



EURALEX XIX
Congress of the
European Association
for Lexicography

Lexicography for inclusion

7-11 September 2021
Ramada Plaza Thraki
Alexandroupolis, Greece

www.euralex2020.gr

**Proceedings Book
Volume 1**

Edited by Zoe Gavriilidou, Maria Mitsiaki, Asimakis Fliatouras

EURALEX Proceedings

ISSN 2521-7100

ISBN 978-618-85138-1-5

Edited by: Zoe Gavriilidou, Maria Mitsiaki, Asimakis Fliatouras

English Language Proofreading: Lydia Mitits and Spyridon Kiosses

Technical Editor: Kyriakos Zagliveris



This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License

2020 Edition

Dictionnaire des francophones - A New Paradigm in Francophone Lexicography

Dolar K.¹, Steffens M.², Gasparini N.³

¹ CREE, Inalco

² Department of Languages, Literature and Communication, Utrecht University

³ Institut International Pour La Francophonie (2IF)

Abstract

Dictionnaire des francophones (DDF) is a general francophone dictionary, the result of an institutional-collaborative project, the goal of which is to provide a new online resource. It aims to cover all varieties of the French lexicon from a descriptive point of view and to highlight the plurality of linguistic norms while endeavouring to treat different linguistic varieties equally. The paper focuses on the dictionary-making process and lexicography technologies used in the project. Some particularly innovative aspects of the DDF are discussed, such as the institutional support and the scientific background in which the project is grounded; the hybrid nature of the dictionary, combining imported resources in a relational database, enriched by a complex speaker-based collaborative input; inclusivity of linguistic variation and the modes of its representation. Taking into account these characteristics as well as some other features of the dictionary lead us to the conclusion that the DDF is a unique object in comparison to existing traditional and collaborative resources, providing a new paradigm in francophone lexicography.

Keywords: *Dictionnaire des francophones*; professional dictionaries; collaborative lexicography; general dictionary; francophone dictionary; linguistic variation of the French language; plurality of norms

1 Introduction

The lexicon, like any part of language, varies across contexts. The diatopic and diastratic varieties of French have been the subject of a long tradition of lexicographic description and analysis. In contrast to the prescriptive ideology of the second half of the 20th century (for example *Refrancisons-nous* by Frère Jean-Ferdinand in Québec in 1951, *Chasse aux belgicisms* by Hanse et al. 1971, etc.), a lot of recent research, often focusing on lexical variation, describes and values linguistic varieties (for example Rézeau 2001; Mercier & Verreault 2002; Thibault 2004 and 2008; Glessgen & Thibault 2005; Bernet & Rézeau 2010; Francard et al. 2015). Research in this field has also led to the production of quality linguistic atlases and glossaries (ALW; Dulong & Bergeron 1980). These resources, mainly focusing on lexical variation, are most often grounded in the field of contrastive lexicography, aiming to describe a well-defined subset of words used in a particular region or by a specific community of speakers. To date, there is no general dictionary of French integrating all its varieties and meeting scientific standards.

The paper focuses on the dictionary-making process and lexicography technologies used in a new online resource pursuing this aim, *Dictionnaire des francophones* (DDF). Its public launch is planned in late 2020 but might be postponed due to the health crisis. The DDF is a general francophone dictionary, the result of an institutional-collaborative project which covers all varieties of the French lexicon from a descriptive point of view. It integrates endogenous norms and highlights the plurality of linguistic norms. The DDF is a hybrid object combining several existing dictionaries and collaborative input (under a free license and open access). Three innovative aspects of the DDF will be highlighted in the present paper: its specific institutional and scientific background, its hybrid structure as well as its inclusivity with regard to linguistic variation.

2 The institutional and scientific context of the DDF

To define the role that the DDF wishes to play in the field of current French and francophone lexicography, it is necessary to present the institutional context of its conception and its descriptive scientific ambition in more detail.

2.1 Institutional background

The DDF is an ongoing lexicographic project led by the *Délégation générale à la langue française et aux langues de France* (DGLFLF), the department of the French Ministry of Culture in charge of language protection and planning, and the *Institut international pour la Francophonie* (2IF, a part of Université Jean Moulin Lyon 3). The conception and designing process was led by the authors of this paper and reviewed by an international scientific committee, chaired by Bernard Cerquiglini.

Benefiting from the financial and technical support of various francophone organizations, the DDF is mostly under an open data license, setting an example of open cultural data, as a part of the proactive policy from the French Ministry of Culture as well as of the European Plan S (<https://www.coalition-s.org/>), aiming to make the research data accessible to the public. The DDF is one of the few institutional lexicographic projects (another example is *LEO*, a multilingual dictionary with collaborative input, designed by Technical University of Munich), that consciously promote the collaborative input of speakers in the description of their own language. To date, no other lexicographic project joins all

of these aspects, which puts the DDF in a unique position.

