Diachronic-information visualization in historical dictionaries

Roberto Therón*
Department of Computer Science and Automation. University of Salamanca. Spain

Laura Fontanillo
Department of Computer Science and Automation. University of Salamanca. Spain

Abstract
The field of computational linguistics has been dealing with the modeling of natural language from a computational perspective since the 1950s. However, the usage of advanced and interactive visualization techniques is very limited. This is especially the case of diachronic linguistics, which is devoted to the study of language change. This work is part of a project that aims to provide novel highly interactive visual solutions to ease the task of lexicographers, ensuring a rigorous treatment of the vocabulary beyond the use of traditional lexicographic sources, and overcoming certain limitations of corpus linguistics.

The paper focuses on the choices made for the design and development of an interactive visual tool that supports different tasks related to the processes of drawing up and consulting historical dictionaries. Particularly, we describe solutions for the exploitation, analysis, and expert-directed validation of the data compiled in the available dictionaries in a manner that is both automatic (provided by computational methods) and intelligent (provided by the experts). We thus describe diachronlex diagrams, an interactive visual solution that facilitates linguistic work related to the understanding of the temporal evolution and the lexical relationships between the different meanings registered in subsequent editions of a dictionary.

Keywords
Diachronic information visualization, diachronlex diagrams, visual analytics, lexicography, timelines, language change.

Introduction
Diachronic linguistics, also called historical linguistics, is devoted to the study of language change; it seeks to investigate and describe the way in which languages change or maintain their structure over time [1].

To the best of our knowledge, our study is the first to initiate an innovative research line in diachronic linguistics; to be more precise, it aims to facilitate an understanding of how the dictionary entry for a given word changes over time by means of the development of interactive visual tools that allow the analysis of temporal phenomena related to the evolution of a lexicon. Our contribution is centered on the Spanish language, but the proposed visual solutions are general and can be used with any language, as long as the necessary historical corpus and dictionaries exist.

Apart from the invaluable effort of lexicographers and experts in related fields, the intervention of data processing technologies will be determinant in the field of diachronic linguistics. Until now, data processing has traditionally meant progress with regard to two subdisciplines of computational linguistics: corpus linguistics and natural language processing (NLP). This progress has been considerable; nevertheless, the work of many traditional linguists would be greatly facilitated with
the aid of truly interactive and visual tools tailored for the different kinds of analysis they conduct. One mundane example of such situation is how the vast majority of dictionaries on the Internet are still presented as lists of meanings.

Our fundamental contribution is therefore the possibility of surpassing the current state of the dictionaries or corpora as mere digital archives and of fostering research in the development of visual and interactive techniques that allow an intelligent approach to the evolution of a dictionary, initially for expert linguists but also for the general public. This involves exploiting both the processing capabilities of computers and the unique cognitive capabilities of humans.

In our case, we wished to contribute to the advancement of the study of the Spanish language. Despite being one of the world languages with the largest number of native speakers, and the fact that The Royal Spanish Academy (Real Academia Española, RAE) was founded in 1713 to preserve it (and publishes a landmark Spanish dictionary for that purpose), a historical dictionary for the Spanish language has not yet been achieved. However, a great effort has been made and the volume and variety of resources already available is such (linguistic corpora, electronic versions of dictionaries, scanned dictionaries and lexicographic material, etc.), that they pose several challenges of analysis. Perhaps if the goal is to make a historical dictionary the first one is to provide lexicographers with a tool that helps them to understand how the meanings of words have changed in the different editions of the dictionary. Furthermore, we wanted to provide a valuable tool for the drawing up of a historical dictionary as well as its subsequent consultation.

We therefore present an interactive visualization tool to explore the evolution of the meaning of words down the centuries. It features a novel kind of diagram (diachronlex) that highlights the temporal patterns of meaning evolution through the different editions of normative dictionaries, as well as having the capacity to highlight historical mistakes and anomalies (e.g., one meaning for a word that disappears in one edition, probably due to a misplacement, and is recovered in the next one). While our tool was produced to explore meanings of the Spanish historical dictionary, the visual encoding introduced here has the potential to be applied to any other language.

We discuss details of these design refinements, together with descriptions of how our visualization tool can be used for different diachronic linguistics and lexicographic tasks.

**Background**

Languages vary from one place to another, from one social group to another, and from one situation to another (synchronic perspective), and obviously over time (diachronic perspective). The question "what is language?" is answered differently under the two Saussurean dimensions: the synchronic perspective (i.e., a synchronic object in the world (independent of history) and the diachronic perspective (i.e., what it means to say that a language changes over time) [2]. Regarding this dichotomy, we subscribe to the proposition that no synchronic fact about language can be successfully explained or fully understood without taking the historical perspective into account [3].

Our motivation is to develop tools that can cope with the huge amount of historical linguistic data while providing suitable means for experts to incorporate their experience and unique human abilities. Although during the process of drawing up the Spanish historical dictionary lexicographers are the main targeted users, we have also considered that other linguists may use the tools for linguistic research as well as occasional and curious dictionary users.
In this section we present the craft of producing dictionaries in the digital era and their ideal use, describe the particularly tough endeavor of producing a historical dictionary, discuss related diachronic linguistics tasks, data sources involved and their structures, and survey existing time-related visualization techniques.

**Electronic dictionaries, historical dictionaries and dictionary-making**

The beginning of the “electronic dictionary era” can be traced back to the late 1960s [4] with the complete transcription of Webster's Seventh New Collegiate Dictionary to punched tape with the objective of carrying out a computer exploration. Since then, the growth in the use of computers has transformed all aspects of dictionary-making, from collecting data on word meanings and uses to creating a set of dictionary entries, and displaying, using, preserving and distributing these entries and the data on which they are based [5]. Thus, in [4] the “Lexicographers' dream in the electronic-dictionary era” is presented by means of a list of 118 dreams.

Our proposal directly addresses some of these characteristics that an ideal dictionary should have. In the appendix, built upon the original 118 dreams, the main dreams of interactive dictionary visualization that this project would allow are listed. They can be summarized as follows: a) efficient automated processing of vast amounts of data that can change over time; b) animated and user-driven presentation of information; c) various means of selection, filtering and access to related and relevant information; d) adaptation to the role, expertise and tasks of the user; and e) they permit hypothesis-driven and exploratory analysis.

Computer science has come to triumph over what seemed to be endemic diseases of lexicography, such as controlling the integrity of the reference system and the lexicon employed in the definitions, which avoids, for example, circularity, reference to suppressed articles, etc. On the other hand it allows the promotion of coherence in works as complex as dictionaries, establishing uniform patterns in the apparatus of abbreviations and marks that are thus reduced to a controlled universe [6]. It turns out to be beneficial even to the organization of the work of editing, since it enables on the one hand the creation of definition models that can be applied to homogeneous groups of words or meanings, while at the same time making possible the specialization of lexicographical work, a need imposed by the multidisciplinary character that is a current trend in editing teams. The time needed for reviewing the dictionary is also considerably reduced by using data processing techniques that help the lexicographer in the proofreading process. Practically all current large dictionaries resort to huge quantities of materials of various origins for the verification of linguistic facts (this is the case of the TLFi (Trésor de la langue française informatisé) [7], for instance).

The drawing up of the Nuevo Diccionario Histórico de la Lengua Española (NDHE) of RAE, which was considered a state project, was expected\(^1\) to be completed in 2018; at the end of this period a draft of the dictionary with about 60,000 entries will be available. The second stage will start in 2019 and will last until 2029: it will produce a dictionary in which, unlike the version completed at the end of 2018, the number of words examined will be considerably increased, most of the pending problems solved, and the word definitions produced by using standardized patterns of definition. Such a project requires complex management to develop hundreds of different tasks,

---

\(^1\) At the time of writing, due to the Spanish financial crisis, the economic support from the government has been suspended, leaving the viability of the NHDE at risk and the timing very unlikely,
both lexicographic and computer-based, all inter-related and undertaken inside or outside the coordination center, so that there are no bottlenecks preventing their integration in the future [8].

