

# In what sense is the OED the definitive record of the English language?

Pius ten Hacken

Keywords: *OED, language, usage notes, dictionaries of record, dictionary use.*

## Abstract

OED (2011) presents itself as “Oxford English Dictionary | The definitive record of the English language”. Superficially, this claim may seem a marketing slogan, but Simpson’s (2000) preface to the third edition shows that it is a reflection of the editors’ understanding of their dictionary, what may be called their ‘lexicographic ideology’. In this paper, I consider the claim from three perspectives. Section 1 presents the foundations of the claim as formulated in the preface. Section 2 analyses the claim with regard to some relevant insights gained in linguistic theory since work on the first edition of the OED started. Section 3 discusses some of the practical reflections of the ideology of recording as opposed to prescribing. Finally, section 4 formulates some general conclusions.

## 1. The OED as presented in its preface

In the *Historical Introduction* to the first edition of the OED, Craigie & Onions (1933) describe the start of the work on the OED in the late 1850s. The Philological Society decided in 1858 that instead of a supplement to an existing dictionary, *A New Dictionary on Historical Principles* should be created. In 1928 the first edition was completed. The third edition is published online with regular updates being uploaded, as described in Simpson’s (2000) preface. I will look here at the description of two aspects of the work, its authoritative nature (section 1.1) and its application of historical principles (section 1.2).

### 1.1. *The OED as an authority*

Simpson (2000: 1) starts his preface with the statement that “[t]he *Oxford English Dictionary* has been the principal dictionary of record for the English language throughout the lifetime of all current users of the language.” Writing in 2000, he implies that most speakers were born after the completion of the first edition and for those that are older, the dictionary had reached a sufficient degree of completion that when they grew up it already counted as the principal dictionary of record. The precise interpretation of the expression *dictionary of record* remains to be determined. We can distinguish at least two aspects of the meaning. One is that the dictionary is descriptive rather than normative in orientation, the other that it has a certain degree of authority.

It is interesting to analyse how Simpson (2000) treats the question of authority. In the starting paragraph of his preface, we find the following remark on the dictionary’s authority:

The Dictionary has come to be regarded as authoritative, and in order to maintain its pre-eminence the Delegates of the Oxford University Press decided in 1990 to authorize a comprehensive editorial programme of revision and updating

According to this remark, the label *authoritative* has been bestowed on the dictionary by the outside world ('has come to be regarded as'). At the same time, it is a qualification that the people behind the dictionary ('the Delegates') are not indifferent to. They 'authorized' a programme, which presumably means dedicating funds to it, with the explicit aim of 'maintain[ing] its pre-eminence'. We can see this as a stylized way to express that they actively seek the authoritative status. Another interpretation is that they are happy with the status and want to make sure the dictionary continues to deserve it.

One of the aspects of authority is the comprehensive lexical coverage of the dictionary. Craigie & Onions (1933: viii) mention the following as the "most important" principles of the work to be undertaken for the OED:

- I. The first requirement of every lexicon is that it should contain every word occurring in the literature of the language it professes to illustrate.
- IV. In the treatment of individual words the historical principle will be uniformly adopted.

These principles were presented in the 'Proposal for the Publication of a New English Dictionary by the Philological Society' presented in 1859. They raise a number of points that have later been the basis for criticism of the OED. The first of these is the 'requirement' that it should contain 'every word'. In this respect, Simpson (2000: 10) shows an awareness of the limitations of what a dictionary can be:

There are a number of myths about the *Oxford English Dictionary*, one of the most prevalent of which is that it includes every word, and every meaning of every word, which has ever formed part of the English language. Such an objective could never be fully achieved. [...] That said, the content of the Dictionary is certainly comprehensive within reasonable bounds.

Simpson does not elaborate on the precise reasons why comprehensive coverage is impossible, a question I will return to in section 2. When we compare Simpson's statement with the first principle quoted by Craigie and Onions, the main difference is that in 1859 the purpose was to cover 'every word occurring in the literature of the language', whereas Simpson refers to 'every word, which has ever formed part of the English language'. In fact, Simpson (2000: 4) mentions that the OED has been 'criticized for its apparent reliance on literary texts'. Simpson addresses this criticism by on one hand claiming that it is overstated, on the other announcing that new, non-literary material has been used in the revised edition.

