

History of the Occitano-Gascon Lexicon

Hélène Carles, Martin Glessgen

DOI: 10.1093/acrefore/9780199384655.013.466

Summary and Keywords

The evolution of the Occitan lexicon between the 4th and the 20th centuries displays a particular character compared with that of the other Romance languages. From a geolinguistic point of view, the Occitano-Gascon dialect group is highly diversified, and the four large areas which constitute this domain (Provençal/Dauphinois, Alverno-Limousin, Languedocian and Gascon) must be considered as a composite group. The process of lexical differentiation in Occitan began under the Roman Empire, increasing from the 8th century onwards, and was further accentuated during the course of the second millennium. The varieties of Medieval Occitan were highly developed and remained pluricentric. The mechanisms of lexical innovation engendered by the development of the various textual traditions, as well as by intertextuality, caused the vocabulary to develop considerably between the 12th and the 15th centuries. During the modern period, from the 16th to the 19th century, the process of elaboration of written culture began to grind to a halt, although Occitan continued to be spoken throughout the territory. The traditional vocabulary thus continued to diversify, parallel to the development of regional literature and the constitution of significant lexical inventories. Thus, during the contemporary period, the diversification of the dialectal varieties of Occitan was highly developed after 15 centuries of lexical evolution. Simultaneously, use of the spoken variety diminished throughout the 20th century as a result of the pressure from standard French. The powerful revival movements of the 19th and 20th centuries were unable to halt the impending demise of the primary varieties of Occitan. This realization highlights the necessity of encouraging linguistic fieldwork, which would allow primary Occitan varieties to be recorded before their extinction, but also lexicological analysis with the object of highlighting the richness of dialectal varieties and the expressivity of contemporary literature.

Keywords: Occitan, Gascon, Latin, lexicon, dialects, linguistic elaboration, discourse traditions, onomasiology, philology

1 Definition of the Object of Study

1.1 General Overview

The vocabulary of Occitan has undergone a fundamentally endogenous evolution within the Romance languages. It largely consists of words inherited directly from the Latin lexicon and it is highly dialectalized. It thus comprises fewer substrate words than Sardinian, for example, with the possible exception of Gascon (cf. §2.3), fewer Germanic loanwords than French (ib.), fewer words of Arabic origin than Spanish (cf. Kiesler, 2006, p. 1653f.), and for the Middle Ages, fewer French loanwords than Italian. Particularly during the Middle Ages, however, Occitan was subject to a high degree of interference from Catalan and Aragonese, as well as from French from the 16th century onwards (cf. §1.2 and §4.1). In addition, contact between eastern and Alpine Provence and northern Italy has been uninterrupted (cf. Trotter, 2006, p. 1779f.). More surprisingly, Occitan contains fewer learned borrowings from Latin than does French, for example, or even Italian. Finally, as the elaboration of Occitan did not progress beyond the regional stage - neither during the Middle Ages nor in the modern period - and as the territory in which it is spoken is of a relatively mountainous nature, lexical diatopic variation is particularly pronounced.

A quantitative account of the evolution of Occitan vocabulary from its constitution during the 7th and 8th centuries up to the 21st century would necessitate the study of approx. 200 000 lexemes. This figure remains approximate (cf. §1.4): for the Middle Ages, the current dictionaries list approx. 33 000 (30 000 entries in the DOM-en-ligne, 23 000 of which belong to LvP; this figure is further increased by the contribution made by the DAGél, which, once completed, will contain 12 000 headwords, most of which will correspond to the headwords of the DOM). For the modern period, Mistral contains approx. 68 000 headwords excluding the many toponyms and anthroponyms - the lexicographical status of such forms is the object of frequent methodological debate. Thus, all in all, the Occitan lexicon may be said to contain between 75 000 and 80 000 lexemes, or 200 000 lexical units if polysemy is taken into account, since a lexical unit is defined as the combination of one word-form with one meaning (cf. Glessgen, 2011).

To analyze a corpus of this size is no easy task as research in the field in general is very patchy; this is particularly true in the case of the lexicon. Far fewer studies have been carried out on Occitan than on neighbouring Romance varieties, particularly since the 1980s which saw a sharp decline in the number of speakers, the exclusion of dialectology from the educational programme offered by French universities, as well as a general loss of interest in research in historical linguistics on a global scale. In France, until the end of the 1970s two-thirds of all third-cycle doctoral theses or state doctoral theses in linguistics were in historical linguistics or dialectology, and a quarter of these were on Occitan. Only 10% of all doctoral dissertations since the 1990s, however, have been written on topics relating to historical linguistics, and those on dialectology or Occitan are almost non-existent.

1.2 The Occitano-Gascon Dialect Group

From a dialectological point of view (cf. §4.2.1) it is necessary to distinguish

(1) the two main 'southern' varieties: Languedocian and Provençal, which represent the most highly-developed varieties both during the Middle Ages and today;

(2) the northern varieties: Auvergnat and Limousin (known collectively in scientific terms as Alverno-Limousin), and Dauphinois, which shares certain traits with Alverno-Limousin, but whose vocabulary is similar to that of Provençal (cf. §4.2.1). Dauphinois also displays genetic similarities with southern Francoprovençal. For the Middle Ages the region known as *Occitania submersa* which today corresponds to the south-western part of the *oïl* domain must also be considered;

(3) the Gascon varieties, which display a relatively high degree of diversification between the Pyrenees (Béarn, Bigorre and the Val d'Aran) and the Atlantic coast (Landes, Bordeaux). The status of Gascon has been the subject of perpetual debate and, from an objective point of view, remains ambiguous; on the one hand, it displays a high degree of phonetic differentiation dating back to the first period (cf. Chambon & Greub, 2002, as well as Baldinger, 1962, p. 331f. or Straka, 1987, p. 408); on the other hand, its vocabulary bears particular resemblance to that of Languedocian (cf. §3.2; §4.2.1 as well as Rohlfs, 1966, p. 104 and p. 179). However, despite the persistence of ideological views to the contrary, the individual character lent to this variety by its internal characteristics can no longer be called into question.

The ontological problem of the relationship of Occitan with the transalpine Romance varieties remains; the similarities, of a primarily lexical nature, between Languedocian and neighbouring eastern Catalan are remarkable (for an in-depth discussion cf. Payrató, 1991 and Trotter, 2006). The similarities between Aragonese and the Occitan varieties are equally pronounced – even more so between Aragonese and Gascon. Nevertheless, Occitan was already distinguishable from Catalan in the Middle Ages as a result of their differing external developmental history. During the modern period, the internal evolution of the two varieties shows a sharp divergence due to contact with the respective predominant languages of French and Spanish. This divergence has been further accentuated by the intensive elaboration undergone by Catalan since the beginning of the 20th century.

Three distinct dialect groups are therefore under examination in the present article: southern Occitan, northern Occitan and Gascon.

1.3 Periodization

The evolution of Occitan - and thus the study thereof - may be considered according to the following periods (for the *oïl*-speaking area cf. Carles & Glessgen 2015, pp. 112–115):

(1) the earliest period (ca 700 – ca 1100), is predominantly oral, and is characterized by the absence of any real vernacular texts; during this period, the nature of linguistic variation in the zone known as *Romania continua* – Gallo-, Italo- and northern Iberoromania – is much more homogeneous than in the later periods, despite the fact that the dialectalization of Occitan had taken place long before (cf. §2);

(2) the Middle Ages saw the appearance and elaboration of a written tradition (ca 1100 – ca 1500). The southern Occitan varieties developed a rich textual tradition, modelled on that of the neighbouring regions which were undergoing the same process, and, in some cases, even earlier than these. The same cannot be said for Gascon and northern Occitan; in these areas, non-documentary texts remained few and far between (cf. §3);

(3) during the modern period, (ca 1500 – ca 1800), Occitan came under the sociolinguistic domination of French, which was adopted as the written language across the whole of Gallo-Romance territory from the 16th century onwards (cf. Brun, 1923, which remains unparalleled). Although firmly established as a written language, French remained a foreign language for the vast majority of the Occitan-speaking population, particularly in rural areas. Moreover, Occitan vocabulary did not undergo the process of elaboration characteristic of the standard Romance languages in the 16th and 18th centuries (standardization and extensive ‘learnèd borrowing’, followed by the development of specialized vocabulary and terminology) (cf. §4.1);

(4) the contemporary period (ca 1800 - 2020) has seen a progressive increase in divergence from the standardized languages. Although Occitan remained in use as a spoken language until the first half of the 20th century, with French, the standard variety, being spoken by a very small minority until the 19th century, it subsequently experienced a sharp decline due to the absence of intergenerational transmission. It may be assumed that Occitan will die out as a ‘natural’ language during the course of the next few decades (cf. Bernissan, 2012). During the 19th and 20th centuries, Occitan underwent substantial elaboration (as did several other primary Romance dialects), as a result of the efforts of several revival movements; however, these were not met with the expected widespread acceptance (cf. §4.2.2).

1.4 Current State of Research

The inventory of medieval and modern Occitan and Gascon vocabulary is relatively well advanced due to the existence of various specialized dictionaries (completed and in progress), as well as the FEW (*Französisches Etymologisches Wörterbuch*; cf. Carles et al. 2019). The following list contains the reference dictionaries for the Middle Ages (cf. the detailed description ib., 135-179, which also offers an introduction to electronic corpora and onomastic resources, including those for medieval Latin; cf. §3.1 for the vocabulary of the troubadours and that of documentary texts):

- Rn = Raynouard, F. J. M. (1838–1844). *Lexique roman, ou Dictionnaire de la langue des troubadours, comparée avec les autres langues de l'Europe latine* (6 vols.): despite its age, this work remains an essential resource;
- Lv = Levy, E. (1894–1924), *Provenzalisches Supplement-Wörterbuch: Berichtigungen und Ergänzungen zu Raynouards Lexique roman* (8 vols.): essential supplement to Rn, particularly from the letter C onwards; [the volumes of LvP contain the nomenclature of both Lv and Rn, with no further documentation];
- DAO = Baldinger, K. (Dir.) (1975–2007). *Dictionnaire onomasiologique de l'ancien occitan* (10 fasc.): uncompleted but very rich for the semantic domains treated;
- DOM = Stimm, H., Stempel, W.-D., & Selig, M. (Dir.) (1996–2020). *Dictionnaire étymologique de l'occitan médiéval*. www.dom-en-ligne.de: in progress (*a-album*), however, the online version allows convenient access to Rn and Lv and includes hyperlinks to the FEW;
- DAG/DAGél = Baldinger, K. (Dir.) (1975–2021). *Dictionnaire onomasiologique de l'ancien gascon* (23 fasc.); Glessgen, M. (Dir.) (2020–). *Dictionnaire d'ancien gascon électronique*. <https://dag.adw.uni-heidelberg.de>: limited to the vocabulary of documentary texts (cf. 3.1), but based on an extensive nomenclature;
- DAOA = Olivier, Ph. (2009). *Dictionnaire d'ancien occitan auvergnat: Mauriacois et Sanflorain, (1340-1540)*: id.

Given their incomplete and complementary nature, however, these resources must be consulted as a whole; the researcher is obliged to combine the information extracted from one with that of the others in order to obtain a complete picture. It is also essential to consult the various bibliographies for the Middle Ages, including that of the DOM - currently the most extensive, and available in full-text searchable format, as well as the 'Levy-Schlüssel', available only in print (LvBibl). A significant amount of work remains to be done on the Middle Ages - not only must the DAGél be completed, which is currently making good progress, but more importantly, the DOM, which is to become the vehicle for the advances in philological research made during the last two decades. It suffices to compare the material included in the several fascicules of the DAO with that of the other dictionaries to gain an awareness of the incomplete nature of the latter. Moreover, the extensive documentary sources which exist for the whole of

History of the Occitano-Gascon Lexicon

the Occitano-Gascon area have only undergone partial exploitation to date, a fact which becomes immediately obvious upon comparison of the material included in the DAOA with that of Rn and Lv.

Lexicological research on medieval Occitan is highly fragmented, and consists mainly of monographs or studies focused on a particular author or text. Studies in the domain are known for their substantial nature and high standards, as becomes obvious upon consultation of the DOM bibliography, whose 4800 entries include editions often accompanied by glossaries, as well as lexicological inventories and studies with a particular focus. All publications between 1975 and 2008 have been critically evaluated by Max Pfister in three dense ten-year 'reports' (1993a, 2000, 2011; cf. also 1993b, 1999 as well as Klingebiel, 2011, pp. 293–330 [medieval and modern vocabulary]). These include some thirty articles and numerous reviews written by Jean-Pierre Chambon, Max Pfister's twenty fundamental studies, such as his remarkable *Lexikalische Untersuchungen zu Girart de Roussillon*, or the glossaries and/or lexicological articles in M. Glessgen's *Thesaur* (ThesSSpG, 1989), Monika Tausend's *Legenda aurea* (LegAurT, 1995), the *Trésor galloroman des origines* (Carles, 2017), to say nothing of the volumes of documentary texts edited by Paul Meyer (DocMidiM, 1909) and Clovis Brunel (ChartPrB and ChartPrB2, 1926, 1952). The considerable lexicological contribution made by this body of work has been incorporated into the revision of letter A- of the FEW (vols. 24 and 25) and the ongoing revision of letter B- (online), as well as into the articles of the DOM and the DAO. These inventories thus provide a glimpse into the unexploited potential of medieval vocabulary, but represent only a very partial coverage of the lexical stock.

