b) the princes divide the kingdom

Doubtless, these are valuable exercises which serve an important purpose in the teaching of English diachronic linguistics (and they are also part and parcel of DIACHRONEX; cf. below). However, they fall short when it comes to preparing students for diachronic corpus data analysis, and in particular for detecting phenomena of change in actual language use. Indeed, actual language use is

complex and multi-faceted; therefore, students often find it difficult to make a transition from the deductive exercises to inductive, corpus-based research which appeals to their identifying, classifying, and generalization skills.[4] It is exactly these corpus exploitation skills that DIACHRONEX wishes to train, and it wishes to do so in an incremental, step-wise fashion. In this way, DIACHRONEX also contributes to students' training in general corpus research skills, which has become an indispensable part of an up-to-date linguistics curriculum.

In trying to overcome the limitations of current deductive exercise material, then, the DIACHRONEX exercises have been devised with the following requirements in mind. They make use of 'realistic', corpus-based material and give sufficient attention to semantic and syntactic changes, in addition to phonological and morphological changes, and to mechanisms of language change. They provide a gradual learning path, from a guided/supervised introduction to older stages of English involving selected, carefully edited corpus data to a learner-centered, data-driven analysis of phenomena and mechanisms of language change. The learning path envisaged thus does not do away with deductive exercises, but views them as foundational for further largely inductive exercises. All exercises include extensive targeted feedback, and can be used in the classroom environment as well as in an electronic environment allowing students to practice at their own pace. The English corpora used in the exercises are PPCEME, PPCME2, LEON, YCOE, COHA, CEMET and CLMETEV.

4   The structure of DIACHRONEX

DIACHRONEX exercises comprise three levels of increasing complexity and of increasing discovery-learning and learner-centeredness. Level 1 exercises teach students to recognize phonological, morphosyntactic, and semantic changes in a guided/supervised or teacher-centered context, which is (only) selectively corpus-based. They provide the essential foundation for the more inductive, learner-centered exercises at Levels 2 and 3. Level 2 comprises corpus-based exercises which, while edited, exploit the students' language analytic skills. It is a crucial level in the entire DIACHRONEX makeup; indeed, it is at this level that students are trained to identify semantic and syntactic phenomena and mechanisms of change on the basis of actual corpus data. Level 3 is the level at which term papers or MA-theses are situated, and where students with limited supervision have to analyze diachronic corpus data themselves.

4.1 LEVEL 1 exercises

This level comprises three modules: the phonological, the morphosyntactic, and the semantic module. The phonological module comprises exercises testing students' knowledge of phonological developments and changes. Examples are in (1)–(3) below. The morphosyntactic module comprises exercises where students have to identify morphological forms or syntactic patterns in short text fragments. Examples are in (4). The semantic component checks students' knowledge of semantic relations and changes by presenting them with different meanings of a particular word in short text fragments. Examples are in (5)–(6).

These Level 1 exercises are, on the whole, deductive in that they are geared towards the application of rules and/or patterns; as such, their data-driven learning effect is limited. Still, the language fragments in which the exercises are embedded (especially in the morphosyntactic and semantic component) are taken from diachronic corpora (e.g., YCOE, PPCME2, PPCEME, COHA); at this level already, then, students are exposed to actual language data, which, it is thought, will help them in making the transition to inductive, corpus-driven exercises.

All Level 1 exercises are offered in multiple-choice format, which allows clear and targeted feedback on each option (distractors as well as correct answer),[5] so that students learn from their mistakes. As the feedback can be seen as taking the place of (the teacher's) guided supervision, the exercises are self-sufficient and can be offered as additional training or self-study. They are made accessible through the university's electronic learning environment and can be solved individually, at the student's own pace.

(1) Which changes, among others, affected PIE <*bheran>?

- *Grimm's law, open-syllable lengthening, vowel reduction*
  Correct! PIE *bh /bh/ was affected by Grimm's law and changed into *b /b/. PIE *e /e/ was retained as OE [e]; this short [e] underwent open-syllable lengthening in ME, resulting in ME [ɛ:]. The vowel in the unstressed syllable was reduced to [ə] in ME.
- *Grimm's law, open-syllable lengthening, Great Vowel Shift*
  This is not correct; the correct answer is: 'Grimm's law, open-syllable lengthening, vowel reduction'. PIE *bh /bh/ was affected by Grimm's law and changed into *b /b/. PIE *e /e/ was retained as OE [e]; this short [e] underwent open-syllable lengthening in ME, resulting in ME [ɛ:]. ME [ɛ:] was not affected by the Great Vowel Shift before /r/.

- *Verner's law, open-syllable lengthening, vowel reduction*
  This is not correct; the correct answer is: 'Grimm's law, open-syllable lengthening, vowel reduction'. PIE *bh /bh/ is not a voiceless stop, and Verner's law only affected the reflexes of PIE voiceless stops. PIE *e /e/ was retained as OE [e]; this short [e] underwent open-syllable lengthening in ME, resulting in ME [ɛ:]. The vowel in the unstressed syllable was reduced to [ə] in ME.

- *None of the above*
  This is not correct; the correct answer is: 'Grimm's law, open-syllable lengthening, vowel reduction'. PIE *bh /bh/ was affected by Grimm's law and changed into *b /b/. PIE *e /e/ was retained as OE [e]; this short [e] underwent open-syllable lengthening in ME, resulting in ME [ɛ:]. The vowel in the unstressed syllable was reduced to [ə] in ME.

(2) What is the Present-day English equivalent of Old English <cnēow>?

- [naʊ]
  This is not correct; the correct answer is [ni:].
  – The pronunciation of OE <ēo> is [e:ə]; this sound shifts to [e:] in ME and is affected by the Great Vowel Shift ([e:] > [i:]) in the EModE period.
  – A [k] before an [n] disappeared during the seventeenth century.
  – Note that OE [w] is not vocalized after [e:ə]; it is simply lost in Middle English.