The execution of the NDHE is a task of great importance insofar as it involves the application of the methods of linguistics and the techniques of philology to texts from the past. Both will facilitate not just achieving coherence in the field of definitions but also establishing the relationship between lexical units throughout history. This work aspires therefore to take advantage of the advances of linguistics in a way that was not possible before. Likewise, it will use data processing techniques both to relate the extensive mass of data that must be used as a starting point, and to permit collaboration among the many teams that should carry out this work.

In the first phase of the project, a series of tasks have been developed which on the one hand aim to obtain the documentary bases from where the lexical data will be extracted to build the dictionary, and on the other hand to establish the basis upon which those data are structured, giving rise to the NDHE. An important part of the work consisted of preparing materials on which the NDHE will be based; among them a very important source is the Corpus del Nuevo Diccionario Histórico del Español (CNDH), which contains 53 million entries deriving mainly from literary texts. However, this work is still in progress [8], and we rely on other corpora and lexical material that is publicly available from the RAE.

What is most relevant for the present work, when we started our collaboration the RAE was building the so called map of dictionaries in which the evolution of words and meanings in these dictionaries over a period of more than two centuries could be traced by comparing several editions of the Academia dictionary [8].

This paper thus describes the interactive visual tool for advanced analysis of the data collected for the map of dictionaries. Our tool was designed and developed after a thorough study of the prototype (described in the Diachronlex: Design decisions section) and the related user tasks, with the aim of supporting the RAE’s lexicographers and future public users.

**Changes in meaning: Tasks of diachronic linguistics**

We summarize here some tasks in diachronic linguistics, some of which are needed to support dictionary-making, and others for a general analysis of changes in meaning that can be visually supported.

It is now well understood that semantic change proceeds by way of an intermediate stage in which both old and new meanings coexist together simultaneously. Furthermore, it seems likely that the vast bulk of many lexemes’ diachronic semantic behavior consists of reconfigurations to the set of existing meanings that rather than of clear linear changes from an old meaning to a new one [9]. Nowadays the emphasis in the lexicographical world is often primarily on corpus building and text encoding. Therefore it is urgent to keep interest in practical lexicography alive, and even more so since part of our historical vocabulary has not yet been sufficiently described [10].

The interest in tracing changes in meaning in a systematic way by analyzing the definitions was expressed decades ago [11] and the difficulty of this task was recognized, as the definitions really consist of combinations of several definitions and synonyms. Nowadays, as in other text analysis problems, attempts of tracking changes in word meaning have resorted to statistical analysis [12]. Accordingly, Pascual [8], affirms for the map of dictionaries, that it will certainly not be possible to obtain a rigorous picture of the real language over a period of more than two centuries, but even if this picture is fuzzy it will provide some guidance as to the evolution of vocabulary over a period when the number of words and meanings has increased considerably, mainly because of the
development of specialized languages. Research on words is therefore not only carried out for the purpose of compiling dictionary entries; their lexical content itself can also be subject to search. Both approaches are often combined [10].

We have stated before that we are interested in two types of users, lexicographers and non-lexicographers. While the former group of users may analyze and change the data, the latter includes both regular dictionary users, that may only use the tool for simple consultation and linguists that may use the tool for another type of linguistic analysis that does not involve any changes in the stored data.

In the following analysis we focus on the more demanding needs of linguists (both lexicographers and non-lexicographers); we have selected four tasks that can be supported by our tool and that will be illustrated in the remainder of the paper. We have divided them into diachronic analysis tasks and dictionary-making (i.e., manipulating the data) tasks. The two diachronic analysis tasks, which can be performed by any type of analyst, are as follows:

- **Visualization of the evolution of meaning and the assessment of particular words.** Conveying the relationships of meanings through different editions of a dictionary is not an easy task since a word can have several tens of meanings in a single dictionary. Beyond the challenge of providing a clear, compact and scalable view, a linguist needs support for further analysis in order to gain insight into the whole evolution of the meanings for a given word. For instance, a considerable number of English words were first recorded in Scottish sources, but thus far this information has been insufficiently documented in dictionaries [10]. By considering an overview of the presence and evolution of words and meaning in a historical dictionary, and in combination with corpus enabled analysis (and possibly other lexicographic sources), an expert would provide more accurate explanations of the evolution of a particular word or group of words.

- **Revealing relationships in the organization of meanings.** Dictionaries in general usually organize the meanings of a lexical entry as a list. Disciplines related to lexicography, e.g. cognitive linguistics, are interested in a different kind of analysis of the meanings and their organization through different editions. Some authors claim that this type of organization not only makes it difficult for the reader to grasp the semantic nuances of these meanings but also hides the real motivated, structured and contextual relationships that exist between these meanings [13]. Having a visual representation of these changes in the listings of meanings can be a very powerful tool.

The two dictionary-making tasks, performed exclusively by lexicographers, are as follows:

- **Validation of computed evolution and eventual editing.** Although computational methods can compare meanings of words in different dictionaries, careful revision is needed for those cases in which although an automatic equivalence has been assigned the degree of difference is high [8]. The lexicographer therefore needs support for the easy annotation and manual modification of the erroneous association of meanings within a particular line of evolution.

- **Cleaning up dictionaries.** It is important to remove inappropriate material that has been accumulated in the course of successive revisions (where inherited definitions were sometimes interpreted more from the lexicographers’ personal perspective of a word than from its actual use) [8]. The tool must provide mechanisms for the easy discovery of such extraneous or otherwise relevant linguistic data or facts.
Visualization of evolution: Related work

Although we are not aware of any previous attempts to visualize the evolution of meanings registered in several editions of a dictionary, there are several works that have dealt with related problems, either facing challenges of linguistic visualization or devoted to the visualization of the evolution of different entities. Moreover many approaches to the interactive visual analysis of time-dependent data have already been introduced.

In [14] a review of computational models for the integration of visual and linguistic information can be found. This is the first article that categorizes the research works that have dealt with the correspondence problem, namely how to associate visual events with words and vice versa. For a review of the main attempts we refer the reader to [15].

In spite of the great achievements of computational linguistics, the introduction of interactive visual tools in linguistics is a very recent tendency and is not yet popular, particularly among professional linguists. At present there are few works dedicated to the interactive visualization of linguistic data; specialized tools for language analysis still make little use of visualizations and visualization tools for language related information (LInfoVis) [16].

The advances achieved are mainly related to document and text mining [17][18], with excellent results such as those available at Many eyes, a website devoted to the popularization of data visualization [19][20][21]. More directly related to the theme of the project, several research works can be mentioned here such as [22], in which software for the visual exploration of a dictionary of Warlpiri is presented, or [23] in which a tool for the presentation and exploration of grammatical trees is proposed, and even successful business examples such as VisualThesaurus and Visuwords in which interactive visual maps (node-link diagrams) are used to visualize word relationships for various languages.

In a related attempt to visualize lexical change, Kempken et al. described several treemap techniques used to visualize the productivity of rule sets in deriving nonstandard spellings in old German texts, among other aspects [24].

In [25] pixel-oriented visualizations were used to cope with a vast corpora from newspapers, looking at the distribution and lexical semantic usage of derivational morphemes across several languages and time.

The visualization of the evolution of different entities has been applied in different fields. One field that has attracted a great deal of study is the visualization of software evolution. An example is [26], in which different visualizations are used to convey how individual lines of code have been added, modified, and deleted as the source code files are changed over time.

