The Illusory Distinction between Lexical and Encyclopedic

In NPD *kea* is defined as ‘a large green New Zealand parrot that usually eats insects but sometimes destroys sheep by slashing the back to feed on the kidney fat: *Nestor notabilis*’. When I reviewed this dictionary for a local newspaper, I said of this definition

>While all the facts are correct, and the technical term *insectivorous* has been avoided for clarity, it seems as odd to define a kea in terms of its sheep-destroying properties as it would be to define it in terms of its windscreen-wiper-destroying properties. These are things that keas do from time to time, not things which they are. A kea that never saw a sheep would still be a kea! (Bauer 2001).

A lexicographer friend, commenting on my review, said that the definition looked like a good one, and that my comment was an unfair one.

I then came across (as a consumer this time, not a critic) the CED’s definition of *rampion:*

>‘a bell-flower (*Campanula rapunculus*) whose root is eaten as a salad…’

As an analyst, I have a great difficulty in deciding how much of this information is lexical and how much is encyclopedic. If the definition had read

>‘a bell-flower (*Campanula rapunculus*); the root of the plant as a vegetable…’

I might not have noticed. As it is, the root being eaten looks like encyclopedic information.

As it happened, it was precisely this information that I required in the context in which I had met the word, so that the definition with this marginally encyclopedic information was much more useful than a straight definition might have been.

Nevertheless, there still seems to be a difference between the two examples. The sheep-killing properties of keas seem rather more peripheral to the nature of the kea (which existed for millennia before the arrival of the sheep in New Zealand!) than the food value of rampion seems to be to knowing about that plant.

Consider as a third example, the relationship between *cuckoo spit* and the *froghopper.* Personally, I am familiar with cuckoo spit, but have never knowingly seen a froghopper. What do I want to know about each from a dictionary? It seems to me that I want to know what cuckoo spit is, namely a froth secreted by a larva, and possibly which larva. Of a froghopper I want to know that it is an insect responsible for cuckoo spit (and possibly, why it got such an unlikely name). From an encyclopedia I might want to know what he froghopper looks like, where it is found, how many different insects are covered by the name, etc. We have here, it seems to me, a classic opportunity for circular definition: Froghopper: insect whose larvae produce cuckoo spit; cuckoo spit: froth from the larvae of a froghopper. Not quite circular because of the ‘froth’ and ‘insect’, but perhaps a bit close for comfort. But this is more or less what we get in most dictionaries. Let’s take NODE as a fairly typical example

**cuckoo spit** … whitish froth found in compact masses on leaves and plant stems, exuded by the larvae of froghoppers.

**froghopper** … a jumping, plant-sucking bug, the larva of which produces cuckoo spit.

We might feel we could improve on some of the details here, but basically that seems fine, and is useful enough (even if we don’t find out where frogs come into all of this — the
insect is supposed to resemble a frog, in that it has a wide head with prominent eyes — or the precise nature of the spit — secreted through the anus of the larva and then frothed up by pumping air through it). But how much of it is definition and how much of it is encyclopedic information? Indeed, is there a clear-cut distinction to be made in cases like this?

The problem is perhaps particularly intense in the case of natural kinds. In one sense we might want to say that the kea is defined by its Linnean classification, but that classification is unhelpful information to the large majority of us who are not biologists. We need not only a label, but something which acts as a definition for the non-expert, perhaps in terms of superordinate and restriction. Thus ‘an alpine parrot of New Zealand’ gives a definition of kea which excludes all possible alternative denotata.

The point I wish to make with the examples I have given so far is that it is not always easy to decide where the boundary lies between lexical and encyclopedic information, and that, in any case, encyclopedic information is often of more use to the reader than a purely linguistic definition might be. Nevertheless, there are cases where we may be justified in criticising actual entries in actual dictionaries by suggesting that they err too far in one direction or another. A classic case — added, I believe, by an editor and not a lexicographer — is surely the definition of overdrive given in HNZD:

\textit{overdrive} noun
\begin{itemize}
\item an extra, higher gear which can be engaged in a car to allow high cruising speeds at decreased engine revolutions, thus saving petrol consumption and engine wear. Overdrive can, in some cars, be applied to the intermediate gears, the extra ratios making for increased driving flexibility.
\end{itemize}

The definition was much abbreviated in the second edition!

A wider point, though, is that it is in principle impossible to make a hard and fast distinction between the lexical and the encyclopedic. First, though, to make this point, we need rather firmer definitions of the differences.

In principle, encyclopedic information is information about the denotatum’s interaction with the real world, while lexical information is such as is required to distinguish the denotatum of the relevant word from the denotata of all other non-synonymous words in the language. Even there, you will immediately see problems. The fact that keas live in New Zealand seems like a fact about the world, and yet it is required to distinguish keas from other parrots in New Guinea (although the kea appears to be the world’s only parrot whose prime habitat is alpine: Diamond and Bond 1999, 2). Only by the fiat of giving the requirements of lexical definition priority can we overcome such a problem; then we can say that encyclopedic information is extra-lexical information, with lexical defined as before.

