

A sweetshop for lexicographers

Jaap Parqui



Jaap Parqui studied psychonomy and Italian linguistics and literature at the University of Leiden. At the beginning of his studies he saw a note calling for trainees to work on an Italian-Dutch dictionary project and knew that this was what he wanted to do. A couple of years later he entered this project as a trainee, eventually rising to become an editor. He then moved on to Van Dale, first as an editor and then as a publisher. He is still fascinated by the idea that in a dictionary the whole world is collected and neatly ordered. jaap.parqui@vandale.nl

Eatable marshmallow pillows
Lickable wallpaper for nurseries
Hot ice creams for cold days
Cows that give chocolate milk
Fizzy lifting drinks
Square sweets that look round
(Roald Dahl, *Charlie and the Chocolate Factory*, 1964. NY, Knopf)

Since the outbreak of the economic crisis in 2008, sales of reference books have fallen considerably in Holland. Electronic dictionaries are meanwhile flourishing. Print dictionaries have been perfected over the centuries within the tight limits of the printed page. It is hard to introduce new ideas to them. Electronic dictionaries, however, are a recent innovation and there are still numerous ways of improving them. There are so many platforms to develop, so many functionalities to invent, so many more sources from which to harvest material, so many new ways to get in touch with the public...

For both publisher and user this is a sweetshop of new possibilities. There are so many interesting, useful and inspiring possibilities that the biggest challenge now is to find a focus and pick the right ideas. Because of the myriad of options offered by the electronic revolution, it is hard to predict which road lexicography in general will take. I will now discuss some of the new developments Van Dale has chosen to work on.

Update cycle

Is it the changing view of the lexicographer or do new words arise (and disappear) faster nowadays? Although the word *hype* was invented in the nineteen twenties, it now seems more alive than ever. With online subscription models, customers expect their data to be right up-to-date. This expectation is challenging. What should be the policy for admitting new words into dictionaries? And will the back office be able to maintain this speed for so many dictionaries simultaneously?

In the *Dikke Van Dale*, our monolingual flagship, we used to include only established words. In practice, this meant that a word needed to have been used in general sources for at least three years. And even now, in the online era, we do not want our dictionary to be cluttered with ephemera and flashes in the pan, especially those that have already

become dated. That is why we decided to give new words a temporary status. New words are accompanied by the message 'temporarily added' plus the year of their appearance. After a couple of years we decide which of these 'temporary' words have survived the test of time.

Only a couple of years ago, if we wanted to publish a new edition of a book, we needed at least three months to perform all the checks. The file would then go to a typesetting company, which would take another couple of weeks, and it is only then that printing would start. In order to publish more often, we therefore needed to increase efficiency greatly. We transformed our data files into XML, made them more uniform and invested in automatic typesetting from XML. Most of the checks are now done along the way. We are still fine-tuning the process. This year our most important online dictionaries will be updated twice and we are working towards an even higher frequency.

Crowd sourcing

With better access to all kinds of corpora the quality of our dictionaries has improved considerably. It can still happen though that we are slow to spot a new word, or that there is an even better translation or definition. We do not cover highly specialised terminology. And you can never have enough phraseology.

We are developing an option that will allow users to add new words, new phrases and new translations or definitions to our dictionaries. These will appear, with their own very distinct look and feel, alongside the 'real' dictionary.

I see many advantages. It could be a very good way to include much more information in our dictionaries. It is a nice way of developing a closer tie with our users. And it can be a good source for the editors, who will promote information from the 'people's dictionary' to the official dictionary.

It is also a leap in the dark. Will the information the crowd adds make sense? Will users actually make use of this option? How will they perceive the difference between official and non-official content? That is why we will give this *Van Dale Wiki* a soft launch. We will start with just one dictionary and a selected group of participants. And from there we will slowly expand this exciting experiment.

Search functionality

We polled our users. The first reason they consult our explanatory dictionary is to look up the meaning of a word. The second reason is to look up the spelling. At present you can only find a word in our online dictionaries if you already know how it is spelled. That is why we are working on a 'suggestion module' that will offer a short selection of words for which you may have been looking. This will be similar to what Google or spell checkers currently do.

It is already possible to look for phrases, by entering one or more words without worrying under which headword you should look. Many other smart-search options are also imaginable.

