ITALIAN HISTORICAL DICTIONARIES: FROM THE ACCADEMIA DELLA CRUSCA TO THE WEB

Pietro G. Beltrami, Università di Pisa e Direttore CNR-OVI, Firenze (beltrami@ovi.cnr.it) and Simone Fornara, Università di Vercelli (simofornara@libero.it)

Abstract

A brief review of the Italian lexicographical tradition, from the first printed works centred on the language of a few representative authors, through the canon of authors cited in the various editions of the Dictionary of the Accademia della Crusca and others that follow suit, ending with the two major works of the 1800s and the 1900s (Tommaso-Bellini and the Grande dizionario della lingua italiana – GDLI or ‘Battaglia’), brings to light the central role of quotation from authors. The cult of written authority drawn from literature is an essential factor to be taken into account in the study of the history of lexicography in Italy. The final part of this article examines the Tesoro della Lingua Italiana delle Origini (TLIO), the natural evolution of historical lexicography in the days of Internet, an exclusively electronic dictionary poised between language dictionary and corpus-based dictionary.

1. Birth and development of lexicography in Italy: a brief survey

1.1. Literature, lexicography and grammar: three fields in close contact

One of the remarkable features of Italian literature is that almost from its very beginning it was able to pride itself upon three authors of absolute genius, who have demonstrated in their works, in the most wonderful way, the infinite potentialities of the new language: Dante, Petrarch and Boccaccio, also known as the Tre Corone (‘Three Crowns’). Their importance runs no risk of being overestimated: they exercised a decisive influence on the development of Italian literature and on the language of the following centuries. Their role as author-guide was already very obvious to Humanist and Renaissance men of letters, judging by the ever-increasing interest in their masterpieces and the language used therein.

One of the first tangible results of this interest is its legacy to lexicography: the first monolingual dictionary of Italian was published in Venice in 1526, with the rather significant title: Le tre fontane di Messer Nicolò Liburnio in tre libri.
divise, sopra la grammatica, et eloquenza di Dante, Petrarcha, et Boccaccio (Della Valle 1993: 32), which could be translated as ‘The three fountains of Master Nicolo Liburnio, in three volumes, following the grammar and eloquence of Dante, Petrarch and Boccaccio’. These three great authors (the three fountains, or sources of eloquence) are used as models, in keeping with the position expressed by the lexicographer in one of his previous works, Le vulgari elegantie (‘Elegant Italian’), in 1521. Liburnio’s choice is part of a broader picture, a consequence of the presence of Pietro Bembo in Venice at the beginning of the sixteenth century, and of his collaboration with the Venetian printer Aldo Manuzio. Together they published their famous ‘Aldine’ editions of the works of Petrarch (1501) and Dante (1502), and then, in 1505, came Bembo’s Asolani, written in a faithful imitation of Boccaccio’s prose. In the following years, Bembo worked on his masterpiece, Prose della volgar lingua (‘Prose about Italian language’), which was published only in 1525, in Venice. It is one of the most influential works in Italian literature from any period. In these three volumes, Bembo theorises on Italian Classicism, citing Petrarch and Boccaccio as the absolute references for style and language for poetry and prose respectively, whilst expressing reserve about Dante, finding him ‘guilty’ of over-realistic use of language in certain parts of the Commedia. The ‘Tre Corone’, although to differing degrees because of these reserves about Dante, are therefore positioned as in the inescapable reference standard for any writer with literary ambitions. At this point, it becomes clearer how well Liburnio’s dictionary fits into this framework: it too is one of the first concrete results of the respect paid to the ‘Tre Corone’, sanctioned by Bembo’s theories, although his reserves about Dante are not shared by Liburnio. Della Valle (1993: 32) points out that ‘the direction taken by sixteenth century lexicography from then on follows Bembo’s example, to which Liburnio also subscribed’. However, Liburnio did retain a certain independence of judgement.

It is impossible to ignore the fact that the birth of Italian grammar is also placed under the sign of the ‘Tre Corone’. Apart from the little-known, unpublished, fifteenth century Grammatichetta vaticana by Leon Battista Alberti, which is based on the living language of Florence and not on that of any authors, the first printed grammar is Fortunio’s Regole grammaticali della volgar lingua (1516), which is based on numerous examples from literature, almost exclusively from Dante, Petrarch and Boccaccio. What Fortunio’s grammar sought to show was that Italian, like Latin, had grammatical rules, the truth of which could be proved by the unquestionable quality of the works of these three authors, comparable to those of the classical Latin masters. This explains the frequent use of quotations, the cornerstone on which the work was built, following a tradition which, while gradually adding to the list of canonical authors, continued for several hundred years, starting with the grammar to be found in the third volume of Bembo’s Prose, until more modern grammars appeared in the nineteenth century. Italian grammars and dictionaries have therefore followed very similar paths, above all at their origin in the sixteenth
century, often converging voluntarily. Italian grammars theorised about language modelled on that of the ‘Tre Corone’, while the first dictionaries created a true and proper repository of their language. Reference to authors, although initially limited to the three giants of literature from the fourteenth century, is therefore the starting point in both cases, which inevitably leads both to be closely linked to literature. The foundation of Italian grammatical and lexicographical tradition is therefore firmly rooted in ‘the cult of authors’.

The development of lexicography in the following years confirmed this tendency. As Della Valle (1993: 33) explains, the years between 1526 (Le Tre Fontane) and 1543 (Alunno’s Le Ricchezze) see the spread of a type of lexicographical collection limited to the language in the works of just a few authors, held up as models to be imitated, in complete accord with contemporary grammar theories which sought to restore the canonical fourteenth century model. Such works are Minerbi’s Vocabulario (Venice, 1535), which relies solely on Boccaccio, and also Francesco Alunno’s Le Osservazioni sopra il Petrarca (Venice, 1539) and Le ricchezze della lingua volgare sopra il Boccaccio (Venice, 1543).

Alunno also published La Fabrica del mondo (Venice, 1548), the first methodical dictionary of the Italian language, in which the canon of authors is extended, to include Ariosto, Sannazaro, and Alunno himself. It is worth mentioning Fabricio Luna’s Vocabulario (Naples, 1536) which introduces a canon of sometimes conflicting author models, for besides the ‘Tre Corone’ are writers from different geographical areas, together with the language of grammarians and the language of the lexicographer himself, and above all Alberto Acarisio’s Vocabolario, grammatica et ortographia de la lingua volgare (Cento, 1543), which is not content merely to follow its predecessors, but rather proposes criticism and adjustments and introduces major innovations in dictionary writing, such as much greater precision in word definitions, for example.

1.2. The Dictionary of the Accademia della Crusca

In 1591, a major change occurred in Italian lexicography: from that date onward, the Accademia della Crusca (founded in Florence in 1582) started to devote itself to the creation of a dictionary, the Vocabolario. The criteria for its compilation were greatly influenced by a most illustrious member of the Academy, Leonardo Salviati, who, before his death in 1589, had drawn up a list of authors to be cited, which the lexicographers followed closely, with some additions. Their models were those fourteenth century authors who had expressed themselves in Tuscan and particularly Florentine. The big difference with regard to Bembo’s position was the number of authors considered to be authoritative: not only the Tre Corone (who were however the core reference), but also minor authors, little known or even unpublished, who had written in Tuscan, quoting from manuscripts owned by the academicians themselves, and
thus not readily accessible for most readers to verify. The aim was to
demonstrate the continuity from ancient to modern Tuscan; in this way, the
living Florentine language was documented with quotations from ancient
authors. The academicians therefore referred back to previous lexicographical
tradition, modelled on the Tre Corone, which had developed outside Tuscany,
but also distanced themselves from this tradition and Bembo’s stance,
attempting to ‘bring home to Florence the seat of lexicographical research,
above all other regions of Italy’ (Marazzini 1993: 177). As regards modern
authors, it is a well-documented fact that the dictionary gave preference to
archaic works, at least in the first two editions: there may not perhaps have been
unanimous agreement on this aspect, yet the exclusions are noteworthy,
foremost among them Tasso, on whose language weighed the express
condemnation of Salviati.