2.2 A francophone dictionary with a plurality of norms

The body of research on variation in the French language across time, space and society (Gadet 2003; Völker 2009), led by lexicologists, lexicographers, dialectologists or sociolinguists, highlights the specificities of different communities of speakers. Although, considering the entire lexicon (i.e. what varies but also what is common to all French speakers) has proven to be useful for both historical linguistics (Chambon 2006; Greub 2002) and synchronic linguistics (Baronian & Martineau 2009; Courbon 2012): it seems that a rigorous definition of units is necessary both for understanding the mechanisms of language evolution over time and the description of its current functioning on phonetic, morphological, syntactic, lexical and pragmatic levels. Nevertheless, the description of the relation between varieties and common lexemes is often not well developed (Poisson 2002; Violette 2006; Guérin 2008).

The contrastive approach favoured so far was driven by a real descriptive ambition which valued the varieties of French language and enabled the collection of precious linguistic data. However, this approach itself isolates and confines a lot of French-speaking communities' uses at the margins. Regionalisms and sociolects in general dictionaries are treated the same way, marked as a deviation in regard to the provided linguistic norm. The fact that regional lexemes are integrated in the same lexical networks as "standard lexemes" is not visible. If the existing lexicographic description of regionalisms helps to exhibit the living heterogeneity of the French language (Bavoux 2008: 17), it fails to give an entire and faithful picture of the actual use to the public at large and de facto maintains what Robillard calls the "platypus syndrome" (Robillard 2008: 325, our translation): language is presented as a juxtaposition of badly assembled parts.

Thus, when speakers commonly use regional or socially marked lexemes, sometimes without any awareness of the deviation from the norm conveyed by dictionaries, the referential resources may give them a devalued or folkloric image of their own linguistic practice. In Quebec, such considerations led to the creation of *Usito*, an online reference dictionary defining the endogenous North American French standard. The recent transition of this resource to free online access is an important step to a larger dissemination of scientific descriptions of francophone varieties.

Among the existing collaborative resources, only *Wiktionnaire*, the French part of *Wiktionary*, offers the theoretical possibility of covering the entire French-speaking field, but the chosen mode of collaboration, based on consensus (new contributions can modify and overwrite the existing ones), tends to privilege the variety of the majority of the actual contributors, in this case Metropolitan French. This predilection is especially visible in the definitions: for example, if the Quebec pronunciation of *tofu* [tofy] is mentioned in *Wiktionnaire*, although without specific labelling, it is the form *soja* (and not *soya* used in Quebec) which is mentioned in the definition (Vincent 2016 and 2017). The description of diatopic and diastratic varieties in such resources might be promising, but it still lacks homogeneity, precision and reliability. We need a general dictionary of French, one that meets scientific standards and is fully grounded in the pluricentric French-speaking world (Lüdi 2012), where all regional varieties are included and equally valued. The DDF aims to fill this gap with an open crowdsourced approach.

3 DDF – A hybrid object

Generally speaking, the scientific resources face a major problem in trying to keep their content up to date. The process of integrating new data provided by the scientific community is long and complex and the input from speakers is most often very poor or practically non-existent. From that perspective, the data input in the DDF, as well as its structure, is particular: it networks existing resources, which can be enriched through a collaborative speaker-based input¹. We will briefly present the data model, the structure and the resources of the DDF.

3.1 Data model of the DDF

Combining several existing resources and opening the dictionary to the speakers necessarily implies a reflection on its form. All dictionary articles have an underlying formal structure (see Atkins and Rundell 2008; Renders 2015), the choice of which conditions and constrains the conveyed information (Mazziotta 2016). The development of native electronic dictionaries (such as *Usito*) may lead to conceptualizing dictionaries differently, for example by adding graphs to simple tree structures directly imposed by the medium of print (Heiden 2004; Měchura 2016). This graph structure allows multiple and personalized access, without traditional constraints (Steinlin et al. 2004), by designing an adapted modular consultation interface and its corresponding mobile applications.

The DDF does not function as a simple resource portal, gathering and displaying search results from different separate resources, but consists of a structured database integrating and networking different sources of data and content. The first step was to define its structure in order to be able to adequately tag different imported resources and enable a collaborative input. As in many other digital dictionaries, the Ontolex Lemon model (McCrae et al. 2017), a standard model in lexicography and terminology, and its lexicographic module Lexicog, seemed to be the best choice for the creation of the RDF database (Resource Description Framework, W3C 2004a, Měchura 2016). Some changes have been applied to the model in order to obtain a more fine-grained labelling of language varieties and to point out semantic relations between lexemes. For instance, the property Place (location) was added to each written form and definition since one of the main objectives of the DDF is to highlight the diatopic variation of French. The model used in the DDF is open and reusable (see full model in Steffens et al. 2020).