- [noʊ] or [nəʊ]
  This is not correct; the correct answer is [ni:].
  – The pronunciation of OE <ēo> is [e:ə]; this sound shifts to [e:] in ME and is affected by the Great Vowel Shift ([e:] > [i:]) in the EModE period.
  – A [k] before an [n] disappeared during the seventeenth century.
  – Note that OE [w] is not vocalized after [eːə]; it is simply lost in Middle English.

- [ni:]
  Correct!
  – The pronunciation of OE <ēo> is [e:ə]; this sound shifts to [e:] in ME and is affected by the Great Vowel Shift ([e:] > [i:]) in the EModE period.
  – A [k] before an [n] disappeared during the seventeenth century.
  – OE [w] is not vocalized after [e:ə]; it is simply lost in Middle English.

- *None of the above*
  This is not correct; the correct answer is [ni:].
  – The pronunciation of OE <ēo> is [e:ə]; this sound shifts to [e:] in ME and is affected by the Great Vowel Shift ([e:] > [i:]) in the EModE period.

– A [k] before an [n] disappeared during the seventeenth century.

– Note that OE [w] is not vocalized after [e:ə]; it is simply lost in Middle English.

(3) Which are the Old and Middle English equivalents of Present-day English <wheat>?

- *OE <whǣte> [wæ:tə] and ME <whete> [we:tə]*
  This is not correct; the correct answer is: 'OE <hwǣte> [hwæ:tɛ] and ME <whete> [wɛ:tə]'.

  – In OE, <h> preceded <w> and both sounds were pronounced. In ME, this changed: <w> preceded <h> and <h> became silent (first in the southern dialects).

  – As the spelling of the vowel in PDE <wheat> is <ea> instead of <ee>, it is more likely that it developed from ME [ɛ:].

- *OE <hwǣte> [hwæ:tɛ] and ME <whete> [wɛ:tə]*
  Correct!

  – In OE, <h> preceded <w> and both sounds were pronounced. In ME, this changed: <w> preceded <h> and <h> became silent (first in the southern dialects).

  – As the spelling of the vowel in PDE <wheat> is <ea> instead of <ee>, it is more likely that it developed from ME [ɛ:].

- *OE <hwǣte> [hwæ:tɛ] and ME <whete> [we:tə]*
  This is not correct; the correct answer is: 'OE <hwǣte> [hwæ:tɛ] and ME <whete> [wɛ:tə]'.

  As the spelling of the vowel in PDE <wheat> is <ea> instead of <ee>, it is more likely that it developed from ME [ɛ:].

- *None of the above*
  This is not correct; the correct answer is: 'OE <hwǣte> [hwæ:tɛ] and ME <whete> [wɛ:tə]'.

  – In OE, <h> preceded <w> and both sounds were pronounced. In ME, this changed: <w> preceded <h> and <h> became silent (first in the southern dialects).

  – As the spelling of the vowel in PDE <wheat> is <ea> instead of <ee>, it is more likely that it developed from ME [ɛ:].

(4) Consider the following text fragment from *Rolle, Epistles* (a1450 (?1348))

For he þat mykel lufes, hym lyst oft syng of his luf, for joy þat
he or scho hase when þai thynk on þat þat þai lufe, namely if
þair lover be trew and lufand.

> **Translation:** Because he who loves much, him [it] pleases to sing often of his love, out of joy that he or she has when they think on the one that they love, especially if their lover be true and loving.

(4a) This text contains many personal pronouns. Can you infer from these in which dialect this text was written?

- *Northern dialect*
  Correct! The third-person singular feminine pronoun is *scho* and *th*-forms in the third-person plural pronouns occur in all three cases (nom. *þai;* gen. *þair*).
- *East Midlands dialect*
  This is not correct; the correct answer is 'the Northern dialect'.
  The third-person singular feminine pronoun in the East Midlands is *schē* or *shē*, but not *scho*, which is typical of the Northern dialect. As well, the East Midlands dialect uses the *h*-form (and not the *th*-form) in the genitive of the third-person plural pronoun.
- *West Midlands dialect*
  This is not correct; the correct answer is 'the Northern dialect'.
  The third-person singular feminine pronoun in the West Midlands is of the *he/ho* type. As well, the West Midlands dialect uses the *h*-form (and not the *th*-form) in the genitive of the third-person plural pronoun.
- *Southern dialect*
  This is not correct; the correct answer is 'the Northern dialect'.
  The third-person singular feminine pronoun in the Southern dialect is of the *he/ho* type. As well, the Southern dialect uses the *h*-form (and not the *th*-form) exclusively in the third-person plural pronoun.

(4b) Which of the following statements about *hym lyst* is correct?

- hym *is the subject of* lyst
  This is not correct; the correct answer is: *'lyst* is an impersonal verb'. *hym* is a personal pronoun in the dative case which takes the semantic role of 'experiencer'.

- lyst *is an impersonal verb*
  Correct! *hym* is a personal pronoun in the dative case which takes the semantic role of 'experiencer'.
- lyst *is a reflexive verb*
  This is not correct; the correct answer is: *'lyst* is an impersonal verb'. *hym* is a personal pronoun in the dative case which takes the semantic role of 'experiencer'.
- None of the above
  This is not correct; the correct answer is: 'lyst is an impersonal verb'. *hym* is a personal pronoun in the dative case and takes the semantic role of 'experiencer'.

(4c) In this text, the word *þat* occurs 4 times. Which of the following statements is correct?

- þat *functions 2 times as a demonstrative pronoun, and 2 times as a relative pronoun*
  This is not correct; the correct answer is: *'þat* functions 3 times as a relative pronoun, and 1 time as a demonstrative pronoun'. *þat* only functions as a demonstrative pronoun in '*þat* þat þai lufe'.
- þat *functions 3 times as a relative pronoun, and 1 time as a demonstrative pronoun*
  Correct! *þat* functions as a demonstrative pronoun in '*þat* þat þai lufe'.
- þat *functions 3 times as a relative pronoun, and 1 time as a definite article*
  This is not correct; the correct answer is '*þat* functions 3 times as a relative pronoun, and 1 time as a demonstrative pronoun'. *þat* functions as a demonstrative pronoun in '*þat* þat þai lufe'.
- *None of the above*
  This is not correct. The correct answer is '*þat* functions 3 times as a relative pronoun, and 1 time as a demonstrative pronoun'.