Flow diagrams are also related to the problem under discussion. As an example of this, and also on the subject of textual data, Viegas et al. proposed History Flow diagrams [27], which use shaded connections between corresponding segments on adjacent revision lines in Wikipedia as a means of visually linking sections of text that have remained the same between consecutive versions. Although different in conception, the resulting representations in our proposal are visually reminiscent of history flow diagrams. Another approach for interactively exploring complex flow scenarios are Sankey diagrams [28]. More recently, in [29] Alluvial diagrams have been proposed; these diagrams highlight and summarize the structural changes between two given times. However,
this kind of diagrams does not deal with linguistic data and is more appropriate for groups of entities that evolve together.

Another group of related works are devoted to the visualization of the evolution of data over time. Aigner et al. [30] provide a categorization of visual methods for the analysis of time-oriented data. Even though many approaches to this task have been published in recent years, most of them are specific to only a particular analysis problem; it is extremely difficult to consider all aspects involved when visualizing time-oriented data.

Finally, although with quite specific functionality, our proposal could be seen as a type of Parallel Coordinates Plots [31], in the sense that a set of dictionaries can be seen as parallel axes where the meanings move up and down according to their position in each dictionary. Variations and extensions of PCP have been previously used for understanding evolution by means of linked-views and animated PCPs [32], linguistic corpus analysis (with each polyline conveying n-gram and its associated data [33]), or for providing valuable overviews of a document collection and visualizing changes over time [34][35].

However, due to the complexity of the structure of historical dictionaries from the perspective of meaning level, the existing techniques are limited to showing only certain aspects of a historical dictionary without providing users with a comprehensive understanding of the evolution of meaning.

**Diachronlex: Design decisions**

Lexicographers require several types of information in order to make assessments on the historical evolution of the meanings of a word. The main task can be summarized as the process of organizing the different synchronic sequences of the Dictionary to explain how the current situation was reached, and comparing the different representations of meanings through the history. RAE's map of dictionaries was built to provide a means of analyzing such evolution and requires the gathering, processing and representation of several types of information including: (1) the order of meanings in the different editions of the RAE's normative dictionary, (2) the actual linguistic content of each meaning, (3) relationships between meanings that can vary over the different editions of the dictionary (splitting, merging, expanding, etc.).

In order to provide a way of visualizing this trace of evolution, the RAE had produced a prototype (see Figure 1) that used a grid structure in which each row pertained to a line of the meaning of evolution (with colors that convey the similarity of meaning). While the possibility of studying the appearance and disappearance of words or meanings and how a secondary meaning becomes a main one or vice versa was extremely useful, the prototype also exposed certain issues.

So as to gain a deeper understanding of the enormous quantities of data stored in the different corpora and dictionaries of the RAE we started a collaboration aimed at providing interactive visual tools. Methodologies originating in fields such as information visualization, statistics, artificial intelligence, data mining, computational linguistics, graphic design, the psychology of perception, and human-computer interaction can be used to solve parts of the specific problem. They are nevertheless insufficient by themselves (a good algorithm for the computation of the similarity of meanings can be designed, but it is not very useful without a good representation of the relationships computed), and problems without an appropriate solution in the visualization of the map of dictionaries therefore still remain. This is the reason why the design of our proposal, the
diachronlex diagram, is based in the combination of the previous fields with a common effort based on the holistic approach of computational information design (CID) [36].

Figure 1. Tool for the visualization of the map of dictionaries available as a public resource for the RAE’s NDHE. Several problems arose even in the simple case of the word *gabinete* (‘cabinet’), with just 13 meanings in the actual edition, this approach was only able to show 2 of them without the use of the vertical scroll.

After several meetings with the people working in the making of the NDHE, including both lexicographers and computer scientists, we examined the workflows and the tools they utilized. For the problem in hand, the design of an interactive tool for the exploratory analysis, the editing and validation of the map of dictionaries, the starting point was the first prototype they were using (at the moment of writing, the prototype has been publicly released as a Web based tool[^2], see Figure 1). The main challenges became obvious at an early stage: (1) exploiting interactivity, (2) avoiding loss of context (even for a simple example, the amount of information conveyed is too large and would force the user to scroll up and down, and lose sight of part of the information), (3) providing connections to other available RAE data sources when possible, (4) avoiding the arbitrary use of color encoding (cognitive memory reinforcement is minimum and the retention of the majority of the information shown does not last), (5) patterns and anomalies should be easily perceived, and (6) the tool should be usable not only by specialists (that is, providing intuitive representations). Thus, instead of just redesigning the prototype we decided to start everything from scratch.

In the following sections we expound the design decisions we took.

**Data sources**

RAE had already undertaken the digitalization, pre-processing and annotation of the dictionaries, storing all the data in its databases. In order to perform this task, they had selected some

[^2]: [http://www.frl.es/Paginas/Mapadicionarios.aspx](http://www.frl.es/Paginas/Mapadicionarios.aspx)
representative dictionaries: the 1780 edition (which is essential since it was the first abridged edition of the Academia dictionary), the 1817 edition (in which many definitions were improved and much attention was paid to correcting scientific and technical terms), the 1884 edition (easier to digitize than the initially chosen 1843 edition), the 1925 edition (in which many provincialisms and Americanisms were introduced), the 1992 edition (very convenient, since it was already in digital form) and the 2001 edition (no special justification was required for this choice) [8]. For its preparation it has therefore been necessary to digitize four of the six dictionaries listed, which meant recording 832,000 meanings and 139,000 graphic variants, corresponding to more than 100,000 lemmas.

RAE provided us a subset of words (corresponding to the letter ‘g’) that was used to populate our database. On top of that we have carried out some of these tasks in order to validate the adequacy and scalability of our design choices, by manually including in our database examples of the meanings of words meanings from up to 27 RAE normative dictionaries extracted from the digitized images of dictionary pages available in the Nuevo tesoro lexicográfico de la lengua española (NTLLE)\(^5\). Our interactive tool performs database queries to retrieve all the necessary information to be represented.

**Computing similarity among meanings**

For the analysis of the evolution, having a set of ordered lists of meanings pertaining to each edition of the dictionary is not sufficient. It is necessary to process the actual linguistic content of those meanings and to compute the similarity among different texts. The output of this process is therefore the set of relationships among meanings in the different editions. Note that between two editions meanings may disappear, appear, split, merge and so forth. The multidisciplinary team of the RAE has developed a semi-automated algorithm for the computation of these relationships [8] and we relied on it during the first stage. The particular details of such algorithm were not disclosed to us and are beyond the scope of this paper.

Visual encoding is actually independent of the algorithm used to compute the similarity among meanings. It is obvious that a fully automated method would not be 100% accurate, as the similarity of meanings is subject to very complex issues such as disambiguation and validation from lexicographers will always be needed. We were more interested in providing a way of understanding evolution through the visualization of the algorithms' output, by means of visual hints so that the experts could recognize when and where a manual correction of a relationship between meanings was needed, and making that task easy through interaction.

In order to prove this independence from the particular computational linguistics algorithm used, and also in an attempt to provide the experts with a different view on the evolution of meanings, we also included an implementation of the NIST [37] and BLEU [38] metrics, which were originally intended to compare machine translation output with expert reference translations in terms of the statistics of short sequences of words (word n-grams).