In this RDF database, every entry is based on the written form of the lexeme related by the LexicalSense property to

¹ A similar method was used in the early years of *Wiktionnaire*, where the 8th edition of *Dictionnaire de l'Académie française* (1932-1935) and *Dictionnaire de Littré* (1883) were used as a substrate.

linked data. Each entry is attributed to an author. The DDF opted for a classic principle in data modelling: each entry in the database corresponds to a form, a meaning, an example and a set of labels related to the form by its meaning. If one form has two or more definitions, different entries can be related through the SenseRelation property indicating semantic relations between definitions. The same set of relations is used to organize semantic relations between different lexical items. An entry could also have no definition, only relations to other entries, as flecational forms for instance. The specificity of the linked data used for the DDF is that each information has a uniform resource identifier (URI) and all data is integrated in a large network connected with other networks that contain lexicographical or conceptual data. The structure is not hierarchical but a bit more verbose than in some other databases. Another way to access the data, beyond the public interface, is through the SPARQL request language. SPARQL is derived from SQL language and allows complex queries in the database, for example a list of words with a specific sequence of letters, a synonym, at least two examples and a geographical indication. Results of SPARQL queries could be displayed as tables or maps, another innovative aspect of the DDF.

3.2 Existing dictionaries included in the DDF

In the DDF database, essentially thanks to the support of *Agence universitaire de la francophonie* (AUF), these existing resources are aligned following the same data model:

- Inventaire des particularités lexicales du français en Afrique noire* (Équipe IFA 1988) gives lexical equivalents between French spoken in Benin, Burkina Faso, Cameroon, Central African Republic, Ivory Coast, Mali, Niger, Democratic Republic of the Congo, Rwanda, Senegal, Chad and Togo on the one hand and French spoken in France on the other. The resource was digitized and tagged; the content is richly described, but sometimes outdated, and an update is necessary.
- Dictionnaire des belgicisms* (Francard et al. 2015), as its name indicates, is a dictionary collecting the particularisms of Belgian French. It has been digitized and tagged before being imported in the DDF.
- Base de données lexicographiques panfrancophone* (<http://www.bdlp.org/>) is an online database which gathers in one place several scientific dictionaries with lexical items used in different varieties of French around the world. Since this resource was already digitized, it only needed to be tagged.
- Wiktionnaire* is a part of *Wiktionary*, a multilingual collaborative dictionary, hosted by Wikimedia Foundation that also hosts *Wikipedia*, placed under a free license; only the part describing French is included in the DDF. This resource is not a traditional dictionary but its quality is more than sufficient to be included in the DDF since it is rich in neologisms and quotes from various sources, though it remains insufficient in description of usages, as already mentioned above. Since it is being constantly enriched, appropriately tagged updated copies will be uploaded regularly to the DDF.

The scientific committee of the project has the mission to gather a maximum of lexicographical resources describing different diatopic varieties of the French language. Other resources will be added to the DDF database in the near future. One should note that the source of all data gathered in the DDF database is always clearly stated. The content of these resources cannot be deleted or modified: the contributors can enrich and update the data only by adding new information.

3.3 Collaborative input

Following the long tradition of sociolinguistic inquiries, the description of linguistic variation and of the common lexicon based on a crowdsourced approach has proven to be highly pertinent. The relevance of asking speakers to identify and map linguistic uses has been present since the beginning of the 20th century through linguistic geography studies (Swiggers 1999; Lauwers et al. 2002; Leemann et al. 2016). With the development of the Internet, lexicography has been liberalized and democratized: a computer or a telephone connected to the Internet is enough to take part in the dictionary-making process. Bilingual or monolingual language dictionaries, but also encyclopaedias (such as *Wikipedia*) and linguistic and cultural resources in general, have thus become widely accessible to the public at large who can consult them, but also contribute in different ways. Data on lexical units selected by researchers are traditionally collected by means of closed-ended questions (for instance, “What do you call a pastry containing chocolate: *chocolatine* or *pain au chocolat*?”) or by means of images, usually with a choice of set replies (see the project *Français de nos régions*, Avanzi et al. 2016). Since collaborative resources offer open unstructured fields that allow speakers to share unexpected data on lexical units chosen by themselves, new or very specific words or uses as well as less known aspects of the language, such as the variation of norms, can be documented (Which words are acceptable in which regions? Which words are criticized? By whom?). Online collaborative lexicography has been developing substantially over the past twenty years and it now appears to be essential for lexicography in general. Within the framework of this publication we adopt the following definition of *collaborative lexicography*: activity which integrates the contributions of a community and creates through the Internet a virtual space in which the contributors participate, collaborate and support each other in writing of dictionary articles and the dictionary-making process (Dolar 2017a; Cotter & Damaso 2007; Meyer & Gurevych 2012; Granger & Paquot 2012). Collaborative lexicography forms a vast and diverse field of linguistic description: the technical methods of data collection, the types of data collected as well as their representation vary greatly from one collaborative resource to another. The collaborative resources available online today are constantly changing and range from blogs and forums (such as the *Babel Project*) to more structured resources focusing on spoken French (for instance *Blazz*, *Le Dictionnaire de la Zone*, *La Parlure* or *Urbandico*) and *Wiktionnaire*, whose form and content is close to professional dictionaries. The advantage of these resources is that they include data provided directly by speakers, but their major drawback is that they do not meet the official scientific standards, since they are not organized into a homogeneous structure that would make using the resource more efficient, and they are not user-centric or designed in accordance with

modern standards.