(5) Consider the following instantiations of the adjective *sad*:
1. 1810 (COHA): But I, unhappy I, alone and <u>sad</u>, Just vegetate, and dwell with melancholy.
2. 1918 (COHA): Now that you've heard the <u>sad</u> story of the poverty-stricken senior, I call for a change of subject.

Which semantic relation exists between *sad* in example 1 and *sad* in example2?

- *Strengthening*
  This is not correct. If strengthening had occurred, the resulting meaning would have been something like 'extremely sad', which is not the case here. Furthermore, it is usually neutral words that undergo strengthening (as, for instance, in euphemisms).
- *Figurative similarity/Metaphor*
  This is not correct. A sad person and a sad story both belong to the concrete, physical domain.
- *Amelioration*
  This is not correct. The meaning of *sad* in example 1 is not more negative than that in example 2.
- *Metonymy*
  Correct! A word naming an internal psychological state is used to refer to an external object evoking that state. The two meanings can be said to be in a 'contiguity' relationship of effect and cause.

(6) Consider the following instantiations of the adjective *pathetic*

1. 1817 (COHA): Independently of these general considerations, the history of Ruth, in connection with that of Naomi and Orpah (sic), has been always regarded as singularly interesting: it is a most pathetic (arousing sadness or pity) tale, illustrative of the operation of the tenderest of the domestic affections, in unison with genuine religion...

2. 2000 (COHA): The house itself is at once a Plantation manor (casa grande) and family home. We discover later that the stable is a pathetic little structure, weather-beaten and rusty, like the ones we can find next to the Black slave sheds...

Which semantic change took place between *pathetic* in example 1 and in example 2?

- *Figurative similarity/Metaphor*
  This is not correct. Figurative similarity often involves a concrete and a more abstract domain. Here, both meanings are equally abstract.
- *Pejoration*
  Correct! The meaning of *pathetic* in example 1 is more positive (it is a "singularly interesting" story, "illustrative of the operation of the tenderest of affections") than that in example 2.

- *Strengthening*

  This is not correct. If strengthening had occurred, the resulting meaning would have been something like 'arousing great sadness or pity'.
- *Specialization*

  This is not correct. If specialization had occurred, the new meaning would be a subcategory of the original meaning. This is not the case here.

4.2   LEVEL 2 exercises

Level 2 takes up a central position in DIACHRONEX. Building on students' acquaintance, acquired in Level 1, with phonological, morphosyntactic, and semantic changes in the history of English, Level 2 exercises aim to train students in identifying and classifying patterns of language change on the basis of an extensive set of corpus-based language data. The ultimate goal is for students to become sufficiently at ease with diachronic language material when analyzing data for a term paper or MA-thesis, or in other words, to successfully isolate the data relevant for a particular research question in the wealth of data a corpus offers (Level 3). In working from corpus data, the exercises are inductive in orientation; at the same time, they exploit the students' theoretical knowledge on language change (as a result of prior formal instruction), and in that sense have a deductive component as well. The exercises are obviously far more learner-centered than Level 1 exercises, but there is still a degree of guidance/supervision: the data are (to various degrees) teacher-edited such that unnecessary variants or irregularities have been weeded out, and, more importantly, such that the data have a particular diachronic story to tell. In other words, unlike unedited corpus data, the data making up Level 2 exercises *all* contribute to the identification of the pattern/mechanism of change under investigation. The exercises are therefore somewhat artificial, but this seems justified when one keeps in mind that they are a necessary stepping stone to Level 3 exercises.

Because of their extensive and more complex nature, these exercises are preferably used in a classroom context, whereby students can work in groups on finding a correct analysis, after which the exercise is discussed in class; alternatively, they can also be used as a take-home assignment at the end of the course. Some of these exercises could also be posted in the electronic learning environment together with a model answer, so that students can practice individually.

4.2.1 LEVEL 2 semantic change exercises

The first component contains exercises dealing with semantic change. A typical exercise will consist of a substantial number (e.g., 40) of corpus fragments

instantiating (selected) usages/senses of a particular lexical item, as they occur over a particular period of time. Students are then instructed:

(i)   to identify the various senses of the word in the corpus data, with particular attention to polysemous examples, as they may serve as 'bridging examples' between two senses;

(ii) to trace the development of the various senses (which new senses developed out of which older ones? which senses became more dominant/fall out of use? – this question entails paying attention to the frequency distribution of senses across the various periods);

(iii) to identify the mechanisms of semantic change at work in each of the changes.

Successful completion of this type of exercise imposes particular requirements on the presentation of the corpus fragments. For one, to make sure that students correctly identify an item's sense, several fragments illustrating the same sense will be necessary. Further, senses with a longer life cycle should be illustrated at more intervals throughout the period covered than senses with a shorter life cycle; in a similar vein, older senses giving rise to newer senses should also occur first in the chronology of corpus fragments. Finally, dominant/prototypical senses will need a higher frequency of occurrence (i.e., a higher number of corpus instantiations) than less dominant senses. This type of design, then, will also give students the necessary cues when dealing with regular corpus material (first and last attestations of a sense, frequency distribution of a sense, etc.)

The language material used in these exercises covers the period Early Modern English until now; using corpus fragments from earlier periods makes the exercises needlessly complicated as these fragments often require translation, which can be a semantic interpretation in itself, thus compromising the exercise. The corpora used, then, are PPCEME, CLMETEV, and COHA. An example of such an exercise (in abbreviated format), together with the instructions can be found in (7).

While the corpus fragments offered are still filtered (or edited), the exercises involve (inductive) discovery learning, in that students themselves have to trace the development of a lexical item's various senses by identifying the meanings of the words in the fragments, by checking first occurrences of senses, and by taking frequency information into account. As such, these exercises might be ideal intermediate steps towards full-fledged corpus research at Level 3.

(7) A corpus-based semantic change exercise: The development of *smart*

Corpus examples

1. 1817 Your son takes wing, you send various bands in pursuit; they meet, a <u>smart</u> brush takes place, and your son escapes – Ar. Count Arandez Yes; after maiming my servants; that I will never forgive.