The difference between the earlier and the later wording is a sign of a shift in the perception of what an authoritative dictionary should record. Comprehensive coverage was understood as based on literary works and seen as a requirement when work on the first edition of the OED started. Nowadays, the impossibility of comprehensiveness in an absolute sense has been recognized and the preponderance of literary works moderated.

### 1.2. *The historical principles and their application*

The status of a dictionary is based on the quality of its collection of evidence, the strength of the methodology used to analyse this material, and the quality of the analyses. From its origins, the OED has been a dictionary compiled on historical principles, as formulated in

principle IV in the quotation from Craigie & Onions (1933: viii). In line with this tradition, Simpson (2000: 1) describes the current work of updating the dictionary as follows:

Each entry already published is being comprehensively reviewed in the light of new documentary evidence and modern developments in scholarship, and further entries are being added both to fill gaps in the historical record and to record changes in the language today.

An aspect that deserves further discussion is the use of corpora. Simpson (2000: 2) emphasizes that “[n]ew material (whether historical or modern) is included on the basis of the documentary evidence available to the editorial staff.” If we define a corpus as “a collection of language material selected according to certain criteria and used as a basis for linguistic investigation”, as proposed by Svendsén (2009: 43), the reading programme set up for the first edition of the OED, as described by Craigie & Onions (1933) makes the OED a corpus-based dictionary. Indeed, Svendsén (2009: 41) also mentions reading programmes as a method of data collection.

A different view of the nature of corpora is taken by Atkins & Rundell (2008). They do not mention reading programmes and adopt the following definition of a corpus from Sinclair (2005), quoted from Atkins & Rundell (2008: 54):

A corpus is a collection of pieces of language text in electronic form, selected according to external criteria to represent, as far as possible, a language or language variety as a source of data for linguistic research.

Compared to Svendsén’s definition, this one is more restrictive in two respects. First, Sinclair takes the electronic form of the content as a criterion for calling something a corpus. Secondly, he requires that the corpus is representative of a language or language variety. The issue of representativity is a theoretical one that I will come back to in section 2. The question of whether a corpus is electronic or not may seem circumstantial at first. However, it has to be seen in connection with the way the corpus is used and the result of corpus queries interpreted. Only with an electronic corpus can we adopt a data-driven methodology of the type advocated, for instance, by Krishnamurthy (2008).

For the work on the third edition of the OED, Simpson (2000: 5) claims that “traditional ‘reading’ is still, in most cases, the most efficient method of making th[e] initial identification” of a new word or sense. This is also the reason for maintaining “four major reading programmes” (2000: 4). Tools such as concordancers are used to extract evidence for the historical development of an entry once it has been identified in the traditional way. We can understand this attitude in the context of the description of the current work of updating the dictionary, as quoted at the start of this section. The human intelligence used in processing data an identifying what is ‘striking’ or ‘interesting’ cannot easily be matched by computer routines processing corpora. It is therefore worth using both resources in a complementary way.

## 2. Language, corpus, and speech community

Since 1858, when the work on the first edition of the OED started, thoughts about some crucial linguistic concepts have developed considerably. It is worth highlighting a number of these developments that are relevant to the interpretation of the work on the OED.

First, when the project of the first edition of the OED started, the field of linguistics was dominated by the comparative-historical approach. As Robins (1967: 169) states, “linguistics in this [i.e. the 19th] century was concentrated on the historical study of Indo-European languages”, which in this period “was almost the preserve of German scholarship”. In Britain, it was represented by Max Müller (1823-1900), a German expatriate who moved to Oxford in 1848 and from 1868 to 1875 was Professor of Comparative Philology (Wikipedia, 2012). It is not surprising, then, that the linguistic ideas underlying the lexicographic policy of the OED were inspired by comparative-historical linguistics. Indeed, the historical emphasis of the OED fitted very well with this scientific environment. August Schleicher (1821-1868) developed an influential theory of languages as organic entities that are born, develop, and die, and that can be grouped into a genealogical tree (cf. Collinge, 1995). Max Müller’s work reflects the influence of Darwin, whose *The Origin of Species* appeared in 1859 (cf. Harris & Taylor, 1989: 185-195). This is the context in which the idea arose to collect and describe all words of English throughout its history.