For the contemporary era, a much larger proportion of the dialectal vocabulary of the Occitano-Gascon area is to be found in the FEW, which thus far exceeds the *Tresor dóu Felibrige* compiled by Frédéric Mistral (1879–1886), and the *Dictionnaire béarnais ancien et moderne* compiled by Vastin Lespy and Paul Raymond (1887), which constitutes the only inventory of Gascon from a panchronic perspective (cf. the remarks made by J.-P. Chambon pertaining to the current state of modern and contemporary Occitan lexicography [2010a, p. 881f.]). The FEW contains the material from approx. 500 glossaries and lexicological inventories - all of which are listed in the *Bibliographie des patois galloromans* (cf. Keller, Geuljans & Wartburg, 1969, pp. 255–300 [Occitan] and pp. 301–316 [Gascon] - as well as all the material contained in the ALF and ALG. To this can now be added the THESOC database (cf. Brun-Trigaud, 2016), as well as several high-quality lexicological repertoires for Gascon (J. Coromines, *El parlar de la Vall d'Aran*, 1990; J. Boisgontier (éd.), Félix Arnaud, *Dictionnaire de la Grande-Lande*, 2 vols., 2001; J.-L. Massourre, *Dictionnaire du gascon des vallées de Luz, de Barèges et de Gavarnie*, 2016). However, the rich lexical material contained in the revised version of the articles for etyma beginning with the letter B- only serves to highlight the rigorous though necessary choices made by Wartburg when compiling the FEW, as well as the fact that the consideration of all available resources results in a far more diverse and in-depth overview of Occitan vocabulary.

Despite this rich corpus of reference works, which, although it contains many imperfections, nevertheless provides a considerable wealth of information, it is surprising to note that there has been no synthetic overview of and/or systematic work either on medieval or modern Occitan vocabulary (cf. the extremely succinct overview by Lafont (1991) on the history of Occitan vocabulary in the LRL). Rare examples of systematic studies include the monograph *Word-Formation in Provençal* by Edward L. Adams (1913), pertaining to the medieval period, which is, however, devoid of any philological and dialectological considerations, as well as, from a dialectological perspective, the chapters on *Les mots dérivés* ('Derived forms') and *Les mots composés* ('Compound forms') in Jules Ronjat's *Grammaire istorique* (1937, vol. 3, pp. 330–482); a complement to the latter is provided by Rohlfs' repertoire of 80 Gascon suffixes (1931; cf. also 1970, pp. 225–231). These works are nevertheless dated and emphasize the formal aspects of word structure, and with the exception of Rohlfs, they display no real consideration of the role played by semantics, historical development, textual genres and onomasiological categories.

According to G. Hilty, Walther von Wartburg had intended to write a history of Gallo-Romance vocabulary from Antiquity to the present day after completing the FEW:

“On the basis of the wealth of material contained in the FEW, he intended to write a history of Gallo-Romance vocabulary from the Romanization of Gaul to the present day. This overview, consisting of one or two volumes, would have represented the crowning glory of the work.” (Hilty, in Baldinger, 1971, p. 34)

Unfortunately, this project never materialized. In a certain manner, K. Gebhardt's thesis on Occitan loanwords in French (1974), based on the material of the FEW, may be considered to represent a systematic contribution to Wartburg's intended project, although it only considers Occitan as a source language and not as a target language (cf. also Swiggers, 1998, pp. 70–72; Trotter, 2006, pp. 1780–1783). Christian Schmitt's attempt to retrace the development of the Gallo-Romance linguistic areas in his thesis of 1974 is more general. However, the assumption that the geographical distribution of modern dialectal forms reflects that of the 3rd and 4th centuries renders it problematic (cf. §2.4). More relevant from a methodological point of view and focused on the Occitano-Gascon area is the doctoral thesis of S. Montigel (cf. Montigel, manuscript in preparation b). Based on lexicographical data contained in the DAG and DAO in particular, it aims to compare several onomasiological categories of medieval Gascon with the Occitan varieties in order to establish the convergences and divergences, using a typological approach (cf. §3.2). Wartburg's plan to write various syntheses using the material collected in his *Thesaurus* and focusing on semantic and/or derivational evolution, regional development, spatial solidarities or loanwords thus remains relevant and one of the many desiderata in the field of historical Romance linguistics.

2 Formation and Emergence of the Vernacular (ca 200-ca 1100)

2.1 Evidence from the Earliest Period

The formation of Occitan vocabulary, like that of the other future Romance languages, begins in earnest from the 5th century onwards. The significant diversification of proto-Romance can be explained by the near absence of written texts and the disintegration of the educational institutions whose efforts to maintain standard Latin would have counteracted the centrifugal forces created by the evolution of the spoken variety. The vocabulary contained in Latin texts prior to the 6th century, however, offers very little evidence of diatopic differentiation. After a systematic gathering and reviewing of such evidence, James N. Adams was able to identify a total of 73 lexemes either specific to or in frequent use throughout the territory of Gaul (2007, pp. 276–369). Upon further examination, roughly half of these lexemes turn out to be pan-Gallo-Romance (39/73), sometimes crossing the Pyrenees and/or the Alps (Carles, 2017, p. 205f.), thereby confirming the relative coherence of Gallo-Romance vocabulary.

Only approx. 10 of these lexemes may be said to be characteristic of the southern regions from the end of the Empire onwards, and consistent with the Romance data. These words, already identified by the FEW, are only present in five sources (Polemios Silvius, 448/49, Lyon; Marcellus, 5th c., Bordeaux (?); Eucheria, 5th c.; *Lex Salica* 507/11; Anthimus, beg. 6th c.). They mainly represent names of plants and animals, and therefore reflect regional realia:

lacrimusa N.F. ‘small grey lizard’ (Polemios Silvius, Adams, 2007, p. 297; cf. FEW 5, 122b, LACRIMUSA, supposedly of pre-Latin origin): Frp. (on French territory)/Dauph./Prv.

darpus N.M. ‘mole’ (Polemios Silvius, ib. p. 297; cf. FEW 3, 13b/14a, **darbo*, supposedly of pre-Latin origin): Frp./Dauph./NPrv.

leuaricinus N.M. ‘pollan, white-fish’ (Polemios Silvius, ib. p. 296; cf. FEW 5, 286b): Frp./Prv.

sofia N.F. ‘dace (*leuciscus vulgaris*)’ (Polemios Silvius, ib. p. 298; cf. FEW 12, 23a/b, SOFIA): Frp./Prv./Lgd.

ripariol N.F. ‘bank-dweller swallow (*hirundo riparia*)’ (Marcellus, Bordeaux(?), ib. p. 291; cf. FEW 10, 417a, occ. *ribeirola*): Lgd., WPrv.

craxantus N.M. ‘toad’ (Eucheria, ib. p. 336; cf. FEW 2/2, 1295b, gaul. *craxantus*): SOcc. [+Cat.]

malum ingenium N.N. ‘trickery’ (*Lex Salica*, ib. p. 316; cf. FEW 4, 685b, INGENIUM): SOcc.

tecco N.M. ‘young salmon’ (Polemios Silvius, ib. p. 331; cf. FEW 13/1, 148b/149a, [Gall.] *tecco*): Auv.-Lim., Béa..

cracatius N.M. ‘sturgeon’ (Anthimus, ib. p. 330f.; cf. FEW 2/2, 1266a, CRAGACUS, (supposedly of Gallic origin): Lgd., Gsc.

trucantus N.M. 'gudgeon' (Polemios Silvius, ib. p. 331; Anthimus, FEW 13/2, 324b/325a, Gall. *trucantus*): Lgd., Gsc.

sitrus N.M. 'hellebore' (Anthimus, ib. p. 334; cf. FEW 11, 661b, SITERUS, supposedly of pre-Indo-European origin): Gsc.

There are thus very few visible traces of differentiation, and the majority of these words are of Gallic or pre-Latin origin, and as such are particularly prone to regional diffusion (cf. §2.2 (3)). These findings suggest a lack of homogeneity throughout the Occitano-Gascon dialect area as the ten words concerned have all undergone dialectalization; none are pan-Occitan. The same conclusions may be drawn for the *oil* and Francoprovençal areas (24/73).

2.2 Evidence from the Merovingian and Carolingian Periods

Few texts from the Merovingian period have survived. Rare traces of the vernacular mainly concern onomastic evidence found on coins, which provides a significant contribution to historical phonetics, for proto-Gascon in particular (Chambon & Greub, 2000; 2002). It does not, however, provide any lexical evidence.

Parallel to its reform of written Latin, the Carolingian Renaissance, on the other hand, served as a catalyst for the elaboration of vernacular writing within the context of Latin. On the basis of such documents it is possible to examine the nature and scale of lexical changes which had already occurred in previous centuries (Carles, 2011; 2017). Documents produced during the period preceding the appearance of fully-fledged vernacular texts – between ca 800 and ca 1100 – included vernacular or partially Latinized lexemes, new lexical meanings and delexical toponyms accompanied by the article (cf. Chambon, 2014).

This type of fragmentary elaboration is found exclusively in Latin documentary texts predominantly from the southern Occitan area (amounting to almost 900 original documents in the ARTEM corpus), but also from the Alverno-Limousin area (175 documents), and from the region known as *Occitania submersa*.

Dauphiné and Gascony are absent from this corpus (Carles, 2017, p. 17). Due to the particular subject matter contained in documentary texts, the vernacular vocabulary identified necessarily focuses on the areas of property management, law, agriculture and the Church. The data examined by the *Trésor galloroman des origines* (TGO) thus includes 279 lexemes and delexical toponyms from the Occitan area (Carles, 2017, p. 55f.). It must be assumed that the analysis of a corpus expanded to include the documents in the *Chartae Galliae* only transmitted in the form of copies would yield over one thousand lexemes (cf. the example in Carles, 2011, p. 309f.; pp. 327–330).

By investigating the elaboration of Occitan vocabulary within the context of Latin it is possible to gain insight into its etymological composition as well as its geolinguistic evolution. The following conclusions may be drawn on the basis of the analysis carried out by the TGO:

History of the Occitano-Gascon Lexicon

(1) In the 10th and 11th centuries the *oïl* and *oc* dialect areas still shared almost two-thirds of all lexical types with the other two Gallo-Romance languages, and almost one-sixth with one of them, at least partially; only one-sixth of the lexemes are limited either to Occitan or to the *oïl* area (55/315 lexemes for Occitan and 48/291 for the *oïl* area. The trajectories of words attested in French but shared by Occitan should also be included in this calculation, which slightly increases the total; Carles, 2017, p. 197f.).

(2) Pan-Gallo-Romance words are worthy of observation as they allow better insight, *ex negativo*, into the chronology and the nature of the processes of lexical innovation at work in Occitan (cf. *ib.* pp. 150–153; p. 129). In fact, they constitute a large percentage of the categories listed below:

- inherited etyma attested in Latin before 500, particularly non-derived and short forms; less frequently, derivatives with semantically transparent affixes (*ib.* p. 150):

AQUA, ARCUS, CARCER, CELLA, CUPA, DENS, EREMUS, FASCIA, FRUMENTUM, GRANDE, HABERE, MANDARE, MERCATUS, MESSIO, MONACHA/-US, MONS, OPERA, ULMUS, HORDEUM, PALLIUM, PARS, PASSUS, PINUS, PLANUM, PODIUM, PONTE, PORTA, PUTEUS, PRATUM, QUIETUS, SACCUS, SANCTUS, SECALE, SERO, SAGMA, TERRA, TURRE;

derivatives: AQUOSUS, CONSUECUDINE, DECANUS, DECIPERE, DIMIDIUS, DISJEJUNARE, INFANS, LINTEOLUM, MINISTERIALIS, MOLINARIUS, NATIVUS, ORATORIUM, PARTIRE, PASTIONE, SEPULTURA, SEXTARIUS, STRATA, TORCULUM, VERVACTUM, VICARIA, VIRIDIARIUM.

It is immediately apparent that these words all reflect concepts central to the daily life of the period, and are therefore words of very high frequency;

- inherited etyma attested in Latin between ca 500 and ca 700 (*ib.* p. 129; p. 150), part of the Latin lexical stock and displaying low variance as regards gender, part of speech or radicals: CAPPÀ, GURGUS n.m., IENUARIUS, MANSUS n.m., PARETE, RACIMUS, RIUS) as well as derived forms attested at the same period and pertaining to daily life (DIURNALIS, EXCLUSA, EXSARTUM, GRANICA, PULLICENUS, PULVERATICUM, ULMETUM and VULPICULUS; only a series of denominations for indirect taxation, almost certainly linked to the Merovingian authorities, is limited to French (PONTATICUS, PORTATICUS, RIPATICUS, ROTATICUS);

- loanwords of ODch origin before ca 700, some of which are attested as early as the 6th or 7th centuries (late lat. ALODE ‘freehold tenure’, BANNUS ‘territory under the jurisdiction of a sovereign’, MARISCALCUS ‘groom’, LOBIA ‘arbour’, WADDIUM ‘pledge’, WARDIA ‘guard’, WERPIRE INF. ‘abandon, throw’), while others only appear during the 9th and 11th centuries (ODch *bord ‘board, plank’, *fehu ‘cattle’, *friskinga ‘young pig’, *want ‘glove’, *sal ‘hall house’, *skara ‘captain’, *thwahlja ‘towel’, *waddi ‘pledge’, *wārjan INF. ‘guarantee’, *werra ‘confusion, discord’; Grm. *alino ‘ell’, *bosk ‘bush’ and *wardôn INF. ‘observe’, OHG. *sazjan INF. ‘place’).