...

6. 1835 In his youth she had laboured to persuade his father to send him to Cambridge, but the old man uniformly replied that Ralph "was a <u>smart</u> lad on the farm, and steady, and by that he knew he was no genius."

...

11. 1852 They can bear to be not quite so smart as other folks. Mr. Haye said he never saw such a pair of young men; and I guess he didn't. "Winifred sighed and still looked into the box, with a face that said plainly she would like to have them <u>smart</u>." O well, mamma, "she said presently," I guess they will look pretty nice, with all those new things; and the socks are nice, aren't they?"

...

16. 1873 Voltaire had had himself an early taste of what might be expected by any intermeddler. He had been caned publicly by the servants of a young sprig of nobility, who had been offended at a <u>smart</u> repartee which had come from his plebeian mouth.

...

21. 1908 It was quite plain that Bob, with his extreme gallantry of manner, his <u>smart</u> clothes, his high ways and his unconquerable gayety, had supplanted him on the pedestal where he had been the year before, just as somebody, somewhere – his sister, perhaps – had supplanted Miss Anne.

...

26. 1937 If she's <u>smart</u> like her mother she'll grow up and marry a rich man like her mother did.

...

31. 1965 Since when did you get to be such a <u>smart</u> mouth?

...

35. 1984 "Someone in the family?" Toby repeated. "Who in the family?" "I would guess Pat Traymore," Philip said sarcastically. "Don't get <u>smart</u> with me," Toby snapped. "I want to know who owns that place now, and which relative is using it."

...

40. 2007 Does this mean you should give up the pleasures of homegrown tomatoes and fresh-picked bouquets? Not if you set up your garden to avoid the

strain, use <u>smart</u> tools designed to lighten the work, and take time to prepare your body.

Questions

This exercise concerns the semantic development of the word *smart* from the beginning of the nineteenth century until now. The following aspects must be discussed:

- Which senses of *smart* can be distinguished in these examples? For this part of the exercise, you can make use of the OED-based definitions of *smart*, given below. Note that there does not need to be a perfect match between your set of senses and the set of OED-based senses (e.g., some instances might convey a sense not included in the list). Pay close attention to instances that show polysemy. It is not necessary to discuss every corpus example as some instances might convey the exact same senses.
- Trace the development of the various senses: Which new senses developed out of which older ones? Which senses became more dominant/prototypical over time, or fell out of use – use frequency information to support your answer? Which senses had a long/short life cycle?
- Identify, and justify, the mechanisms of semantic change underlying the development of new senses of *smart*. Draw up a diagram of how the different senses have developed from (and relate to) each other.

Write a coherent text of a maximum of **1000 words** in standard English discussing these three questions.

Definitions (based on the OED)
a. (said of objects, like whips) painful, hard, sharp
b. (said of blows, strokes,...) hard, fast, painful
c. (said of wounds, bruises,...) painful
d. (said of words, criticism,...) sharp, offensive
e. with intensity, force, speed, abruptness
f. considerable
g. pert, impudent, insolent
h. healthy
i. clever, capable
j. neat, fashionable

4.2.2 LEVEL 2 syntactic change exercises

The second component of Level 2 comprises exercises on changing syntactic patterns and mechanisms of syntactic change. We will discuss four types of exercises developed so far, each requiring students to infer syntactic patterns and changes from ample corpusdata:

–   construction of a mini-grammar on the basis of (an) extensive corpus fragment(s);

–   comparison of syntactic features in parallel texts in different periods;

–   identification of change in syntactic patterns on the basis of an extensive set of corpus fragments across various periods;

–   confrontation of corpus data with particular proposals made in the literature on an instance of syntactic change.

The corpora from which the fragments are extracted include the YCOE and the Penn Parsed Corpora.

In the first type of inductive syntactic change exercise, students are offered (an) extensive corpus fragment(s) from a particular period, and they are asked to construct a mini-grammar which focuses on particular (morpho-)syntactic features. An abbreviated example can be found in (8), in which students have to set up a mini verb-grammar of Middle English. While this exercise is, in the first instance, synchronic in that it concerns language patterns in a specific period rather than the development of certain patterns, it is also diachronic in that it cannot be solved without comparing the syntactic phenomena in the older language period with Present-day English, and is thus likely to raise issues of syntactic change. In offering unedited corpus fragments, this type of exercise trains students not only in pattern recognition (in a particular domain), but also in separating relevant data from (a wealth of) irrelevant data – an important corpus skill. In completing this exercise, students will automatically make use of prior teacher-based (or textbook-based) instruction on, or acquaintance with, the morphosyntactic patterns of the relevant period (tested at Level 1); to the extent that students apply that knowledge, the exercise has a deductive angle. At the same time, it has a learner-centered/inductive angle in that students, in setting up the mini-grammar, have the opportunity to independently work out the comparison between the older and Present-day English stage. This type of exercise works well in class discussions, including supervised roundup and teacher-led feedback sessions, or as a take-home assignment. In each case, it is essential to make model answers available through the electronic learning environment.

(8) A mini verb-grammar of Middle English
Read the following text (from a fourteenth-century English Bible) and iden-
tify all the verb forms. Using the information in the text:

   –   **describe the inflectional morphology of verb conjugations;**

   –   **describe periphrastic verb forms (passive, progressive, future,**
        past);

   –   **describe finite and non-finite verb forms;**

   –   **discuss verb-related word order (inversion, negative placement,**
        clause-final position)

On the basis of this information, design a concise verb grammar of four-
teenth-century English in comparison with Present-day English verb gram-
mar. Focus on all aspects mentioned above.