While most of the existing collaborative resources are based on a single mode of data collection, the objective of the DDF is to offer the possibility of contributing in different ways and thus to obtain a maximum amount of relevant information, gathering content, semantic relations, examples, pronunciations (in the future maybe also audio files)² and comments. The methods and modes of contribution to the DDF are based on best practices of existing collaborative dictionaries (see Dolar 2017b and Steffens 2017). As already pointed out, in the DDF, data can only be added (additive type of contribution), but not modified or merged (aggregative type). The DDF allows several modes of contribution; one can

- add a new lexical entry via the provided contribution form (see figure 1 below), including its written form, place of use, definition and example as mandatory information (other non-mandatory information can be provided in the appropriate fields),
- add a new definition to a form already present in the database, thus creating a new lexical entry,
- add an example or other types of information to existing lexical entries – such as grammatical categories, usage labels and semantic relations.

The purpose of the contribution form with strictly defined lexicographic fields is to obtain structured lexicographic information which is directly integrated into the dictionary structure.

Figure 1: Contribution form (structured data)

Discussion on forums and validation of content are also included. These two further forms of contribution are designed as follows:

- As the DDF pays special attention to discussions and negotiations that may arise around certain formal or historical aspects and norm-related topics, it offers the possibility of posting comments about written form and pronunciation, etymology and usage, including notes on linguistic policies and norms, etc. The comments do not have a predefined form (open debate, non-structured data) and follow a forum-like flow, close to a chat or a conversation, which enables the users to share information that was not included in the contribution form. If this kind of information is already provided in the imported resources, it will also appear in the discussion forums. In this case, the first message of the conversation is the indication of source and it cannot be edited, but contributors can write other comments, replies to comments and vote for other users' contributions. A preview of the comment with the most votes is displayed on the main page of the lexical entry.
- The contributors are thus also encouraged to validate existing contributions and comments. Three options are available in the DDF: [√ Je valide] validates existing contributions and comments and brings them higher on the list of results, [! Je signale] draws attention to a problem. Inappropriate contributions that contravene French laws (racism, incitement to violence, etc.) can be deleted by the administrators ([- À supprimer] initiates the process of deletion). The possibility of votes involves the users in the process of sorting out the contributions. Information is then not deleted but ordered according to the number of votes, ideally prioritizing the most relevant one.

Since it is assumed that the contributors are beginners without any particular lexicographic skills, they are guided through the contribution process. Contributing is divided into micro-tasks and the DDF offers pedagogical and technical support during the process. For instance, short definitions accompany linguistic terms, didactical inserts explain each type of linguistic information and advice and tips are given dynamically. Other dynamic elements will be implemented, namely a human-machine dialogue, where questions will adapt according to the given answers. This tool will help contributors to participate without introducing any technical or scientific vocabulary.

Given the structure of the dictionary and data collected from the contributors, the DDF is actually a doubly hybrid resource, combining structured data and talk/forum discussions on the one hand, and professional resources and collaborative input on the other. To ensure quality and transparency, the source of each information always remains visible and clearly stated. Contributions in the DDF are submitted to a proofreading process carried out a posteriori by peers.

4 DDF – An inclusive dictionary

² The DDF aims to be a multimodal object and recorded pronunciations will be added in collaboration with the Lingua Libre project (<https://lingualibre.org>). The possibility of adding other pronunciations (geographically tagged) is being implemented.

The main interest of the DDF is its inclusiveness both in terms of nomenclature, since all varieties of French can be present and adequately represented without restriction, and in terms of accessibility to all users, including learners. Despite the growing trend among educational resource developers to integrate online tools, the great potential of collaborative lexicography and crowdsourcing initiatives is insufficiently exploited in teaching programs and methods (Sabou et al. 2012; Steffens 2016). Among other users of the DDF, French language learners are able to discover the specificities of regional varieties of French in order to achieve a better understanding of the interactions between varieties, for instance by identifying the lexical causes of communication failures (the use of regional lexicons).

The collaborative approach, based on voluntary contributors acting as witnesses of linguistic usage and feeding the database in a dynamic and continuous manner, seems to be an effective mean of covering all varieties of the language, including those which are not represented in the corpora of traditional lexicographers. However, the project faces some major challenges: ensuring the scientific quality of the collected data by accompanying the contributors through the process, making it accessible to everyone (see 3.3), but also designing a visualization mode of data that allows the variation to be represented in its integrity, without eclipsing or isolating it.

