Adapting the dictionary entry structure and DWS configuration for creating a dictionary aimed to be published on paper, online and as electronic dictionary software for PC and mobile

Anna Rylova

ABBYY / Moscow State University

E-mail: Anna_Ry@abbyy.com

Abstract

Publishing one and the same dictionary on paper, as a desktop application, online and mobile application is a common practice in most of publishing houses. But the type of media the dictionary will be published on must be taken into account on the very first stages of developing the dictionary concept, the Style Guide and defining the entry structure. The formats of dictionary publishing give a lot of new challenges to lexicographers who have to keep in mind the dictionary users' needs, apply linguistic theory and always think about how the dictionary is intended to be published. The factor of dictionary publishing media influences not only the editorial decisions, but also a Style Guide and in the same time it affects the Dictionary Writing System design. The DWS must be adopted for the creation of the dictionary entry that will have different structure and visualization in paper and electronic dictionary. In this paper we describe how dictionary data and a dictionary writing system can be adapted for the various formats in which dictionaries can be published today. Our research is based on our own experience of creating dictionaries within our team of lexicographers, developing a professional dictionary writing system ABBYY Lingvo Content and electronic dictionary ABBYY Lingvo (since 1989) for PC, mobile and online.

Keywords: dictionary writing system; dictionary publishing; entry structure

When creating a dictionary a lexicographer has to keep a lot of factors in mind - these are the linguistic theory issues and the dictionary end-user needs. Every time when the entry is being edited, a lexicographer tries to make it relevant and accessible to the dictionary users. Linguistic theory and end-user needs affect every aspect of the lexicographic work and play an important role in a dictionary Style Guide creation process. But in the era of electronic lexicography even more factors come into play. One of the important factors that have a great impact into process of dictionary making is the type of media on which the dictionary is intended to be published.

Publishing one and the same dictionary on paper, as a desktop application, online and mobile application is a common practice in most of publishing houses. But the type of media the dictionary will be published on must be taken into account on the very first stages of developing the dictionary concept, the Style Guide and defining the entry structure.

The common examples of differences between electronic and paper versions of the dictionary are as follows: 1) the electronic dictionary normally include collocations and idioms as separate entries to facilitate the search process while the paper dictionary usually includes the collocations in the entry with a main word; 2) the electronic dictionary can include more examples while the paper dictionary has a limited number of examples because of space limitation; 3) the electronic dictionary includes the full forms of labels while the paper dictionary includes abbreviations; 4) the electronic dictionary usually has the full headword forms in the entry while the in the paper dictionary the headword is normally replaced with the tilde; 5) dictionary entries in electronic and paper dictionaries have different visualization.

Though we are mentioning the electronic dictionaries in general, every type of electronic dictionary has its own specifics. For example electronic desktop applications often include some features for language learners e.g. special functions for words memorizing; online dictionaries have functions of logging the users' queries thus helping publishers to update the dictionary content, and surprisingly the mobile dictionaries have the problem of space (the same as paper dictionaries) because they have limited memory size and mobile gadgets often have a small screen.

So when creating a dictionary entry aimed to be published on paper, online and on handheld devices we have to think the entry structure over very thoroughly. The principle "one database – many dictionaries" give a lot of new challenges to lexicographers who have to keep in mind the dictionary users' needs, apply linguistic theory and always think about how the dictionary is intended to be published.

The factor of dictionary publishing media influences not only the editorial decisions, but also a Style Guide and in the same time it affects the Dictionary Writing System design. The DWS must be adopted for the creation of the dictionary entry that will have different structure and visualization in paper and electronic dictionary. The dictionary makers and DWS developers must work together to adjust the DWS for the needs of publishing the same dictionary on different media. It is preferable when DWS, electronic dictionary software for PC and mobile devices and online dictionary are provided by the

same software developer as it reduces the time spent for negotiations and software customization.

In this paper we describe how dictionary data and a dictionary writing system can be adapted for the various formats in which dictionaries can be published today. Our research is based on our own experience of creating dictionaries within our team of lexicographers, developing a professional dictionary writing system ABBYY Lingvo Content and electronic dictionary ABBYY Lingvo (since 1989) for PC, mobile and online.