God made mankynde aftur his owne ymage & lyknesse, & put hym in paradys
þat was a lond of blysse, & ȝef hym þat lond to haue y-woned þer-ynne euer
more & neuer to han be ded, so þat he were buxom to hym & dyde what he bede
hym and kepte trewliche his heste as he was y-holde by pure kynde. For man ne
hadde no þing of hymself, bote al þinge þat he hadde it was of Goddis ȝefynge,
boþe his owne beynge & his wonynge; and al þynge þat hym neded God graunt-
eded hym, & forbede hym no þing, but þat he ne schulde noȝt eten of a tre þat
was a-mydde paradys, þat was y-cleped a tre of knowynge boþe good & yuel. …

A second type consists of exercises in which syntactic patterns are compared in
parallel corpus texts of different periods. An abbreviated example, again, is
given in (9). The analysis of the syntactic factors constraining or facilitating the
use of *do*-support is a synchronic exercise in each of the individual texts, but the
comparison will yield insights in syntactic change, and is thus diachronic. As in
the previous exercise, this exercise combines the (deductive) application of (pre-
viously acquired) theoretical information with learner-centered/inductive acqui-
sition of knowledge. Again this type of exercise works well both in class and as
a take-home assignment. With regard to corpus skills, it practices pattern recog-
nition and the selection of relevant data from a wealth of undifferentiated data.

(9) Comparing syntactic patterns in different periods: The use of *do*-support in Boethius' *Consolation of Philosophy*

The following 2 text samples are translations from the Latin version of Boethius' *Consolation of Philosophy* to English. The first is written by George Colville in 1556 and the second by Richard Graham in 1695.

Look for uses of *do* in the two texts. Which changes do you observe in the syntactic factors constraining or supporting the use of *do* in the two samples? Are there verbs that never seem to combine with *do*-support?

1556 (George Colville)	**1695 (Richard Graham)**
Perceyuest thou not those thynges	Dost thou perceive these things, said she, and do they sink into thy Mind ?
Why wepest?	Why dost thou weep?
Why shedest thou teres?	Why do thy Tears overflow?
hide it not.	And if thou dost expect Help from the Physician, truly discover thy Distemper.
Doith not the crueltye of fortune suffyciently appere agaynst me nedeth it anye other instruction?	Need my Sorrows then be repeated; and do not the Severities acted by Fortune against me, appear enough of themselves, without these Admonitions?
Doth not this place ... shewe the?	Doth not the very Face and Horror of this Place move thee?
wherein thou dydst chose the sureste seate for thy selfe in our scoles	Is this the Library which thou didst choose In which ..., thou didst skilfully read upon all Divine and Human Learning?
Had I than so vyle habyt, and suche sorowful countenaunce? When I dyd serch out with the, the secret causes of nature?	Was this my Habit? Was this my Look, when with thee I penetrated into the Secrets of Nature? when thou traced'st out to me the several Motions of the Stars?
When thou dyddest declare vnto me the course of the sterres with a Virge or rodde of geometry? When thou dyddest forme my maners and dysposicion of lyfe to that symylytude and lykenes of the heuenly order?	when thou didst shew me how to form my Life and manners by Divine Rule and Order? Certainly thou didst deliver this Sentence as an Eternal Sanction by the Mouth of Plato that those Commonwealths are most happy, who are gouuerned by Philosophers, or by those who study to be so.
if it happen that the rulers of commennaltyes do studye wysedome…	By the same Person also thou didst advise wise and discreet Men to take upon them the Government of their Coun

A third type of exercises involves the identification of syntactic change on the basis of an extensive set of corpus fragments across various periods of the English language. The fragments are selected such that 'they tell a story', and in that sense involve some degree of teacher guidance/supervision. They are learner-centered in that students by studying the corpus fragments have to draw conclusions on how the syntactic patterns have developed and what mechanisms of change might have been at play. Again, a deductive angle is present in that the student embeds the syntactic changes in particular processes/mechanisms of change he/she has previously acquired. An example of this type is an exercise on auxiliary development, in particular from OE *cunnan* to PDE *can* – an abbreviated version can be found in (10).

In working through this exercise, students again learn important corpus skills: when examining the change of a grammatical item such as *cunnan*, they should have an eye not only for the morphological changes the item underwent, but, importantly, for the changing syntactic patterns in which it occurs, as well as for semantic changes. The corpus data will thus be set up such that they reflect these key changes. For instance, the data will show that in the earlier stages of the English language, the various verb forms of *cunnan* patterned with NP objects as well as infinitives, and that later they diverged morphologically, syntactically, and semantically, with *can* (*could*) + infinitive reserved for auxiliary usage (compare, for instance, Old English sentence 3. with Modern English sentence 11. in exercise (10) below). Assuming previous instruction on the grammaticalization process, this exercise can easily test students' capability for recognizing grammaticalization. This type of exercises can both be used in class or as an assignment. Other exercises of this type include the grammaticalization/ development of the progressive, which manifests itself mainly through its use in more and more syntactic contexts (passive progressive, progressive of stative verbs,…) and the grammaticalization of participles to conjunctions or prepositions (*considering*, *regarding*).

(10) The development of Old English *cunnan* to Present-day English *can*. Analyze the following examples focusing on changes in the morphological characteristics, syntactic potential and meaning of the verb forms of *cunnan* and *can*. What process of syntactic change are these changes indicative of?

Old English

1. Leofre ys us beon beswungen for lare þænne hit ne **cunnan**. (*Ælfric's colloquy*)

    Dearer is us be flogged for learning than it not know.

    (We would rather be flogged for learning than not know it.)

...

2. þa ondswarede he & cwæð: Ne **con** ic noht singan; & ic forþon of þeossum gebeorscipe uteode, & hider gewat, forþon ic naht singan ne **cuðe**. (*Bede's History of the English Church* (o2))

    Then answered he and said, "I cannot sing; for that was the reason why I left the entertainment, and retired to this place because I could not sing."

...

3. Se druncena ne gecnæwð naðer ne fæder ne modor, ne freond ne feond, ne he gescead ne **can** betwyx gode & yfele. (*Chrodegang of Metz* (o4))

    The drunk knows not neither father nor mother, not friend nor enemy, he knows not the difference between good and evil.

Middle English

4. I ne **can** ne I ne mai tellen alle þe wunder ne alle þe pines ðat he diden wrecce men on þis land. (*The Peterborough Chronicle 1070–1154*)

    I cannot, no I may not tell all the wounds non all the pains that he did cause men on this land.