In the 20th century, comparative-historical research lost ground within the broader field of linguistics, largely due to the influence of Saussure (1916). Saussure distinguished diachronic and synchronic approaches and argued that the synchronic study is prior because a proper diachronic study should compare synchronic stages of a language rather than describing the development of individual elements (1916: 246-50). This insight was widely accepted and led to a shift in emphasis from historical studies to the study of how the contemporary language system works.

Saussure proposed to study how linguistic elements are related to each other in a particular stage of the language. Instead of following, for instance, the development of a particular sound through the history of a language, Saussure proposed to study the position of the sound in the phonemic system of the language. Historical research is then only possible when we compare the (entire) phonemic system of one stage of the language to the (entire) system of another stage. In the case of words, the considerations can be similar. The system of words is of course much less restrictive than the system of phonemes, but Cruse (1986: 112-135) gives an overview of what was achieved in the study of what he calls *lexical configurations* or, more traditionally, *lexical fields*. An important representative of this approach was Jost Trier (1894-1970). For a large dictionary such as the OED, however, it is hardly feasible to change the lexicographic policy so radically as to incorporate such insights. It would have implied either a discrepancy between earlier and later volumes or a significant reworking of the earlier ones.

A second issue that has raised a lot of discussion in 20th century linguistics is the status of the data used as evidence by a dictionary and their relation to the language they represent. In distinguishing *langue* and *parole*, Saussure separated the collective, social aspects of language from the individual ones. For him, the *langue* is a system realized in the brains of a community of speakers (1916: 30). The *parole* is the way individual speakers select their words to form utterances. Thus, the lexicon belongs to the *langue*, but the formation of sentences to the *parole*.

There are problems with the Saussurean concept of *langue*, because it requires that there is an entity, the language system, which is realized in the brains of several individual speakers. Saussure claims that the *langue* is not fully realized in the brain of an individual speaker. However, how can an entity be realized in such a way? When studied in more depth,

it is also easy to recognize that speakers of the same speech community do not share exactly the same system. As Bloomfield (1933: 45) put it, “[i]f we observed closely enough, we should find that no two persons [...] spoke exactly alike.”

Chomsky (1980: 217-25) argues for a radical solution to these problems. In earlier work, he had argued for a distinction between *competence* and *performance* (e.g. Chomsky 1965: 3). Competence differs from Saussure’s *langue* in that it is individual. Chomsky adopts this concept of competence, realized in individual speakers’ minds, as the basic, ‘real’, and empirical sense of language, while relegating the sense of language as shared by a speech community to the status of an epiphenomenon. This does not mean that speech communities do not exist, but only that there is no theoretical need to delimit them and their languages rigorously. For Chomsky, linguistic theory should not be concerned with the language of a speech community as a notion to describe and explain, because there is no empirical way to verify whether, for instance, a particular word is ‘really’ a word of English. We can in principle determine whether it belongs to a particular speaker’s vocabulary, because the vocabulary is knowledge stored in the speaker’s brain and the brain is an empirical entity. There is no non-arbitrary way, however, to determine whether this word is part of English.

Named languages such as English are not empirical entities. To the extent that they exist, they are set up on the basis of speakers’ competence. There are many historical examples of this process of setting up a named language. A relatively well-documented older example is French. When Vaugelas (1647) embarked on the project of compiling a grammar and dictionary for French, he did not find a language *French* clearly laid out for him. Instead, he had to start by determining which speakers he found good speakers of French and then to use them as informants for setting up a standard. A more controversial case was the establishment of Greek. As described by Mackridge (2009), discussions on whose Greek was the proper Greek went on for several centuries. A more recent case is Serbo-Croatian. As Greenberg (2008) describes, both the issues of the unity of this language in Yugoslavia and the number and nature of languages corresponding to it after the split of this country triggered a lot of debate. What these examples illustrate is that named languages do not present themselves as ready-made entities, but only come into existence by people setting them up as such. As a consequence, a corpus of texts can as a matter of principle not be representative of an empirical concept of language.