The computation of the similarity of meanings can be summarized as follows: 1) meanings are mapped to arrays with meaning number 1 as the first element of the array; 2) the most recent edition is used as a reference dictionary (typically, the 2001 edition); 3) Using the reference dictionary, for each dictionary in reverse chronological order, and for each meaning in the dictionary, the distance to all meanings in the reference dictionary is computed (using the NIST or

\(^5\) http://buscon.rae.es/ntlle/SrvltGUILoginNtle
BLEU metric); 4) Each meaning is associated to the least distant (with highest distance score) meaning in the reference dictionary; 5) If the score is lower than a threshold (we use 0.001 by default) the distance computation is repeated using the next most recent dictionary as a reference; 6) This process is repeated until the meaning can be associated to a meaning in a previous dictionary (score > threshold) or considered to be unrelated (score < threshold).

The NIST and BLEU algorithms provide slightly different results (with NIST providing more accurate associations because a list of non-informative n-grams is used when calculating how informative a particular n-gram is) and can also be compared to the RAE's algorithm; this that can enrich the lexicographic analysis.

Diachronlex diagrams

Following the approach of Aigner et al. [30], before deciding to use an existing technique for the visualization of diachronic lexical information or developing a new one, we considered these three questions: 1) What are the characteristics of the time axis? 2) What is analyzed? 3) How is it represented?

The first question is related to the temporal primitives. In our case the time axis is made up of time points rather than time intervals (with duration). Aigner et al. [30] suggest that data values are only valid at certain points in time (dictionary editions in our case) to avoid misinterpretations. The structure of time is linear (as opposed to cyclic or branching time).

The second question corresponds to the data. The same authors distinguish two possible frames of reference: abstract vs. spatial data. The linguistic nature of our dataset clearly falls into the first category, meaning that a priori no spatial mapping is given. A second criterion to take into account is the number of time-dependent variables, which in the case of the map of dictionaries is just one (meanings), i.e., the goal is to represent univariate data. However the representation of the full text of all meanings may lead to overcrowded and cluttered displays. In this case, deriving data abstractions that reflect the interests and needs of users (the ordinal number for the meaning is a perfect candidate in the map of dictionaries) and driving overview+detail interfaces is suggested.

Finally, the third question addresses visual representation itself. One criterion for answering this question is time dependency, in which the representations are still images of change to convey time dependency. The problem in question is an instance of the first case, as we want the lexicographers to concentrate fully on the data and to compare different parts of the time axis. The final criterion is whether to use 2D or 3D presentation spaces. We opted for the 2D space, as we wanted to offer the more familiar type of representation to our target users.

Given this design study, we felt that a time diagram was a very natural encoding and it also has the advantage of being a classic method for visualizing temporal change. Furthermore, for simple data and basic analysis tasks, these approaches outperform specialized techniques, because they are easy to learn and understand [30]. However, the classical time diagram does not fit our problem exactly, since starting from the most recent edition of a dictionary a meaning may have branching relating to none or several meanings of any of the previous editions. We had to come up with a novel time diagram, specially tailored for the evolution of meanings in historical dictionaries, that we have called diachronlex diagrams.

Following the well known visual information-seeking mantra of Shneiderman [39], and its most recent update for visual analytics [40], we built an interactive interface that by means of automated processes analyses the data and provides an initial overview of the data set, while enabling
subsequent focusing on regions of interest and access to details on demand, which in turn enables further analysis.

Figure 2. Scheme of a basic diachronlex diagram. The first meaning of the most recent (leftmost) edition of the dictionary is related to the first meaning of the remaining editions; it is also however related to the second meaning in dictionaries 3 and 5. In both cases those meanings were merged into the first meaning of the next edition. However, in the case of the second meaning of Dictionary 3 it evolved from a split of the previous edition, while in the case of the second meaning of Dictionary 5 it appeared in this very edition. The case of the third meaning of Dictionary 1 is also interesting: it is related to Dictionary 4 although it did not appear in dictionaries 3 and 2, i.e. that meaning was missing in two editions for unknown reasons (a dashed segment connecting the two meanings is used to convey the gap).

Figure 2 shows a simple example of the idea behind diachronlex diagrams. This kind of diagram is intended to be used as overview of the evolution of meaning. In order to draw a diachronlex diagram a grid structure is used and the following procedure is followed:

1. Dictionary editions are placed in the columns (right to left in chronological order)
2. Meanings are placed on the rows (top to bottom in ascending order).
3. Meanings are labeled with the ordinal number of the leading meaning (leftmost dictionary edition). In order to do this the relationships between meanings must be pre-computed.
4. A line is drawn connecting all the related meanings. In the case of a bifurcation or merging of meanings, the resulting line will have branches.

One could argue that the natural direction of time should be from left to right and that the ordering of meanings should follow an ascending order from the bottom to the top of the Y axis. In the first case we have followed the practice of RAE lexicographers; nonetheless we have provided an option to invert the time direction. In the second case it is a common and widely adopted practice to use an ascending order in the list of meanings of any dictionary (very often the most used or prevalent meaning appears first, especially in the case of old dictionaries).

An actual example of a diachronlex diagram as compared to an overview of the same word, gabinete (‘cabinet’), produced with the RAE tool is shown in Figure 3.
Figure 3. Left) Diachronlex diagram for the word *gabinete* (‘cabinet’). By avoiding the presentation of unnecessarily detailed information the temporal patterns are easily perceived. Right) Complete view of the same word using the RAE tool for the map of dictionaries: the snapshot was taken by using a 15% zoom in the browser at the cost of losing text readability.

Besides the fact that it is not possible to obtain an overview with the RAE tool due to its use of the vertical scroll (see Figure 1, the overview in Figure 3 is a mockup), even understanding the evolution followed in a single row is quite difficult. As the original position of meanings is changed, a pair of ordinal numbers is provided on top of each meaning (see for instance in Figure 1 how the second row is divided into 2 cells for the 1992 edition; the bottom cell shows two numbers, 2 (in green) and 8, conveying the fact that the second meaning of the 1992 edition is related to the eighth meaning of the 2001 edition). At first sight, one notices that the diachronlex diagram is quite different: instead of using the same row for a computed line of the meaning of evolution, the actual position of each meaning in its edition of the dictionary is maintained. It is thus easily perceivable whether a meaning (as in the previous example of meaning 8) has lost prominence (going down from right to left), as this is also the case of meaning 13 (blue line), for instance.

More benefits of our approach come from the developed user interactions and visualization strategies that allow the analysis. The reminder of the paper explains such mechanisms and their benefits.
Enabling change analysis

The complexity of the evolution of meanings varies from word to word: it ranges from very simple cases (see Figures 4 and 5), with typically less than 10 meanings per dictionary that mostly have either a stable pattern (meanings maintain a very similar order in all editions) or a down-like pattern (meanings appearing in old editions tend to go down and down in the meaning lists as new editions appear), to entangled cases, with more than sixty meanings (see Figure 8) in the latest editions as meanings emerge and submerge, split, merge, etc.).

![Figure 4](image)

**Figure 4.** Overview of the evolution of meaning of the word *mena* (‘ore’). Note the use of line width to convey the degree of uncertainty in the computed similarity between meanings. Some temporal patterns therefore arise and with further analysis (including corpus analysis), the lexicographer is able to assess the evolution of the word. Also, the focus is on the first evolution of meaning (yellow), which appeared in 1869, and the text of each meaning is shown inside a box; since the other lines do not interfere in the view the user has chosen not to send the other evolution lines to the background.

Dealing with uncertainty

If we want to support the analysis of diachronlex diagrams one important aspect is the conveying of uncertainty. We decided to use edge transparency based on the scores of our similarity of meanings algorithm, a solution that has been used in the work of Collins et al. [41]. However, in combination with the background grid and the wide range of hues in the diagrams, transparency may not work well in some cases. Optionally edge width can be used to map similarity (see Figure 4). In addition we provide a slider to control the mapping from scores to alpha or width values, as sometimes the default mapping function provides unsuitable degrees of transparency such as too faint to print. Finally, the similarity scores (between 0-different and 1-same) can also be shown.
Avoiding misinterpretations

We were concerned about the scalability and readability of the diagrams; this is related both to the number of dictionaries and the number of meanings.