...

5. Ac ure helend saweð his holie word hwile þurh his hagen muð hwile þurh his apostles. and oðre lorðeawes þe **cunnen** holie boc-lore.(*Trinity Homilies* a1225)

    But our saviour preserved his holy word as well through his own mouth as well trough his apostles and other spiritual teachers that know holy book-lore.

...

6. How ofte, whanne þu hast be in þy preyeres, haþ he yrauyssched þe in-to so heiʒ desir þat þu **canst** not telle hit? (*Aelred of Rievaulx's De Institutione Inclusarum* c1400)

    How often, when you have been in your prayers, has he transported you into so high desire that you cannot explain it.

...

7. O what false touches **con** he / how **can** he stuffe the sleue wyth flockes. (*Caxton's History of Reynard the Fox* 1481 (PPCME2))

    O what false tricks knows he/ how he can stuff the sleeve with tufts of wool.

Modern English

8. If Cobham did practise with Aremberg, how **could** it not but be known in Spain? (*A complete collection of state-trials, and proceedings for high-treason, and other crimes and misdemeanours, commencing with the eleventh year of the reign of King Richard II, and ending with the sixteenth year of the reign of King George III* 1600)

...

9. **Could** not this man, which opened the eyes of the blinde, haue caused that euen this man should not haue died? S(econd Oxford Company *The Holy Bible*. 1611)

...

10. **Canst** thou deny but that all ill Men deserve Punishment? (Boethius, *Of the consolation of philosophy. in five books* 1695)

...

11. But my Lord Antrim has cut of his hear, and got one of those new fassioned perewks, which have so much hear in them that a good one **cant** cost les then pound. (*Letters by Hatton, Alice E.* 1699–1700)

In a fourth type of exercises, selected corpus material instantiating a particular change is confronted with earlier proposals in the literature on that change. An example of this type is on the development of Old English *lician* to Present-day English *like*. The exercise assumes knowledge or previous instruction about the mechanisms of 'reanalysis' and 'anology' and about impersonal constructions.

The different steps in the exercise are as follows. In a first step, students are familiarized with a particular view on the development of the Old English *lician*-construction to the Present-day English *like*-construction. In the example at hand, the view selected is that by Jespersen, who in his 1927 work distinguishes four stages. In the first stage, English was a relatively free word-order language, in which a pattern such as (11a) was not uncommon: in (11a), the experiencer, in the dative case, precedes the verb and the verb agrees in number with the nominative *peran* (expressing cause). In the second stage, case markings are reduced, leaving only the verb morphology to indicate which constituent is the subject; see (11b). But when, in the third stage, the verb morphology is no longer a distinguishing factor, a sentence such as (11c) is structurally ambiguous: "[b]ecause *the king* was in the position normally reserved for subjects by this time, it was interpreted ... as the subject. When this happened, pronouns came to appear in the nominative case" (Allen 1986: 376); cf. (11d). To summarize, Jespersen saw the development of the *lician* > *like*-construction as a clear instance of reanalysis.

Students are then asked to confront this view with the data. The data will be selected (by the teacher) such that they challenge essential claims of Jespersen's view. As such, the data will reflect:

– that OVS word order was much less common than SOV or SVO word orders, even in Old English (cf. 12);

– that personal pronouns, which retained case marking (cf. 12), were far more frequent than full noun phrases in the experiencer role, and that structural ambiguity was far less likely;

– that impersonal constructions continued to exist long after the case system had largely disappeared (cf. 13).

A successful design of these corpus data, then, will teach students that they should cast their nets wide in syntactic developments, and be attentive to (the frequency of) variants of the patterns under investigation (e.g., different word orders, distribution of word orders over periods, differential case loss in nouns and pronouns).[6]

(11)   a.   Đam cynge licodon peran.
              the-DAT king-DAT pleased-PL pears
         b.   The king likeden peares.
         c.   The king liked pears.
         d.   He liked pears.

(12)   þa ongan he þencan hwæðer hit hire **licode**. (*Mary of Egypt*)
         Then began he to think whether it pleased her.

(13)   Hit may **like** your good Grace to be advertised that I have this nyght, after that the Kings Grace had souped, presented and distinctely redde un to his Highnes as well your Grace's Lettre dated xxj=th= day of this present Septembre addressed un to my selfe ... (A letter: *sir Thomas More to cardinal Wolsey* 1523)

A sample of the *lician > like* data is presented in (14).

(14) *lician > like*
Old English
1. Sua eac Dauit, ðe folneah on eallum ðingum **Gode licode**, sona sua he ða byrðenne næfde sua monegra earfeða, he wæs mid ofermettum gewundad… (*Cura Pastoralis* c894)

> So also David, [ðe Gode licode] in nearly in all things, as soon as he did not have the burden(s) of many troubles, he was struck with arrogance

...

2. Witodlice Noe ana wæs rihtwis betweox eall manna cyn, & he for his rihtwis-nesse Gode **licode**. (*Vercelli Homilies*)

> Certainly Noah alone was righteous between all races of men and [he Gode licode] with his righteousness.

...

3. Ac me nu þynceð & bet **licað**, þæt swa hwæt swa þu oðþo in Romana cirican oðþo in Gallia oðþo in hwylcre oðerre hwæt þæs gemætte, þæt ælmeahtegum Gode ma **licie**, þæt þu bihygdelice þæt geceose ond in Ongolðeode cirican fæstlice to healdenne gesette, seo nu gena is neowu in geleafan. (*Bede's History of the English Church* c897)

> But it seems now to me and better [(me) licað, þæt] if you have found any-thing, either in the Roman, or the Gallican, or any other church, which pleases Almighty God more, you (may) carefully choose that, and firmly establish it in the English church to (be) observe(d), which at present is still new to the faith.

Early Middle English
4. þe pit tineð his muð ouer þe man; þe lið on fule synnen **þe him wel likeð** and ne wile hem forleten. (*Trinity Homilies* a1225)

> The pit closes its mouth over the man; who lies on foul sins [þe him wel likeð] and (he) does not want to abandon them.