Against this background, I argued in ten Hacken (2009) that dictionaries should not be seen as descriptions of a language. Although Simpson (2000: 1) calls the OED “the principal dictionary of record for the English language”, there is no empirical entity corresponding to ‘the English language’ for which the OED could be taken to be a description. In this context, it is necessary to keep in mind the distinction between formal claims and informal, pre-theoretical use of an expression such as *English*. When Chomsky & Lasnik (1995:33) state that “[i]n English, generally only objective Case-assigning verbs can occur in the passive,” they do not imply that there is an entity called *English* that has this property. Therefore, when Simpson (2000: 6) states that “[f]rom its base in Britain, the English language has expanded over the centuries to become a world language,” it is not absolutely inevitable to take this as evidence for an essentialist view of *English*. In both cases, the statements can be interpreted as using *English* in an informal sense.

In ten Hacken (2009), I propose an alternative interpretation of dictionaries, in which they are not descriptions of a language, but tools by which users of the dictionary solve problems of a particular type. The nature of the dictionary determines which types of problem it can solve. This is particularly obvious for learner’s dictionaries. The title of the *Oxford Advanced Learner’s Dictionary of Current English* clearly indicates the scope of the problems

it intends to solve by means of ‘Advanced Learner’ and ‘Current English’. The current edition does no longer has ‘Current English’ on the cover, but this should probably not be taken as a sign that the scope of the dictionary has changed.

It is equally straightforward to frame the purpose of general monolingual dictionaries of the type for which van Sterkenburg (2003) uses the expression ‘*the*’ dictionary in terms of problem solving. The use of ‘the English language’ in the title of Collins (1986) can be interpreted as indicating the range of problems it can be consulted for. When we look up words that exist both in English and in German, e.g. *computer* or *kindergarten*, we expect to find different information in Collins (1986) than in Wahrig (1997). Whereas Collins (1986) gives information on the use of these words in English (e.g. the shortening of *kindergarten* to *kindie*), Wahrig (1997) gives gender and inflection class information for these words in German.

It is much more difficult to apply the same type of interpretation strategy to the OED if we accept that it is “the definitive record of the English language”. One aspect that may be used as a basis for interpreting this claim is the opposition between recording and prescribing. This might be taken as an intended systematic difference between the OED and general dictionaries of English.

### 3. Recording usage versus formulating usage notes

The distinction between descriptive and prescriptive lexicographic approaches is most clearly reflected in the frequency and formulation of usage notes. Usage notes instruct the language user to use a word in a particular way or to refrain from using it in a particular way. Although Atkins & Rundell (2008: 233-235) concentrate only on usage notes in learner’s dictionaries and in bilingual dictionaries and Svensén (2009: 250) mentions them only in the context of synonyms, they are also widely used in general dictionaries for native speakers. If we take the OED as a dictionary of record, it should not include usage notes. It is therefore interesting to compare whether and how information expressed in usage notes in general dictionaries is expressed in the OED. In order to explore this, I selected a number of usage notes from COD (2011) and compared them to the corresponding entries in OED (2011). In order to exclude idiosyncrasies of COD (2011), I also compared the entries with Chambers (1998) and Collins (2000). The usage notes can be divided into three broad categories, pertaining to cultural, grammatical, and etymological sensitivities.

#### 3.1. *Issues of cultural sensitivity*

Usage notes that belong to the category of cultural sensitivity are cases where the dictionary alerts to potential offensive use. Two examples in COD (2011) are *dumb* in the sense of ‘not able to speak’ and *oriental* as a term denoting people from Asia. In the case of *dumb*, COD (2011) gives the following usage note:

In the sense meaning ‘not able to speak’, **dumb** has been overwhelmed by the sense ‘stupid’ to such an extent that the use of the first sense is now almost certain to cause offence. Alternatives such as **speech-impaired** should be used instead.