In short, pan-Gallo-Romance diffusion primarily concerns the words which were part of Latin written tradition in Imperial times, but also those which appear to have been formed in Latin

between the end of the 4th and the end of the 5th centuries, and which had been sufficiently integrated in the usage of the period to be employed in written texts. To the latter may be added the Germanic non-derived loanwords which had succeeded in establishing themselves throughout the whole of Gallo-Romance territory. For at least a century and a half after the fall of the Roman Empire, Latin provided lexical continuity across the territory of future Gallo-Romania. This suggests that Latin displayed a high degree of homogeneity as well as horizontal continuity in its spoken form, which is surprising for this period (Carles, 2017, p. 151).

(3) Very few of the Latin words which were already part of written tradition underwent differentiation within Gallo-Romania. All in all, only 11 Occitan words of the 153 analyzed (i.e. 7%) are the result of selection from an established stock of words (cf. ib. pp. 127–133):

pan-Occitan words: *mallol* ‘land newly planted with vines’ < MALLEOLUS, *montel* ‘small natural elevation above the surrounding terrain’ < MONTICULUS, *pinna* ‘pine cone’ < PINEA, *pon-tell* ‘small bridge’ < PONTICULUS, *val* ‘valley’ < VALLE (p. 155);

dialectalized Occitan words: SOOcc. *figeira* ‘fig-tree’ < FICARIA and *verdre* ‘to run’ (applied to liquids) < VERTERE, Lgd./Auv. *quadros* ‘hewn blocks of stone’ < QUADRUS, Lgd./Auv./Gsc. *devesa* ‘reserved area of land’, as well as the two lexemes found only in Latin context which survive in the form of toponyms accompanied by the article: OAuv. *monsteriole* ‘small monastery’ < MONASTERIOLUM and *sindocius* ‘hospital, hospice’ < XENODOCIUM (pp. 165–167).

Only two categories of Latin words formed before ca 700 display more pronounced differentiation within Gallo-Romance territory: that of words of Gallic or pre-Latin origin, which are always transmitted through the medium of Latin, and that of derivatives which are unattested before 700.

(i) In the first of these categories, out of the twelve words of Gallic origin and the six pre-Latin words in the TGO, just over half are pan-Gallo-Romance (10/18, e.g. Gall. **balma* ‘cave’, **brogilos* ‘ground, terrain’), while the others display regional diffusion either within the three Gallo-Romance dialect areas or across part of each of the areas belonging to two or three languages (ib. p. 131f.). Occitan has reflexes belonging to the following six lexical types (i.e. 37.5% of the 18 forms observed within this category):

NOOcc./Frp. *raxia* ‘long and narrow plot of land (particularly for vines)’ < pre-Lat., *vern(i)a* ‘alder’, part of a toponym, < Gall. *verno-* ‘alder’ and *vorze* ‘group of willows’, id., < Gall. **worra* ‘willow’, OAuv./OGsc. (found only in Latin context) *nava* ‘plain surrounded by hills or mountains’ (?), id., < Gall. **nawa* ‘hollow’, SOOcc. *patu* ‘communal pasture’ < pre-Lat., eastern Gallo-Romance *sainna* ‘marshy field’, part of a toponym, < pre-Lat.

Of these six forms, four have survived as part of a toponym (*nava*, *sainna*, *vern(i)a* and *vorze*), almost certainly due to the fact that they were limited to the domain of spoken language, both

in Antiquity and during the early Romance period. The advanced state of dialectalization displayed by these lexemes can be ascribed to the fact that they underwent spatial differentiation before the three linguistic areas had become autonomous - considering that the regional diffusion of Gallic words may already have taken place during the Latin stage of their development.

(ii) For proto-Romance derivatives (formed before 700, but not attested until later), the base is just as likely to be Latin as Germanic or pre-Latin. The TGO lists no less than 109 examples of evolution of this type, only half of which are pan-Gallo-Romance (e.g. *AFFIDARE < FIDERE, *ALODARIUS < ODch **alôd* 'freehold tenure (land)', *BRUCARIA < Gall. **bruco* 'heather'). The patterns of diffusion displayed by the other half are limited to a particular language or dialect; there are 17 proto-Romance derivatives absent from the *oïl* varieties and Francoprovençal, and which experienced differing degrees of diffusion and varied chronology throughout the Occitan and Gascon areas in both the medieval and modern periods (up 15.5% of the 109 words observed, cf. ib. pp. 133–138):

In the case of the following forms, relatively wide-spread diffusion can be assumed: OOcc. (Lat.) *barta* 'brushland' < **bárr-ata*, *conobrarre* 'to exploit land', *flaxata* 'woollen covering', *francal* 'free from taxes (applied to land)', *nogairetus* 'area where walnut trees grow', *par-ran* 'enclosed garden in proximity to a dwelling', *quartairada* 'agricultural unit of measurement', *resclausa* 'sluice, lock', *rival* 'stream', *toron* 'spring', *vaissa* 'hazel' (ib. p. 155f.);

the following forms are characterized by a clear regional diffusion: NOOcc. /Prv. *acaptar* 'to acquire feudal rights', OAuv. Lat. *arzilerius* 'terrain containing clay deposits', OLgd. Lat. *cabanile* 'rudimentary shelter or storage construction' as well as the following onymized forms: OAuv. *pedrucia* 'stony ground' (?), SEOOcc. *salella* 'small seigneurial residence' and OLgd. Lat. *vadello* 'ford' (ib. p. 165).

The geographical differentiation displayed by these proto-Romance derivatives is probably a consequence of their late formation: they may perhaps never have existed beyond the regional level or they may simply have been words of low frequency.

(4) These results for the early period are in stark contrast with those gleaned from the analysis of later derivatives which probably date from the Romance period, i.e. after 700. These are derivatives for which there is no evidence of pan-Gallo-Romance diffusion as a result of direct continuity. The issue remains a sensitive one; however, based on the results of detailed lexicological analyses carried out by the TGO, there is no reason to assume continuity of 'horizontal' communication between the different Gallo-Romance languages after 700. Across all categories, the TGO contains 60 innovations ranging from the 7th to the 11th century, 35 of which are characteristic of Occitan (up 60%; cf. ib. pp. 133–144; pp. 156–162; pp. 165–167). One third of these 35 display evidence of fairly wide, possibly even pan-Occitan diffusion (12), while the others show extensive regional diffusion to a higher or lower degree (23):

History of the Occitano-Gascon Lexicon

OOcc. *acapta-mentum* 'acquisition of feudal rights', *berl-etus* 'area where cress grows' (?), *boat-aria* 'area where cattle are gathered (field, market)' (?), *con-obre* 'exploitation of possessions', *flex-aria* 'area covered in bracken', *formi-mentus* 'costs relating to funerals', *gur-picione* 'ceding of a possession', *mai-a-das* 'scattering of branches (over an area)', *mej-aria* 'tenanted farm', *peir-at* 'area covered in stones', *pog-et* 'hill', *vineogol-ia* 'tax on the exploitation of vineyards' (p. 156);

NOOcc./OLgd. *abs-eddat* 'uncultivated land', *bec-eira* 'area where birches grow' toponym, *preveir-il* '(fiefdom accorded to a parish priest)',

NOcc. *cap-mas* 'principal dwelling of a domain', *mas-ale* 'country dwelling, farm', toponym, *mol-eta* 'area containing several mills' (?), toponym, *parc-eria* 'proportion of produce ceded as rent', *ug-on-encos* (eponymous adj. of uncertain provenance), *vec-eria* 'area where hazel grows', toponym;

SOOcc. *afronta-ciones* 'territorial boundaries', *amas-adus* 'collected, grouped together', *figuair-eda* 'area where fig-trees grow', toponym, *meisson-enc* 'at harvest time', *melgoires-ense* 'coined in Melgueil (applied to currency)', *mercad-al* 'of large capacity (applied to measures)', *mergol-ies* 'coined in Melgueil (applied to currency)', *porc-ada* 'unit of measurement: area of land sufficient for the pasture of one herd of pigs', *ross-els* 'specific to Roussillon (denotes a currency)', *trescol* 'crest of a hill', toponym;

OAug. (in Latin context) *aprad-az* 'land converted into meadows', *claus-atge* 'enclosed plot of land', OPrv. (in Latin context) *braz-aria* 'small tenancy exploited without the aid of animals', toponym?, OLgd. (in Latin context) *albenqu-eta* 'terrain characterized by soil of a whitish hue', toponym? (ib. pp. 165–167)

Based on the testimony of the TGO, the diatopic differentiation of vocabulary within the Gallo-Romance area is therefore relatively limited but clearly present even before 700, and it is particularly pronounced between the 8th and early 12th centuries.

A follow-up study of a quarter of the entries of the TGO (Montigel, Manuscript in preparation a) also shows that half of the words identified as pan-Gallo-Romance are shared by Italo- and Ibero-Romania (often with the exclusion of Portuguese), whereas the other half are divided equally between (1) Gallo- and Italo-Romance words, (2) Gallo-Romance words encompassing Catalan and/or northern Italian and (3) exclusively Gallo-Romance words (ib. chap. 4.3.1). Gallo-Romance vocabulary as a separate entity is therefore recognizable, but it is still very much a part of 'central' Romania. Of the ten lexemes found by the TGO to be exclusively Occitan, three are indeed only found in this variety, while the remaining seven are also found in Catalan and/or in northern Italian (ib. chap. 4.3.2/4.3.4). Once again, there is recognizable differentiation, which, however, is still coherent with the broader geolinguistic framework.

2.3 Evidence from Reconstruction and Toponymy: Substrates and Superstrates

Traditional written sources provide a general orientation which can be considerably enriched by the contributions of comparative reconstruction (cf. Chambon, 2007; DÉRom I, II) and of toponymic sources (cf. the methodological articles collected in Chambon, 2017, pp. 943–1149 and the perspective by Carles, *ib.*, pp. 939–942). The articles of the DÉRom in particular allow a more in-depth approach to the question as to the lexical identity of Gallo-Romance and Occitan (cf. the maps of proto-Romance in DÉRom II, pp. 107–162 [J. Delorme]).

Ancient Greek forms which entered Latin as regionalisms and which were later preserved in Occitan are few, but easily identifiable (nautical: *cau* ‘rope’, fishery: *broumet* ‘bait’, horticulture and viniculture: *empeltar* ‘to plug’). According to Wartburg (1952), these date back to *Magna Grecia*. The precise dating, however, remains the object of discussion, as there is a distinct possibility that they constitute later loanwords from the Mediterranean (cf. Felixberger, 2003, p. 599). The true provenance of the few loanwords considered to be of ‘Aquitain’, ‘Iberian’ and ‘Ligurian’ origin, and which entered Occitan as Latin regionalisms, remains unclear; these forms should be considered in the light of other pre-Latin or pre-Indo-European words which have mainly survived as toponyms. For the whole of the Gallo-Romance area, this varied group amounts to the reflexes of no more than a few hundred etymological bases (*ib.*; for Gascon cf. the inventory in Rohlfs, 1970, pp. 40–59, consisting of approx. one hundred words of pre-Latin origin, which include denominations for plants and animals, terms belonging to pastoral life and territorial denominations, some of which are however Basque loanwords).

The most significant ‘substrate’ for Gallo-Romance geolinguistics is Gallic, as it was a contact language of Latin for centuries until the fall of the Roman Empire (Ternes, 1998, p. 281). Most Gallic loanwords belong to specific lexical fields, that of rural life in particular. Due to the current state of research, the number of these loanwords is as yet unclear, particularly as the potential afforded by toponymy as a methodological tool has been consistently underexploited (cf. the exemplary article on Gall. *attegia* ‘caban, hut’ [Adams, 2007, p. 316] by Billy & Chambon, 1990). In the FEW, 240 etyma are explicitly labelled as Gallic (Müller 1982), a figure which can be increased to approx. 330 words by the inclusion of lexemes whose label is unclear or whose classification remains uncertain. However, a corpus of this type would require an in-depth study, the result of which would most likely be a further reduction in volume (cf. Felixberger, 2003, p. 597). The celtisms in Occitan have not been the subject of specific studies, unlike those identified in French (cf. the references in Felixberger (2003) as well as the complementary lists of 134 confirmed celtisms in French and 60 Latin words of Gallic origin in Lambert, 2003, pp. 188–206. These lists also include several Occitan words such as SOcc. *agragnoun* (‘plum; type of black grape’) < **agran(i)o* ‘id.’ FEW 24, 268b). However, the geolinguistic analysis carried out by Müller (1982) provides a coherent general orientation. Müller identifies several areas

History of the Occitano-Gascon Lexicon

throughout the Gallo-Romance territory which display clusters of loanwords; the Francoprovençal area in particular, as well as the neighbouring northern Occitan domain, including northern Languedoc and the Pyrenean area of Gascony, which is generally rich in substrates. For southern Languedoc, Provence and the other regions of Gascony, on the other hand, the percentage is much lower, as is the case for the *oïl*-speaking areas (ib. p. 617, map 3).