Perhaps the most obvious cases of a lack of distinction between lexical and encyclopedic information are to be found in gender systems. In a language like Russian, gender is determined by the inflectional class of the noun, so that nouns which end in an \textit{–a}, like \textit{kniga} ‘book’ are feminine, words that end in \textit{–o}, like \textit{vino} ‘wine’ are neuter, etc. But there are some exceptions to this very general rule. The exceptions are words like \textit{djadja} ‘uncle’ and \textit{deduška} ‘grandfather’ which look as though they should be feminine nouns, but which are masculine on account of their denotata. Similar rules give us masculine \textit{agricola} ‘farmer’ in Latin, although the noun declines like a feminine noun (Corbett 1991, 37). Now we might

\begin{itemize}
\item This should not be read as suggesting that a Linnean classification should be avoided in dictionaries; such labels are useful in a number of ways, in particular in that they allow us to check whether different labels refer to the same plant or animal. Nevertheless, they do not help lay-people to understand what the word means in terms of their own experience of the world.
\end{itemize}
argue fruitlessly about whether it is the lexical fact that an uncle is defined as ‘a male sibling of one’s parents’ that is the determinant of this masculine gender or the fact that uncles in the real world tend to show male characteristics such as having beards, being married to women, wearing certain clothes, and so on. But in the final account, the lexical fact is determined by the encyclopedic fact of particular real-world features. Thus the lexical information is derivable from and not independent of the encyclopedic information, and the two cannot be easily separated.

A second place where it is impossible to distinguish between lexical and encyclopedic information is where a word makes direct reference to some feature of a real-world entity. Words which involve colour terms are clear examples of this: greenstone, red lead, robin red-breast, yellowhead; but there are other examples as well such as bald crow, gold ring, hard hat, slow march. Indeed, in the light of such examples, we also need to think about compounds such as bear cub, earthquake, lighthouse or silkworm. In each case they are so named because of a real-world link with bears, the earth, a light and silk respectively rather than because the semantics associated with bear happens to fit quite well with the semantics associated with cub (and so on). This is not a necessary feature of compounds, as we can easily illustrate: monkey business has nothing to do with real-world monkeys, if fox-gloves have anything to do with real-world foxes it is no longer clear what, and cat gut does not come (and has never come) from cats. Thus there are compounds whose elements do not refer directly to the real world, while others do, and with this second group (the words like silkworm), too, it is impossible to separate encyclopedic reference from lexical semantics.

Let us return to the question of lexical information and definition. The case of the kea is perhaps slightly dubious, so let us instead consider its distant cousin, the kakapo. The kakapo is

• The most endangered parrot in the world (there are about 86 left at the time of writing — see <http://www.kakaporecovery.org.nz/then/recovery.html>.
• The largest parrot in the world.
• The heaviest parrot in the world (male individuals can weigh up to 4.5kg).
• The only flightless parrot in the world.
• The only nocturnal parrot in the world.
• The only parrot to use the ‘lek’ breeding system.

But if we are trying to define the kakapo for a dictionary, which of these features should we take as the defining one, and which should be taken as encyclopedic information? I take it that we can agree that the rareness of the parrot is a fact about the world, not a fact about the parrot (since another parrot could at any time become more endangered than the kakapo, without that affecting how we would wish to define the kakapo), but any of the others seem to be good candidates for a definitional statement. How do dictionaries actually deal with this?

the New Zealand owl-parrot, large-winged but almost flightless (CED)
a ground-living nocturnal parrot, Strigops habroptilus, of New Zealand, resembling an owl (CoED)
a large, almost flightless, and now rare, nocturnal parrot, Strigops habroptilus, of New Zealand (MQ3)
a large flightless New Zealand parrot with greenish plumage, which is nocturnal, ground-dwelling, and now endangered (NODE)
a chiefly nocturnal burrowing New Zealand parrot with green and brown plumage: *Strigops habroptilus*. (NPD)

In principle, all of these dictionaries are giving encyclopedic information about the bird, not least that it is a New Zealand bird. But in any text, the apparently encyclopedic information may be more useful than the apparently lexical information, and in any case, it is not clear what information belongs in which category. Consider the following actual text (written in the nineteenth century):

The ferrets have not got among the birds on the river, evidently, as Kakapo's are squealling [sic] about in hundreds. Will have to tie up the dog if I don't want a camp full of corpses in the morning.

Pascoe, John (ed.), *Mr Explorer Douglas*. Wellington: Reed. 1957, 123

Obviously the rarity of the bird was not in question at the period, although Douglas foresaw its fate. What would really help here is the clearly encyclopedic information that the kakapo has no idea of self-defence (despite having a strong beak and strong claws). It is strong enough to crush a ferret and see off a dog, but it does neither — fine in the mammal-less world of pre-settlement New Zealand, but not conducive to survival today.

Those of us who use dictionaries, and those of us who write them, know that however much science there is in lexicography, there is also a great deal of art, and that a scientifically accurate definition is not necessarily ipso facto a good definition. It seems to me that we must not denigrate the art side of writing definitions: the lexicographer may have a better feel for a good definition than a strict scientific approach would provide, and certainly may be better able to provide a definition which is more helpful to the poor human reader. Of course, when we give lexicographers the freedom to provide these slightly encyclopedic definitions, we also expect them to behave responsibly, since it is easy to go too far — something which is not useful for the publisher (it uses up too much space) or for the reader (the definition ceases to be enlightening, fails to distinguish fact from myth, or reads like an advertisement). Somewhere in the variation that this allows is the leeway that lets the consumer prefer the definitional style of one dictionary over another. That, surely, is a good thing. But I hope to have demonstrated that any hard and fast theoretical distinction between lexical and encyclopedic information would not only be (in the dichotomous terms of *1066 and All That* — Sellar and Yeatman, 1930) a ‘Bad Thing’ but would lead the lexicographer into an untenable position.

Lexica

References