The most exciting experience of adapting the dictionary data we gained while publishing ABBYY Lingvo Comprehensive English-Russian Dictionary that was published electronically in 1990 and then permanently updated by ABBYY team of lexicographers with a new electronic version (desktop and online) published every one or two years. This dictionary was available on handheld devices (Pocket PC and Palm) since 2003 and on mobile phones since 2006. And it was only in 2007 when the first paper edition has been released. The lexicographers had a lot of new tasks while adapting the electronic dictionary to the paper publishing and the DWS developers helped them to configure the DWS for their needs.

In the presentation we describe the following details:
1) how the entry structure can be adapted for the different kinds of publishing – paper, online and mobile;
2) which DWS functions and which mark up should be used to avoid a lot of manual work and process the dictionary data automatically while preparing it for publishing;
3) what is the best way to make a comfortable visualization of dictionary in the DWS interface so that lexicographer can see how the entry will look like in electronic and paper dictionary.

All the three tasks were solved successfully while preparing ABBYY Lingvo Comprehensive English-Russian Dictionary for paper publication and the other dictionaries (about 40 of them) that were created by ABBYY lexicographers or in collaboration with publishing houses and individual professionals.

The Requirements for the dictionary writing system were worked out in close collaboration with editors, lexicographers, the developers of the electronic versions of the dictionary for different platforms, and the staff of the ABBYY Press publishing house.

First, the key features of each version of the dictionary were identified, which led us the following list of questions:
- ✓ What is the optimal size of dictionary entries (space is crucial)?
- ✓ How can we make the layout more user-friendly?
- ✓ How users will interact with dictionaries in different formats?
- ✓ What opportunities each publishing format provides?

In answering the above questions, we developed different sets of styles and export settings for each entry section.

When creating a new dictionary or preparing an existing dictionary for publishing, the following choices are made for each entry section that is described in the Style Guide. These choices affect how the corresponding section is displayed in the various formats and how entries are exported from the dictionary writing system.

1) How the name of the section will displayed (e.g. "Idiom" for the electronic version and "◇" for the printed version)
2) How section levels will be numbered (numbers for homonyms, senses etc.)
3) How the sections will appear on-screen and on the printed page (fonts, line breaks, colours)
4) How many same-level sections will be allowed (e.g. no more than 5 examples in the printed version)
5) Which sections will make it into each version of the dictionary (e.g. the etymology section may be excluded from the mobile version)
6) How headwords will be displayed in the body of the entries (they can be displayed in full or replaced with a swung dash)
7) How headword alternatives will be handled (e.g. separate entries in the mobile version and one entry with two headwords in the printed version)
8) How labels will be displayed (e.g. abbreviated in the printed version and spelled out in the electronic version)
9) Which entries will make it into each version (the paper or mobile version may contain fewer entries than the electronic version)