The great innovation of the Crusca dictionary, apart from the corpus of texts
selected for reference, lies in its methodology: the organisation, efficiency and
coherent methods of work put into practice by the compilers are unprecedented
in the Italian lexicographical tradition. Perhaps because of these characteristics,
the academicians, who undertook the work as veritable ‘dilettantes’, untried at
similar enterprises, succeeded in maintaining unity amongst themselves, and
completed the greater part of the work in a relatively short space of time, from
1591 to 1595. They then faced the problem of financing their work: a serious
problem for a private association like the Accademia della Crusca, which could
not count on public support. They resolved on self-financing, which allowed
them to retain substantial editorial freedom, as they were not required to answer
to a higher authority. In all probability, economic reasons lie behind the decision
to publish the dictionary in Venice and not in Florence. The first edition of the
Vocabolario degli Academici della Crusca or Dictionary of the Accademici della
Crusca, in a single volume, was published by the printing house of Giovanni
Alberti, in Venice, in 1612. This did not prevent the Dictionary from fulfilling
the intentions of the academicians, thus giving Florence the role of leading city
in Italian lexicography.

The great leap forward in quality compared to previous works, which had
rarely been more than lists of words, becomes crystal clear on examination of
the thousand-page dictionary. As an example, consider the definition of ‘word’:

(1) Vocabolo. Voce, con la quale son chiamati i nomi particolari di
ciascuna cosa. Lat. vocabulum. Bocc. nov. 79.32. Il Medico, che a Bologna
nato, e cresciuto era, non intendeva i vocaboli di costoro. E nov. 2.9. Quasi
Iddio, lasciamo stare il significato de’ vocaboli, ma la ’ntenzione de’
pessimi animi non conoscesse. Dan. Par. 8. Pigliavano ’l vocabol della
stella. Conv. 6. Vedemo nelle città d’Italia, ec. molti vocaboli essere spenti,
e nati, e variati. È da vocabolo vocabolario, che è questo libro. (‘WORD.
Term which is used to indicate the particular names of all things. Lat.
 vocabulum. Bocc. Story 79.32 The Doctor, who was born in Bologna and
grew up there, did not understand their words. Also Story 2.9. As if God, apart from the meaning of words, did not know the intentions of evil souls. Dan. Par(adise). 8. They took the words of the stars. Conv(ivio). 6. In the cities of Italy, many words can be seen to die out, to be born, and to change. And also from word (Vocabolo) comes dictionary (Vocabolario) which is what this book is.

The microstructure is simple and linear: first is given the meaning of the word; then a Latin equivalent if one exists; then come illustrations of usage from authors, with direct quotations, in this example all from Boccaccio and Dante; finally, the word vocabolario is indicated to be derived from the primary term, with a self-referential definition. The lack of precise definitions for certain words, whose meaning is taken for granted, is one of the limits of the first editions of the dictionary. It has been much discussed, and is seen as being inherited from previous dictionaries, and remained current practice for another two hundred years, until the end of the nineteenth century. The most famous examples are those pertaining to the fields of botany and zoology, like the following:

(2) Cammello. animal notissimo (‘CAMEL. A very well-known animal’)

(3) Cipolla. agrume noto, e ne sono di più maniere (‘ONION. Well-known tangy vegetable, with many varieties’)

(4) Cipresso. arbore noto (‘CYPRESS. A well-known tree’)

(5) Elefante. animal noto (‘ELEPHANT. A well-known animal’)

The first edition of the dictionary did not contain many technical terms, a deliberate choice on the part of the academicians who did not deem it necessary to include words which were not in common usage. In keeping with their compilation criteria, they instead recorded words present in Tuscan and Florentine dialects, such as danaio for denaro (‘money’) or uguanno for quest’anno (‘this year’).

The Crusca Dictionary had considerable success, not only in Tuscany, but at the same time was also the object of numerous criticisms regarding the editorial line followed by the academicians. Such notable critics as Paolo Beni and Alessandro Tassoni pronounced against the Dictionary in no uncertain terms, criticising its archaic character and too strongly marked Florentine orientations, and demanded that pride of place be given to modern authors, of the sixteenth century, above all Torquato Tasso, the most famous author to be excluded from the academy’s selection.

Despite all this, the second edition of the Dictionary, still published in Venice, and in a single volume, in 1623, was fairly faithful to the first edition, limiting its scope to minor adjustments and corrections. A great leap forward, both in quality and quantity, occurred with the third edition, which was published in
1691, this time in Florence, in three volumes, bearing the mark of the Accademia della Crusca publishing house. The fact that this new edition was delayed by more than thirty years, bears witness to the academicians’ desire to improve on the previous results. One of the deciding factors in this improvement was undoubtedly the cultural status of some of the collaborators. The presence of two such scientists as Francesco Redi and Lorenzo Magalotti may well explain why the third edition devoted so much space to technical and scientific terms, which had previously been ignored. As proof of this, were now included many seventeenth century authors of science, first among whom was Galileo Galilei, who had the advantage of being both a scientist and a Tuscan, thus safeguarding the basic selection criteria of the Dictionary. Obviously the greater reference to modern authors, no longer only those of the fourteenth century, no longer only those from Tuscany, is the sign of a change in outlook by the compilers, who were no longer so rigidly bound by Salviati’s canons, which nonetheless remained a firm point of reference. This change in outlook explains the inclusion of authors such as Castiglione, Machiavelli, Guicciardini and, above all, Torquato Tasso, which compensated in some small way for the now-outdated condemnation which had led to their exclusion from previous editions. In short, the third edition of the Crusca Dictionary ‘reconciled the continuing respect of the archaic tradition (following Salviati’s guidelines) with an open-mindedness towards new things shown by the inclusion of such names as Castelvetro, Buommattei, and Bartoli amongst the grammatical references’ (Marazzini 2002: 315).

Other significant changes are the presence of words unattested by examples from authors (like accannellare ‘avvolger filo sopra i cannelli’ (‘to wrap thread around reels’) which was not defined in the first edition and simply linked to the word cannella (‘reel’) and greater use of the abbreviation V.A. (voce antica ‘ancient word’) to indicate those words whose use, in the opinion of the academicians, was unadvisable, being too archaic. The presence of these indications was an attempt to render more easily accessible the language of authors of past times, and not to offer a model to be imitated, as the Academy pointed out in the introduction to the Dictionary (see Crusca 1691: I, 16). Some words, which in the previous editions were not headwords, such as vocabolario, in example (1), were now given independent entries, and a more precise definition:

(6) VOCABOLARIO. Raccolta di vocaboli, notati colla spiegazione del lor significato, quale appunto si è questo libro […] (‘A collection of words, with indication of their meaning, which is what this book is’)

However the custom of not defining some words was maintained, as examples (2)–(5) show.

The Accademia della Crusca pursued its lexicographical activity through the following centuries: from 1729 to 1738, the fourth edition of the Dictionary, in six volumes, was produced, without any major changes, apart from an
enlargement of the corpus. It remained faithful to the guiding principles of the first editions, yet, when compared to the more liberal spirit of the third edition, showed a certain rigidity as to scientific terms, which were treated once more with circumspection. Finally, in 1863, work began on the fifth edition, which ended in 1923, by decree, signed by the minister Gentile, when the eleventh volume had been reached, with the word ozono.