The visual representation of lexicographic data plays a key role in the accessibility and comprehension for the public at large. To optimize the readability of the data, six criteria were used: 1) intuitiveness of the interface, 2) clarity and comprehensibility of the information, 3) univocity of colour codes used in the interface, 4) simple and easy-to-access functions, 5) accuracy of the data, 6) representation of the required types of data. Furthermore, to cope with the “difficulty of combining portability and small device size with a comfortably large display” (Lew 2010: 299), the DDF interface was developed primarily for mobile phones.

4.1 Geography and diatopic variation

The geographic subsets of data are clearly delimited by a colour code (see figure 2) in order to highlight regional lexical networks and to avoid confusion (Vincent 2011 and 2016). This systematic way of presenting data helps to give an accurate picture of the distribution of the described units (Are they common to speakers of all varieties of French, everywhere in the world? Is their use limited to France alone or another particular French-speaking region?). First and most visible is thus the diatopic variation: data display follows geographical location – data is specified at city level and is displayed at region/country level.

The order of display of different meanings in the result list is based on several criteria. The most salient one is the geographical adequation between the reader and the data. For example, if the user is located in Dakar, the search results that are tagged as geographically nearest will appear higher in the result list. The user can specify a city and the interface infers in which region and country it is situated since linked data give access to Geonames (<http://www.geonames.org/>), a database of locations defined with relations to each other. This is set by default but it is also possible to personalize the search criteria by indicating a preferred semantic domain, and more options of personalization are planned for the next versions of the DDF. The order of definitions is further based on votes and on semantic relations between definitions. The latter is still being implemented, but ideally in the case of a definition Y with a specified relation to X (*Y by hyponymy of X*), Y would be displayed after X.

The screenshot shows the search results for the term "savoir". At the top, it says "Terme recherché : savoir" and "[Définition]". Below this, there are four colored boxes representing different regions/countries:

- Belgique** (pink box): verbe. Pouvoir — **Note** : Dans le sens « avoir la capacité de ». Il peut être à l'indicatif, dans une phrase positive. Source :
- Monde** (grey box): verbe. Avoir dans la mémoire. Source :
- Rwanda** (yellow box): verbe, transitif. Pouvoir, être capable de. Source :
- Sénégal, Côte d'Ivoire, Burki...** (orange box): verbe, transitif. Connaitre. Source :

On the right side, there is a section titled "Discussion sur l'étymologie" (orange box) with the following text: "Du latin populaire *sapere*], en latin classique *sapere*, « avoir de la sagesse », avec influence de *sapiens* « sage », d'où « être perspicace », « comprendre », puis « savoir », et élimination du classique *scire* « savoir ». Très ancien français : *savoir* (Serments de Strasbourg), puis *savoir*, et enfin *savoir*. Pendant très longtemps, du moyen français jusqu'au XVIIIe siècle, le mot s'écrivait *savoir* par fausse régression au latin classique *scire* (« savoir »). Il fallut attendre 1740 pour que l'Académie française enregistrât, en la troisième édition de son dictionnaire, le mot sous sa graphie actuelle."

Figure 2: Search results highlighting the diatopic variation

4.2 Usage labels

As the main objective of the DDF is to represent and document all varieties of French including their different registers and uses in specific social contexts, the project is based on an inclusive and descriptive perspective, far from any

prescriptive goals. In order to describe precisely the conditions of use of lexical items, sociolinguistic labels were implemented in the DDF. The aim of this labelling is to document, rather than to legitimize, particular uses: the labels are based on facts directly observable by the contributors (Who says that? In what context?).

Each meaning of a given form can thus be linked to different diastatic and diatopic labels. The inventory of labels aims to reflect and include different lexicographical traditions (Hausmann 1989). The main and most common lexicographic labels, integrated to controlled vocabularies, are present, but some minor editorial choices have been applied (for example, replacing *populaire*, pejorative and outdated, by *très familier*). The labels also aim to be user-friendly and accessible to contributors – during the contribution process they appear as a closed-ended list, accompanied by short definitions.

Both the inclusive approach and descriptive labelling are essential for teaching French, in particular to non-French-speaking learners, for whom it is necessary to provide information on actual usage in various French-speaking varieties by giving them information on the context in which an expression could – or should – be used. The goal is not to impose a certain use but rather to reflect the diversity and thus to allow both Francophones and learners of French as a foreign or second language to interact with a wide range of examples and uses and to adapt their linguistic practice to the given circumstances.