...

5. Iwisliche þa clennesse iwelt alle unþeawes and halt gode þeawes þe gode **likiað** and monnan. (*The Lambeth Homilies* a1225 (c1200))

> Certainly purity subdues all vices and and holds good virtues [þe gode likiað] and men.

...

6. Aȝeanes ðat Adames hierte was i-attred ðurh dieules meneinge, swa ðat him baðe wel **likede** hit and ec teiþede, ðolede Crist ðat me þurh-stong his hierte; (*Vices and Virtues* a1225 (c1200))

Whereas Adam's heart was poisoned through the devil's incitement, so that both [him wel likede hit] and (he) also yielded, Christ suffered that men pierced his heart;

...

7. ðis is mi leue sune, him me **likeð** swiðe. (*Vices and Virtues* a1225 (c1200))

This is my dear son, [him me likeð] much.

Late Middle English

8. And al be it so that youre emprise be establessed and ordeyned by greet multitude of folk, yet thar ye nat accomplice thilke ordinaunce but **yow like**. (*The Tale of Melibee* c1390)

And al be it that your enterprise is ordered and ordained by a great multitude of people, yet dare you not accomplish this order, but [yow like].

...

9. Thanne seyden they with o voys, "Worshipful lady, we putten us and oure goodes al fully in youre wil and disposicioun, and been redy to comen, what day that it **like** unto youre noblesse [...], that we mowe fulfille the wille of yow and of my lord Melibee." (*The Tale of Melibee* c1390)

Then, they said with one voice, "Worshipful lady, we put us and our goods all fully in your will and disposition, and are ready to come whatever day that [it like unto youre noblesse]..., that we may fulfill the will of you and of my lord Melibee."

...

10. ..., and bituene ham hade ordeynede þat one of ham shulde haue Kyng Leir to soiourne al his lif tyme, with xl knyȝtes and heir squyers, þat he might worshipfully gone & ryde whider þat he wolde, and into what contre þat him **likede**, to playe and to solacen. (*The Brut* or *The Chronicles of England* c1400)

, and between them (they) had ordained that one of them should have to live (with) King Leir al his life time, with forty knights and their squires, that he might worshipfully go and ride wherever that he wanted, and into whatever county [þat him likede], to amuse and to entertain.

...

Early Modern English

11. If 't may **like** your Grace by your honorable Letters subscribed with your hand to bind your Grace for the accomplishment of this desire, trusting that yow will depeach our said Pursevant immediatly, for the long delay of so honorable a Journey wee think should sound to your dishonor. (*lord Surrey's second letter of challenge to king James the fourth of Scotland* 1513)

...

12. I praye yow at tyme conveniente recommende me to my goode sonne Johan More. I **liked** well his naturall fashion. (Thomas More 1521)

...

13. I **like** not this Jury for our purpose, they seeme to be too pitiful and too charitable to condemne the Prisoner. (A state-trial 1554)

Because this exercise is quite complex, we believe that it will benefit students most if discussed in class, with appropriate teacher supervision/guidance; it is thus less suitable as a take-home assignment. As always, posting a model answer is indispensable.

An important advantage of this type of exercises is that students learn to confront and compare theoretical linguistic background with actual language data and their analyses of the data. After successfully completing such exercises, students will be more prepared to tackle individual corpus-based research on the basis of some theoretical background. Exercises that require students to do individual corpus studies belong to the category 'Level 3 exercises', and will be discussed below.

4.3   *Level 3 exercises: Corpus research*

The corpus-based exercises as described above should facilitate students' corpus research at later stages of their career (when writing term papers for MA-courses or when writing an MA-thesis on English diachronic linguistics). As a preliminary, students need to get acquainted with corpus exploration in practice (query syntax, search strings, the use of tags and wildcards, etc.). This can be achieved by working with easy-access on-line corpora such as TIME, COCA, and COHA. As well, they need instruction about processing and reporting on results: the difference between relative, absolute, and normalized frequencies, how to calculate statistical significance, etc. The next step is then to tackle a term paper topic or write an MA thesis. The information provided at Level 1 and especially Level 2 should have improved their pattern recognition skills, their ability to target relevant data in the wealth of corpus data, to detect processes of semantic and syntactic change, and confront corpus data with their hypotheses. In other words, it should have facilitated the step from deductive (textbook type) exercises to 'live' corpus research.

5   Conclusion

The three levels of exercises described above provide students with a gradual learning path towards individual diachronic corpus-based linguistic research. The deductive Level 1 exercises allow students to apply theoretical concepts (such as 'case' or 'metonymy') to real language. They confront students with language from older stages in the history of English and train them to analyse short corpus instances. The more inductive Level 2 exercises confront students with the analysis of more extensive language material and allow them to discover and analyse linguistic patterns, changes in those patterns and the mechanisms behind those changes. In the exercises at Level 3, students tackle corpus research in groups or on their own starting with easy-access on-line corpora with a lot of practical guidance from the teacher. It is our belief that after this three-step programme, students will be maximally prepared for individual corpus research. Moreover, the exercises clearly portray a learner-centered approach that might motivate students more to embark on this kind of individual research in the context of a master thesis and even to aspire to do linguistic research on a professional basis.

Notes

1. The development of the corpus-based teaching materials reported on in this article was made possible by a grant from the Teaching Council (Onderwijsbeleid) of the University of Leuven (project no. OWP 2010/19). The final version of this paper was written during the second author's research stay at the Freiburg Institute for Advanced Studies (FRIAS) of the University of Freiburg. The second author would like to thank the FRIAS and the Alexander von Humboldt Stiftung for their generous financial support, and the Research Foundation Flanders (FWO) and the University of Leuven for granting and financially supporting his sabbatical leave.
2. Compare with the domain of foreign language learning and teaching (and in particular, English as a second language), where corpora have not only significantly informed language description (e.g., corpus-based learner's dictionaries such as the *Collins Cobuild English dictionary for advanced learners* (Sinclair 2001); corpus-based grammars such as *The Longman grammar of spoken and written English;* Biber *et al.* 1999) or underlie teaching materials (e.g., corpus-based vocabulary tools such as COLLEX-BIZ; Blanpain, Heyvaert and Laffut 2008), but where the data-driven application of corpora in the classroom (cf. Aston, Bernardini and Stewart 2004; Aijmer 2009) also finds increasing recognition.