1.3. Legacy of the Accademia della Crusca and new requirements

From 1612 onwards, any mention of Italian lexicography inevitably meant speaking of the Crusca Dictionary. Although it was often questioned and even opposed, it remained an irreplaceable reference point. Any opponents, proposing other models of language, were almost unavoidably bound to adopt at least the basic structure and organisation of the Dictionary. Reference to authors, in the form of quotations from their works, became the accepted practice. Those who attempted to introduce other methods, without reference to authors, were almost always isolated examples, as was the case for Adriano Politi’s Dittionario toscano (1614), which proposed Siennese and not Florentine, with examples from proverbs and sayings. Others tried to raise various names to a level of authority attaining that of the Accademia della Crusca, entering into open controversy with the academicians, as did Girolamo Gigli, who compiled the Vocabolario cateriniano (published from 1717 onward), based on the writings of St Catherine of Sienna. As Della Valle writes (1993: 54), such attempts ‘inevitably ended as lost causes, fighting hopelessly against a dictionary, such as that of the Academy, which had succeeded in imposing a model of literary language which, for the written word at least, meant the waning of regional variety.’

It is above all in the treatment of technical and scientific lexis that the Crusca Dictionary caused discontent, giving rise to the publication of remedies against the exclusion of this sector by the Academy, in response to an ever more pressing demand, with ‘a distinction between general lexis and technical terms. This sign of the times can also be found in the Dictionnaire de l’Académie française, which published for the first time in 1694 and reserved for itself common and general lexis, leaving specialised terms for another publication, the Dictionnaire des arts et des sciences, published in the same year’ (Matarrese 1993: 166). This was also the case in other encyclopaedias published abroad, such as Ephraim Chambers’ Universal Dictionary of Arts and Sciences, published in England in 1728 and translated into Italian in 1749. In Italy, one of the most significant examples is Antonio Vallisnieri’s Saggio alfabetico d’Istoria medica e naturale, published posthumously in Venice in 1733, in a collection of writings selected by his son. In the preface, the author expresses his admiration for the glossaries of technical and scientific terms published in other European countries; logically enough, the body of the text includes numerous foreign terms, adapted and even in their original linguistic forms.
Nevertheless, Vallisnieri cannot avoid continual direct reference to the Crusca Dictionary for the meaning of many words.

A form of compromise is reached between Accademia della Crusca criteria and the requirements of modern word meanings (as its title shows) in Bergantini’s *Voci italiane d’autori approvati dalla Crusca nel Vocabolario d’essa non registrate con altre molte appartenenti per lo più ad Arti e Scienze che ci sono somministrate similmente da buoni Autori* (‘Italian words from authors approved by Crusca Academicians but excluded from that Dictionary, and many others belonging to the fields of Art and Science found in the works of good authors’) (Venice, 1745). The author, whilst accepting the Crusca canon of authors, decides to broaden his scope, including non-Tuscan authors in order to cover scientific terms, particularly in the fields of medicine and anatomy. The basic criterion of quotations from authors is strictly adhered to, even if it is adapted to meet the new requirements.

Towards the middle of the eighteenth century, lexicography turned with growing interest to the study of language in everyday use, even if not attested to by quotations from authors. The century closed with the start of the publication of a work that was the best example of this new interest, and of the influence of the great European encyclopaedic works, particularly the English and French (above all Diderot and d’Alembert’s Encyclopédie): Francesco D’Alberti di Villanuova’s *Dizionario universale critico enciclopedico della lingua italiana* (published in Lucca from 1797 to 1805). One of the most significant innovations of this work is the combination of the living language, often collected by surveys conducted by the author himself, with the traditional Crusca-approved author quotation method. As Serianni (1989:63–64) remarks, ‘Author examples, mostly those of the Crusca academy, but also from a small selection of new authors, were still present, but were no longer the indispensable passport for a word to be accepted in a dictionary.’ Thus D’Alberti succeeded in including in his dictionary words used by craftsmen, for which no written authority existed. ‘We thus turn more to everyday language than to authors, to the language of educated, moderately well-read men [...] The very expression ‘vulgar (= popular) tongue’ shows that we must follow the usage of the people, for they use the language of the century, and it is difficult to prove that all ancient words are better than modern ones’ (Alberti di Villanuova 1797: XI). The Universal Dictionary represents a first step in the direction of a true general dictionary, containing all the lexical heritage of Italian.

Despite this beginning, the nineteenth century opened with a return to more traditional methods, bearing witness to the force of the canon created by the Crusca criteria. From 1806 to 1811, in Verona, the Abbot Antonio Cesari, a leading figure in the defence of Purism, published the so-called *Crusca Veronese*. In keeping with his theories, and firm belief in the absolute validity of the fourteenth century model of language, without any concessions to modernity, he compiled a veritable Cruscan-type dictionary, even more rigid in its criteria. Cesari thus included in his dictionary words which had been
removed from that of the Academy, going back to the authors of the ‘Golden Age’, mostly from the fourteenth and some from the sixteenth century, eliminating almost all traces of modernity and confirming the aversion to technico-scientific language. Another work which closely followed Cruscan methods was Vincenzo Manuzzi’s *Vocabolario della lingua italiana già compilato dagli Accademici della Crusca ed ora nuovamente corretto ed accresciuto*, published in Florence from 1833 to 1840. Although Manuzzi closely followed Cruscan traditions, he was less rigid than Cesari, and included quotations from modern authors such as Leopardi.

While these works confirm the lasting impact of the Crusca Academy, arguments were developing as to what criteria were essential for dictionary work. The most important result of this debate was the publication in Milan, from 1817 to 1824, with an appendix in 1826, of the *Proposta di alcune correzioni ed aggiunte al Vocabolario della Crusca* (‘Suggested Corrections and Additions to the Crusca Dictionary’), the fruit of the collaboration of a team of scholars guided by the classicist, Vincenzo Monti. Their intention was to underline the inadequacy of the Crusca Dictionary, considered not only to be outdated, but also characterised by too rigid and restrictive a view of language. A few years later, with his *Voci e maniere di dire italiane additate a’ futuri vocabolaristi* (Milan, 1838–1840), Giovanni Gherardini, from Milan, distanced himself from the Cruscan ethic, accepting the canon of authors, but confirming the need to take into account the living, spoken language.

These controversies gave Italian lexicography a salutary jolt: the nineteenth century saw dictionaries flourish in the country as never before. For many, the Crusca remained a starting point, even if they often deviated from its criteria quite substantially. This is the case for the *Vocabolario universale italiano* or *Universal Italian Dictionary* from the Tramater printing house, published in Naples from 1829 to 1840. It contains a very rich word base, including many more technical and scientific terms than the Crusca. The most important innovation, however, is the abandonment of pre-scientific definition, which had been seen not only in the Crusca, (see examples (2)–(5)) but also in Alberti’s Dictionary, and in Cesari’s Crusca veronese. In Tramater’s dictionary, for example, *cane* ‘dog’ is not defined as ‘a well-known animal’, but in a more precise and accurate way: ‘Species of domestic mammal belonging to the carnivore family etc.’ Terms in fields such as zoology and biology bear witness to the amount of care taken to create a proper definition, such as the following example:

(7) **Macaone. (Zool.) Sm. Specie d’insetti del genere farfalla, e propriamente de’ papiglioni, dell’ordine de’ lepidotteri, che abita sopra diverse piante ombrellifere e sulla ruta; ha le ali caudate, gialle da ambe le parti, egualmente colorite con un contorno bajo-bruno, e con istrisce lunate gialle; all’angolo della coda e delle ali, evvi una striscia rossa di fuoco. Il
suo nome deriva dall’essere stata questa bella specie dedicata a Macaone. Lat. *papilio machaon* (‘SWALLOWTAIL (Zool.) N. m.: Species of insect of the butterfly type, or rather papilionidae, of the lepidoptera order, that lives on Umbelliferae; it has yellow, tailed wings with a dark brown outline, and yellow crescent shapes; where the wings join the body, there is a red line. The name comes from the dedication of this beautiful species to Machaon. Lat. *Papilio machaon*’)

This change in method is the solution to a need for compilers to compensate for a serious defect of the Cruscan tradition, and editors of Tramater openly comment on their perplexity, in the preface to this dictionary: ‘we did intend to remove the Academy’s definitions when they are not definitions; […] well-known herb, well-known fruit, type of stone, such expressions tell us nothing’. (Tramater 1829: XIII). Example (7) shows how Tramater was one of the first to give grammatical indications (Nm for Noun, masculine) and field (Zool. for Zoology). The great stride forward in the quality of Italian lexicography must be noted, as Marazzini (2002: 387) remarks: ‘The Tramater dictionary became the best available on the Italian market, because of its wealth of information and readiness to innovate, and it would have retained primacy, had it not been surpassed by the Tommaseo-Bellini, which sealed the fate of all its competitors’.