Müller's overview confirms the conclusions reached by the TGO; namely, that the words of Gallic origin fall into two categories: those already attested in written Latin and which therefore underwent subsequent wide diffusion throughout the Gallo- and Italo-Romance areas, and those which were limited to regional spoken Latin in Antiquity and have thus existed as regionalisms in Occitan from the earliest period.

The Germanic superstrate, on the other hand, has been studied in depth (Pfister, 1978; 1998; Felixberger, 2003, pp. 602–604). While interest in this question was originally fuelled by pan-Germanism (E. Gamillscheg, *W. v. Wartburg*), the treatment of Germanisms in the FEW (vols. 15–17) remains exemplary and has even been surpassed by further studies since (M. Pfister, W. Haubrichs as well as all those involved in the completion of the FEW as successors to Wartburg). It is likely that superstrates played a more significant role in the formation of Gallo-Romania than substrates.

The impact of these on internal differentiation, however, primarily concerns the *oïl*-speaking areas; conclusions for Occitan can only be drawn *ex negativo*. The predominant contact language of Gallo-Romania, Old Low Franconian (or Old Dutch, in accordance with current terminology) has strongly permeated the *oïl* dialects (particularly in the northern areas of Picardy and Wallonia), as well as the Francoprovençal dialects to a certain extent. Only those Old Dutch loanwords already present in Latin are found in the *oc*-speaking area, along with other borrowings from Old Germanic - these are the same words which later became pan-Gallo-Romance, as is also evident from the TGO (cf. the precise geolinguistic summary of Pfister, 1973, which definitively identifies 51 Old Dutch words belonging to the latter category as pan-Gallo-Romance, “diffused by the Merovingian chancellery” in his opinion, although this is perhaps too restrictive an interpretation; ib. pp. 145–148). No Old Dutch loanwords are specific to Occitan within the Gallo-Romance area, apart from a certain number of derivatives formed at a later period (e.g. *salella*, §2.1 (3)).

Possible Burgundian loanwords, which sometimes spread to Occitan varieties, are extremely rare and should be considered with extreme caution (according to Pfister, 1998, p. 234f., only a dozen bases can be considered as plausible; cf. Felixberger, 2003, p. 603).

Finally, Gothic also exercised an influence on Occitan, as the Visigoths occupied a large part of the southern territory of Gaul throughout the 5th century. After defeating Clovis in 507, they held sway over Languedoc and eastern Provence until 711. However, it is necessary to examine potential Gothic loanwords in Occitan in detail, as they are often referred to (cf. Pfister, 1987,

pp. 333–336; Pfister, 1998, pp. 235f.; Felixberger, 2003, p. 603), but have not been re-evaluated since Gamillscheg's *Romania Germanica* (1934, vol. I, pp. 297–398, including pp. 363–398). A fresh understanding of the issue can be gained on the basis of the entries of the FEW, bearing in mind that Wartburg may have overestimated the impact of Gothic (cf. Carles, 2017, p. 121; a Gothic etymon is retained for OOcc./OPoit. *vaissa*, rejected for *gardia* ('tax collected as payment for seigneurial protection') and *lobia* ('arbour') (both pan-Gallo-Romance) and is uncertain in the case of OOcc./OFrp. *guurpire*).

To summarize, the languages in contact with Latin which may have had a significant effect on the formation of a particular Occitano-Gascon character and which merit further research are Gallic, Gothic (to a lesser extent), as well as the pre-Latin languages in the case of Gascon.

2.4 Observations on the Basis of Reconstruction: The Latin Etyma

The selection from Latin made by the different Romance languages is illustrated, at least in part, by the inventory of C. Schmitt (1974), based on vols. 1–23 of the FEW. Although Schmitt's dating of diatopic divergence between the 1st and the 4th century is too early, his inventory and onomasiological classification remain valid. He thus provides a repertoire of vocabulary exclusive to each of the Gallo-Romance idioms according to the respective selection of Latin bases. For the Occitan and Gascon areas, this includes nearly 550 etyma whose reflexes are absent from French and Francoprovençal (Schmitt, 1974, pp. 130–179), but unfortunately omits concepts relating to regional realia (ib. p. 133). He also provides no systematic evidence concerning diffusion within these areas. Schmitt's analysis of the material is too concise and he fails to consider the medieval period. Due to the fact that the inventory is based on the Latin etyma rather than on the countless forms resulting from semantic and formal evolution within the etymological trajectories examined, the study focuses on a selection made primarily from an inherited Latin stock rather than on later developments belonging to the 8th and 9th centuries.

If one were to subtract all loanwords contained in vols. 15–20, all substrate words and Hellenisms as well as the new etyma in vols. 24 and 25 from the entries of the FEW, the number of Latin etyma would amount to approx. 14 000. The 550 reflexes which are exclusively Occitan or Gascon thus correspond to approx. 4% of all Gallo-Romance lexical stock inherited from Latin. This figure would increase if one were also to take into account the 6000 Latin derivatives and, more particularly, the numerous proto-Romance formations which Wartburg classified under the main etyma as opposed to treating them as etyma in their own right (cf. Carles et al. 2019, p. 44f.). These forms are in fact absent from Schmitt's inventory. A calculation based on the combination of the above parameters would therefore appear to confirm the figures cited by the TGO regarding the percentage of Latin words selected within the Occitano-Gascon dialect area (ca 12%, cf. §2.2). The primary value of Schmitt's inventory derives from the fact that it provides a large quantitative basis for this type of investigation; much could be gained from further exploitation along these lines.

2.5 Summary

Various quantitative analyses have shown that the Gallo-Romance idioms still displayed relative lexical coherence around the year 1000. The process of spatial differentiation began under the Roman Empire and continued after its fall. An intensification can be identified from the 8th century onward with the collapse of horizontal communication throughout Gallo-Romania. From this point onwards, the lexical identity of Occitan was clearly recognizable when compared with that of the neighbouring Romance languages, although it continued to undergo diversification throughout the second millennium (cf. Carles, 2017, p. 198).

3 The Late Middle Ages (ca 1100 - ca 1500)

3.1 The Development of Genres

1100 saw the beginning of the development of a rich tradition of vernacular writing in the Occitan area, catalyzed by the expansion of European culture, to which it contributed. The southern Occitan area represents one of the epicentres of this intellectual and artistic revolution, due to its pivotal geographic position between the Mediterranean and Northern Europe, as well as its proximity to al-Andalus, which was culturally highly developed, and to the city-states of the Italian peninsula. During the 12th century, written Occitan established itself in three areas:

- (1) the lyric poetry of the troubadours, whose first epicentre was almost certain to have been in the Limousin area (catalyzed by the abbey of Saint-Martial de Limoges), and subsequently in Languedoc (cf. the evolving reference database BEdT, the electronic editions CAO/RIALTO and the concordance COM-2);
- (2) scholastic legal texts (through the translation of the *Codex Iustinianus* as early as 1149/60 [Gard], cf. Codi) and, at the same time, legal practice, primarily in Languedoc (cf. Brunel, 1926; 1952; Glessgen, 2018);
- (3) religious texts (translations of the New Testament etc.) of which the Limousin area was also the first epicentre (Glessgen/Pfister, 1995a/b).

Between the 13th and the 15th centuries, Occitan vernacular writing was primarily concentrated in Languedoc and Provence, particularly in the areas of secular and religious literature (Chambon, 2010; Menichetti, 2016) as well as documentary texts (property management, inventories, accounting). It also included other genres (medicine and pharmacognosy, astronomy, technical texts), but to a much lesser extent than in French or Italian (cf. the inventory of Brunel, 1935, containing 357 entries, to be replaced by Menichetti [manuscript in preparation] containing the description of 500 non-documentary mss.; cf. DiTMAO for medico-biological glossaries in Hebrew script). The development of written culture stagnated in the Auvergne

History of the Occitano-Gascon Lexicon

and Gascon areas, and by this stage the Limousin had also fallen behind; written culture in these regions remained largely limited to documentary texts, which primarily developed during the 14th and 15th centuries (cf. Chambon & Olivier, 2000; Chambon, 2012 for Auvergne; Glessgen & Pfister, 1995b for Limousin; Baldinger, 1962 and the approx. 1000 references in the DAGBibl; Allières, 1995 for the Gascon area). This was all the more true for the Dauphiné where textual production, even that of documentary texts - remained limited (cf. DocMidiM).

The elaboration of written culture obviously had an immediate effect on the development of the lexicon; indeed, the two processes are inseparable. Occitan thus acquired a range of poetic, religious and legal and administrative terminology from a very early stage, and developed denominations for a large series of specific concepts in the domains of the emotions, courtliness and court life, feudalism and chivalry, religion and worship, law, wealth management, agriculture and trade. In each of these domains, the words employed in the texts naturally belonged to the spoken language of the period, as is evident from the classic example of the author Chrétien de Troyes, writing in the *oïl*-speaking area (Carles, 2013). However, these texts also contain new word-formations resulting from the process of translation (particularly in religious and legal texts), thereby enriching the lexical stock of the spoken language through both semantic change and derivation. The expansion of the lexicon occurred in a similar manner in the case of terminology belonging to the semantic fields of medicine and biology, although the denominations of primarily plant-based pharmaceuticals were heavily influenced by Mediterranean trade.

The mechanisms of lexical innovation based on textual traditions and intertextuality thus led to the considerable enlargement of Occitan vocabulary between the 12th and 15th centuries.

3.2 Factors Affecting Internal Cohesion and Differentiation

The grapho-phonetic and morphological physiognomy of the *scriptae* belonging to the Occitan and Gascon areas is highly regionalized, thus facilitating the identification of the affiliation of a particular text with a large area such as Languedoc, Provence, Dauphiné, Gascony or the Limousin. The Occitano-Gascon dialect group thus constitutes an example of polycentric codification, with no single centre of high prestige. The extent to which this regional diversification is reflected in the lexicon, however, remains wholly unappreciated. At first glance, the lexical variation apparent in the various textual traditions does not show a high degree of spatial differentiation, in contrast to the pronounced regional nature of grapho-phonetic variation and the high level of dialectalization characterizing the spoken varieties. Each individual text belongs to a tradition and therefore relies on earlier models, upon which the vocabulary specific to the dialect area from which it originates exerts minimal influence.

The *Dictionnaire de l'ancien occitan auvergnat* (DAOA) by Ph. Olivier, consisting of approx. 12 000 entries, allows an estimation of the degree of internal differentiation displayed by Occitan in documentary texts. The DAOA constitutes a systematic lexical inventory of words extracted

from texts belonging to the field of administrative practice in Auvergne between the mid-14th and early 16th centuries (Mauriac and Saint-Flour/Cantal, 1341–1537). The completed section of the DOM (*a- album*) may be used for the purpose of comparison (given the varying quality of textual production, edition and lexicological analyses for the Occitan and Gascon area, the DOM may be considered to be representative of southern Occitan written tradition). The entries from *a – album* correspond to the first 200 entries of the DAOA. The results of the comparison, aided by the FEW, show a high degree of similarity between the vocabulary of Languedoc and Provençal in the 12th and 14th centuries and that of the documentary texts of the Auvergne region in the late Middle Ages. This comparison does not take grapho-phonetic variation into account, which clearly reflects dialectal differences (e.g. the systematic opposition /ka/ ~ /tʃa/ as evidenced by OAuv. *achabamen* vs SOOcc. *acabamen* ('summit; completion; perfection; accomplishment'), or variations specific to a particular region such as *aguada* vs *aigada* ('flood; high tide') [< AQUA], *agar* vs *egar* ('to irrigate') [< AEQUARE] or *aprial* vs Lgd. *abril/abriu* ('April') [FEW 25, 59b, APRILIS]). Lexematic variation, i.e. variation involving radicals, derivatives or semantics, is very limited and is to be found in only 10% of the entries of the DAOA:

- 4 cases of semantic specialization: it is not yet known whether they are simply infrequent forms, or determined by the particular discourse tradition, or the product of dialectal differentiation (*aconhadet* 3SG-accommodate-PRF 'put (a dwelling) at s.o.'s disposal' 1387, *aisimen* 'useful equipment, fittings (in a dwelling)' 1378 [poss. < Fr.]), *aysinas* 'id.' 1381, *alberc* 'tax for accommodation in a seigneurial dwelling' ca 1397
- a single derivative from a particular dialect with a specific meaning: *ayresselh* N.M. 'parsley' (1427), attested only in the work of Daudé de Pradas (Rgt) and in NMOcc. (FEW 3, 239a, ERICIUS); a second derivative also appears to be geolinguistically marked: *afrontador* N.M. 'barrier against wind or snow' (1429); cf. OOcc. *afrontar*
- 3 (+ 2) word-formations shared by the *oil* area, which can be assumed to be loanwords or simply the result of spatial contiguity:

acomplisement M. 'achievement' (1428); cf. Fr. *accomplissement* (since 1384, JMeun, FEW 2, 982a, COMPLERE)

affretz M.PL. 'periodic payment' (1409); cf. Mid. Fr. *frais*, OAuv. *freiz* (1400, Cantal, FEW 3, 755a, FRANGERE), with the subsequent derivatives *afretar* INF-TR. 'pay, distribute (tax)' (?) (1387) et *affretacios* F.PL. 'payment, distribution (of a tax)' (?) (id.)