In the corresponding entry from OED (2011), it is hard to retrieve this information. The sense ‘foolish, stupid, ignorant’ occurs as sense 7b. The only indication that the first sense, ‘destitute of the faculty of speech’ is no longer common is that it has no citations after 1884. Arguably, this is entirely reasonable in a historically oriented dictionary. It should be noted that Chambers (1998) and Collins (2000) also give the sense ‘stupid’ as the last one, marked “orig *US*” by the former and “informal” by the latter, without any explicit usage note.

In the case of *oriental*, COD (2011) gives the following usage note for the nominal sense “a person of Asian, especially East Asian, descent”:

The term **oriental** is now regarded as old-fashioned and potentially offensive as a term denoting people from the Asia. In US English **Asian** is the standard accepted term; in British English, where **Asian** tends to denote people from South Asia, specific terms such as **Chinese** or **Japanese** are more likely to be used.

In this case, OED (2011) is somewhat more explicit. It has several nominal senses for *oriental*, and “[a] native or inhabitant of the Orient, esp. East Asia” is sense 2b. As the last example for this sense, dated 1999, we find “You can’t say Orientals any more. It’s considered insulting—like Negro or stewardess.” As with *dumb*, Chambers (1998) and Collins (2000) do not have a usage note here, and they do not even include any markers for this sense.

What we observe, then, is that COD (2011) seems to be more explicit on issues of cultural sensitivity than Chambers (1998) and Collins (2000). OED (2011) approaches the issue by means of the selection of examples. Arguably, this is the most fitting solution for a dictionary that considers itself a record of the English language.

### 3.2. *Issues of grammatical sensitivity*

A second category of usage notes concerns what I will call here grammatical sensitivity. They concern a clash between grammatical prescriptions and developments in the language. Two examples are the use of *out* as a preposition (e.g. *out the door*) and the use of *they* as a sex-neutral pronoun referring to an indefinite pronoun (e.g. *anyone*).

COD (2011) gives prepositional *out* as “non-standard contraction of **OUT OF** (sense 1)” and adds the following usage note at the end of the entry:

The use of **out** rather than **out of** as a preposition, as in *he threw it out the window*, is common in informal contexts, but is not widely accepted in standard British English.

OED (2011) gives four senses of *out* as a preposition, two of which are marked as obsolete. The first sense, “[f]rom within, away from”, is described as “[f]ormerly *poet*. Now *regional* and *nonstandard*.” The second sense, “[o]utside, beyond”, is described as “[n]ow *nonstandard*.” Examples for both senses extend into the 1990s. The use of an explicit label *nonstandard* is difficult to reconcile with a purely descriptive approach. Chambers (1998) marks these senses of *out* as “*colloq* or *NAm*” and “now *rare*”, respectively. Collins (2000) gives a usage note specifying that the prepositional use of *out* is “common in American English, [but] regarded as incorrect in British English”. We can observe two parameters in these usage notes, one geographic and the other based on register. Collins (2000) and Chambers (1998) note the geographic distinction, which OED (2011) does not mention

explicitly. It can be inferred from ‘regional’, but this label is very unspecific. Chambers (1998) and COD (2011) mention register. We might interpret OED’s (2011) use of ‘nonstandard’ along these lines, but the expression as such is of a much more normative nature. Whereas the usage notes in general dictionaries describe when *out* can be used as a preposition, the labels used by OED (2011) only suggest that it should not be used this way.

The case of singular *they* is different and has been discussed widely in the literature on language and gender. Miller & Swift (1980: 43-58) discuss it as a solution to what they call “The pronoun problem”. COD (2011) gives the following usage note on *they*:

It is now widely held that the traditional use of **he** to refer to a person of either sex is outdated and sexist; the alternative, **he or she**, can be clumsy. It is now generally acceptable, therefore, to use **they** (with its counterparts **them**, **their**, and **themselves**) instead. This is especially the case where **they** follows an indefinite pronoun such as **anyone** or **someone** (*anyone can join if they are resident*). In view of the growing acceptance of **they**, it is used in this dictionary in many cases where **he** would have been used formerly.