5 Conclusion

The social impact of having a collaborative resource integrating different varieties of French from a holistic perspective can be seen at different levels. A complete description of different varieties of the language, of their sociolinguistic relations, of the objective norm (how one really speaks) and of the endogenous local norms (what is considered acceptable in a given linguistic community) has many benefits, not only because it averts situations of deep linguistic insecurity, but also because it preserves the Francophone linguistic heritage worldwide by describing and promoting its diversity. From a social perspective, the free, online, collaborative and dynamic DDF creates a space for people, from the very young to the elderly, to share their lexical usages and their linguistic and cultural knowledge and competences.

The DDF is a unique project, presenting many innovations in comparison to existing traditional and collaborative resources. The modular platform of the DDF offers a new model for accompanied collaborative lexicography, seeking to exploit all potentials of the Internet by making a wide range of linguistic but also cultural data accessible to everyone. The platform gives access to an up-to-date, constantly renewed image of French varieties spoken in different parts of the world. In the present paper we outlined three main innovative features of DDF: the institutional support and the scientific background in which the project is grounded (plurality of linguistic norms and equal treatment of linguistic varieties of French); the hybrid nature of the dictionary, which combines imported resources in a relational database, enriched by a complex speaker-based collaborative input; inclusivity of linguistic variation and the modes of its representation.

Other innovative aspects of the DDF should also be mentioned. The DDF has great potential as a teaching and learning resource. Since there is a real need of useful tools for teaching communication in various French-speaking regions (Steffens & Baiwir 2020), games, mobile applications and other didactical materials are in preparation in collaboration with several francophone organizations. One should also point out the user experience in regard to both, reading (ergonomics of the pages, numerous displaying options) and the contribution process (technical support and pedagogical tools). The DDF will also provide some basic sociolinguistic information about the contributors (via the log used for the proofreading process) and readers (via the metric tool Matomo). This type of data will not be fully publicly accessible for GDPR compliance, but available for the scientific community.

The position of this new object in relation to differential lexicography, online dictionaries and collaborative resources cannot yet be fully described and defined. However, due to the scientific objectives of this institutional-collaborative project (creating a general dictionary, integrating variational aspect, plurality of linguistic norms and the common lexicon), its accessibility and features, the structure of the database and its complex input, it is safe to suggest that the DDF offers a new paradigm in francophone lexicography.

6 References

- ALW = *Atlas linguistique de la Wallonie* (1953-). Liège: Presses universitaires de Liège.
- Atkins, S. & Rundell, M. (2008). *The Oxford guide to practical lexicography*. Oxford: Oxford University Press.
- Avanzi, M., Barbet, C., Glikman, J. & Peuvergne, J. (2016). Présentation d'une enquête pour l'étude des régionalismes du français. In *Actes du 5^e Congrès mondial de linguistique française (Tours, 4-8 July 2016), SHS Web of Conferences*, 27. Accessed at https://www.shs-conferences.org/articles/shsconf/pdf/2016/05/shsconf_cmlf2016_03001.pdf [30/05/2020]
- Baronian, L. & Martineau, F. (2009). *Le français d'un continent à l'autre. Mélanges offerts à Yves Charles Morin*. Québec: Presses de l'Université de Laval.
- Bavoux, C. (2008). *Le français des dictionnaires. L'autre versant de la lexicographie française*. Bruxelles: De Boeck Supérieur.
- BDLP = *Base de données lexicographiques panfrancophone*. Accessed at <http://www.bdlp.org> [30/05/2020]
- Bernet, C. & Rézeau, P. (2010). *Dictionnaire des expressions quotidiennes - On va le dire comme ça*. Paris: Poche. Blazz. Accessed at <http://www.blazz.fr/> [30/05/2020]
- Chambon, J.-P. (2006). Lexicographie et philologie: réflexions sur les glossaires d'éditions de textes (français médiéval et préclassique, ancien occitan). In *Revue de linguistique romane*, 70, pp. 123-141.
- Cotter, C. & Damaso, J. (2007). Online Dictionaries as Emergent Archives of Contemporary Usage and Collaborative Codification. In *QMOPAL - Queen Mary's Occasional Papers Advancing Linguistics*. Accessed at <http://linguistics.sllf.qmul.ac.uk/linguistics/media/sllf-migration/departement-of-linguistics/09-QMOPAL-Cotter-Damas>