1.4. *Nicolo Tommaseo’s Dictionary*

As Serianni points out (1990: 69), only two historical dictionaries were produced in Italy after it became a united nation (1861), the fifth, incomplete edition of the Crusca Dictionary, which has already been discussed here, and Nicolo Tommaseo and Bernardo Bellini’s *Dizionario della lingua italiana*. Before examining this great work in more detail, it is useful to recall that the post-union period saw the birth of the first dictionaries of Italian intended not only for scholars and literary men, but also for ordinary people, based on modern language, used in everyday communication. The *Novo vocabolario della lingua italiana secondo l’uso di Firenze* by Giorgini and Broglio (completed in 1897), was a prime example, compiled according to new criteria, in complete contrast with the Crusca principles, and inspired by the linguistic ideas of the novelist Alessandro Manzoni, who died in 1873 and therefore did not see his ideas put into practice. Manzoni’s model was the *Dictionnaire de l’Académie française*, and one of the most innovative characteristics of the Giorgini-Broglio dictionary was the absence of quotations from authors. Almost all the examples are in fact anonymous phrases from everyday life, or sayings, or proverbs; archaisms are almost completely banished. Giorgini-Broglio dictionary was not a great success. Two other much more successful works were compiled on similar principles, but were less of a break with previous lexicographical tradition: they were the *Vocabolario della lingua parlata* by Rigutini and Fanfani (1875) and the *Nòvo dizionàrio universale della lingua italiana* by Policarpo Petrocchi (1887–1891).
The moment has now come to examine the real masterpiece in nineteenth century Italian lexicography, Tommaseo’s dictionary, also known as the Tommaseo-Bellini, as, although it was started by Tommaseo, it was completed by Bellini. The work was encouraged by Pomba, the publisher, who was based in Turin, not Florence, and work on it began in 1857. The first part was published by the UTET (Turin Editions Typographic Union) in 1861, in the same year that Italian unity was accomplished, a significant coincidence. The edition was completed by 1879. This dictionary was of far superior quality to others in the nineteenth century, and is rightly considered to be the first true historical dictionary of Italian. One of its main qualities is its ability to reconcile synchronic and diachronic aspects of language: although it gives priority to modern usage, ample space is devoted to past usage, documented by quotations from authors. It was a veritable milestone in Italian lexicography, in which ‘literary tradition, technico-scientific terms, and everyday language achieve a balance never before attained’ (Della Valle 1993: 82). Amongst the numerous innovations of the work, it is worth mentioning the abandonment of too rigid a distinction between archaic and modern words, following a conviction that the former may well come back into use, possibly with a different meaning, at the hands of writers, and also a completely new organisation for definitions. Different senses of a word were now organised ‘following logical criteria, from the most common and universal meaning, with a hierarchical structure for different senses, numbered individually’ (Marazzini 2002: 388). For all its indisputable qualities, Tommaseo-Bellini is not without flaws and incongruence, as many scholars have noted. Apart from the order in which certain definitions are given, etymology is not always faithfully recorded, and it can be difficult to decode the numerous abbreviations, not all of which are listed in the table. Finally, there is Tommaseo’s much discussed ‘subjectivity’: he is a ‘lexicographer who more than anything else was passionate about his own work’ (Marazzini 2002: 388). Many words are coded ‘T’, for Tommaseo, indicating that what follows is his own personal opinion on the subject. This is clear in the fields of politics, civics and literature, as the following examples show. They are also coded with two crosses, to mark dead words, or those whose use is inadvisable:

(8) †† COMUNISMO. [T.] S.m. Istituzione sociale, o piuttosto Sogno d’istituzione, in cui i beni materiali fossero tutti ugualmente distribuiti ad arbitrio de’ capi della società. Parola e idea esotica (‘COMMUNISM. [T.] N.m. A social institution, or rather a dream of an institution, where material assets will all be equally distributed in accordance with the wishes of the leaders of society. Exotic word and idea’)

(9) †† SOCIALISMO. [T.] S.m. Il solito ismo di mal augurio: ci viene di Fr.; e con questo nome denotasi la dottrina (se dottrina è, non appetito nè sogno) che propone un che tra il Comunismo e il Despotismo, cioè di rimpastare tutta la società civile senza norme morali determinate,
rompendo tutte le tradizioni e le consuetudini [...] (‘SOCIALISM. [T.] N.m. The only ‘ism’ which augurs ill: from Fr.; this term designates the doctrine (if there is a doctrine, rather than a hunger or dream) between Communism and Despotism, which seeks to reorganise society without any fixed moral code, abolishing all traditions and customs [...]’)

(10) †† PROCOMBERE. [T.] V.n. Cadere dinnanzi o Cadere per, dal lat. pretto, l’adopta un verseggiatore moderno, che per la patria diceva di voler incontrare la morte: Procomberò. Non avend’egli dato saggio di saper neanco sostenere virilmente i dolori, la bravata appare non essere che rettorica pedanteria (‘SUCCUMB. [T.] V.n. to fall before or to fall for, from Lat., used by a modern poet, who said he wanted to meet death for his country: ‘Procomberò’. Since he did not prove to be able to endure pain in a manly way, such bragging appears to be nothing more than pedantic rhetoric’)

As is seen here, Tommaseo’s irony is far from veiled: while (8) and (9) show his aversion for both ideologies, example (10) is an accusation implicitly addressed to Leopardi, although he is unnamed, for his use of the word procomberò in the song All Italia (‘To Italy’).

The Tommaseo-Bellini remains a major work, still of use today, thanks to the work of an intellectual of such stature as Nicolo Tommaseo, who dedicated himself to the dictionary until his death, drawing upon his own vast cultural knowledge, and his reference cards, which were of greater precision than ever before. The value of his work is demonstrated by the fact that a century after its publication, another lexicographical enterprise of similar ambition began, leading to the creation of the greatest historical and general dictionary of the Italian language: the GDLI.

2. An author-based dictionary

2.1 The Grande Dizionario della Lingua Italiana Great Dictionary of the Italian Language

The Grande dizionario della lingua italiana (GDLI), published in Turin, by the Utet publishing house, from 1961 onwards, is also called the ‘Battaglia’ after its founder. It is the largest and most important historical dictionary of Italian: it is composed of several volumes, with a total of 22,504 pages, and is the fruit of over forty years of labour (the final volume was published in 2002) and the collaboration of numerous scholars, first among whom was Salvatore Battaglia, who was replaced when he died by Giorgio Barberi Squarotti. The 2004 Supplement was recently published, under the direction of Edoardo Sanguineti, containing all the new words that have appeared in the years since the first volumes were published, bringing to 23,277 the total number of pages. The size
and quality of the work make it a reference point for any student of Italian language and literature. Its role is even more important when the GDLI is considered not only as an exceptional tool for reference and study, but also as a way of reaching a better understanding of the culture rooted in the past, from the earliest centuries of the birth and development of Italian as a language.

2.2. The structure of the GDLI

By selecting two words as examples to illustrate the structure of the dictionary, it will soon become clear why the GDLI has acquired such a prominent place, without precedent in Italian lexicographical tradition, and why it has become, over the years, an indispensable instrument both for the linguist and for the student of literature. The two examples are taken respectively from the first and the last volume (XXI) because, with a work covering so broad a span (more than forty years), it is necessary to demonstrate the inevitable changes and adjustments which have been introduced over the course of time.