The growing acceptance referred to in this usage note is visible from the usage notes in Collins (2000) and Collins (1986). Collins (2000) gives the following usage note:

It was formerly considered correct to use *he*, *him*, or *his* after pronouns such as *everyone*, *no-one*, *anyone*, or *someone*, as in *everyone did his best*, but it is now more common to use *they*, *them*, or *their*, and this use has become acceptable in all but the most formal contexts: *everyone did their best*.

There are two differences between the usage notes in Collins (2000) and COD (2011). First, Collins (2000) leaves open the use of *he* in highly formal contexts, whereas COD calls it generally “outdated and sexist”. Secondly, Collins (2000) restricts the use of *they* to cases where its antecedent is an indefinite pronoun, whereas for COD (2011) it is only “especially the case” in such contexts. Only COD (2011) opens up the possibility of using *they* in contexts such as “When I see a student, they often complain about ...”. Still, Collins (2000) is much more liberal in this respect than Collins (1986), which simply labels the use of *they* to refer to an indefinite antecedent as “*Not standard.*”

OED (2011) divides the article for *they* into three main sections, for personal pronoun, demonstrative and possessive pronouns, and adverbs. The relevant sense is given as sense 2 in section I and described as follows:

Often used in reference to a singular noun made universal by *every*, *any*, *no*, etc., or applicable to one of either sex (= ‘he or she’).  
See Jespersen *Progress in Lang.* §24.

This definition is purely descriptive and does not refer to any of the 20th century discussion. There are eight examples, covering the period 1526-1874, and the reference to Jespersen concerns a book published in 1894. Perhaps this is an entry that has not been updated in the third edition, although for other senses there are examples up to 1981.

In cases where general dictionaries have grammatical usage notes, OED (2011) does not use examples to convey the same information. In the case of *they*, a descriptive definition is given. Although neither the definition nor the examples address developments of the past

century, the policy is one of pure recording. In the case of *out*, the use of the label *nonstandard* diverges from a purely descriptive approach. Compared to general dictionaries, OED (2011) is less informative and more prescriptive in this entry.

### 3.3. *Issues of etymological sensitivity*

A final category of usage notes considered here concerns etymological sensitivity. Whereas there is always a sense of pedantry about usage notes, in this case it is particularly prominent. An example is the use of *decimate*. Etymologically, the word includes *decimus*, the Latin ordinal number ‘tenth’. In the Roman army, decimating was a collective punishment of a legion for mutiny, in which every tenth soldier would be put to death. By regular sense extension, it has come to mean ‘reducing considerably’. This development is reflected in the usage note given by COD (2011):

The earliest sense of **decimate**, ‘kill one in every ten of’, has been more or less totally superseded by the more general sense ‘kill or destroy (a large proportion of)’. Some traditionalists argue that this latter sense is incorrect, but it is clear that this is now part of standard English.

It is interesting to see that Collins (2000) gives a usage note for this entry, but not on the same issue:

One talks about the whole of something being *decimated*, not a part: *disease decimated the population*, not *disease decimated most of the population*.

The issue treated in the COD usage note is treated only implicitly by Collins (2000), in that the first sense is the more general one and the etymological sense has the label “esp. in the ancient Roman army”. The Collins (2000) usage note bars a further semantic development, which is not treated at all in COD. The earlier edition, Collins (1986), does not include any usage note. Chambers (1998) takes a more etymological approach in its presentation of the senses, listing senses with ‘tenth’ first and labeling the extended sense as “loosely”.