haynos ADJ. 'hostile' (1439); possibly related to the type Fr. *haineux*

- a series of 5 prefixations *a/ad*:

ablanchesir INF-TR. 'coat with lime' (1395); cf. Frp. *blätseyi* (FEW 15/1, 141a, Grm. **blank*); the type preceded by *a-* is attested sporadically throughout the Gallo-Romance area (Mid. Fr. *ablanchir* 1453, Alais *ablanqui*, ib., 142a/b), which suggests a polygenetic origin

acoregir INF-TR. 'to correct' (1494); cf. OOcc. *corregir*

History of the Occitano-Gascon Lexicon

acreysensa F. 'enlargement' (1476); cf. OOcc. *creisensa* / OFr., Mid. Fr. *accroissance*

adaygrezir INTR. 'turn into vinegar' (1383); absent from FEW 24, 97b/98a, ACER; cf. OOcc. *aigrezir*

adhaguada F. 'sudden increase in the volume of a river' (1418); absent from FEW 24, 134b, ADAQUARE; cf. OOcc. *adagar*

• 2 variants with extended radicals:

aconplinnament M. 'action of completing a task' (1388); cf. EOPrv. *acompliment* (1488, Pans, Avignon, FEW 2/2, 981a, COMPLERE)

agrandesir TR. 'to enlarge' (1468); cf. aocc. *agrandir*

• 2 isolated Latinisms due to insufficient mastery of written Latin:

acsequitoria F. 'document drawn up for the execution of a court decision' (1391); cf. lat. EXECUTORIUS adj.

adjurpacion F. 'usurpation' (1463); cf. Lat. USURPATIO

• 2 entries which are difficult to analyze; one due to the inadequate treatment of letter B- in the FEW, the other due to its being a hapax:

abonnar INF-TR. 'subject s.o. to a specific tax' (1475); *abonnat* M. 's.o. who is subject to taxation' (id.); cf. Oil *aboner*, Prv. *abourna*, Lim. *abou(r)na* (FEW 1, 466a, Gall. **botina* 'boundary marker']]

agiala F. 'type of vat' hapax (1509) (cf. FEW 4, 123b, GERULUS)]

These findings are not easy to interpret and one might be tempted to consider Medieval Occitan as a language poor in variation, characterized by a high degree of spatial cohesion. This could in turn lead to the view that, at least in the late 14th and 15th centuries, a more or less homogeneous written standard, or 'koinè', developed. This hypothesis is still to be found in recent textbooks (e.g. Swiggers, 1998, p. 68 or Weth, 2014, p. 494), but is, however, fundamentally flawed. The appearance of cohesion arising from the above comparison can be explained much more satisfactorily in terms of the stereotypes determined by the discourse tradition in question: the terminology specific to the fields of agriculture and property management has undergone a highly-developed process of textualization, and, as part of spoken language, has been shaped by multiple exchanges amongst different regions. In other words, this vocabulary belongs to a specific lexical field and to a particular discourse tradition.

A systematic analysis of poetic, religious or medical vocabulary comparing two series of texts from different periods and regions would lead to similar results. Medieval vocabulary is inextricably linked to textual traditions. Each tradition uses only a small number of words (cf. Glessgen & Dallas, 2019) and thus, once a model has been established, it tends to remain stable throughout the centuries (cf. the classic example of the judicial sentence, Krefeld, 1987). The same textual traditions are operational throughout the whole of the Occitan and Gascon area.

History of the Occitano-Gascon Lexicon

This leads to the establishment of what might cautiously be called textual 'standards' in terms of discourse structure, syntax and vocabulary. Due to the fact that most genres originated in southern Occitan, the vocabulary of this region must necessarily have played a significant role in the constitution of various discourse traditions. Nonetheless, this should not be interpreted as evidence for the existence of a wide-reaching Medieval Occitan standard, particularly as this apparent homogeneity was counterbalanced by maximal variation in the domains of grapho-phonetics and morphology.

The nature of medieval lexical variation is currently being pursued in depth by S. Montigel, who is conducting an investigation into the convergences and divergences between Gascon on the one hand, and southern and northern Occitan on the other. Using a typological approach, this enquiry is based primarily on the vocabulary treated in the DAG and DAO, which possesses the two-fold advantage of being organized by lexical fields and – (nicht kohärent: manchmal kurz, manchmal lang; nächste Seite z.B. lang) in the case of the DAG, at least – of being based on vocabulary extracted from editions, and therefore from primary sources. On the basis of the early results of this work in progress, the comparison between the DAOA and the DOM can be refined. An analysis of the vocabulary by lexical field gives rise to three additional observations (cf. Montigel, manuscript in preparation b):

(1) for core vocabulary corresponding to concepts which are fundamental from a cognitive point of view (such as parts of the body, wild animals, etc.) Gascon, southern Occitan and northern Occitan display a high degree of similarity and are in some cases identical.

(2) vocabulary elaborated in the context of written culture (e.g. law, medicine, religion), and/or in the field of trade and its associated products, also shows little geolinguistic variation within the Occitano-Gascon area. However, the considerable diachronic evolution undergone by these words has contributed greatly to the differentiation of Occitan from the neighbouring languages.

(3) Finally, words pertaining to daily life, primarily linked to regional culture and fundamentally shaped by the spoken language (land, agriculture, kinship names) are very highly diversified (e.g. 9 of the 12 kinship names found in OGsc. are absent from the Occitan varieties).

Furthermore, similarities between Gascon and Languedocian are much more marked than those between Gascon and Limousin, despite the fact that both regions border on Gascony. It may thus be concluded that the degree of internal differentiation displayed by medieval vocabulary is inextricably linked to the lexical fields represented. The apparent homogeneity within written discourse traditions masks the profound differentiation which characterizes the evolution of the Occitan varieties.

The numerous studies and reviews written by J.-P. Chambon regarding the vocabulary of a particular text or group of texts also bear witness to the dialectalized nature of these varieties. His identification of hundreds of diatopically-marked lexemes in all genres only serves to highlight the omnipresence of dialectal diversity throughout the whole of the Occitan and Gascon area. As the written variety tends to show a predilection for lexemes with a wide areal diffusion – and thus for pan-Occitan forms – dialectal words of limited areal diffusion are not often present in written texts. The consequence of this is that former research has vastly underestimated the quantity of such forms (Glessgen, 2016; Carles & Glessgen, 2020; Carles, 2020). If a quarter of all non-learnèd medieval vocabulary found in written texts from the *oïl*-speaking area may be considered to be regional (cf. Glessgen, 2016), this figure could well prove to be even higher for the Occitano-Gascon dialect group; elaboration during the medieval period remained pluricentric, with no single variety functioning at a supra-regional level.

4 The Modern and Contemporary Period (ca 1500 – 2020)

4.1 The Modern Era (ca 1500 – ca 1850)

The modern era, from the 16th to the middle of the 19th century, represents the period of evolution of Occitan about which the least is known (cf. Chambon, 2010a, p. 880). Following the realization that the regional Occitan *scriptae* had been replaced by French between the end of the 15th and the middle of the 16th centuries (Brun, 1923), the following centuries were generally ignored by research until the revivalist movements of the second half of the 19th century. It is true that the modern era is characterized primarily by the slowing down of the process of elaboration of the Middle Ages. Nevertheless, Occitan remained the mother tongue of almost all speakers of the region (cf. §4.2.1) and its vocabulary continued to evolve, contrary to the prevailing theories (e.g. Lafont, 1991, p. 21). The traditional vocabulary continued to diversify, alongside the development of regional literature and the compilation of significant lexicographic inventories.

Although the production of secular literature and religious texts did not compare to that of either French or Italian, these two ensembles underwent real development during the 17th and 18th centuries, particularly in regions possessing a strong tradition of written culture, such as Languedoc, where Gascon authors were also active, and Provence (cf. Brun, 1927; Lafont & Anatole, 1970, pp. 267–482; Lafont, 1974; Gardy, 1985 and 1997; Boyer & Gardy, 2001, pp. 119–220 [for religious language: Eygun, 2001]; Bernsen, 2006, pp. 1988–1992; Courouau, 2008, 2017 and, above all, 2015 which boasts a large inventory of mss. and 18th century Occitan printed editions, *ib.* pp. 489–542; cf. also the 160 works written between 1600 and 1830, collected in the Digital Library *Tolosana*). By then, French had established itself as the standard

and Occitan texts began to display more similarities with the local dialects; their linguistic interest thus increased. In addition to erudite baroque poetry, therefore, minor genres are particularly well-developed, such as pastorals and *noëls* (devotional poems or songs in the vernacular), songs, carnivalesque farces and popular theatre.

The 18th century saw a real effort in the field of linguistic elaboration, to which a series of dictionaries pertaining to the Rhodanian area bear witness: the *Dictionnaire provençal et français* by Father Sauveur-André Pellas (Avignon, 1723; ca 10 000 entries; cf. Stéfani, 1964), the dictionary of the same name by Father Pierre Puget (Aix-en-Provence, ca 1730/40, inedited), the *Dictionnaire languedocien-français* by Father Pierre Augustin Boissier de Sauvages (Nîmes, 1756; 2 vols., 1785 (2nd edition); cf. Lieutard, 2010; ca 10 000 entries), followed by the *Dictionnaire de la Provence et du Comté Venaissin* by Claude-François Achard (Marseille, 2 vols., Provençal-French, French-Provençal, 1785; ca 12 000 entries) and lastly, a rich anonymous dictionary, still in manuscript form (Arles, 6 vols., ca 1770; ca 34 000 entries, cf. Thomas, 2013). Although Pellas's work is very French-orientated, that of Boissier, despite its declared purpose of facilitating the learning of French, is based on a nomenclature which consciously steers clear of Gallicisms (cf. the rich lexicography of the 18th century, also encompassing northern Occitan and Gascon (Schlieben Lange, 1991, particularly pp. 115–121, and Courouau, 2015, pp. 307–350 [D. Fabié]).

The development of texts written in modern dialect reached its peak at the time of the Revolution; political texts and translations were accompanied by a great many scattered offerings based on folklore (cf. Boyer et al., 1989; Merle, 1990 and particularly Pic, 1989, which contains an inventory of about 230 mss. and editions, as well as a bibliography; for the Auvergne region cf. Chambon & Olivier, 2000). Significant editorial work remains to be carried out on the Occitan texts of the modern period. This would provide a foundation upon which to base lexicological analyses with the aim of highlighting the diversity and creativity inherent in this heritage.

Only one aspect of modern linguistic history has been studied more specifically to date: this concerns the interaction between French and Occitan, which led on the one hand to the formation of a regional French variety based in part on the dialects in these areas, and on the other hand, to the formation of Gallicisms in Occitan varieties. The latter forms part of the process of dedialectalization by substitution; however, it also highlights the complex nature of the interaction between the two languages.

Several varieties of regional French formed within the Occitano-Gascon area from the 16th century onwards, with an intensification in the 15th and 19th centuries (cf. Swiggers, 1998, p. 74 and particularly Chambon, 1997; Chambon & Carles, 2007 and Chambon & Grémois, 2007). Typologically, they are characterized (1) by loanwords from wide-spread dialect words, (2) by syntagmatic, semantic and formal innovations in French and (3) by the maintenance of French

lexemes fallen out of use elsewhere (cf. Carles, 2020). These interference varieties developed amongst bilingual speakers, primarily the urban bourgeoisie (approx. 5% of the population at the time of the French Revolution). The use of regional French at the expense of the Occitan dialects intensified sharply around 1900, following the introduction of compulsory schooling which led to generalized bilingualism (Glessgen, 2020).

Parallel to the introduction of Occitanisms into French through the channel of regional French, the Occitan dialects absorbed an ever-increasing number of French words, which often replaced dialectal forms (cf. Chambon, 2010c). Regional French varieties developed throughout the whole of the territory starting from the urban centres. The ‘Gallicisms’ in Occitan, however, are the result of a general north-south movement, particularly pronounced along the axis of the Rhône, which emerges very clearly from certain maps of the ALF (e.g. 712 *jardin*, ALF 738 *loup*, 850 *meunier*). Numerous encounters - primarily due to commercial exchanges which took place via the principal channels of communication - served as a catalyst for the francisation of the dialects, particularly from the 18th century onwards. In addition, the phenomenon known as *parachutage* (‘air drop’) led to periodic lexical substitutions, seemingly random from a geographical point of view, also frequent in ALF maps.

4.2 The Contemporary Era (ca 1850 – 2020)

4.2.1 Diversification of Traditional Vocabulary

The internal differentiation of the vocabulary of the dialectal varieties of Occitan reached its peak during the 19th century after fifteen centuries of evolution. As a result of the absence of the systematic development of a supra-regional standard between the 16th and 19th centuries, Occitan contains a relatively small proportion of Latinisms, as well as a limited number of loanwords from neighbouring languages, except in border regions and without considering French. The pronounced diversification of the Occitan varieties has been favoured by the absence of physical geographical unity and the lack of geopolitical autonomy throughout the area in which they are spoken. This renders Occitan of particular interest amongst the Romance languages, both from an epistemological and a typological point of view.

It is against this backdrop that the above-mentioned inventories of traditional lexical forms were compiled by dialectologists between the end of the 19th and the second half of the 20th century (cf. §1.4), culminating in the *Nouveaux atlas linguistiques de la France* (NALF). The six atlases pertaining to the South further improved coverage of the Occitano-Gascon area. The systematic reorganisation of this material in the THESOC database (in progress), as well as the digitization of the FEW (also in progress), will provide a new foundation for future interpretative research.