3. In addition to the development of diachronic exercises for English, which are reported on in this paper, the project involves a component for the development of French diachronic exercises.
4. Indeed, when confronted with corpus data, students often point out that they do not see the wood for the trees.
5. For the purpose of this paper, feedback is shown for the four options at once. In the interactive electronic learning environment, students only see the feedback on the option they chose (which explains why there is some repetition in the feedback).
6. This particular exercise has only challenged Jespersen's (1927) view. It may, however, be a stepping stone to a more advanced exercise which confronts Allen's (1986) views on the development of *like* with corpus data (involving the competition with OE *cweman* and PDE *please*, and the change from a lexical case system to a structural case system).

References

Aijmer, Karin. 2009. *Corpora and language teaching.* Amsterdam: John Benjamins.

Allen, Cynthia L. 1986. Reconsidering the history of *like*. *Journal of Linguistics* 22: 375–409.

Aston, Guy, Silvia Bernardini and Dominic Stewart. 2004. *Corpora and language learners*. Amsterdam: John Benjamins.

Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad and Edward Finegan. 1999. *The Longman grammar of spoken and written English*. Harlow: Pearson Education.

Blanpain, Kristin, Liesbet Heyvaert and An Laffut. COLLEX-BIZ: A corpus-based 'lexical syllabus' for Business English. *ITL – International Journal of Applied Linguistics* 155: 77–93.

Brinton, Laurel J. and Leslie K. Arnovick. 2011. *The English language: A linguistic history.* Oxford: Oxford University Press.

Hasenfratz, Robert and Thomas Jambeck. 2005. *Reading Old English: A primer and first reader.* Morgantown, WV: West Virginia University Press.

Hogg, Richard. 2002. *An introduction to Old English*. Edinburgh: Edinburgh University Press.

Hopper, Paul J. and Elizabeth Closs Traugott. 2003. *Grammaticalization.* Cambridge: Cambridge University Press.

Horobin, Simon and Jeremy J. Smith. 2002. *An introduction to Middle English*. Edinburgh: Edinburgh University Press.

Jespersen, Otto. 1927. *A modern English grammar on historical principles, Vol. 3*. London: Allen and Unwin.

Johansson, Stig. 2009. Some thoughts on corpora and second language acquisition. In K. Aijmer (ed.). *Corpora and language teaching*. 33–44. Amsterdam: John Benjamins.

Pollington, Stephen. 1999. *First steps in Old English*. Norfolk: Anglo-Saxon Books.

Sinclair, John *et al*. (eds.). 2001. *Collins COBUILD English dictionary for advanced learners*. London: Harper Collins.

Singh, Ishtla. 2005. *The history of English: A student's guide*. London: Hodder Arnold.

Corpora

CEMET: *The Corpus of Early Modern English Texts*, compiled by Hendrik De Smet, Department of Linguistics, University of Leuven, Belgium. Available from http://wwwling.arts.kuleuven.be/fll/hdesmet/links.htm.

CLMETEV: *The Corpus of Late Modern English Texts (extended version)*, compiled by Hendrik De Smet, Department of Linguistics, University of Leuven, Belgium. Available from http://wwwling.arts.kuleuven.be/fll/hdesmet/links.htm.

COCA: Davies, Mark. (2008–) *The Corpus of Contemporary American English: 425 million words, 1990–present*. Available online at http://corpus.byu.edu/coca/.

COHA: Davies, Mark. (2010–) *The Corpus of Historical American English: 400 million words, 1810–2009*. Available online at http://corpus.byu.edu/coha/.

LEON: *Leuven English Old to New*, compiled by Peter Petré, Department of Linguistics, University of Leuven, Belgium. Available from http://wwwling.arts.kuleuven.be/fll/ppetre/index.htm

Penn Parsed Corpora of Historical English. http://www.ling.upenn.edu/histcorpora/. (PPCME 2 and PPCEME)

TIME: Davies, Mark. (2007–) *TIME Magazine Corpus*: 1*00 million words, 1920s–2000s*. Available online at http://corpus.byu.edu/time.

YCOE: *The York-Toronto-Helsinki Parsed Corpus of Old English Prose*. http://www-users.york.ac.uk/~lang22/YCOE/YcoeHome.htm.

Synchronically, the structure of the English lexicon consists of common, literary, and colloquial layers. Leonhard Lipka further divides a literary layer into scientific, foreign, and archaic words. He includes technical, slang, vulgar, and dialectal words in the colloquial layer. The major part of the lexicon Lipka identifies as a common layer--lexemes, which are present in all the varieties of English. The following diagram provided by Lipka illustrates the structure of the English lexicon (2002, p.17). We believe this diagram should include neologisms, as well; they can appear in all the la Diachronic linguistics is the study of a language through different periods in history to see the shift of a language through time. Dr. Richard Nordquist is professor emeritus of rhetoric and English at Georgia Southern University and the author of several university-level grammar and composition textbooks. Updated July 07, 2019. Diachronic linguistics is the study of a language through different periods in history. Diachronic linguistics is one of the two main temporal dimensions of language study identified by Swiss linguist Ferdinand de Saussure in his Course in General Linguistics (1916). The other is synchronic linguistics. The terms diachrony and synchrony refer, respectively, to an evolutionary phase of language and Request PDF | On Jan 1, 2012, Katrien Heremans and others published DIACHRONEX: Corpus-based exercises for English diachronic linguistics | Find, read and cite all the research you need on ResearchGate. The article proposes a hypothesis that attempts to explain the reasons for these difficulties based on the idea, already anticipated by Elaine Tarone, that the origin of the problems [Show full abstract] shown by students whose L1 is English to write accent marks on appropriate syllables is not so much the lack of success when locating the stressed syllable of the word, but it is in fact the very mechanism of identification of the problematic syllables and the action of segmenting words into syllables. Read more. Article.