Flexion forms

Many users will need conjugations of French verbs, the flexion of German nouns or plurals of Dutch nouns. Print dictionaries have very little space for flexion forms and usually limit themselves to key forms, or to a reference to a model verb or noun. In electronic dictionaries we have abundant space. Producing all these forms in so many languages is a major project. We do this portion by portion and try to find the cleverest balance between automatic expansion and handiwork to complete the paradigms.

In paper dictionaries flexion forms are typically given after the headword. When a user looks up a translation in a foreign language, however, that is typically the moment when he or she needs to see its inflected forms. We will therefore place a small button next to the translation that will pop up a window with the forms. We will do something similar for spoken pronunciation.

Images

It is a cliché, but sometimes a picture does say more than a thousand words. Some print dictionaries do have pictures, of course, but once again electronic dictionaries offer much more space. The production of pictures is very expensive and so we are considering placing links to pictures on the Internet. And we are not limiting ourselves to stationary images. For words that describe movement – dances, for example, or a horse's gait – film is much more appropriate.

Language platform

Translators often tell us that they look up a word in a bilingual dictionary and then check that translation in an explanatory dictionary. They might also want to check specialised dictionaries or read paragraphs on usage. I am doing it myself, as I write this article in a language that is not my own. It is therefore important that our users

should have a choice of many products and be able to switch effortlessly between them. We built an online service, into which it is relatively easy for us to incorporate new products. We are developing new products ourselves, but we also collaborate with other publishing houses – *Oxford English Dictionary* and *Oxford English Thesaurus*, for example. In this way we are building a complete online platform that will provide users with many different kinds of reliable language information.

Books

Does this mean we will not make books anymore? We love the possibilities of electronic dictionaries, but we also love the smell and the romance of books – and the fact that you can use them without your computer. And so do many of our users. Once you have invested in the data, it is not a huge investment to publish it in print too. I expect we will publish books for many years to come.

Conclusion

So where will we be in ten years? As the saying goes, it is hard to make predictions, especially about the future. If we reason along the lines I have sketched here, in ten years time we will have high quality dictionaries on many devices and platforms, with lots of data, both official and user-generated, not just words, but also sounds and images that are very easy to search for on many devices and platforms. I think dictionaries will then be more reliable, richer and easier to use than ever before.

Van Dale Publishers

It is more than 150 years since Johan Hendrik Van Dale died a slow and horrible death of smallpox just before finishing his *Van Dale Groot woordenboek der Nederlandse taal*. It was finally completed by his student, Jan Manhave. If only Johan Hendrik had known the status his *Dikke Van Dale*, 'Fat Van Dale', would go on to attain in Holland and Belgium! In the nineteen seventies a publishing house was established around this title and a series of large translation dictionaries evolved. Over the years the portfolio has grown, and nowadays Van Dale publishes dictionaries in many shapes and sizes as well as grammars, usage books and books for children. In the last couple of years Van Dale has also started in-class language courses and a translation service. <http://vandale.nl/>

X International School on Lexicography – Life Beyond Dictionaries

Ivanovo State University is celebrating the tenth anniversary of its International School on Lexicography by holding the current event in two different places simultaneously: Ivanovo (Russia) and Florence (Italy). In Florence we plan to discuss questions connected with cultural tourism and heritage dictionaries, cultural aspects of learner lexicography and new dictionary projects, while in Ivanovo participants will focus on dictionaries for special purposes, historical lexicography and problems of the modern lexicographic scene. Online video conference will give us the opportunity to have personal interaction without being at the same location.

Florence was chosen to co-host this conference in view of the cooperation that our university has had since 2004 with the local Romualdo Del Bianco Foundation – Life Beyond Tourism, which promotes international events, including our annual international student workshops. This year the workshop participants will present their projects as part of the conference.

Invited speakers include Linda Mugglestone (Oxford University) and Sandro Nielsen (Aarhus University) in Ivanovo, and Paolo del Bianco (Romualdo Del Bianco Foundation), Janet De Cesaris (Universitat Pompeu Fabra) and Geoffrey Williams (Université de Bretagne Sud).

<http://en.lex2013.16mb.com/>

Olga Karpova

olga.m.karpova@gmail.com