In this case, we expect the OED to take a historical approach. It gives four senses, of which the etymological sense is number 3 and labeled “*Milit.*”. The first two senses are obsolete and refer to taxation and to the introduction of a decimal system. The order is not historically motivated, however, because the oldest citation is for sense 3. Sense 4 is marked “*transf.*” and divided into 4a and 4b. Sense 4a has “one in every ten” and 4b, marked “*rhetorically of loosely*” has “a large proportion of”. It seems, then, that OED is particularly concerned to indicate how far removed the extended sense is from the original. This impression is reinforced because, unusually, sense 4a is not accompanied by any examples. The only reason for it to be mentioned seems to be to make the link between senses 3 and 4b more gradual.

In the case of *decimate*, a historical description of the development of senses is expected in the OED. This is what we find. However, some details of the way the entry is set up suggest a certain uneasiness with the extended sense, resulting in a bridging sense without citations.

### 3.4. Usage notes and the OED

If we assume that general dictionaries are sources of information and the OED is a historical description of the English language, we expect usage notes in general dictionaries and not in the OED. The frequency of usage notes in general dictionaries varies considerably. COD (2011) is quite generous in this respect, Collins (2000) uses them much more sparingly, and Chambers (1998) limits itself largely to usage labels.

In the case of cultural sensitivities, OED (2011) makes good use of examples to avoid explicit usage notes. This is the ideal way to handle such issues for a historical dictionary, because it records the change of attitudes without making any explicit judgements. For etymological sensitivities, the same approach can be used, although in the example of *decimate* it is reinforced by the use of explicit labels indicating the relations between senses. For grammatical sensitivities, it is much harder to use purely descriptive, historically oriented material to express the problem. Here, OED (2011) uses labels in a way that can hardly be seen as purely recording the developments.

## 4. Conclusion

The question raised at the beginning pertains to the interpretation of the OED's claim to be "the definitive record of the English language". Each of the elements of this phrase deserves some comments.

In section 2 it was shown that from a linguistic perspective, "the English language" cannot refer to an empirical entity. When Craigie & Onions (1933: viii) state that "[t]he first requirement of every lexicon is that it should contain every word occurring in the literature of the language it professes to illustrate", this cannot be interpreted as a statement about English as an empirical entity. Simpson's (2000: 10) statements that "fully comprehensive coverage of all elements of the language is a chimera" and that the idea that words are not English unless they are in the dictionary "may be acceptable logic for the purposes of word games, but not outside those limits" are more realistic in this respect. *English* must be interpreted either, in the same way as Chomsky uses it, in a pre-theoretical sense, or, in the same way as Vaugelas does for French, as an entity to be set up in the course of recording. In both cases, it is worth noting that a dictionary can only be a partial record, because a full record will also include a grammar.

One of the central aspects of being a "record of the English language" is a descriptive rather than prescriptive approach. Section 3 gives some examples of how information that in other dictionaries is encoded as prescriptively oriented usage notes is encoded in OED (2011). Two main tools are used, the selection and content of examples and labels. The labels represent judgements, which in particular in the case of *nonstandard* are not crucially different from prescriptions.

A record is definitive if it has an unmatched authority. In section 1, we saw that this authority was an aim of the dictionary, not just an attribute assigned to it by its users. The claim to achieve this aim makes it more difficult to admit the conceptual problems involved in recording the English language.

In sum, the OED's claim to be "the definitive record of the English language" is problematic. This is not caused by any flaws in the dictionary, but only by the unrealistic

nature of the aim. As there is no empirical entity to be referred to as *the English language*, there cannot be a purely descriptive account of it. This does not mean that the OED can only be interpreted as a prescriptive dictionary, setting up the standard of English. An alternative interpretation is available, in parallel to the one I proposed for learner's dictionaries and for general dictionaries. In this interpretation, the OED provides information about English words, where *English* is taken in the pre-theoretical sense. This information can be used to solve problems. The range of problems is what distinguishes the OED from other dictionaries. It includes questions of the development and use of words over time. Obviously, this argument extends also to corresponding dictionaries for other languages.