- [o.pdf](#) [30/05/2020]
- Courbon, B. (2012). Représenter la diversité linguistique dans un dictionnaire monolingue: de la “traduction interne” à l’intégration sémantique. In M. Heinz (ed.) *Dictionnaires et traduction*. Berlin: Frank und Timme, pp. 153-196.
- Dictionnaire de la Zone*. Accessed at <http://www.dictionnairede lazone.fr/> [30/05/2020]
- Dolar, K. (2017a). Les dictionnaires collaboratifs en tant qu’objets discursifs, linguistiques et sociaux. PhD thesis. Université Paris Nanterre, Paris, France.
- Dolar, K. (2017b). Les dictionnaires collaboratifs non institutionnels dans l’espace francophone: éléments de typologie et bilan. In *Repères – Dorif*, 14. Accessed at http://www.dorif.it/ezine/ezine_articles.php?art_id=380 [30/05/2020]
- Dulong, G. & Bergeron, G. (1980). *Parler populaires du Québec et de ses régions voisines: Atlas linguistique de l’Est du Canada*. Québec: OLF.
- Francard, M. (2011). L’intégration des régionalismes dans les dictionnaires de référence du français. Le cas des belgicisms. In *Français et Société*, special issue, pp. 13-25.
- Francard, M., Geron, G., Wilmet, R. & Wirth, A. (2015). *Dictionnaire des belgicisms*. Bruxelles: De Boeck.
- Gadet, F. (2003). La variation: le français dans l’espace social, régional et international. In M. Yaguello (ed.) *Le grand livre de la langue française*. Paris: Le Seuil, pp. 91-152.
- Glessgen, M.-D. & Thibault, A. (2005) (eds). *La lexicographie différentielle du français et le Dictionnaire des régionalismes de France*. Actes du Colloque en l’honneur de Pierre Rézeau, Strasbourg: Presses Universitaires de Strasbourg, pp. III-XVII.
- Granger, S. & Paquot, M. (2012) (eds.). *Electronic Lexicography*. Oxford: Oxford University Press.
- Greub, Y. (2002). Les régionalismes lexicaux du moyen français et la formation des français régionaux, d’après l’exemple d’un corpus de farces (1450-1550). PhD thesis. Université de Neuchâtel, Neuchâtel, Switzerland.
- Guérin, É. (2008). Le “français standard”: une variété située? In *Actes du Congrès Mondial de Linguistique Française - CMLF’08*. Paris: Institut de Linguistique Française, pp. 2303-2312. Accessed at <https://www.linguistiquefrancaise.org/articles/cmlf/pdf/2008/01/cmlf08250.pdf> [30/05/2020]
- Hausmann, F. J. (1989). Die Markierung im allgemeinen einsprachigen Wörterbuch: eine Übersicht. In F. J. Hausmann *et al.* (eds.) *Wörterbücher/Dictionaries/Dictionnaires. Ein internationales Handbuch zur Lexikographie/An International Encyclopedia of Lexicography/Encyclopédie internationale de lexicographie*, I. Berlin/New York: Walter de Gruyter, pp. 649-657.
- Heiden, S. (2004). Interface hypertextuelle à un espace de cooccurrences: implémentation dans Weblex. In *JADT 2004: 7es Journées internationales d’Analyse statistique des Données Textuelles*, pp. 577-588.
- IFA = Équipe IFA (1983). *Inventaire des particularités lexicales du français en Afrique noire*, AUPELF. Paris: Edicef.
- La Parlure*. Accessed at <http://www.laparlure.com> [30/05/2020]
- Lauwers, P., Simoni-Aurembou, M.-R. & Swiggers, P. (2002). *Géographie linguistique et biologie du langage: autour de Jules Gilliéron*. Leuven: Peeters.
- Leemann, A., Kolly, M.-J., Purves, R., Britain, D. & Glaser, E. (2016), Crowdsourcing Language Change with Smartphone Applications. In *PLoS One*, 11(1). Accessed at <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0143060> [30/05/2020]
- Lüdi, G. (2012). French as a pluricentric language. In M. Clyne (ed.) *Pluricentric Languages. Differing Norms in Different Nation*. Berlin: De Gruyter, pp. 149-178.
- Mazziotta, N. (2016). Représenter la connaissance en linguistique. Observations sur l’édition de matériaux et sur l’analyse syntaxique. Habilitation à diriger des recherches (Habilitation of Conducting Research). Université Paris Nanterre, Paris, France.
- McCrae, J., Bosque-Gil, J., Garcia, J., Buitelaar, P. & Cimiano, P. (2017). The OntoLex-Lemon Model: development and applications. In *Proceedings of eLex 2017*, pp. 587-597. Accessed at <https://elex.link/elex2017/wp-content/uploads/2017/09/paper36.pdf> [30/05/2020]
- Měchura, M. (2016). Data Structures in Lexicography: from Trees to Graphs. In A. Horák, P. Rychlý & A. Rambousek (eds.) *Proceedings of Recent Advances in Slavonic Natural Language Processing*, pp. 97–104. Accessed at: <http://www.lexiconista.com/raslan2016.pdf> [30/05/2020]
- Mercier, L. & Verreault, C. (2002). Opposer français “standard” et français québécois pour mieux se comprendre entre francophones? Le cas du *Dictionnaire québécois français*. In *Le Français moderne*, 70(1), pp. 87-108.
- Meyer, C. & Gurevych, I. (2012). *Wiktionary: A new rival for expert-built lexicons? Exploring the possibilities of collaborative lexicography*. In S. Granger & M. Paquot (eds.) *Electronic Lexicography*, Oxford: Oxford University Press, pp. 259-291.
- Poisson, E. (2002). Français en usage au Québec et dictionnaires. In C. Verreault, L. Mercier & T. Lavoie (eds.) *Le français, une langue à apprivoiser: textes des conférences prononcées au Musée de la civilisation (Québec, 2000-2001) dans le cadre de l’exposition Une grande langue: le français dans tous ses états*. Québec: Presses de l’Université Laval, pp. 93-111.
- Projet Babel*. Accessed at <http://projetbabel.org/index.php> [30/05/2020]
- Renders, P. (2015). L’informatisation du Französisches Etymologisches Wörterbuch. Modélisation d’un discours étymologique. Strasbourg: ELiPhi-Éditions de linguistique et de philologie.
- Rézeau, P. (2001). *Dictionnaire des régionalismes de France*. Bruxelles: De Boeck/Duculot.
- Robillard, D. de (2008). Revendiquer une lexicographie francophone altéritaire constructiviste pour ne plus saler avec du sucre. In C. Bavoux (ed.) *Le français des dictionnaires*. Bruxelles: De Boeck Université, pp. 321-335.
- Sabou, M., Bontcheva, K. & Scharl, A. (2012). Crowdsourcing research opportunities: lessons from natural language processing. In *i-KNOW '12 Proceedings of the 12th International Conference on Knowledge Management and*