Once the appropriate settings are selected for each entry section, the styles and export settings are selected to suit the output format.
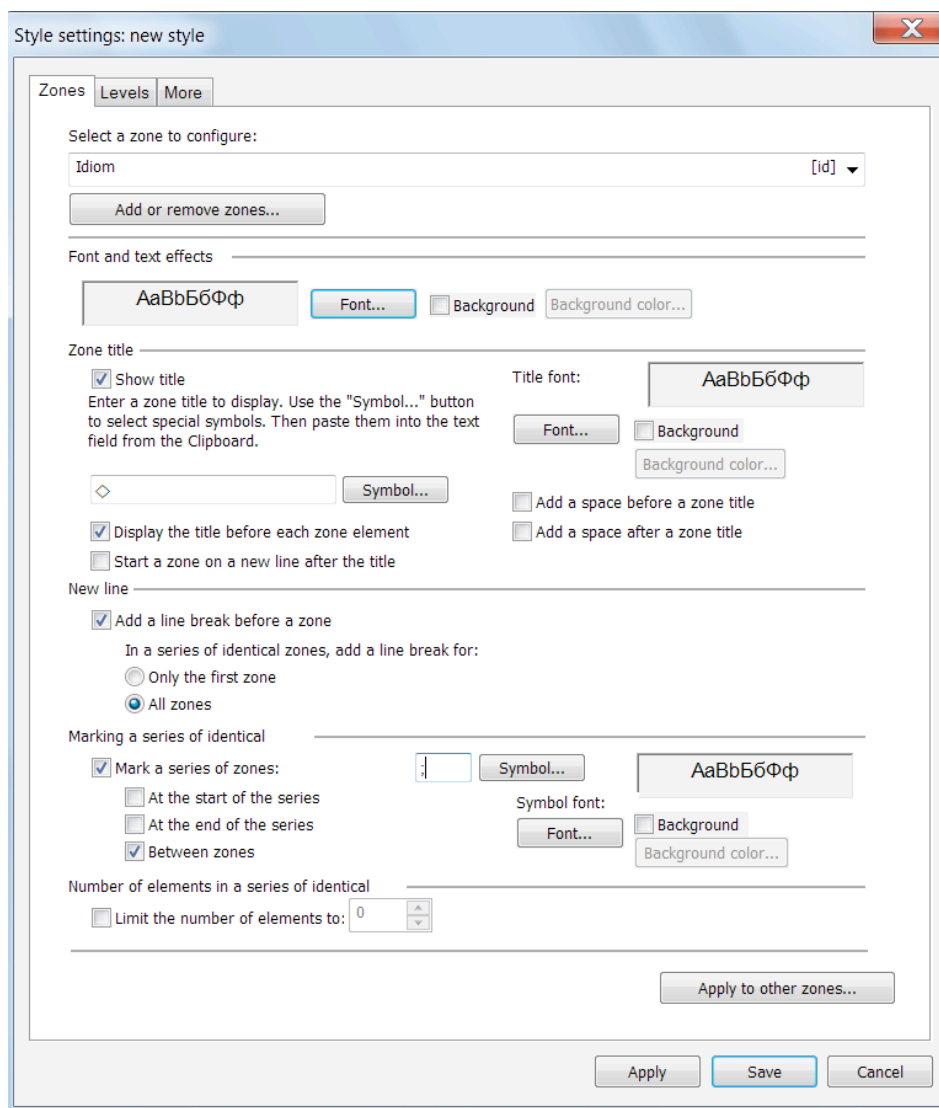
Figure 1. An example of DWS styles settings interface.

Once the lexicographer or editor has selected the settings for each entry section, they can proceed to create dictionary entries in a dedicated window of the dictionary writing system. This window displays the structure of the entry, which the lexicographer needs to fill with content. At any moment, the lexicographer can invoke one of the preview windows and view the entry as it will appear in the printed, mobile, CD, or online dictionary. The lexicographer can easily switch among the preview windows and adjust the display and export settings if required. The lexicographer can also at any moment export an entry into the printed format for proofreading purposes. There is also an option to export the entire dictionary or selected entries into the Lingvo electronic format.
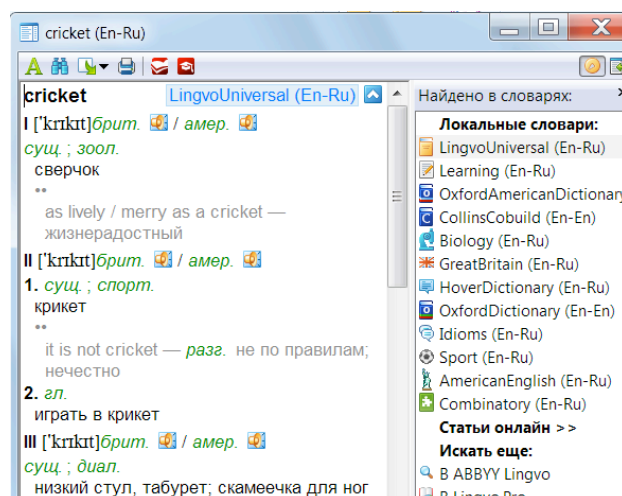

Figure 2. ABBYY Lingvo electronic dictionary

cricket¹ ['krɪkɪt] *n зоол* сверчок ♦ as lively / merry as a ~ жизнерадостный
cricket² ['krɪkɪt] I *n спорт* крикет ♦ it is not ~ *разг* не по правилам; нечестно
II *v* играть в крикет
cricket³ ['krɪkɪt] *n диал* низкий стул, табурет; скамеечка для ног

Figure 3. ABBYY Lingvo paper dictionary

Thus, the lexicographer can easily see how an entry will appear in the various formats while still working on the dictionary.