From a geolinguistic perspective, H. Goebel’s dialectometric transformation of the material contained in the ALF allows researchers to establish the dimensions of the various linguistic areas

History of the Occitano-Gascon Lexicon

of the Occitano-Gascon domain on a quantifiable empirical basis. These areas have already benefited from a detailed description (cf. the overview in Ravier, 1991, essentially based on phonetic criteria). Using the dendrographic maps, (471 for the lexicon and 1117 for phonetics), the distribution of vocabulary and phonetic traits, as well as the degree of convergence amongst the different areas can be investigated. This tool renders geolinguistic variation visible by distinguishing a given number of entities termed 'choremes' (the fewer the number of distinguishable entities, the greater the degree of differentiation between them). The following observations may be made with regard to the lexical differentiation of Occitan within Romania based on the number of choremes, which range from 2 to 20:

- (1) Unsurprisingly, the principal distinction amongst the Gallo-Romance varieties is that between the Occitano-Gascon varieties on the one hand and the French and Francoprovençal varieties on the other (2 choremes). 3 choremes separate Franco-Provençal from French;
- (2) of a more unexpected nature is the fact that the lexical distance between the north-eastern varieties of French on the one hand, and the central and western varieties on the other (4 choremes), is more pronounced than that between the Gascon/Languedoc area on the one hand, and NOcc. (which, surprisingly, includes the Bordelais region), and the Dauph./Prv. areas on the other (5 choremes); the distinction between the Gascon/Languedoc area and NOcc. represents the maximum lexical opposition within the Occitano-Gascon dialect group; this is further confirmed by the phonetic maps (4 choremes), which thus unite Gascon and Languedocian in opposition to the other varieties of Occitan;
- (3) the second distinction within the Occitano-Gascon dialect group, subordinate to the first, is that between the northern group (Bord./Lim./Auv.) and the eastern group (Dauph./Prv.), while the Gsc./Lgd. group show convergence (6 choremes); it should be mentioned that the character of eastern Pyrenean Catalan shows more similarity with the Dauph./Prv. group than with all other varieties, as they have tended towards conservation of the same lexemes;
- (4) subordinate to the above distinctions, Gascon vocabulary (still with the exception of the Bordelais) is distinct from Languedocian (7 choremes);
- (5) after determination of the four major Occitano-Gascon areas (NOcc., Dauph./Prv., Lgd., Gascon), any further distinctions (19 choremes in all) mainly concern the *oïl* area (for which eight groups can be identified: Pic., Wall., Lorr., Frcomt./Bourg., the West, the Ile-de-France, the Centre, the South-West) and Francoprovençal (further divided into 3 groups: France, Jura/Vaud, Valais/Aosta Valley);

(6) further sub-categories of Occitan include: the distinction between Alpine Prv. and the Dauph./Prv. area (10 choremes), that of ELgd. and WLgd. (15 choremes), Auv. and Lim./Bord. (16 choremes) and finally, the distinction between the northern third of WLgd. and the central and southern parts (20 choremes); lexically, the Béarnais displays solidarity with western Gascon, as does the Dauphiné with Provençal.

The phonetic maps differ from the above findings on several significant points:

- (1) the rupture between Gascon and Languedocian is more significant (5 choremes instead of the 7 for the vocabulary) and occurs earlier than that between NOcc. and Dauph./Prv. (8 choremes instead of 6); the Bordelais, on the other hand, displays similar traits to Gascon;
- (2) for NOcc. the distinction between Lim. on the one hand (12 choremes), and Auvergnat and Dauph. on the other results more clearly (the latter with only 18 choremes);
- (3) Finally - contrary to the results of the lexical analysis - there is a separation between EGsc. (including Béarnais) and WGsc. (including the Bordelais and the Landes) which begins at 11 choremes and occurs before that between Lim. and Auv. (16 choremes), Dauph. and Prv. (19 choremes) and before the internal differentiation of Lgd. (17 choremes), which distinguishes a north-eastern area from a central-southern area.

Once progress has been made on the digitization of the THESOC database, it will be possible to further refine these results. A first dialectometric analysis of the 232 maps contained in the six atlases collected in THESOC, however, confirms (1) the fundamental divergence between the four large areas of NOcc., Dauph./Prv., Lgd. and Gascon, (2) the composite character of Lgd., divided into three groups (N., S. and EOcc.) and (3) the transitional character of Bordelais on the one hand, and of Dauph. on the other (Brun-Trigaud, Malfatto & Sauzet, ms.).

In short, a systemic view of Occitan can be gained from both lexicological and phonetic analyses, which reveal profound divergences between the four major areas. Any approach to the Occitano-Gascon varieties must thus take into account the composite nature of the group.

An in-depth investigation with the scope of refining this macroscopic geolinguistic inquiry should take the form of an onomasiological approach, since the differentiation of Occitan is firmly rooted in lexical fields. It is thus difficult to evaluate the foundational study carried out by Gerhard Rohlfs (1970) within the general framework of linguistic variation and change, as the nomenclature, consisting of some 700 lexemes, either exclusively Gascon or belonging to both the Gascon and the Ibero-Romance area (ib., pp. 38–115), is not organised by lexical field. A more coherent picture of the lexical autonomy of Gascon is to be gained from an approach focused on a particular conceptual domain such as the denominations of plants, for example, which have been studied in depth by Jean Séguy (1953) and Gerald Bernhard (1988). The ety-

mological analysis of ca 350 names of Aranese plants carried out by Bernhard under the direction of Johannes Hubschmid shows the clear diversification of Gascon in this particular conceptual field, both in the Middle Ages and during the modern period (Bernhard, 1988, pp. 122–134).

Differentiation could also be investigated from the point of view of word-formation. Although the suffixes primarily found in the Romance languages are also present in Occitan, they are not used in the same manner throughout the Gallo-Romance area. Thus the exemplary studies carried out by Franz Rainer on the suffixes *-men(t)*, *-aria* and *-on(e)* show that Occitan behaves quite differently to the neighbouring varieties of Francoprovençal, Catalan and French in this respect (Rainer, 2018a/b; Rainer, forthcoming). Studies of this type can be greatly enriched by the consideration of variation within the Occitano-Gascon dialect group itself.

The issue of geolinguistic differentiation represents one area amongst many, all of which merit in-depth study. The present considerations, however, highlight the real potential which resides in diachronic and variational lexicological studies on the basis of the Occitan varieties, but also in the analysis of the relationship between the composition of the vocabulary at any given time and place and the manner in which it is exploited in Occitan and Gascon texts.

4.2.2 The Effects of Linguistic Elaboration

The contemporary era is characterized by two significant developments, the first of which involves lexical enrichment, the second impoverishment. The strong renaissantist movements experienced by Occitan - the Félibrige (founded in 1854), followed by Occitanism emanating from the literary journal *Oc* (founded in Toulouse in 1923) and the Institut d'Études Occitanes (Institute for Occitan Studies, founded in 1945) served as a catalyst for significant poetic creativity. Although this facet of Occitan history remains unknown to most French speakers, it has been the object of excellent research led by Occitanist sociolinguistics, perhaps at the expense of other equally worthy topics (cf. the overviews by Teulat, 1979; Bec, 1991; Boyer & Gardy 2001, pp. 221–462; Tomàs, 2006; Weth, 2014).

Once again, it may be said that lexicological aspects have been treated as the ('poor relation', despite their significant involvement in the process of elaboration. Frédéric Mistral's *Tresor dòu Felibrige* thus remains the richest lexicographical inventory of modern Occitan, despite certain defects (cf. Chambon, 2010b, pp. 208–210). An analysis of the lexical choices in *Mireio*, Mistral's masterpiece, reveals the level of attention paid to the choice of words, taken from the variety he aimed to establish as standard. The geolinguistically composite character of his vocabulary is made apparent by Jean-Claude Rivière's thesis (Rivière, 1985; cf. also Rostaing, 1971). An analysis of the first 200 verses, using the FEW (cf. Carles & Glessgen, forthcoming) identifies approx. thirty diatopically-marked lexemes. 40% of these (13/31) belong, as expected, to Provençal or Dauph./Prv., and less frequently to Rhodanian:

History of the Occitano-Gascon Lexicon

- Prv.: *aloubati* ADJ. 'greedy (wolf-like)' 21, *abrama* PST-PTCP 'inflamed by desire' 24, *terreiròu* M. 'burrow' 133, *estoublo* F. 'fallow' 139, *spelido* PST-PTCP-F 'hatched' 166, *esvalir* (REFL.)-INTR. 'dissipate', *batudo* F. 'labour' 181, *escàfi* F. 'mockery' 191;
- Dauph./Prv.: *vege* M. 'willow' 119, *cano* F. 'reed' 135, *eigagno* F. 'dew' 102; 169;
- WPrv./Rhd.: *canestello* F. 'basket' 42; 97, *barrulaire* M. 'ploughshare' 114.

One-third of these words (9/31), however, are characteristic of east. Lgd., or of both this region and (west.) Prv.

- ELgd.: *falabrego* F. 'nettle tree', *espalanca* INF. 'break (applied to branches)' 22, *jitello* F. 'branch' 25, *afrescoulido* ADJ. 'cool' 168; Lgd.: *magagno* F. 'pain' 170;
- WPrv./ELgd.: *oustaloun* M. 'small dwelling' 38, *capistello* interjection. 64, *magnan* M. 'silkworm' 101, *pesqui* M. 'fishpond' 87.

These first investigations lead to the conclusion that Mistral constructed his lexical model on the basis of WPrv. and (E)Lgd., thus fusing the two main southern varieties possessing the most highly-developed tradition of written culture. Furthermore, a tendency to prefer pan-Occitan words can be noticed (with the exception of Gascon), as well as towards the incorporation of archaic Gallicisms (mostly dating back to the 18th century and often characteristic of Prv.):

- pan-Occitan words (excluding Gascon): *clafi* INF. 'fill' 69, *fedo* 81, *baudufo* F. 'spinning top' 188;
- Gallicisms: *chato* F. 'young girl' 1, *chatouno* F. 'liddle girl' 101, *ventoulet* M. 'light wind' 31, *aubrage* M. 'clump of trees' 85, *estampa* PST-PTCP. 'fashioned' 122;
- cf. also a recent Arabism (Algerian loanword): *ràfi* M. 'farmhand' 145.

The lexical character of *Mirèio* thus reflects a composite variety of southern Occitan quite distinct from both Gascon and the Alverno-Limousin group, but not averse to already well-integrated loanwords from French.

This remains an isolated example. Modern Occitan literature constitutes an extraordinary linguistic observatory due to the fact that it is doubly-rooted within a long tradition of written culture and varied dialectal heritage. The intertextual, stylistic as well as lexical choices which characterize it are thus charged with significance and are endowed with a poetic function. This is emphasized by the studies of Philippe Gardy in an exemplary manner, particularly as regards the work of Max Rouquette (Gardy, 1996) and by those of Jean-Pierre Chambon on the work of Jean Boudou (Chambon, 2017, pp. 677–703; p. 749s.; pp. 829–935).

Lexical elaboration undertaken by the renaissantist movements (whose principal actors have often been activists, writers and linguists, such as P. Bec or R. Lafont) represents part of a broader desire to endow Occitan with a standard variety. Occitan and Gascon have thus experienced several attempts at standardization on the basis of Rhodanian, Toulousian and Gascon. These efforts, intended to counteract the decline of intergenerational language transmission,

experienced very limited social acceptance, with many speakers even remaining unaware of their existence. The resulting semi-artificial and therefore secondary varieties were unable halt the impending demise of the primary varieties of Occitan.

Nonetheless, they dominate contemporary school books as well as the media (newspapers, music, radio, television, blogs) and are practiced by neo-speakers who naturally imitate the semantic and formal models provided by French. This realization highlights the necessity of encouraging linguistic fieldwork, which would allow primary Occitan varieties to be recorded before their extinction, but also lexicological analysis with the object of highlighting the richness of dialectal varieties and the expressivity of contemporary literature.

5. Conclusions and Perspectives

The trajectory of lexical development in Occitan represents a special case amongst the Romance languages, rendering it ideal for methodological studies. From a chronological point of view, this development essentially takes place between the 4th and the 20th century. From a geolinguistic point of view, the Occitano-Gascon dialect group is highly diversified and it must also be observed within the context of its neighbouring varieties: Catalan, Aragonese, Poitevin, Francoprovençal, Alpine and north-eastern Italian. Contrary to other minority Romance languages currently in existence, Occitan developed a written tradition very early on and underwent four centuries of significant lexical elaboration, making it one of the principal Romance languages of the Middle Ages.

The Occitano-Gascon varieties are very well documented: with the exception of documentary texts, medieval texts are relatively well-edited and rich lexicological inventories of high quality exist for the modern dialects. Only the period between the 16th and the 18th century still requires a systematic editorial effort. Nevertheless, the lexical material pertaining to both the medieval and modern period has not yet been integrated into the existing lexicographical reference works. The current patchy state of lexicography hinders systemic lexicological inquiry. Thus, only a limited number of interpretive studies has been carried out on the Occitan lexicon to date.

