

C14

CHAPTER 14

 REGIONAL VARIATION IN
 THE FRENCH OF FRANCE

MATHIEU AVANZI

C14S1

 14.1 INTRODUCTION

C14P1

A range of types of linguistic system coexist in France.¹ There is the reference variety of French, which is the national variety that is considered standard (see 14.1.2).² Then there are regional varieties of French, which are particular forms of French spoken in specific regions and may have lexical, grammatical, phonetic, or other characteristics that distinguish them from Standard French (14.1.3). There are also regional languages that belong, like French, to the family of Gallo-Romance languages (such as Normand, Picard, Provençal, or Franco-Provençal) or to other linguistic families, including the wider Romance family (Catalan, Occitan), other Indo-European families (Celtic: Breton; Germanic: Flemish, Alsatian), and the non-Indo-European language, Basque (see Chapter 20, this volume).

C14P2

The terms *patois* ('patois') and *dialecte* ('dialect') are often used by lay speakers to refer either to regional varieties of French or to the regional languages and their own different varieties (see Chapter 20, this volume). From a technical linguistic perspective, the difference between 'patois', 'dialect', and 'regional language' is not straightforward, since these terms are often used interchangeably. What really distinguishes them is the prestige status of the idioms which they designate. 'Patois' is a derogatory term typically used to refer to a variety spoken within a provincial village that is not, or barely, understandable by people who are not from that village (Boyer 2021). The term 'dialect' is used to refer to a variety that is not standardized and is used only in informal contexts, regardless of the size of the community of people who speak it. Finally, 'regional language' refers to a dialect that has somewhat a

¹ I would like to thank André Thibault, the three anonymous reviewers, and the editors of this handbook, who provided helpful comments on earlier drafts of the manuscript. Any mistakes or flaws that might remain are mine and mine alone.

² On the external history of French and the development of a standard language, see Chapters 4, 5, and 6, this volume.

level of political recognition, for example because it is used in the media, taught in school, or has some kind of official status at the regional level.

C14S2 14.1.1 Demographics

C14P3 In metropolitan France (including Corsica), French is the only official language recognized by the state (Article 2 of the French Constitution states, ‘La langue de la République est le français’ (‘The language of the Republic is French’)).³ Other regional languages, deriving from Gallo-Romance (Oïl dialects spoken in the north, such as Picard, Normand, or Poitevin; Oc dialects spoken in the south, such as Gascon, Languedocian, or Provençal; Franco-Provençal, spoken in south-eastern Alpine areas, such as Lyonnais and Savoyard) or belonging to other language families (Celtic: Breton; Germanic: Alsatian and Mosellan; Basque or Corsican), are recognized as part of the country’s linguistic heritage but have no official status (Article 75-1 of the French Constitution states, ‘Les langues régionales appartiennent au patrimoine de la France’ (‘Regional languages are part of France’s cultural heritage’)).⁴

C14P4 According to the 2018 report published by the *Organisation internationale de la Francophonie* (OIF),⁵ an estimated 66,060,000 people speak French in France (including French Overseas Departments and Territories), which represents 96.5% of the entire population. Among the most recent and reliable censuses on French regional languages, Bernissan’s (2012) study estimates that 267,872 people speak Occitan (Oc dialects spoken in the southern half of France) in the area where Occitan dialects were ancestrally spoken in France (excluding Italy and Spain), which represents 1.8% of the population. As for Breton, Broudic (2013a) states that it is spoken by 5.3% of the population living in Celtic Brittany (which represents some 206,000 speakers). Statistics for other regional languages are no more encouraging (see Kremnitz 2013 for a comprehensive list of references; see also Chapter 20, this volume). Nowadays, regional languages are facing extinction: among the speakers of a regional language, none are monolingual, and most are above age sixty. This can be explained by the fact that the intergenerational transmission process has critically slowed down since the early-twentieth century (Filion 2010).⁶

³ *Texte intégral de la Constitution du 4 octobre 1958 en vigueur*, <https://www.conseil-constitutionnel.fr/le-bloc-de-constitutionnalite/texte-integral-de-la-constitution-du-4-octobre-1958-en-vigueur> accessed 7 Jul. 2022. On language legislation in general, see Chapter 5, this volume.

⁴ <https://www.legifrance.gouv.fr/loda/id/LEGITEXT000006071194/#75-1> accessed 7 Jul. 2022.

⁵ *La Langue française dans le monde. Synthèse 2018* (<https://observatoire.francophonie.org/2018/synthese.pdf> accessed 7 Jul. 2022).

⁶ In some areas, such as Hauts-de-France or southern France, the coexistence of Gallo-Romance regional languages and French led to some hybrid systems, such as the variety of Picard strongly influenced by French described in Hornsby (2006). In this chapter, in contrast to Hornsby, I do not use the expression ‘regional French’ to refer to a variety that mixes the properties of a regional language with those of French; I use ‘regional French’ to characterize the linguistic properties (lexical, grammatical, phonetic, or phonological) that French overlays locally.