The first example is the word *Amore* (‘Love’). After grammatical information (sm = sastantivo maschile ‘N m = Noun, masculine’), the first definition is given, ‘Intense effect which strives to possess its object and to join with it, so that it seeks to preserve it and endeavour to do good to it’, followed by 20 quotations from 16 authors and works belonging to Italian literary tradition, from Dante to Luzi, that is to say from the fourteenth century to the early twentieth century. Authors and works are indicated with abbreviations, the key to which is given in a separate volume accompanying the work (*Indice degli autori citati* or *Index to authors and quotations*). There follows, under the first definition, a more restricted meaning based on ancient usage, ‘Ant. Instinct (the scholarly distinction between instinctive or natural love, common to all animals, unconscious and outside moral law, and love by choice, found only in humans, the expression of intelligence and will, and as such good or evil [...]’), also illustrated by 5 quotations from authors (from Dante and Savonarola). A similar structure is found in the following definitions, numbered from 2 to 19, covering a total of 15 printed columns. As an example, meaning number two is defined thus, ‘Attraction towards a person of the opposite sex, concentrating interests, thoughts and actions around it (and in it might prevail either the physical or the spiritual)’, with four other specifications, confirmed by 133 quotations, from 81 authors and works, from the works of Giacomo da Lentini to Cesare Pavese. Meanings are organised according to a criterion of ‘currentness’ best explained in the words of the initiator and founder of the dictionary, Salvatore Battaglia, in *Criteri di lavoro* in 1951: ‘organising the various meanings of a word [...] from the most usual and frequent, as it is the first to spring to mind for present-day speakers, and rising (or better still, falling) through the less common meanings, more specific, more limited in meaning, to end with the ancient, outdated, historical uses’ (Tommaseo-Battaglia 2002: 72). As can be seen, these words recall the methods of the Tommaseo-Bellini, although the coherence is
much greater in the GDLI. At the end of the definitions, after the final meaning and related examples, comes the etymological derivation.

The veritable fulcrum of the dictionary is the massive presence of text quotations from authors. To this end, Salvatore Battaglia, before starting on his great work, wrote: ‘Our Dictionary will be supported by a broad and continuous documentation of ‘examples’: all from authors. It was in Italy that the modern lexicographical tradition of ‘quotations’ was first introduced. The value of Tommaseo-Bellini is chiefly in the wealth and variety of its quotations. The Dictionary that we are starting to compile will have the same quality: it will be as rich as its collection of quotations. […] A reader will feel the value of a word, of a meaning, of a saying, if he sees it in use in an authentic and unique expression. As Italian has a homogenous literary tradition, and is present from the first centuries in a similar phonetic and morphological form to that of the present day, it is possible to introduce author quotations from the very beginning of our literature’ (Tommaseo-Battaglia 2002: 72). At the same time, however, Battaglia also sought everyday references, in order to ‘bring Italian lexicographical tradition up to date’, and to satisfy the needs of a vast category of potential dictionary users: ‘Readers of newspapers or reviews quite logically expect to find in a modern dictionary any word to be met with in their daily reading: from sport to legal terminology, from economics to politics, from mechanics to pure science’ (Tommaseo-Battaglia 2002: 73–74). Despite the fact that the GDLI is based to a great extent on literary references, from the very first volumes, the reference corpus did not only contain literary texts, but also many other documents, such as registers, accounts and dictionaries.

The second example, Viaggio (‘voyage’), is taken from the last volume, and contains nine main senses, with 112 quotations. Although the basic structure of the definition is the same (grammar specifications; main meaning; examples of the main meaning; other numbered meanings with examples if possible; etymology), a more thorough examination of the references to works quoted reveals the most relevant change since the first volume: the reference corpus has grown enormously (Marazzini 2002: 59), whether through the extension of literary reference (to authors who were already included from the start, and to newly added authors) or through the addition of texts of other types, such as modern newspapers (the Corriere della Sera is quoted as an example under meaning number six, to document the use of the word amongst drug addicts) or ‘the grey literature produced by companies and businesses’ (De Pasquale 2004: 300).

It can be reliably affirmed that the reference corpus of today is about four times bigger than it was for the first volume, a testimonial to the unceasing work of modernisation and amplification carried out by the different project directors over the years.
The GDLI, whilst remaining in every sense a historical dictionary, also allows quite effective searches for everyday language, chiefly because of the noteworthy increase in modern and contemporary content in the corpus. An example of this may be seen when examining the word *Virus*, whose fourth and final definition, gleaned from the field of computing, is explained without reference to texts: ‘Computer Science. Elements of instruction which, either directly introduced, or hidden inside an apparently innocuous programme, are intended to damage memory data by copying.’

Inevitably, documentation of contemporary language is destined to become outdated with the passage of time, as the least recent volumes show, although taken individually they stand as a historical survey of the forty years between the publication of the first volume and of the last one. Only the continuous publication of volumes of updates, like the 2004 Supplement, could overcome these limits, just as it would overcome another of the disadvantages unavoidable in a historical dictionary of this size, the lack of documentation of new words and terms of foreign origin recently imported into Italian which begin with the first letters of the alphabet. For example, *computer*, now readily accepted in Italian, is not present in the third volume of the GDLI, published in 1964, when this word was as yet unknown in Italy. To find the definition of this word, it is necessary to consult the recent 2004 Supplement, which records as the first attested use the *Corriere della Sera* of January 29, 1966, which agrees with the date indicated by the GRADIT and antedates the use registered in the DELI (1968, in the magazine *Domenica del Corriere*).

When considering these limits, it might be objected that such goals are beyond the scope of a historical dictionary, which is not intended to record contemporary usage (for which purpose it is necessary to use an instrument like the GRADIT), but rather to document the evolution and the use of Italian over the centuries, in all periods of history. It is therefore worthwhile trying to understand what the real goals were at the outset, by examining what the founder had to say on the subject. The GDLI, as it has already been pointed out, is today unanimously recognised as the foremost historical dictionary of Italian; it may come as a surprise therefore that in 1951, Salvatore Battaglia wrote these words: ‘We should not forget that our Dictionary seeks not to be an archive of language, but a living documentary of language as it is spoken and written, to express contemporary reality’ (Tommaseo-Battaglia 2002: 72). The primary objective appears therefore to coincide with that of a general dictionary, which is to show the current state of the language, present and contemporaneous.

Further on, Battaglia expressed his position even more clearly (see Tommaseo-Battaglia 2002: 73): historical documentation appears as a consequence, a type of corollary to the first objective, which is still to bear witness to living use of the language. With hindsight, almost half a century after these words, it may be asserted that, with the passage of time and the lengthy work of completing the dictionary, a natural and inevitable transfer has occurred, so that the secondary
has become the prime objective, and vice versa. However, the publication of the 2004 Supplement has restored to some extent the function of documenting the most recent and innovative use of language, although it inevitably lessens the ease of use of the dictionary.

2.4. A complete dictionary

The GD LI represents a deliberately unattainable end, but will nevertheless remain a milestone for the future, in keeping with the view of lexicography as a continuum without end. The GD LI is not without its flaws, ‘the too literary style, the uncertain philological system, […] and a lack of sureness in the classification of examples according to the different word senses’ (Mengaldo 1994: 26), but it cannot be denied that it covers almost every aspect of Italian lexicography: everyday and literary language, archaic words (e.g. sagitta, teristro), dialect and regional language (e.g. zurlare, zabetta), interjections (e.g. titic-titoc), vulgarity (e.g. cazzo), specialised language (e.g. telereferendum for journalism, telequiz for television, western for the cinema, trequartista and tie-break for sport), technical and scientific language (e.g. glutamminico, oculocompressore, zygadite), neologisms (e.g. tangentopoli, wojtilismo) and foreign words (e.g. ticket, zaruk) uniting in a single work the progress and characteristics of preceding lexicographical tradition.