Although the RAE is working with only 6 editions of the dictionary (see the data sources section), we validated our proposal with a greater number of dictionaries. In Figure 5 an example with 27 dictionaries is shown, where a curious up and down pattern (a pink line for the fourth evolution of meaning) stands out. Such an evolution line also includes a meaning that disappears for 23 years (1927 to 1950, note the dashed pink line and the red dots in the fourth position in the lists of meanings of the dictionaries in-between), which disappears again after the 1950 edition for 39 years (until the 1989 dictionary). The publication years of the various source dictionaries are not evenly spaced in time. Although, by default, diachronlex diagrams show each subsequent dictionary as evenly distant from the previous one (as this results in a well-packed display), in order to avoid misinterpretations as to the pace of linguistic change, we provide a chronological view in which the user can assess the actual separation between dictionaries and their meanings (see Figure 6).

![Figure 5](image)

**Figure 5.** A) Example of analysis allowed by the diachronlex diagram with the word *humillar* ('to humiliate'). The analysis begins with the overview of evolutions of meaning for 27 RAE dictionaries. The user has focused on the fourth line, which exhibits an interesting pattern with meanings disappearing between editions (1936 to 1947), and curious ups and downs from 1950 to 2001. The diagram uses a focus+context strategy. B) Interactions permit to ascertain overlapping lines (first row: snapshot where pink and red lines overlap; second row: hovering over the red line; third row: hovering over the pink line)

Depending on the specific word under analysis, and taking into account that the user may be aware of some external circumstances related to particular editions of the dictionary (such as a particular interest in technical jargon or the presence of particular people among the academics in charge of the edition [42]), we provide a way of hiding selected dictionaries. This is the case of Figure 7, in which nine dictionaries have been hidden in order to obtain a cleaner representation. Since the user is interested in the evolution of the fourth meaning, and considering that the rest of
the meanings show a linear evolution, i.e., no change in the actual textual content of the meaning and no change in their positions, it is safe to hide dictionaries that would otherwise use screen real estate without adding any valuable information.

**Figure 6.** In most cases, the analysis can be carried out with the default (even) separation among editions. However, under certain circumstances an accurate perception of the time separating the meanings is needed. Diachronlex diagrams can show the evolution lines following a timeline. The transitions between different arrangements (i.e., among this and Figure 5) are subtle animations in order to reduce the cognitive load of the user.

Diachronlex diagrams can also be refined by showing/hiding the background grids and changing their color, with the intention of improving the readability of the diagrams both for analysis or presentation purposes. Different examples of this feature can be seen in Figures 6, 7 and 13 where the background grids have been hidden for the sake of a better result in print.

**Figure 7.** During the analysis and even for presentation purposes, the analyst may decide that some editions are not adding valuable information. Any edition of the dictionary can be hidden/unhidden at the user’s choice, cleaning up the diagram. The relationships between meanings are recomputed taking just the visible editions into account. In this case, nine editions have been hidden as compared to Figures 5 and 6.
**Color encoding**

Improvements to the basic representation can make it clearer and more visually engaging [36]. At this stage the color scheme was decided upon. Perhaps the best references were the works of Tufte [43][44][45] and Ware [46][47]. Every decision was taken bearing in mind that the users would be lexicographers in the beginning and generic users in the future, once our tool is deployed as an open service. In both cases we assume that the users are not familiar with advanced visualization techniques.

![Figure 8](image.png)

**Figure 8.** Overview of a complex evolution of meaning for the word *guerra* (‘war’) with the default color scheme. Note that in this case the whole user interface is shown.

The use of color was a most delicate decision to make. Several authors (Stone [48], Borland et al. [49] and Munzner [50] are three examples) have warned about poor design decisions taken even by the visualization community. The best color map to use depends on the data type. In our case the type of data is ordinal (depending on the ordered list of meanings of the last edition). However, this order is not retained in all editions, and the use of the recommended gray-scale color map (perceptually ordered with luminance contrast) does not work well in most cases. Inspired by Ware's recommendation [46] therefore, we use a selection of 60 different colors based on
luminance and saturation variation from twelve base colors: the six opponent-channel colors (red, green, yellow, blue, black, white) plus pink, cyan, gray, orange, brown and purple.

Figure 9. Overview of a complex evolution of meaning for the word guerra ('war'). In this case the user has chosen to analyze the patterns in the evolution and the lines of meanings that follow a similar pattern over time are of the same color. However the diagram is still complex, with branching (dark blue and orange lines) and enough line crossings to cause difficulties in the analysis.

This is the default color map, that is better suited for analysis of individual evolution lines as it makes possible to visually separate the dots corresponding to the most recent dictionary and then follow each line (see Figures 3 to 8); however, the user is given the option to use a rainbow color map, a blue scale or a green scale, which provide good results for cases in which groups of ordered meanings follow a similar pattern. Since the color of lines is assigned following the order in the list of meanings of the most recent dictionary, groups of meanings evolving similarly would have a similar color, making this evolution group visually evident (see Figure 13, where a rainbow scale is used for the same example as Figure 8).

Finally, we must note that we do not rely exclusively on color for the analytical tasks: the provided interactions and filters (described in the next sections) help solve the problem when the static representation is not clear enough.
Disentangling complex diachronlex diagrams

In the complex cases, the overview provided by the diachronlex diagrams can be cluttered and the color encoding may not be handled properly by human perception (see Figure 8).

On the first hand, as it can be seen in most figures, we provide detail in context and manipulation, making focusing on particular evolution lines possible while keeping the general evolution of the word on the background. This is coupled with a coordinated views approach that helps the user perform further analysis once the diachronlex overview is obtained. These features are discussed in the section devoted to further interactions. We now focus on the features that help the user disentangle complex diachronlex diagrams, maintaining the overview.

We have implemented two algorithms that: 1) compute the similarity of the evolution lines, and 2) simplify the branched evolution lines.

On the first case, based on the length and shape (ups and downs) of the evolution lines we identify evolution patterns. This can be exploited in two ways: a) instead of using the default color scheme, we can use the same color for lines that exhibit a similar temporal pattern (see Figure 9, brighter colors are used for clusters); and b) we can convey clusters, by computing the average position of an evolution pattern, so we draw just a line per pattern, with a width depending on the number of meaning evolution lines they represent (see Figure 10, the width factor can be customized by the user). Also, hovering over a clustered line, a tooltip indicating the number of grouped lines is shown.

---

6 We are grateful to the reviewers of the paper, which contributed to the final design of the algorithm with insightful suggestions.
Figure 10. Next in the analysis started in Figures 8 and 9, is to cluster the evolution lines that present a similar pattern. The average positions of the lines with similar evolution patterns are computed, and the width of the lines is increased according to the number of lines they represent.

In the second case, by removing the branching of lines, the resulting diagram is clearer. We do that by using the top positions of the branches in order to draw the simplified evolution line (see Figure 11). A complex evolution line with many branches can clutter the overall diagram and, in order to gain insight into that particular line, a detailed analysis must be performed (we have implemented filters and other interactions to enable it). We believe that these lines can be temporarily simplified without compromising the understanding of the overview. Furthermore, each branch can be expanded/contracted individually when the user needs to be reminded about the actual relationships of a particular meaning.
Figure 11. A third step in the same analysis of Figure 8 is to simplify the branching lines, as they are not the focus at the moment. Notice how the dark blue and orange (on top) meaning evolutions now are conveyed as not branched lines (a down arrow icon is used in the contracted points to indicate that the line had branches), simplifying the diagram. Subtle animations can be used, back and forth, both in the simplification and clustering steps, so the analyst does not get lost or to avoid reaching to misleading conclusions.