C14S3 14.1.2 Reference French

C14P5 Like any other standardized language that is widely distributed throughout the world, French varies in several ways, leading to some linguistic usages or forms that are recognized as standard, normative, academic, central, etc. and others, conversely, that are considered non-standard, incorrect, popular, peripheral, etc. When dealing with regional variation in the French of mainland France—i.e., how this language varies from a geographical point of view—the notion of ‘Reference French’ (henceforth RF) is commonly used to characterize this normative or standard variety, against which Metropolitan French speakers assess the linguistic legitimacy of their speech or that of their contemporaries (Francard et al. 2000).⁷

C14P6 From a theoretical point of view, RF is a complex concept, even difficult to grasp, since it can be defined negatively as the sum of the uses of French that are recorded without diatopic markings in the bestselling dictionaries (from publishers such as Larousse and Robert) and grammars (Riegel et al. 2004; Grevisse and Goosse 2016) in Europe (Poirier 1995: 26). RF is also defined as the variety described in handbooks of French as a foreign language (Lyche 2010), leading Laks (2002) to conclude that RF is a sort of ‘doxic’ construction.⁸

C14P7 Geographically speaking, RF is often associated with the city of Paris and the surrounding Île-de-France region, either in the representation of speakers (Moreau et al. 2008; Detey et al. 2016a), or by scholars aiming at defining the national norms of pronunciation (Léon 1966; Martinet and Walter 1973; Carton et al. 1983; Morin 2000; Armstrong and Pooley 2010), or recording words and expressions that do not belong to ‘general’ French (Rézeau 2000). Reasons for this include the fact that, within France, Greater Paris plays the role of a ‘centre’ (in the sociological sense, defined in opposition to the concept of ‘periphery’; see Singy 1996 for an application to French dialectology).⁹ In this context, one can understand why Parisian innovations so frequently deregionalize and remain very rarely confined to their area of origin, and why Parisian regionalisms are so rare (Simoni-Aurembou 1991).¹⁰

⁷ However, internationally—both inside and outside Europe—what is considered RF may vary depending on the focus adopted (see Chapters 15, 16, 17, and 18, this volume; Thibault 2020a; Lüdi 2020; Chalier 2021). In studies dealing with diastratic or diamesic variation, the notion of ‘Standard French’ is usually preferred to that of RF (Gadet 2007, also see Chapters 10 and 12, this volume).

⁸ ‘Doxa’ is a term used in philosophy that refers to a set of generally accepted popular assumptions designated positively or negatively.

⁹ In France, the most prestigious universities are located in Paris and its suburbs; it is in these institutions that most national politicians study and learn to speak in public. In addition, all national media outlets are primarily headquartered in Paris, from where they disseminate their use of French through radio stations, television channels, and newspapers all across the country. Paris is also the home of the *Académie française* (‘French Academy’), an institution created in the seventeenth century to control the use of French through its dictionary and grammar (see 4.7; 5.2.1; 6.2.3; 28.3, this volume). Even though none of the Academicians currently are linguists, the institution still has a certain impact on French linguistic policies and RF (Cerquiglini 2021 still mentions the institution in his book on the evolution of *participe passé* (‘past participle’) rules). On the academy today, see 5.3.2.1; 5.3.5.2, this volume.

¹⁰ This role of Paris and the Île-de-France as a ‘centre’ disseminating its innovations and defining, at the territorial level, what is standard against what is peripheral is as old as the diffusion of the first *scripta* of this variety (Glessgen 2017; see also 15.2.1, this volume).

C14S4 14.1.3 What is a regionalism?

C14P8 A regionalism is traditionally defined as a linguistic feature that exists only (with a certain frequency) in the French spoken within a defined region of the territory and that is not (or barely) used outside this region. In other words, a regionalism is the specific use of a given feature which is not (or hardly) widespread throughout the territory (Arrivé et al. 1986: 597). In this sense, for instance, the use of the adverb *rien* ('nothing') to negate the verb *avoir peur* ('to be afraid'), as in the sentence *il a rien peur* ('he is not afraid of anything'), is a regionalism, because it is employed within a bounded region in the centre-east of France and is not known elsewhere (Avanzi 2020a: 127).¹¹