3. From paper to the web: TLIO

3.1 Historical outline

The Tesoro della Lingua Italiana delle Origini (TLIO) is a historical dictionary of old Italian (from its origins until the end of the fourteenth century) that is currently being developed by the Opera del Vocabolario Italiano (OVI, ‘The Italian Dictionary’), a CNR National Research Council Institute with headquarters at the seat of the Accademia della Crusca, in Florence. The project began in the sixties. The Crusca Academy, after the suspension of the fifth edition of the Dizionario degli Academici by the Italian government in 1923, proposed a completely new dictionary project, on which work started in 1965, thanks to funding from the CNR. The work, scheduled for completion in 2021 (the seven hundredth anniversary of Dante’s death), was intended to be composed of two parts: an exhaustive Treasure of the ancient language up to 1375 (the death of Boccaccio), along the lines of the Thesaurus Linguae Latinae, and a Historical Dictionary from 1375 to the present, presumably modelled on the Oxford English Dictionary (but no detailed editorial project was produced for either the Tesoro or the Dizionario storico)?

It was decided to use computer technology, creating an electronic version of all Italian texts prior to 1375 and a selection of those after that date. At the time, this was a courageous and far-sighted decision. From this project, for the
Tesoro, there is a digital record of 14 million partially lemmatised occurrences, which after a long and painstaking procedure of conversion to a more modern system, became in 1995 part of a database of ancient Italian which today contains 21 million occurrences.

In 1972 it was decided to continue work on the Tesoro alone, postponing work on the dictionary of language through successive periods until after its completion. In 1985, after a period of difficulty, the work group was transferred to the CNR and the ‘Centro di studi’ OVI was created, which has now become an Institute. Until the end of 1995, work was exclusively devoted to preparation of the database of ancient Italian texts necessary for the creation of the TLIO. Work on individual entries began in 1996. As of June 2004, some 12,000 entries have been compiled, of the 45,000 for the completed work. From 1999 onwards, roughly 2,000 entries have been written every year.

3.2 The organisation of the TLIO

The TLIO is an Internet-based dictionary (cf. below par. 3.2.1) which lies balanced, or suspended, somewhere between a language dictionary and a corpus dictionary; in some ways it is a corpus dictionary, a glossary of ancient Italian texts, with a philological bias, but with a method which seeks to describe ancient Italian lexis other than simply by text commentary.

The TLIO has adopted as far as possible the system of definition of a monolingual dictionary; definitions are not translations, from ancient to modern Italian, but are definitions in there own right whenever possible. As an example, consider a complex verb like aprire open:

1. Rendere possibile il passaggio (attraverso un luogo) o l’uscita (di qsa da un contenitore) agendo su ciò che ha la funzione di impedirlo (porta, serratura, coperchio) (‘to enable the passage (through a place) or the release (of something in a container) by acting on whatever prevented this (a door, a lock, a cover)’).
2. Rendere accessibile (un luogo nel quale altrimenti non si può entrare) (‘to make accessible (a place where otherwise entrance was impossible)’)
3. Creare un varco, una spaccatura, una fessura in qsa di coerente (‘to create a passage, a split, a fissure in something coherent’)
4. Portare in posizione distesa (un corpo o un organo ripiegato); allargare (le braccia) (‘to put into an extended position (a body or folded object); to spread wide (arms)’)

Although it is a historical dictionary, the TLIO does not give definitions in chronological order of attested meaning, but rather seeks to show the relationships between meanings (moving from the more general to the more specific, and from the literal to the figurative). Another reason justifying this choice is that attested chronology is not necessarily the true chronology of the appearance of the word and its meaning in the language; if this were the case,
documentation would have to include unpublished works, and surveys would have to be performed on mid-Latin texts too, as many words appeared for the first time in such texts: this would be extremely time-consuming (sometimes such notes on mid-Latin attestations are included, consulting works like Pär Larson's *Glossario Diplomatico Toscano avanti il 1200*). It is possible to deduce documented chronology not only for the very first documented use of a word (section 0.3 of the entry), and for the first documented use in regional varieties of language (section 0.4), but also that of all the different meanings, as the modern definition is followed by the oldest example of the word in that sense.

An idea of the philological position of the TLIO, as a corpus dictionary, can be found in the fact that editors have to ask themselves if a particular detail might impinge upon the interpretation of the evidence; in the case of texts translated from Latin or French, which make up a sizeable part of the corpus, the presentation of an unclear example under a definition is justified by giving the quotation from the original. Furthermore, in entries for which there are also examples where the meaning is not understood, if it appears that the meaning differs from those defined for the other examples (all the more so if there is only one piece of evidence), the standard definition adopted to allow an automatic search (with the possibility thereby of obtaining suggestions from readers) is 'unconfirmed meaning'. For example, it has not been established what a *curiosa* might be, from the only occurrence in the corpus, in *Capitoli dei Disciplinati di Santa Caterina di Città di Castello* (*Testi trecenteschi di Città di Castello e del contado*, ed. by Francesco Agostini, Florence, Acc. della Crusca, 1978, pp. 138–57):

[CURIOSA] s.f.

[.....]

1 Signif. non accertato (sorta di edificio?) ('unconfirmed meaning (type of building?)')

[1] Stat. castell., XIV sm., pag. 151.34: E che lla detta fratenita non possa mai acquistare possesion fore del luogo nostro, né hedificare churiose e alt(r)e case se non solamente a capacitá e a bastança di conpanni, con ciò sia cosa che ‘l nostro Singnore (Gesù) (Cristo) descendesse della gloria de Dio Padre, volse a sé despensare la pove(r)tà ('And that the fraternity may not acquire possessions outside our place, nor build churiose and other houses save to shelter members of the fraternity, for Our Lord Jesus Christ came down from the glory of God the Father, to live in poverty')

And there is also an example for the word *cova* which cannot be understood from the given meanings (in the *Carte della divisione della compagnia di Bernardino Ugolini*, in Mario Chiaudano, *Studi e documenti per la storia del Diritto commerciale italiano nel sec. XIII*, Turin, Istituto Giuridico
dell’Università, 1930, pp. 95–113); it may also be noted that for the third meaning, doubt is expressed by a question mark:

[COVA] (1) s.f.

1 Contenitore a doghe di legno per la conservazione del vino (‘Container made of wood for wine conservation’)

2 Misura di quantità per materiali coloranti (‘Measure of quantity for colour dyes’)

3 Fras. *Essere messo in cova*: essere rinchiuso, essere incarcerato? (‘Idiom: *Essere messo in cova*: ‘to be put in a barrel’; to be locked up?’)

4 Signif. non accertato (‘unconfirmed meaning’)

[1] ? Doc. sen., 1281-82, pag. 95.11: [Item] XVIJ lib. li arnesi de la butigha ciò deschi et soprese et soprano et la chasa et un leto fornito cho leçuola ** et avaci di su le soprese et chove de pani et d’altri arnesi che so’ ne la butigha. || Essendo i *pani* del ms. chiaramente dei ‘panni’, è possibile che la parola *chove* sia da collegare con *coverta* piuttosto che con *cova* || As *pani* in the manuscript clearly means ‘panni’ clothes, it is possible that *chove* should be linked to *coverta* rather than to *cova*.

Examples are chosen according to linguistic rather than historico-literary criteria. If the date or linguistic origine ar the same, editors are invited to favour non-literary examples, from legal documents, statutes, accounts, etc., which are more reliable as linguistic evidence, and in general less often referred to in lexicographical tradition. The TLIO, beyond its quality as a dictionary, and in spite of an unavoidable provisional nature, displays in a simple manner an ample harvest of interpreted examples of language in use, most of which were not present in lexicographic works produced until now, thus furnishing a service that a databank could not provide.