## References

### A. Dictionaries

- Chambers 1998.** *The Chambers Dictionary*. S. Anderson et al. (eds.). Edinburgh: Chambers Harrap.
- COD 2011.** *Concise Oxford English Dictionary*. (12th edition.) A. Stevenson and M. Waite (eds.). Oxford: Oxford University Press.
- Collins 1986.** *Collins Dictionary of the English Language*. (Second edition.) P. Hanks (ed.). Glasgow: Collins.
- Collins 2000.** *Collins English Dictionary*. (5th edition), Glasgow: HarperCollins.
- OED 28 October 2011.** *Oxford English Dictionary*. (Third edition), J. Simpson (ed.). <http://www.oed.com/>.
- Wahrig 1997.** *Deutsches Wörterbuch*. (6th edition), R. Wahrig-Burfeind (ed.). Gütersloh: Bertelsmann.

### B. Other literature

- Atkins, B.T. S. and M. Rundell 2008.** *The Oxford Guide to Practical Lexicography*. Oxford: Oxford University Press.
- Bloomfield, L. 1933.** *Language*. (1st British ed. 1935). London: Allen & Unwin.
- Chomsky, N. 1980.** *Rules and Representations*. New York: Columbia University Press.
- Chomsky, N. and H. Lasnik 1995.** 'The Theory of Principles and Parameters.' In N. Chomsky (ed.), *The Minimalist Program*. Cambridge (Mass.): MIT Press, 13–127.
- Collinge, N.E. 1995.** 'History of Comparative Linguistics.' In Koerner, E.F.K. and Asher, R.E. (eds.), *Concise History of the Language Sciences: From the Sumerians to the Cognitivists*. Oxford: Pergamon / Elsevier Science, 195–202.
- Craigie, W. A. and C.T. Onions 1933.** 'Historical Introduction', *A New English Dictionary on Historical Principles, Supplement*. Oxford: Clarendon, vii–xx.
- Cruse, D. A. 1986.** *Lexical Semantics*. Cambridge: Cambridge University Press.
- Greenberg, R. D. 2008.** *Language and Identity in the Balkans: Serbo-Croatian and its Disintegration*. (Updated edition.) Oxford: Oxford University Press.
- ten Hacken, P. 2009.** 'What is a Dictionary? A View from Chomskyan Linguistics.' *International Journal of Lexicography* 22.4: 399–421.
- Harris, R. and T. J. Taylor 1989.** *Landmarks in Linguistic Thought I: The Western Tradition from Socrates to Saussure*. London: Routledge.
- Krishnamurthy, R. 2008.** 'Corpus-driven Lexicography.' *International Journal of Lexicography* 21.3: 231–242.

- Mackridge, P. 2009.** *Language and National Identity in Greece, 1766-1976*. Oxford: Oxford University Press.
- Miller, C. and K. Swift 1980.** *The Handbook of Nonsexist Writing: For writers, editors and speakers*. New York: Harper Perennial.
- Robins, R. H. 1967.** *A Short History of Linguistics*. (Second edition, 1979.) London: Longman.
- de Saussure, F. 1916.** ‘Cours de linguistique générale.’ In C. Bally & A. Sechehaye (1981, eds.), *Édition critique préparée par Tullio de Mauro*, Paris: Payot.
- Simpson, J. 2000.** ‘Preface to the Third Edition of the OED.’ 28 October 2011. <http://www.oed.com/public/oed3preface/preface-to-the-third-edition-of-the-oed>.
- van Sterkenburg, P. 2003.** ‘“The” dictionary: Definition and history.’ In P. van Sterkenburg (ed.), *A Practical Guide to Lexicography*. Amsterdam: Benjamins, 3–17.
- Svensén, B. 2009.** *A Handbook of Lexicography: The Theory and Practice of Dictionary-Making*. Cambridge: Cambridge University Press.
- de Vaugelas, C. F. 1647.** *Remarques sur la langue françoise*. In J. Streicher (1934, ed.). Paris: Droz.
- Wikipedia** ‘Max Müller.’ 6 March 2012. [http://en.wikipedia.org/wiki/Max\\_Müller](http://en.wikipedia.org/wiki/Max_Müller).