It is, however, certain that a thorough investigation into the history of Occitan vocabulary would yield a significant typological contribution to the study of the mechanisms of lexical variation and evolution which characterize the Romance language family as a whole.

References

The abbreviations used in this article correspond to those used in the FEW and the DOMBibl.

1 Dictionaries and electronic databases

ALF = Gilliéron, J. & Edmont, E. (1902–1910). *Atlas linguistique de la France* (13 vols.). Champion.

ALFnum = Goebel, H. (2019). *Dialectométrie: Atlas linguistique de la France*. <http://dialektkarten.ch/dmviewer/index.fr.html>

ALG = Séguy, J., Allières, J., & Ravier, X. (1914–2000). *Atlas linguistique et ethnographique de la Gascogne* (vols. 1–4) / *Atlas linguistique de la Gascogne* (vols. 5–6) (6 vols.). Centre National de la Recherche Scientifique / Institut d'études méridionales de la Faculté des lettres.

ARTEM = AA.VV. (2010). *Chartes originales antérieures à 1121 conservées en France*. <http://www.cn-telma.fr/originaux>

BEdT = Asperti, S. (Dir.), *Bibliografia Elettronica dei Trovatori*. http://www.bedt.it/BEdT_04_25/index.aspx

CAO = Careri, M. (Dir.) (in preparation). *Corpus dell'antico occitano*, <http://www.ri-alto.unina.it/Cao>

Chartae Galliae = AA.VV. (2014). *Chartae Galliae*. <http://www.cn-telma.fr/chartae-galliae>

Codi = Duval, F. (2019). *Miroir des classiques. Section Corpus juridique* (Codex, Justinianus I / Sommes au Code / *Lo Codi occitan*) (2019 version). http://elec.enc.sorbonne.fr/miroir_des_classiques

COM2 = Ricketts, P. (Dir.) (2005). *Concordance de l'Occitan Médiéval – The Concordance of Medieval Occitan. Les Troubadours. Les textes narratifs en prose*. CD ROM.

DAG = Baldinger, K. (Dir.) (1975–2021). *Dictionnaire onomasiologique de l'ancien gascon* (23 fasc.). Niemeyer; De Gruyter.

DAGBibl = (2021). *Bibliographie du DAGél*. <https://christine.adw.uni-heidelberg.de/siglum/SiglumIndexPage?4>

DAGél = Glessgen, M. (Dir.) (2020–). *Dictionnaire d'ancien gascon électronique*. <https://dag.adw.uni-heidelberg.de>

DAO = Baldinger, K. (Dir.) (1975–2007). *Dictionnaire onomasiologique de l'ancien occitan* (10 fasc.). Niemeyer; De Gruyter.

History of the Occitano-Gascon Lexicon

DAOA = Olivier, Ph. (2009). *Dictionnaire d'ancien occitan auvergnat: Mauriacois et Sanflorain, (1340-1540)*. Niemeyer.

DÉRom I = Buchi, É., Schweickard, W. (2014). *Dictionnaire Étymologique Roman (DÉRom) 1. Genèse, méthodes et résultats*. De Gruyter.

DÉRom II = Buchi, É., Schweickard, W. (2016). *Dictionnaire Étymologique Roman (DÉRom) 2. Pratique lexicographique et réflexions théoriques*. De Gruyter.

DiTMAO = Bos, G., Corradini, M. S., & Mensching, G. (in preparation). *Dictionnaire des termes médico-botaniques de l'ancien occitan*. In collaboration with the *Istituto di Linguistica Computazionale* of the CNR.

DOM = Stimm, H., Stempel, W.-D., & Selig, M. (Dir.) (1996–2020). *Dictionnaire étymologique de l'occitan médiéval*. Niemeyer. <http://www.dom-en-ligne.de>

FEW = Wartburg, W. v. et al. (Dir.) (1922–2003). *Französisches etymologisches Wörterbuch. Eine Darstellung des galloromanischen Sprachschatzes* (25 vols.). Zbinden; Schroeder.

FEWCompl = Chauveau, J.-P., Greub, Y., & Seidl, C. (2010). *Französisches Etymologisches Wörterbuch. Complément*, Éditions de linguistique et de philologie.

FEWél = Provisional version of the FEW published online. <http://www.atilf.fr/few>

Jonas = IRHT (Romance section) (2013), *Répertoire des textes et des manuscrits médiévaux d'oc et d'oïl* (version 2013). jonas.irht.cnrs.fr

Lv = Levy, E. (1894–1924), *Provenzalisches Supplement-Wörterbuch: Berichtigungen und Ergänzungen zu Raynouards Lexique roman* (8 vols.). Reiland.

LvBibl = Baldinger, K., & Dieckmann Sammet, D. (1983), *Complément bibliographique au Provenzalisches Supplementwörterbuch de Emil Levy*. Slatkine.

LvP = Levy, E. (1909). *Petit dictionnaire provençal-français*. Winter.

RIALTO = (2019) *Repertorio informatizzato dell'antica letteratura trobadorica e occitanica* (2019 version). <http://www.rialto.unina.it>

Rn = Raynouard, F. J. M. (1838–1844). *Lexique roman, ou Dictionnaire de la langue des troubadours, comparée avec les autres langues de l'Europe latine* (6 vols.). Silvestre.

THESOC = Brun-Trigaud, G., Carpitelli, E., Casagrande, S., Dalbera, J.-P., Del Giudice, P., Georges, P.-A., Malfatto, A., Molinu, L., Oliviéri, M. Ranucci, J.-C. (1992), *Thesaurus occitan 'THESOC'. Base de données linguistiques*, Laboratoire «Bases, Corpus, Language», UMR 6039, Université Nice Sophia Antipolis/CNRS. <http://thesaurus.unice.fr>

Tolosana = (2020). *Bibliothèque numérique patrimoniale des universités toulousaines*. SICD. <http://tolosana.univ-toulouse.fr/fr/corpus>

2 Monographs and articles

Adams, E. L. (1913). *Word-Formation in Provençal*. Macmillan.

Adams, J. N. (2007). *The Regional Diversification of Latin 200 BC – AD 600*. Cambridge University Press.

Allières, J. (1995). Les scriptae occitanes V. Gascogne, Béarn. *LRL*, II/2, 450–466.

Baldinger, K. (1962). La langue des documents en ancien gascon. *RLiR*, 26, 331–362.

Baldinger, K. (Ed.) (1971). *Walther von Wartburg (1888-1971): Beiträge zu Leben und Werk, nebst einem vollständigen Schriftenverzeichnis, ZrP*, 87 [= special issue].

Bec, P. (1991). Okzitanisch: norme et standard. *LRL*, V/2, 45–58.

Bernhard, G. (1988). *Die volkstümlichen Pflanzennamen im Val d'Aran (Zentralpyrenäen)*. Gottfried Egert Verlag.

Bernissan, F. (2010). Combien de locuteurs compte l'occitan en 2012? *RLiR*, 76, 467–512.

Bernsen, M. (2006). Geschichte der Literatursprache in der Romania: Okzitanisch. *RSG*, II, 1980–1996.

Billy, D., & Chambon, J.-P. (1990). Pour la mise à jour du «Thesaurus Galloromanicus»: Un article d'essai: ATTEGIA. *RLiR*, 54, 61–91.

Boisgontier, J. (Ed.), & Arnaudin, F. (2001). *Dictionnaire de la Grande-Lande* (2 vols.). Éditions Confluences.

Boyer, H., Fournier, G., Gardy, P., Martel, P., Merle, R., & Pic, F. (1989). *Le texte occitan de la période révolutionnaire (1788-1800)*. Section Française de l'Association Internationale d'Études Occitanes.

Boyer, H., & Gardy, P. (Eds.) (2001). *Dix siècles d'usages et d'images de l'occitan: Des troubadours à l'internet*. L'Harmattan.

Brun, A. (1923). *Recherches historiques sur l'introduction du français dans les provinces du Midi*. Champion.

Brun, A. (1927). *La langue française en Provence de Louis XIV au Félibrige*. Siège de l'Institut Historique de Provence.

History of the Occitano-Gascon Lexicon

Brun-Trigaud, G. (2016). Vers un renouveau des atlas linguistiques régionaux. *Géolinguistique*, 16, 7–20.

Brun-Trigaud, G., Malfatto, A., & Sauzet, M. (ms.) (2017). Essai de typologie des aires lexicales occitanes: regards dialectométriques (Conference, 12th Congress of the AIEO, Albi 10-15. July, 2017).

Brunel, C. (1926; 1952). *Les plus anciennes chartes en langue provençale: recueil des pièces originales antérieures au XIII^e siècle*, and id. *Supplément*. Picard.

Brunel, C. (1935). *Bibliographie des manuscrits littéraires en ancien provençal*. Droz.

Carles, H. (2011). *L'émergence de l'occitan pré-textuel: Analyse linguistique d'un corpus auvergnat (IX^e-XI^e s.)*, Éditions de linguistique et de philologie.

Carles, H. (2013). L'innovation lexicale chez Chrétien de Troyes. *Romania*, 131, 281–337.

Carles, H. (2017). *Trésor galloroman des origines – Les trajectoires étymologiques et géolinguistiques du lexique galloroman en contexte latin (ca 800–1120)*. Éditions de linguistique et de philologie.

Carles, H. (2020). Les trajectoires du français régional en domaines occitan et francoprovençal. id. & M. Glessgen (Eds.), *L'écrit familial au début du XX^e siècle. L'apport de la correspondance des Poilus*. Éditions de linguistique et de philologie.

Carles, H., & Glessgen, M. (2015). La philologie linguistique et éditoriale. In W. Schweickard & C. Polzin-Haumann (Eds.), *Manuel de linguistique française* (Manuals of Romance Linguistics, 8) (pp. 108–130). De Gruyter.

Carles, H., Dallas, M., Glessgen, M., & Thibault, A. (2019). *Französisches Etymologisches Wörterbuch: Guide d'utilisation*. Éditions de linguistique et de philologie; Société de Linguistique Romane.

Carles, H., & Glessgen, M. (2019). L'élaboration scripturale du francoprovençal au Moyen Âge, *ZrP*, 135, 68–157.

Carles, H., & Glessgen, M. (2020). Le vocabulaire régional dans la langue d'oïl orientale et sud-orientale au Moyen Âge. In id. (Dir.), *Dictionnaire des régionalismes du français médiéval*. Éditions de linguistique et de philologie.

Carles, H., & Glessgen, M. (Manuscript in preparation). Le lexique de *Mirèò*.

Chambon, J.-P. (1997). Les emprunts du français moderne aux dialectes ou patois: une illusion d'optique en lexicologie historique?, *Lalies*, 17, 33–53 [= Chambon, 2017, I, pp. 381–401].

Chambon, J.-P. (2010a). Développement et problèmes actuels des études occitanes. *Comptes rendus des séances de l'Académie des Inscriptions et Belles-Lettres*, 154, 877–889.

Chambon, J.-P. (2010b). Brèves remarques sur le *Tresor dóu Felibrige* de Frédéric Mistral. *RLiR*, 74, 199–214.

Chambon, J.-P. (2010c). Prolégomènes à l'étude historique des emprunts de l'occitan au français. In A. Thibault (Ed.), *Galicismes et théorie de l'emprunt linguistique* (pp. 45–70). L'Harmattan [= Chambon, 2017, I, pp. 571–595].

Chambon, J.-P. (2012). Histoire de l'occitan (et du français) dans le domaine auvergnat: progrès récents en linguistique et en philologie (bilan et bibliographie). *Bulletin historique et scientifique de l'Auvergne*, 113, 105–123 [= Chambon, 2017, II, pp. 1121–1139].

Chambon, J.-P. (2014). Sur la date de la fixation de l'article défini devant les noms de lieux: sondages en domaine occitan». *Estudis Romànics*, 36, 373–381 [= Chambon, 2017, II, pp. 1141–1149].

Chambon, J.-P. (2017). *Méthodes de recherches en linguistique et en philologie romanes* (2 vols.). Textes choisis et présentés par É. Buchi, H. Carles, Y. Greub, P. Rézeau & A. Thibault, Éditions de linguistique et de philologie.

Chambon, J.-P., & Greub, Y. (2000). Données nouvelles pour la linguistique galloromane: les légendes monétaires mérovingiennes. *Bulletin de la Société de Linguistique de Paris*, 95, 147–182.

Chambon, J.-P., & Olivier, Ph. (2000). L'histoire linguistique de l'Auvergne et du Velay: notes pour une synthèse provisoire. *Travaux de Linguistique et de Philologie*, 38, 83–153.

Chambon, J.-P., & Greub, Y. (2002). Note sur l'âge du (proto)gascon. *Revue des Langues Romanes* 66, 473–495 [= Chambon, 2017, I, pp. 51–73].

Chambon, J.-P., & Carles, H. (2007). À propos du traitement des emprunts à l'occitan dans le TLF, In P. Rézeau (Ed.), *Richesses du français et géographie linguistique* (pp. 313–325). De Boeck.

Chambon, J.-P., & Grémois, E. (2007). Pour la révision de l'article *cheire* du *Trésor de la Langue Française*. In P. Rézeau (Ed.), *Richesses du français et géographie linguistique* (pp. 327–360). De Boeck.

Coromines, J. (1990). *El parlar de la Vall d'Aran*. Curial Edicions Catalanes.

Courouau, J.-F. (2008). *Moun lengatge bèl: les choix linguistiques minoritaires en France (1490-1660)*. Droz.