Knowledge Technologies.

- Steffens, M. (2017). Lexicographie collaborative, variation et norme: le projet 10-nous. In *Repères – Dorif*, 14. Accessed at https://www.dorif.it/ezone/ezone_articles.php?art_id=393 [30/05/2020]
- Steffens, M. (2018). Figement, langage et didactique du français. In *Langues et linguistique*, 37, pp. 120-134. Accessed at <http://www.lli.ulaval.ca/recherche/revues/revue-langues-et-linguistique/index-des-numeros-et-articles/vol-37-2018/> [30/05/2020]
- Steffens, M. & Baiwir, E. (in print). Intégrer la variation diatopique à l'enseignement du français: le rôle des outils numériques, *ÉLA*.
- Steffens, M., Dolar, K., & Gasparini, N. (2020). Structuration de données pour un dictionnaire collaboratif hybride. In: *Terminologie & Ontologie: Théories et Applications. Actes de la conférence TOTh 2019*, pp. 413-426.
- Steinlin, J., Kahane, S., Polguère, A. & El Ghali, A. (2004). De l'article lexicographique à la modélisation objet du dictionnaire et des liens lexicaux. In *Actes de EURALEX'2004*, pp. 177-186.
- Swiggers, P. (1999). La géographie linguistique de Jules Gilliéron: aux racines du changement linguistique. In *Cahiers Ferdinand de Saussure*, 51, pp. 113-132.
- Thibault, A. (2004). Dictionnaire suisse romand: Particularités lexicales du français contemporain. Geneva: Éditions ZOÉ.
- Thibault, A. (2008) (ed.). Richesses du français et géographie linguistique: Recherches lexicographiques sur les variétés du français en France et hors de France (Tome 2). Bruxelles: Duculot.
- Urbandico*. Accessed at <http://www.urbandico.com> [30/05/2020]
- Urban Dictionary*. Accessed at <https://www.urbandictionary.com> [30/05/2020]
- Usito*. Accessed at <https://usito.usherbrooke.ca/> [30/05/2020]
- Vincent, N. (2011). Combien faut-il de dictionnaires pour décrire le français? In O. Bertrand & I. Schaffner (eds.) *Variétés, variations et formes du français*. Palaiseau: Éditions de l'École polytechnique, pp. 389-404.
- Vincent, N. (2016). La prise en compte de plusieurs variétés nationales dans un dictionnaire du français: exercice de lexicographie pratique. In C. Molinari & D. Gavinelli (eds.) *Espaces réels et imaginaires au Québec et en Acadie: enjeux culturels, linguistiques et géographiques*. Milano: Led.
- Vincent, N. (2017). Présence et légitimité des variétés nationales dans les dictionnaires gratuits en ligne. In *Repères – Dorif*, 14. Accessed at https://www.dorif.it/ezone/ezone_articles.php?art_id=379 [30/05/2020]
- Violette, I. (2006). Pour une problématique de la francophonie et de l'espace francophone: réflexions sur une réalité construite à travers ses contradictions. In *Francophonies d'Amérique*, 21, pp. 13-30.
- Völker, H. (2009). La linguistique variationnelle et la perspective intralinguistique. In *Revue de Linguistique Romane*, 73, pp. 27-76.
- W3C (2004). Accessed at <http://www.w3.org/2004/OWL> [30/05/2020]
- Wiktionnaire*. Accessed at https://fr.wiktionary.org/wiki/Wiktionnaire:Page_d%E2%80%99accueil [30/05/2020]