Note that we are aware that the new positions of evolution lines (see, for instance, the new position of the cyan line in the center of Figure 11) after the described operations can be misinterpreted (a group evolution could be mistaken for a single meaning evolution); however, these operations are meant to provide two benefits: reveal interesting patterns that may be occluded in a cluttered diagram and provide a visual hint of a grouped evolution. Examples of both are discussed in the following sections.

Further interactions and coordinated views

Besides the described features related to the overview in diachronlex diagrams, other views and interactions have been implemented in order to support analytical tasks. Yi et al. [51] propose seven general categories of interaction techniques widely used in InfoVis that were relevant in our design process. Figure 8 shows our tool's interface.

A panel can be expanded on the left side of the diachronlex view, comprising a meanings view (Dictionary List), an editions view (Edition List) and a search history. The meanings view can be
seen with the whole textual content for all meanings of the reference dictionary, or just for an excerpt of the textual content. On the other side a fisheye number list is used to provide a quick access to a particular meaning (note that this view can be split into two lists, the first one for evolution lines of meanings present in the most modern edition of the dictionary, and the second one for evolution lines that have not reached the last edition). By selecting a meaning on any of the views, all coordinated views would highlight the data related to that meaning: the meaning textual content in the meanings view, the number of the meaning in the reference dictionary and the evolution line for that meaning including textual boxes containing all the meanings involved in that particular meaning evolution (see Figures 8 to 12 and Figure 14).

The editions view is used to hide/unhide an edition on the diachronlex diagram (see Figure 8) and the search history is used to access words previously analyzed during the current session. With the exception of the editions and search history views, all views make use of a kind of focus+context implementation.

Zooming has been enabled in the overview visualization, and is used once the focus of the analysis is on a particular meaning evolution. If the text of the meanings is long or there are many meanings, the text boxes overlap. By means of the zooming interaction, the space available for the representation of that evolution line increases and the text boxes become more separated. Another example of reconfiguring our representations is that we have enabled an option for changing the direction of time in the visualizations.

Text boxes for the textual content of meanings are shown when the user hovers over number labels or time points in the diachronlex diagrams. Finally, we have implemented animated transitions (slow-in, slow-out) when the analyst simplifies branching, clusters evolution patterns or changes from and to the chronological view, in order to minimize the cognitive load on the users when performing an analytical task.

In addition to the described interactions, we have implemented four filters, which are very valuable to support the particular lexicographers' analytical tasks. These filters were chosen after discussion with the lexicography experts. Thus, upon the user's request, the overview can show only meanings that disappear in any edition, are no longer present in the current edition, emerged in some edition after the oldest one or have branches in any edition of the dictionary.

Finally, the features discussed so far are meant for the exploratory analysis of the evolution of meanings. We have also considered users with the ability to edit in our visualizations. This kind of users can easily change the relationships of meanings by drag and drop operations. These changes are tracked in user sessions and can be annotated. Any change can be undone.

**Use cases**

We have already used some figures to illustrate diachronlex diagrams and their analytical features. The examples shown are actually part of some use cases that we are going to discuss here. We begin with the two diachronic-analysis tasks.

Puche Lorenzo conducted a lexicographical analysis of the mining technical jargon and their incorporation to the academic dictionaries during the 19th century [52]. The presence of neologisms in a language and their eventual incorporation to lexicographic repertories is a confusing matter when the linguistic context is far away from the common language and pertains to a specialized field. By comparing the actual use of mining words (and their chronology) in the corpora and the moment of their incorporation to the academic dictionary, Puche Lorenzo arrives at
interesting lexicographic conclusions. One of the benefits of diachronlex diagrams is that they can be used during the analysis stage and after, once the user wants to present the results. We have analyzed the words studied by Puche-Lorenzo with our tool, and we confirmed his findings. One example can be seen in Figure 4. The analysis would start with a visual inspection of the temporal patterns for the word *mena* (‘ore’) (Figure 4). In the current edition of the dictionary just three meanings are present: the first (yellow), related to the mineral; the second (orange), related to a type of fish; and the third (green), related to a geographic concept. We omit the discussion on the other lost meanings for the sake of clarity. It is easily understood that the meaning related to a fish lost its prominence in the 1869 edition, giving way to the meaning related to a mineral to take the first position. Now, if we focus on the yellow evolution line (Figure 4), we notice a thinner edge between 1869 and 1884, conveying a high degree of uncertainty, i.e., the 1869\(^7\) meaning is quite dissimilar to the present in the 1884\(^8\) edition (the similarity scores are also shown in Figure 4). Puche Lorenzo explains that due to diverse improvements promoted by science, the meanings change according to the extended terminology [52]. This is confirmed by the terms used in both meanings and by the fact that, after 1884, once the terminology has been established, the meaning has not changed at all. Similar patterns would appear for the set of words analyzed in [52]. Diachronlex diagrams cannot provide explanations to these linguistics facts, but can help to highlight interesting situations to the experts. We believe that our tool would have been a time-saving aid for Puche-Lorenzo, who had to go searching in the dictionaries and reading one by one all the entries for every word, before been able to arrive at a conclusion or confirm his hypothesis. Furthermore, the diachronlex diagram could have been also an excellent means for the presentation of the results.

In [53] Morillo-Velarde analyses the treatment of the bull-fighting lexicon in the dictionary of the RAE (DRAE 2001). He identifies 56 verbs taken and marked in the academic dictionaries as specific to the bull-fighting lexicon (including *humillar* (‘to humiliate’). We have shown in Figures 5, 6 and 7 how a particular meaning exhibited an erratic evolution. Indeed, that meaning was referring to ‘lower the bull its head to ram, or as defensive precaution’\(^9\). Morillo-Velarde does not discuss those verbs from a diachronic perspective, however he does point out something that could be further analyzed with our tool: it may happen that the missing meaning recorded in the specialized vocabulary is considered a simple extension of any of the meanings set out as general, i.e. because it is not understood as the use of a special lexicon, but as a special use of the lexicon [50]. He also states that contemporary lexicography is shown to require surpassing the traditional conception, which considers only the basic query of the DRAE’s electronic edition, and the dictionary as a book. Taking into account the circumstances under which each of the 27 RAE’s dictionaries were edited (e.g., more or less attention paid to specialized language), the leaps, ups and downs of the meaning related to bullfighting could be explained. Thus, the benefit provided by the diachronlex diagram is that of shedding more light on those cases where the absence or presence of a particular meaning in a given edition of the dictionary can only be understood when other editions beside the six ones selected by RAE are considered; the addition of data pertaining to those editions of the dictionary not considered in the map of dictionaries, is supported by the

\(^7\) *f. Min. Los minerales ó metales mezclados todavía con la ganga ó con la piedra y tierra de la mina* (“f. Min. Minerals and metals still mixed with gangue or with stone and earth from the mine.”)

\(^8\) *f. Min. Mineral metalífero, principalmente el de hierro, tal como se extrae del criadero, y antes de limpiarse* (“f. Min. Metaliferous mineral, mainly iron, as extracted from the hatchery, and before cleaning.”)

\(^9\) *Bajar el toro la cabeza para embestir, o como precaución defensiva.*
scalability of the solution (we have developed a module for uploading the additional data that is kept per user). Had we used the RAE tool, we would have found that the bullfighting meaning appeared in the 1992 edition in the fourth position and passed to the fifth one in the most current edition. However, the RAE had already registered the meaning as early as in the 1927 dictionary.