A new century is posing new tasks for the lexicographers and new software technologies come to help the lexicographers in their everyday work. The collaboration between practicing lexicographers and software developers is very fruitful and helps both parties to facilitate their work and move further to new creative tasks. Thus the collaboration between ABBYY lexicographic team, ABBYY Lingvo Content software developers, ABBYY Lingvo desktop, mobile and online dictionary developers and ABBYY Press publishing house helped to develop new DWS features and a new methodology of dictionary creation.

References

Anokhina, J. (2010). Lingvo Universal English-Russian Dictionary: Making a Printed Dictionary from of an Electronic One. In *Proceedings of 14th EURALEX International Congress*. Leeuwarden, the Netherlands.

Atkins, B.T.S., Rundell, M. (2008). The Oxford Guide to Practical Lexicography. Oxford: Oxford University Press.

Dziemianko A. (2010). Paper or electronic? The role of dictionary form in language reception, production and the retention of meaning and collocations. *International Journal of Lexicography,* 23(3), pp. 257-273.

Grefenstette, G. (1998) The future of linguistics and lexicographers: will there be lexicographers in the Year 3000? In T. Fontenelle *et al*. (eds.) *Proceedings of the Eighth EURALEX Congress* Liege: University of Liege, pp. 25-41.

Hockey, S.M. (2000). Dictionaries and lexical databases. In *Electronic Texts in the Humanities: Principles and Practice.* Oxford & New York: Oxford University Press, pp. 146-171.

Heid, U. (2009) Aspects of Lexical Description for Electronic Dictionaries. In S. Granger, M. Paquot (eds*.) eLexicography in the 21st century: New challenges, new applications, Proceedings of eLex 2009, Louvain-la-Neuve, 22-24 October 2009*. Louvain-la-Neuve: Presses Universitaires de Louvain.

Kuzmina V., Rylova A. (2009) The ABBYY Lingvo electronic dictionary and the ABBYY Lingvo Content dictionary writing system as lexicographic tools. In S. Granger, M. Paquot (eds*.) eLexicography in the 21st century: New challenges, new applications, Proceedings of eLex 2009, Louvain-la-Neuve, 22-24 October 2009*. Louvain-la-Neuve: Presses Universitaires de Louvain.

Moon, R. (2008). Dictionaries and collocation. In S. Granger, F. Meunier. (eds.) *Phraseology. An Interdisciplinary Perspective*. Amsterdam: Benjamins, pp. 313-336.

Rundell, M. (2009) The Road to Automated Lexicography: First Banish the Drudgery… then the Drudges? In S. Granger, M. Paquot (eds*.) eLexicography in the 21st century: New challenges, new applications, Proceedings of eLex 2009, Louvain-la-Neuve, 22-24 October 2009*. Louvain-la-Neuve: Presses Universitaires de Louvain.

Rylova, A. (2010) Electronic Dictionary and Dictionary Writing System: how this duo works for dictionary user's needs (ABBYY Lingvo and ABBYY Lingvo Content case). In A. Dykstra, T. Schoonheim, (eds.) *Proceedings of the XIV Euralex International Congress*. Leeuwarden, 6-10 July 2010, Leeuwarden: Fryske Akademy.

Verlinde, S., Binon, J. (2010). Monitoring Dictionary Use in the Electronic Age. In A. Dykstra, T. Schoonheim, (eds.) *Proceedings of the XIV Euralex International Congress*. Leeuwarden, 6-10 July 2010, Leeuwarden: Fryske Akademy.