3.2.1 TLIO: a dictionary published via Internet TLIO is published via Internet as a work in progress, at the address www.vocabolario.org (the OVI site) or at http://tlio.ovi.cnr.it/TLIO (first page of the dictionary); only samples are published in printed form, in the annual bulletin of the OVI. This is a structural
characteristic of the TLIO: it is not written to be printed and then transferred to Internet.

The primary consequence of this organisation is that entries do not need to be dealt with in strictly alphabetical order, although most work for the moment has been done on the first letters of the alphabet. The Internet publication by the end of June 2004, of 12,035 words, contained 5006 for the letter A, 1691 for B, 4238 for C, 506 for D and 594 for E to Z; A and B are almost completed (a hundred or so words remain).

In the second place, unlike a printed work, or even a CD publication, an Internet-based dictionary is always under the control of the editors, and therefore entries can be revised or updated at any time; new entries can be added, and some entries can be replaced. In addition to error correction (which can be done without publishing an appendix of corrections, as would be the case for a printed work), updating depends on the database used by the editors; which can be enriched by newly published texts, or new editions of texts already included in the database, with fewer errors than the previous editions.

It is also possible to focus on certain aspects of some entries, leaving other aspects to be dealt with at a later date; at this phase in the creation of the TLIO, a complete list of the meanings of all the words in the reference corpus, with indications about chronology, is being compiled, but observations about syntax, derivation etc. are not systematically included for the time being. The structure of the dictionary, which gives this type of information in a separate paragraph before the definition (see below) means that, during a later phase, such information can be added quite easily for each word.

3.2.2. Dictionary and database

A particularly important characteristic of web publishing is the fact that the dictionary can be connected to a searchable copy of the database used by the compilers. It is thus possible to compile the entries from a selection of examples, as is usual in a dictionary, and at the same time allow readers to consult the entire bank of documentation on which the compilation is based. To this end, at the beginning of a entry is included a list of all the word forms present in the corpus that are lemmatised forms of that headword; the system allows a one-touch advanced search for all forms together, and the presentation in chronological order of all the contexts present in the corpus. For example, when finding the entry digiunare (the verb ‘fast’), the 1055 examples in the corpus which include any form of digiunare can also be displayed (reference here is made to the Internet version as of June 2004), complete with bibliographical indications and with the option to read the entire page of the source text, where the quotation is found. As the search is by form, homographs will also be found, which the compiler discarded when preparing the entry, such as digiuno N.m., ‘fast’, which is a homograph of digiuno ‘I fast’ (but it is also possible that, in this way, errors - made by the compiler when discarding homographs – may be discovered)
3.2.3. TJO is a dictionary of all the linguistic variety of medieval Italy. Separating the dictionary of the origins from that of the following periods may give a more faithful picture of the linguistic situation in ancient times. As has already been stated, the national Italian language was only codified in the sixteenth century, chiefly based on the Florentine of the major authors of the fourteenth century; yet, even if Florentine and to some extent Tuscan had an unassailable position in the fourteenth century, it would not be correct, for the first centuries, to oppose dictionary language (Florentine and Tuscan) and excluded dialects. As a result, the TLIO treats all varieties of Italian language on an equal footing, as they can not yet be called dialects at this period. For example, the first five examples given for the first meaning of the verb *digiunare* ‘to abstain (generally for religious or medical reasons) from the absorption of food’, are, in the order presented, from a text from the Marches (the *Ritmo su Sant’Alessio*, second half of the twelfth century), from a Roman dialect text conserved in a Tuscan manuscript (the *Miracole de Roma*, around the second half of the thirteenth century), from a Lombard text (Pietro da Bescapè’s *Sermone*, 1274), a Milanese text (Bonvesin da la Riva, third from last decade of the thirteenth century), and from a Florentine text (Bono Giamboni’s *Libro dé Vizi e delle Virtudi*, before 1292).

It can be noted that the Battaglia dictionary quotes examples from non-Tuscan texts together with Tuscan texts, for example, for the word *digiunare*, the *Ritmo su Sant’Alessio* (from the Marches) and Iacopone da Todi (from Umbria). Such examples are used only rarely, however, and without specific linguistic interest. When quotation from such texts makes it necessary to indicate variants for the headword, they are qualified as ‘ancient’: this is the case for *deiunare* and *diiunare*, which in the TLIO corpus are only found in non-Tuscan texts. The compilers of the TLIO are on the contrary invited to ensure that all the linguistic variety of the origins is represented in the best possible way in the entries in the dictionary.

The bibliographic abbreviations referring to the texts are created in such a way that both the date and the linguistic origin of the text are immediately clear. In the first section for each headword, is given the earliest attested use of the word (section 0.3) then the first attested use for each variety of Italian (section 0.4): to this end are quoted texts whose linguistic origin is well defined. All the texts quoted in section 0.4 must appear at least once in a definition. Under each definition, the selection of examples is organised without distinction of linguistic variety, in chronological order, from the oldest reference for the meaning in question.

3.2.4 TJO is a dictionary compiled first hand. The main source for the TLIO is the database of ancient Italian, which contains almost all published Italian texts written up to and including the fourteenth century, in whatever variety of Italian language, at least all those that can be considered to be of relevance. The list of headwords is taken from the analysis of the database, chiefly using
lemmatisation, but also sometimes searching by word form. Occasionally one or two words may be added that have been found in dictionaries, and that are documented by texts not in the corpus, or not yet included in it, along with attestations that are documented by dictionaries and not present in the corpus, in which case they are always rigourously marked as being ‘outside the corpus’.

To compile definitions, the dictionary writers examine all the documentation extracted from the database, without any preselection. This method is perfect for low or relatively low frequency words, (less than a few hundred occurrences), but is very labour intensive for high frequency words (a thousand or ten thousand occurrences). Obviously, in the second case, compilers have to adopt a strategy of simplification, but to do this from all examples is more efficient than to work from a selection taken from every single text, but without a comprehensive idea of the whole.

3.3 Some aspects of the macrostructure

The TLIO has adopted a rigidly one-word system of entries (with some exceptions). Expressions for which there exists a single word variant (codified in dictionaries and more prevalent in modern language) are treated under that entry: for example, a pena (adverbial) appears under appena (adverb); a ciò che (conjunctive expression) appears under acciocché (conjunction). In other cases, they are included under a headword contained in the expression: for example adesso che (conjunctive expression) is found under adesso (adverb or conjunction); per mor che, which was in use in Lombardy and Venice with a final or concessive value, is placed under amore N. m. (but permordè becomes a headword, because it appears in prepositional expression permordè de and permordezò likewise becomes a headword, a conjunction with adversative value).

As a general rule, every headword corresponds to a distinct etymology; for example, abituato, adjective from abito ‘dress’ ‘wearing a dress’ is differentiated from abitato, adjective from Late Latin habituari, ‘to get accustomed to doing something’; affettare, verb from fetta, ‘to cut into slices’, is differentiated from affettare, from the Latin affectare, ‘to desire, to want with passion’. There is of course a more complex set of rules, dealt with by compilation norms. A more interesting point, for this paper, is the fact that, in a dictionary built on a plurilingual corpus, the selection of the appropriate form for the headword may pose some problems. For example, in the documentation of the TLIO, the word digiunare is attested, considering only the form for the infinitive, and disregarding other variant spellings: degiunare, deiuinare, deiuinare, dezuinare, dezunare, diggiunare, digiunare, digiunare, diiunare, diiunare, gegiunare, gigiunare, ieuinare, zazunar, zezinare, and zizinare. By this choice, the principle adopted is that the TLIO is the section devoted to ancient language of a future general historical dictionary, which is why the preferred headword is always the form used in modern language, if it can be
found in the ancient documents at least once. If the modern form is not attested, the ancient form is chosen as headword, and links from the modern forms are added, based on dictionaries, that readers might search for. In each case, the web-based dictionary allows searches for any form or spelling, through which process, from any ancient Italian text, it is possible to discover if a word form has been covered, and under which headword.