Courouau, J.-F. (Ed.) (2015). *La langue partagée: écrits et paroles d'Oc: 1700-1789*. Droz.

History of the Occitano-Gascon Lexicon

- Courouau, J.-F. (2017). *Le rococo d'oc: une anthologie poétique (1690-1789)*. Presses universitaires du Midi.
- Eygun, J. (2001). Les langues du sacré en pays occitan (XVIe–XIXe siècles). In Boyer & Gardy, 2001, pp. 177–220.
- Felixberger, J. (2003). Sub-, Ad- und Superstrate im etymologischen Wörterbuch des Französischen. *RSG*, 1, 594–607.
- Gamillscheg, E. (1934–1936). *Romania Germanica* (3 vols.). De Gruyter. (partial revision of vol. 1 (1970). *Die Franken*. De Gruyter).
- Gardy, P. (1985). *L'Écriture occitane aux XVI^e, XVII^e et XVIII^e siècles: origine et développement d'un théâtre occitan à Aix-en-Provence (1580-1730): L'œuvre de Jean de Cabanes* (2 vols.). Atelier de reproduction de thèses; Centre international de documentation occitane.
- Gardy, P. (1996). *L'écriture occitane contemporaine. Une quête des mots*. L'Harmattan.
- Gardy, P. (1997). *Histoire et anthologie de la littérature occitane*, vol. 2. [only published vol. to date] *L'âge du Baroque (1520-1789)*. Presses du Languedoc.
- Gebhardt, K. (1974). *Das Okzitanische Lehngut im Französischen*. Lgd.
- Glessgen, M. (1989). «*Lo Thesaur del hospital de Sant Sperit*»: *Edition eines Marseiller Urkundeninventars (1399-1511) mit sprachlichem und geschichtlichem Kommentar unter besonderer Berücksichtigung des Rechtswortschatzes*. Niemeyer.
- Glessgen, M. (1995). Les scriptae occitanes III.a) Provence. *LRL*, II/2, 425–434.
- Glessgen, M. (2012a). *Linguistique romane* (2nd edition). Armand Colin.
- Glessgen, M. (2012b). Trajectoires et perspectives en scriptologie galloromane. *Medioevo Romano*, 36, 5–23.
- Glessgen, M. (2016). La régionalité lexicale au Moyen Âge: hypothèses, constats et perspectives. In id. & D. Trotter (Eds.), *La régionalité lexicale au Moyen Âge* (pp. 1–34). Éditions de linguistique et de philologie.
- Glessgen, M. (2018). Les 'Documents linguistiques' de la France. Histoire, présent et perspectives d'un projet centenaire. *Comptes rendus des séances de l'Académie des Inscriptions et Belles-Lettres (CRAI)*, 161, 1239–1270.
- Glessgen, M. (2019). L'apport des 'Inconnus' du FEW à la recherche étymologique, *Italia Dialettale*, 80, 633–676.

History of the Occitano-Gascon Lexicon

- Glessgen, M. (2020). Le plurilinguisme en France au début du 20^e siècle – perception et réalité. In H. Carles & M. Glessgen (Eds.), *L'écrit familial au début du XX^e siècle. L'apport de la correspondance des Poilus*. Éditions de linguistique et de philologie.
- Glessgen, M., & Dallas, M. (2019). L'intégration du vocabulaire des *Documents linguistiques gal-loromans* dans le DEAF électronique, *Lexicographica*, 35, 235–267.
- Glessgen, M., & Pfister, M. (1995a). Okzitanische Koine. *LRL*, II/2, 406–412.
- Glessgen, M., & Pfister, M. (1995b). Okzitanische Skriptaformen I. Limousin/Périgord. *LRL*, II/2, 412–419.
- Keller, H.-E., Geuljans, R., & Wartburg, W. v. (1969). *Bibliographie des patois galloromans (1550-1967)* (2nd edition). Droz.
- Kiesler, R. (2006). Sprachkontakte: Arabisch und Galloromania. *RSG*, II, 1648–1655.
- Klingebiel, K. (2011). *Bibliographie linguistique de l'occitan médiéval et moderne (1987-2007)*. Brepols.
- Krefeld, T. (1987). *Das französische Gerichtsurteil in linguistischer Sicht: Zwischen Fach- und Standessprache*. Lgd.
- Lafont, R. (1974). *Baroques occitans. Anthologie de la poésie en langue d'oc (1560-1660)*. Aubanel.
- Lafont, R. (1991). Okzitanisch. Histoire interne de la langue II. Lexique. *LRL*, V/2, 18–23.
- Lafont, R., & Anatole, C. (1970). *Nouvelle histoire de la littérature occitane* (2 vols.). PUF.
- Lambert, P.-Y. (2003). *La langue gauloise*. Éditions Errance.
- Lieutard, Hervé (2010). Le système graphique de l'abbé Boissier de Sauvages: défense et illustration de la diglossie franco-occitane. *Lengas*, 68, 105–144.
- LRL = Holtus, G., Metzeltin M., & Schmitt, C. (Eds.) (1988–2005). *Lexikon der romanistischen Linguistik* (8 vols.). Niemeyer.
- Massourre, J.-L. (2016). *Dictionnaire du gascon des vallées de Luz, de Barèges et de Gavarnie* (Langues et civilisations romanes). Éditions Jean-Louis Massourre.
- Menichetti C. (2016). La traduction intra-romane en contexte religieux: la genèse des Actes des apôtres en occitan vaudois. In S. Maffei Boillat & A. Corbellari (Eds.), *L'aventure du sens. Mélanges de philologie provençale en l'honneur de François Zufferey* (pp. 147–175). Éditions de linguistique et de philologie.

History of the Occitano-Gascon Lexicon

- Menichetti, C. (Manuscript in preparation). *Répertoire critique des manuscrits littéraires en ancien occitan*.
- Merle, R. (1990). *L'écriture du provençal entre 1775 et 1840*. Centre International de Documentation Occitane.
- Montigel, S. (Manuscript in preparation a). Solidarités anciennes de la Galloromania et de l'Italoromania (ca 800-1120). Online publication.
- Montigel, S. (Manuscript in preparation b). Le statut lexical du gascon médiéval. Étude typologique du domaine occitano-gascon. In J. Härmä & L. Schøsler (Eds.), *Actes du 29^e Congrès International de Philologie et Linguistique Romanes* (Kobenhavn, 1-6 juillet 2019). Société de Linguistique Romane/Éditions de linguistique et de philologie.
- Müller, B. (1982). Geostatistik der gallischen/keltischen Substratwörter in der Galloromania. In O. Winkelmann & M. Braisch (Eds.), *Festschrift für Johannes Hubschmid zum 65. Geburtstag. Beiträge zur allgemeinen, indogermanischen und romanischen Sprachwissenschaft* (pp. 603–620). Francke.
- Payrató, Ll. (1991). Katalanisch. Evolución lingüística interna II. Léxico. *LRL*, V/2, 152–162.
- Pfister, M. (1973). La répartition géographique des éléments franciques en gallo-roman. *RLiR*, 37, 127–149.
- Pfister, M. (1978). Le superstrat germanique dans les langues romanes. In A. Varvaro (Ed.), *Atti del XIV Congresso internazionale di linguistica e filologia romanza* (vol. I, pp. 49–97). Macchia-rolì. [= In M. Glessgen & W. Schweickard (Eds.), *Ex traditione innovatio, I, Max Pfister. Scripta minora selecta* (pp. 305–353). Wiss. Buchgesellschaft].
- Pfister, M. (1993a). Rapport sur l'état de la recherche en lexicologie médiévale dans le domaine occitan. In G. Gasca Queirazza (Ed.), *Atti del Secondo Congresso Internazionale della Associazione Internazionale d'Études Occitanes (Torino, 31 agosto - 5 sett. 1987)* (vol. 2, pp. 923–954). Università di Torino.
- Pfister, M. (1993b). L'avenir de la recherche lexicographique de l'ancien occitan. In S. Gregory & D. Trotter (Eds.), *De mot en mot. Aspects of Medieval Linguistics. Essays in Honour of William Rothwell* (pp. 161–171). University of Wales Press.
- Pfister, M. (1998). Germanisch-romanische Sprachkontakte. *LRL*, 7, 231–245.
- Pfister, M. (1999). La lexicographie de l'ancien occitan. *La France latine*, 129, 151–160.
- Pfister, M. (2000). Überblick über die altokzitanische Lexikologie 1986-1998. In A. Rieger (Ed.), *Okzitanistik und Provenzalistik. Geschichte und Auftrag einer europäischen Philologie* (pp. 73–96). Peter Lang.

History of the Occitano-Gascon Lexicon

- Pfister, M. (2011). Bilan de la linguistique de l'ancien occitan. In A. Rieger (Ed.), *L'Occitanie invitée de l'Euregio. Actes du Neuvième Congrès International de l'Association Internationale d'Études Occitanes (Aix-la-Chapelle, 24-31 août 2008)* (vol. 1, pp. 73–100). Shaker Verlag.
- Pic, F. (1989). Essai d'inventaire des textes en occitan de la période révolutionnaire (1788-1800). In Boyer et al., 1989, pp. 13–216).
- Rainer, F. (2018a). Patterns and niches in diachronic word formation: the fate of the suffix *-men* from Latin to Romance. *Morphology*, 28, 397–465.
- Rainer, F. (2018b). La descendencia colectiva, aumentativa e intensiva del sufijo latino *-ARIA* en las lenguas románicas. *Verba*, 45, 225–264.
- Rainer, F. (forthcoming). Instrument nouns in *-ONE* in Latin and Romance. In ##, Cambridge University Press.
- Ravier, X. (1991). Okzitanisch. Les aires linguistiques. *LRL*, V/2, 80–105.
- Rivière, J.-C. (1985). *Sens et poésie. Étude lexicale de l'œuvre poétique de Frédéric Mistral* (2 vols.). CID.
- RLiR = (1925–). *Revue de linguistique romane*. Société de Linguistique Romane.
- Rohlf, G. (1931). Zur Kenntnis der Pyrenäenmundarten. *RLiR*, 7, 119–169.
- Rohlf, G. (1966). *Einführung in das Studium der romanischen Philologie. Allgemeine Romanistik, französische und provenzalische Philologie*. Winter.
- Rohlf, G. (1970). *Le gascon. Études de philologie pyrénéenne* (2nd edition). Niemeyer.
- Ronjat, J. (1930–1941). *Grammaire istorique des parlers provençaux modernes* (4 vols.). Société des Langues Romanes.
- Rostaing, C. (1971). Utilisation des variantes dialectales dans l'œuvre de Mistral, *Travaux de linguistique et de littérature*, 9, 323–328.
- RSG = Ernst, G., Glessgen, M., Schmitt, C., & Schweickard, W. (Eds.) (2003, 2005 & 2008). *Romanische Sprachgeschichte / Histoire linguistique de la Romania* (3 vols.). De Gruyter.
- Schlieben-Lange, B. (1991). Okzitanisch: Grammatikographie und Lexikographie. *LRL*, V/2, 105–126.
- Schmitt, Ch. (1974). *Die Sprachlandschaften der Galloromania: Eine lexikalische Studie zum Problem der Entstehung und Charakterisierung*. Herbert Lang.
- Stéfanini, J. (1964). Dictionnaires provençaux inédits du XVIII^e siècle, *RLiR*, 28, 123–126.

History of the Occitano-Gascon Lexicon

Straka, G. (1987). *Langues et parlers de France au Moyen Âge: quelques considérations sur la liberté de leur emploi*. Ville de Montbrison.

Swiggers, P. (1998). Langues en contact: l'occitan et les langues romanes. *LRL*, 7, 67–82.

Ternes, E. (1998). Keltisch und Romanisch. *LRL*, 7, 266–291.

Teulat, R. (1979). Le problème de l'écrit occitan depuis 1914. In A. Armengaud & R. Lafont (Eds.), *Histoire d'Occitanie* (pp. 883–910). Hachette.

Thomas, J. (2013). À propos de la découverte d'un dictionnaire inédit du XVIIIe siècle, *Lengas*, 73. Reproduction of ms.: <http://tolosana.univ-toulouse.fr/notice/165803770>

Tomàs, J. (2006). *Lingüística e renaissentisme occitan*. Institut d'Études Occitanes.

Trotter, D. (2006). Contacts linguistiques intraromans: roman et français, occitan. *RSG*, II, 1776–1785.

Séguy, J. (1953). *Les noms populaires des plantes dans les Pyrénées centrales* (Monografías del Instituto de estudios Piranaicos. Filología, 18). Consejo Superior de Investigaciones Científicas.

Wartburg, W. v. (1952). Die griechische Kolonisation in Südgallien und ihre sprachlichen Zeugen im Westromanischen, *ZrP*, 68, 1–48.

Weth, C. (2014). Le provençal / L'occitan. In A. Klump, J. Kramer & A. Willems (Eds.), *Manuel des langues romanes / Manual of Romance Languages* (pp. 491–509). De Gruyter.

Zufferey, F. (2008). Histoire externe de l'occitan. *RSG*, 3, 2998–3020.

ZrP = (1877–). *Zeitschrift für romanische Philologie*. De Gruyter.

Hélène Carles

Université de Strasbourg

Martin Glessgen

University of Zurich / ÉPHÉ-PSL