In [13], Ibarretxe-Antunano discusses how cognitive linguistics can contribute to lexicography from a theoretical as well as practical standpoint. He deals with the macrostructure (organization of the set of meanings) of a lexical entry and refers to the problem of lexicographic linearity [54]. The different meanings are linearly ordered, without taking into account their inner structure and possible relationships among them [13]. As Ibarretxe-Antunano states, this fact does not mean that a change in the design of a dictionary entry is needed, since many times it does not depend on the lexicographer but on the dictionary edition. This can be seen in Figure 11, where several meanings of the word guerra (‘war’) drop down together from very prominent positions to low ones, between two editions that are quite close in time (1992 and 2001). Thus, this situation can be inspected from the cognitive linguistic perspective for each of the involved evolution lines (see the green line pertaining to the ‘prisoner of war’[^10] meaning in Figure 12).

![Figure 12](image-url)  
Figure 12. The focus of the analysis is now on a set of meanings that were consistently listed among the first positions in older editions of the dictionary. Particularly, the analyst is interested in the second meaning in 1925, that, suddenly, drops down to the 61st meaning in the most recent edition.

[^10]: Prisionero de guerra.
edition (note both the fisheye list on the right and the meanings list for the focused dictionary (2001), on the left). Similar behaviors can be found in this example. Once this fact becomes evident, it is the task of the analyst to provide an explanation for it.

The ability of diachronlex diagrams to convey this fact can also be seen in Figure 13, in which further analysis for the word guerra (‘war’) is conducted. In this occasion, a rainbow scale was selected in the diachronlex diagram, and the bluish and purple lines drop down stands out.

**Figure 13.** Due to the nature of the entries, that organizes the meanings as ordered lists, and the fact that each edition of the dictionary is strongly based on the previous one, a rainbow scale can be used as a quick way to obtain an overall idea of the present patterns of “group editing”, when the lexicographer decides to move a group of meanings for whatever linguistic or organizational reasons.

As for the dictionary-making tasks, Figure 14 shows an example in which a lexicographer is inspecting the meanings of gavilla (‘sheaf’). A gap (see the dashed line) for the first meaning can be easily noticed. That meaning is related to ‘the gathering of branches of wheat or barley’\(^{11}\), and has remained stable from 1780. This is reinforced\(^{12}\) by two visual hints for meanings that are almost the same (see the leftmost meanings in Figure 14): a) thicker box borders, b) diamond-shaped points. The evolution of the second meaning is clear, but there is a meaning that has not been associated (green dot, first meaning of the 1817 edition). Thus, the lexicographer decides to inspect the content of that meaning that says it is the same as gaviota\(^{13}\) (‘seagull’). The lexicographer then goes to the source (scanned image of the original 1817 dictionary) and finds out the problem: there is a second entry for gavilla that contains only the meaning related to seagull. It becomes clear that the word gavilla has been misspelled (it is listed, alphabetically, after gavillero.

\(^{11}\) *La junta de sarmientos o cañas de trigo o cebada […]*

\(^{12}\) One reviewer of the paper pointed out the benefit of doing so.

\(^{13}\) *Lo mismo que GAVIOTA […]*
and before gavion). Actually, the meaning related to seagull should have been listed under gavina (‘seagull’ in some parts of Spain), as the single evolution line for that word (see the bottom of Figure 14) is in fact related to seagull, but only since 1884. The final step would be to annotate this case with our tool that otherwise would have passed unnoticed.

Similarly, errors in the process of digitalization of the entries are easily discovered; for instance, the first meaning for gracia (‘grace’), in the 1884 edition was apparently “f…”, a nonsensical meaning; the actual dictionary entry was checked and this error was not present.

![Figure 14. Process of discovery of a mistake related to ‘seagull’ that the dictionary editor made in 1817, when it was erroneously listed as a meaning of the word gavilla, instead of as a meaning of gavina.](image)

Finally, for any evolution line, if the computed association of meanings is not correct in the opinion of the lexicographer, he/she can change the connection with our tool using a drag and drop operation and annotate the change. Figure 15 shows the editing view (for the word gavilla), in which the meanings are organized in a grid. Each meaning is conveyed as a circle (or as numbered star, if the meaning appeared in that edition for the first time), labeled with the position of that meaning in its dictionary. Annotations and changes can be kept per user or shared among the group of editors.
Users with editing rights can easily annotate meanings or change associations by dragging any meaning (numbered icon) and dropping it into the desired evolution line (row of the new associated meaning).

**Usability evaluation**

Through the design and development process (a period of time extending, approximately, 24 months) we conducted different kinds of evaluations. At a very early stage, when very little usability evaluation had been performed, we demonstrated our prototype to both lexicographers and software engineers at RAE, when a proof-of-concept was the objective. The initial user’s feedback indicated that a more visual and interactive approach was more effective to support analytical and sense-making tasks related to the study of evolution of meanings than the traditional one, mostly based in tabulated lists of meanings. The feedback they provided influenced the design of the views and their associated interactions in an iterative process. At this stage, we were able to identify additional tasks not considered in our initial design, and we tailored our proposal according to the expressed needs of the people in charge of NDHE. For instance, they were most worried about finding a compact way to represent an overview of the evolution of a word and diachronlex diagrams were very well received. Another feature that resulted in very positive feedback was the possibility of manipulating the evolution lines by drag and drop operations. On the other hand, they were not much interested in exploring the possibilities of automated methods (others than theirs) for the meaning similarity computation.

Also, the design of user interaction in our tool was subjected to heuristic evaluation, in which experts in user interaction design (both experts in our department not involved in the development of our tool and colleagues from other research groups) were asked to assess our tool bearing in mind its purpose as a visual means to explore and analyze evolution of meanings in historical dictionaries; at this stage, they did not follow specific task scenarios; instead they were engaged in “free-play” with the tool. Thus, these assessments determined the strengths and weaknesses of our design choices. Based on these findings a list of recommendations for changes to improve the design was made (including also some valuable suggestions made, over the last 9 months, by both reviewers of internal technical reports and of the present paper). These heuristic evaluations helped to uncover major design problems that were carefully studied and addressed, arriving to the final design described in section 3.
An example of former visual design decisions that were questioned were the use of curved lines and lines with shadows to depict individual meaning evolution. In particular, the experts were concerned with the use of curved lines, as it could imply a data resolution that was not present (it looked like there was a changing trend when in fact it was inexistente). The tool was redesigned avoiding the use of “artificial” curves. The use of shadows on the lines was also discarded, because it just added visual complexity without any additional analytical power.

Finally, the design of our tool was driven by a formative user study. We wanted to evaluate the potential of our tool, both for linguistics experts and novice users, to explore the data stored in RAE’s map of dictionaries. In total, 16 participants were recruited and tested individually. Evaluations were conducted with 6 linguists of different fields (from lexicographers to teachers of Spanish as a foreign language), as well as with 10 non-linguist users. Ages ranged from 24 to 71. Our tool was run at a resolution of 1600x1200.

We built a tasks list that gradually grew in complexity, i.e., the task semantics became more complex, the total number of meanings for the word involved in the task increased (from dozens to hundreds), multiple filters and interactions were required to complete the task, etc.

Examples of simple tasks are: "Which edition of the dictionary contains more meanings for the word gota (‘drop’)?", “Find the patterns of evolution lines for the word copa (‘cup’)?”, and "How strong are the similarities among meanings of the word ritmo (‘rhythm’)?" More complex ones include "What meanings of the word ropa (‘clothing’) are the result of a merge of previous ones?" and "How do the diachronlex diagram obtained with the NIST algorithm differ from the one obtained using the BLEU algorithm for the word mesa (‘table’)?" Some tasks consisted of two or three subtasks. For example, 1) For the word correr (‘to run’), find all the meanings that disappeared before the last edition of the dictionary. Are they related to any evolution line?, 2) Associate those meanings that should pertain to other evolution line according to your knowledge, and 3) annotate the change.