The TLIO differentiates between masculine and feminine; for example, arciera ‘female archer’ is treated as a separate word from arciere s.m. ‘an archer’, asina s.f. ‘female ass’ is treated as a separate word from asino s.m ‘ass’. Some exceptions are made for cases where the same inanimate object is referred to, which may be either masculine or feminine, such as calle, N.m. / N.f. meaning ‘lane, path’.

Notes
1 Sections 1 and 2 are by Simone Fornara; section 3 and appendix by Pietro Beltrami. Carmela Château translated the original Italian text.
2 Statistical comparison between the first booklets containing bibliographical references, and those which gradually completed and brought them up to date, published during work on the dictionary, ending with their recent publication in one comprehensive volume, provides concrete evidence of the gradual growth of the corpus of references: from a 56 page booklet from 1961, for the first volume, through 65 pages (1962, volumes I-II), 80 (1964, volumes I-III), 100 (1966, volumes I-IV), 111 (1968, volumes I-V), 118 (1970, volumes I-VI), 119 (1972, volumes I-VIII), 121 (1973, volumes I-VIII) and 123 pages (1975, volumes I-IX), to which must be added the 27 pages of the 1984 booklet (authors added in volumes X, XI, XII), the 26 of the 1996 booklet (authors added from volumes XIII to XVIII), the 10 of the 1998 booklet (authors added in volume XIX) and the 10 of the 2002 booklet (authors added in volume XX and XXI), for a sum total of 196 pages. The combined index (2004), which also contains all the texts and authors appearing for the first time in the 2004 Supplement, is 298 pages long, but a bigger font was used in the printing.
3 The procedure, which has been partially completed, was to produce concordances for all the tokens in the texts; to lemmatise all the tokens, but to make a selection of occurrences, in order to obtain a set of examples; to produce ‘context cards’ for each entry, containing the examples taken from each text. Before a text was entered, its philological aspect was inspected, by critical examination of its edition; in many cases, above all for documentary texts, the original manuscripts were checked.
4 The idea of compiling a dictionary covering periods from the fifteenth to the twentieth century has not been abandoned, but lack of resources means that work must focus exclusively on the TLIO for the time being.

References
A. Dictionaries


Alunno, F. 1548. La fabrica del mondo. Venezia: Nicolò de Bascarinì.


Gigli, G. 1717. Vocabolario cateriniano.


Luna, F. 1536. Vocabolario di cinquemila vocaboli toscani non men oscuri che utili e necessari del Furioso, Boccaccio, Petrarca e Dante novamente dechiarati e raccolti da Fabrizio Luna. Napoli: Giovanni Sulzbach.


Italian historical dictionaries: from the Accademia della Crusca to the web


B. Other literature


382 Pietro G. Beltrami and Simone Fornara


Appendix - The entry *digiunare*.

The example, *digiunare* v[erb], is presented here in abbreviated form: only a few examples are given, and only part of the list of forms in section 0.1. The sections that precede the definition are as follows:

0.1 - list of variant spellings (see above).
0.3 - first reference in the corpus (bibliographical abbreviations refer to the bibliography published in Internet); the number in bold type indicates the definition under which the text is quoted.
0.4 - first reference by linguistic origin, divided into the following areas: Tuscan, Northern, Middle and Southern, Sicilian. All texts mentioned appear at least in one example under the definition.
0.5 - note on syntax (constr.: construction); final section, alphabetical list of locutions and idioms (if any).
0.6 - O - synonyms and antonyms; D - derivatives.
0.7 - recapitulation of definitions.
0.8 - compiling team and date of delivery (gives the status of the corpus at the time of writing; after revision, the date is updated only if the corpus has been updated again).

**DIGIUNARE v.**

0.1 *çaçuna, çaçunà’, çaçunam, çaçunar, çaçunase*, [+ 116 forms].
0.2 DEI s.v. *diginare* (lat. *desjeuner*). || Meaning 3 is a calque from Ancient French *desjeuner*.
0.3 Ritmo S. Alessio, XII sm. (march.): 1.


In Sicilian texts: Giovanni Campulu, 1302/37 (mess.); Stat. palerm., 1343; Stat. catan., c. 1344.


digiunare di: «non si vuole digiunare pure del cibo», Paolo da Certaldo, XIV sm. (tosc.).

digiunare pron.: «se tu vòi intender como / se zazuna lo rico omo», Anonimo Genovese (ed. Cocito), a. 1311.

Idiom.: *Digiunare in pane e acqua* 1.1.

1 O giunare.

D *digiunamento*, *digiunatore*, *giunare*.

1 To abstain (gen. for religious or medical reasons) from the absorption of food. || See 1 [11].

2. Figurative. To abstain from something (particularly from a sin).

3 To have breakfast.


1 To abstain (gen. for religious or medical reasons) from the absorption of food.


[3] Pietro da Bescapè, 1274 (lomb.), 1895, pag. 64: Quando fo sí ferma la credança, / La pasca fen per alegrança; / Tri dí avevano \[\[ciçunao]/ Per lo segnor ki fo penao, / Ki no mangiaven ni bevevano / Per grande grameça k’ili avevano.


[5] Bono Giamboni, *Vizi e Virtudi*, a. 1292 (fior.), cap. 58, pag. 96.15: Or ecco bella gente che si trae innanzi a battaglia, che qual è magro e afflitto per troppo *digiunare*, agrestando il corpo di molta astinenza, qual è palido nel volto per troppo vegghiare, stando dì e notte in orazione.


- Referring to animals.

[36] Palladio volg., XIV pm. (tosc.), L. 4, cap. 24, pag. 166.2: E allora scegli un di chiaro e bello senza tempestade alcuna di vento, o d’acqua, o d’altro, e piglia i buoi, e mettigli di prima nella stalla. E se egli fossero troppo feroci e aspri, facciansi stare *digiunando* un die, e una notte legati bene…

- Figurative.

[37] Fiore, XIII u.q. (fior.), 34.12, pag. 70: Molto m’atten[n]le ben sua promessione, / Ma non di gioia né di nodrimento: / Ch’e’ di speranza mi dovea nodrire / Insin ched e’ mi desse giug[giamento]. / *Digiunar* me ne fece, a ver vo dire; / Ma davami gran pez[z]le di tormento, / Con salsa stemperata di languire.

[1] Bestiario toscano, XIII ex. (pis.), cap. 42, pag. 63.27: ché si noi desesmo tutto lo nostro avere per Dio a’ poveri, e tutto tempo *digiunassemo in pane e in acqua*, e diesessimo tutte l’ore del di e de la notte, e ricevessimo martirio sopre le nostre carne si como fece sancto Laurenço che fue arostito, si ce le potrebbe lo dimonio si avenenare agevilemente che a pena ce ne vedremo…

[2] Cronica deli imperadori, 1301 (venez.), pag. 218.11: In questo tempo, in lo territorio Tolese una fantulina de XII anni, da po ch’ell’ave recevudo la santa communion del prievevede in lo di de Pasqua, per se mese *pane e acqua decunando*, e dende luogo inanzi da ogni eibo e bevanda per tri anni se retene.

…

2 Fig(urative). To abstain from something (particularly from a sin).


…

3 To have breakfast.

[1] Tristano Cors., XIV ex. (ven.), pag. 75.42: Allo matino, cossì tosto com’ello apare l’alba, se leva meser Tristan e va oldir mesa en una capella de là dentro e puo faxe amantinente mandar per li tri compagnon, per ch’ell’i se *deçuni* algun pocho, ch’ell’i serano plu forti puo a dever cavalchare. E quando ell’i fono a lu venuti, el li faxe *deçunare* alquantu e puo se fano aportar lor arme; si s’armono tuti quatro.