During each of the formative evaluation sessions, we followed a protocol of giving the participants an overview of the evaluation about to be performed, and then giving them a brief tutorial on the use of our tool, spending no longer than 20 minutes interacting with the tool while talking aloud. We timed each user, and took notes on comments they made and any critical incidents and usability issues that occurred during this walk through of the tool. Once a user had successfully figured out how to use the tool, we began having them perform the set of tasks. Next, participants spent 10 minutes inspecting how the tool developed by the RAE for the map of dictionaries convey the meaning evolution for some words involved in the tasks they had completed with our tool. Finally, participants filled out a user satisfaction questionnaire, and provided any feedback they had.

Time to perform the set of tasks ranged from about 2 minutes to more than an hour.

The result of any formative usability study should permit the redesign of those parts of the tool for which issues with the user interface were identified. Some examples of such issues and improvements follow.

One of the first refinements provided by the feedback of the participants (mainly by non-linguist ones) was the possibility to have the oldest editions of the dictionary on the left side and the most recent one on the right (rather than the opposite, which was the layout our tool used). We implemented this more conventional axis of time, while keeping the former one as the default.

Another issue had to do with the abstraction in which dictionaries correlate with points in time. One linguist was concerned with it and stated that “This is not unusual in this field, but it is probably more accurate to associate them with indeterminate intervals in time that overlap. The
reason that this is important is that ortholexical work often comes in waves (several dictionaries being published in close proximity to one another that vary in register or completeness), and these dimensions of variation turn out to be more relevant linguistically than the periods of time that separates them. The equally space columns on which these dictionaries appear don't permit us to have a sense of these bursts of activity or very uneven proximity (temporal or traditional) that relates them.” On the other hand, other participants (both experts and novices) praised the cleanliness of the more compact approach. We addressed the issue by developing a “chronological view” (described in section 3), that could be enabled at user desire, while maintaining the evenly spaced view as the default one.

Regarding the interactions provided for uncluttering complex diagrams, participants were happy to play with the visualizations and tend to ignore the possible linguistic implications of the unrealistic positions of clustered and contracted branches. Linguists, in general, tend to spend more time hovering over the different evolution lines and reading the actual text of the meanings. They were cautious with the clustering and simplification of branches at first, and preferred to reveal patterns by color, although they mainly used the default color scheme. Most of them, appreciated the aforementioned features once they were shown a real example on how to use them.

One clear area of improvement was the alignment of the numbers in the fisheye list with the Y-axis in the diachronlex diagram. This happened when the studied word had few evolution lines and/or many meanings disappearing before the current edition of the dictionary. We addressed this issue by splitting fisheye lists (as explained in section 3), which solved the problem in most cases; however, poor alignment still can occur. If needed, the alignment can be done manually, by using the magnification control. Still we plan to include an automated alignment in the next iteration on the tool.

Finally, as mentioned, users filled out satisfaction surveys at the end of the session. The average user satisfaction scores were quite high for such a novel user interface with so little training (the scale was 1 to 7, with Disagree, Agree as the anchors of the scale). Users liked using this tool (avg. score 6.5) and the system seems to be easy to learn (avg. 5.4). Furthermore, it seems to be easy to discover patterns using color (avg. 5.8) and simplification and clustering (avg. 6.1). The possibility of correction of computed meanings associations and annotation was appreciated (6.2). Because alignment of fisheye lists has not yet been automated, there is still room for improvement (avg. 5.1). The most notable average low ratings centered on the accuracy of evolution lines computed (avg. 4.8), but this is somehow mitigated by the fact that participants found appropriate the way in which uncertainty is conveyed (6.0). Finally, our tool was preferred as a visual means for the analysis and manipulation of the map of dictionaries over the RAE’s tool (6.2).

Conclusions and future work

In this paper we presented a novel visualization technique for the representation, analysis, validation and modification of an historical dictionary. We have proposed diachronlex diagrams as an overview for the evolution of meanings, and we have also presented several use cases. Although all the work has been done with the NDHE in mind, the generality of the visual encoding can be applied to any other historical dictionary.

We have applied the computational information design methodology to the diachronic linguistics domain.
We have been lucky to work together with the multidisciplinary team of computer scientist and lexicographers of the RAE, which provided insightful recommendations while developing our tools. This has been very helpful with the validation of our visual design approach and we believe that it can be easily transferred to other related fields such as corpus linguistics or computational linguistics.

We have shown how our proposal can be integrated in a bigger system that could ease the challenge of building a historical dictionary. It is part of our future work to enable visual analysis in many different tasks of the lexicographers and the possibility of reaching their dream. Many of these analysis possibilities would also be of general interest.

The use cases discussed show how diachronlex diagrams support actual analytical tasks of the lexicographers, leading to the discovery of temporal patterns and anomalies recorded in the academic dictionaries of the Spanish language.

An interesting limitation of diachronlex diagrams is that dictionaries do not supply detailed semantic distance measures between definitions. Due to the nature of the data, it is only possible to show the various definitions as ‘whole step’ increments, i.e., meaning x and meaning x+1 appear to be the same ‘semantic distance’ as meaning x+1 and x+2, even if there is a much greater difference. This leaves an open research line in our future.

This tool represents a significant contribution as it eases the complex tasks of analysis and validation of diachronic lexicography works and it is intuitive enough to be used by regular visitors of an online dictionary.

Acknowledgements
The authors wish to thank Spanish Government project FI2010-16234, the participants of the usability evaluations and the insightful comments of the reviewers of this paper.

References


C Rohrdantz, A Niekler, A Hautli, M Butt and D A Keim. Lexical Semantics and Distribution of Suffixes --- A Visual Analysis. In: EACL 2012 Joint Workshop of LINGVIS (Visualization of Linguistic Patterns) and UNCLH (Uncovering Language History from Multilingual Resources), April, Avignon, 2012, pp 7-15


Appendix

29 dreams of interactive dictionary visualization
(based on the work of G. M. de Schryver [4])

1) Management of huge amounts of data;
2) Combination (and integration) of various reference information sources;
3) Dynamic repository of knowledge;
4) Non-static representations of actions and processes;
5) Seeing particular things in motion;
6) Visualization of semantic and associative fields by means of interactive colored computer graphics;
7) Allowing the user to browse the corpus;
8) Rapid access to large amounts of lexicographical evidence in corpora for skilled users;
9) Access to static lexicographic data (in databases), dynamic lexicographic data (through corpora), non-lexicographic data;
10) Hypertext articles with links to context-sensitive, extralinguistic and culture-specific information, and links to tailored corpus collections;
11) Enabling skilled users to make their own judgment on equivalences by scanning corpus examples;
12) Creating a dynamically growing historical language database;
13) The consultation-reading of a database should give way to exploration/research of that database;
14) Filtered searches;
15) A multitude of new routes to the data - temporal filters;
16) Liberating users from alphabetical order;
17) Inclusion of all inflected forms as lemma signs;
18) Information about the use of every single word;
19) Analysis of any word in context;
20) Definition chaining;
21) Powerful and multi-purpose “navigation” facilities, highly user-friendly interfaces;
22) Customizing the display/presentation;
23) Different levels of interface for different users;
24) Presence of both look-up mode (the user is in search of a specific piece of information) and browsing mode (relaxed reading);
25) Zooming into different levels of granularity, dynamic interactive lookup;
26) Layers of varying complexity in the interpretation of the lexicon;
27) Providing various styles (i.e. different ways to present the information);
28) Adjusting the information level according to the different target groups and their needs;
29) Persistent user sessions.