C14P9 This definition, while quite widespread and commonly used, is not always straightforwardly applicable. First, many regionalisms are observed not within one geographically compact and well-demarcated zone, but within several zones. This is why reference will be occasionally made to the French spoken in other countries despite the fact that this chapter focuses on regional variation in the French of mainland France. For example, the use of the word *sucre* ('sugar') to name what dictionaries record as a *serre-fils* ('cable clamp'), also commonly called *domino*, is a regionalism used in Belgium, Switzerland, and the entire south-eastern part of France. Yet such regions are not contiguous (Avanzi 2020a: 10). A second issue is quantitative. What should be the baseline percentage of territory covered to allow one to consider that a variant is regional? It is true that the local particularities of French most often extend over relatively small areas on the scale of French-speaking Europe (see, for example, the word *couque* in Figure 14.5). But this is far from a rule, and there are many cases of features spanning areas that may include up to 75% of the territory where French is spoken in Europe. For example, the verb *beugner* (derived from the noun *bugne*, the name of a local type of doughnut in the Lyon region, this verb is used to refer to getting in a minor collision) is so widespread that it is quicker to describe its geography by pointing to the areas where it is not used. In this case, it is only in the north-western quarter of France and in Wallonia that this verb is not known to speakers (Avanzi 2020a: 11). Third, a final problem, resulting directly from the previous one, concerns the status of Paris and the Île-de-France region in the demarcation of linguistic phenomena that can be defined as regional or not. Since it may be agreed that Paris and the Île-de-France are where everybody 'speaks' RF, theoretically we should consider uses outside of this area to be regional, even if they are in the majority. For example, most dictionaries record the pronunciation of the word *persil* ('parsley') without final *-l* as the standard pronunciation [pɛʁsi], because the area where speakers do not pronounce this consonant includes Paris and the Île-de-France. However, at the level of French-speaking Europe, these speakers are not in the majority: in the entire periphery of Francophonie in Europe (south, west, Brittany, northern regions, eastern fringe, French-speaking Switzerland, and the Lyon region), people say [pɛʁsil], making sure to articulate the final *-l* (Avanzi and Thibault 2021: §5.4).

C14P10 To sum up, a regionalism is a linguistic feature employed by people living in compact, geographically bounded areas which, regardless of size, exclude Paris and Île-de-France.

¹¹ In RF, the syntax is different: *il (n')a peur de rien*.

C14S5 14.1.4 Where do regionalisms come from?

C14P11 It has often been stated that French regionalisms are remnants of regional languages which are nowadays facing extinction or have already disappeared.¹² In fact, regionalisms that were borrowed from Gallo-Romance dialects or other regional linguistic systems spoken in France constitute just one category of regionalism (14.1.4.1). Some regional variants of French represent surviving linguistic structures that used to be more widespread in RF: these are called archaisms (14.1.4.2). Third, peripheral speakers can be responsible for innovations in French. Nevertheless, because of the monocentric status of French in France, such innovations are unlikely to spread widely outside the area where they were first used (14.1.4.3).¹³

C14S6 14.1.4.1 Borrowings

C14P12 This category of regionalism consists of linguistic items originally foreign to French and borrowed locally following a situation of more or less prolonged contact with another linguistic system (which would be considered a patois, dialect, or regional language, depending on its status, see Chapter 20, this volume).¹⁴ Before the diffusion of RF throughout the whole country (a process that took centuries, see 15.2.1, this volume), people used linguistic systems that were perfectly adapted to describe the world around them, whether regarding agricultural tools and techniques, everyday work, meteorological phenomena, or the ceremonies that punctuate the twelve months of the year. No wonder, then, that when dialect speakers started speaking French, they adapted their dialectal words to the French language to talk about those realities for which the language of the country they were living in did not always have a name. The number of regionalisms that belong to the categories of dialectalisms (i.e. borrowings from regional languages) is without a doubt largely exaggerated (Chambon 1997a, 1997b). Among the confirmed ones, there are some widespread verbs such as *péguer* ('to stick lightly', adapted from the Old Occitan verb *pegá*, with the same meaning, formed on the noun *pega* ('pitch')), which is used throughout the south of France (Avanzi 2020a: 182). Another common word is *drache* (from Flemish *draschen*, referring to a sudden heavy shower), known in the entire northern fringe of France and Belgium (Avanzi 2020a: 60). Many other regionalisms are used in more restricted areas. In Savoy, the word *monchû* (from the patois of the area of Chamonix, corresponding to French *monsieur*) is used to refer to a tourist, especially a Parisian one (Avanzi 2020a: 144).

C14S7 14.1.4.2 Archaisms

C14P13 Evolution and changes in French do not take place everywhere at the same time, nor at the same pace. In fact, some innovations that occurred in the Île-de-France have not yet necessarily taken place in the rest of the country. As a result, the name of an object, concept or idea, or indeed a sentence structure or pronunciation, could have changed in one place

¹² Tuailleon (1983: 10); Chaurand (1985: 367); Walter (1988: 176); for a critical point of view, see Chambon (2006).

¹³ See Poirier (1995) for the origin of this classification.

¹⁴ On the process of borrowing in general, see 19.3.1, this volume.

but not necessarily in another. ‘Archaisms’ are those linguistic items that belonged to RF at a certain time but which, for various reasons, have gradually disappeared from RF yet survive in certain regions. One well-known example is the use of the expression *à cette heure* (lit. ‘at this hour’), often written *astheure* (a spelling closer to the pronunciation), as a synonym of *maintenant* (‘now’). The locution was first attested at the end of the fifteenth century with the spelling *à cette heure*, and as early as 1530 with the spelling *astheure* (see Wartburg 1922–2002, vol. 4: 468). In literature, we find it in the works of many writers (see Grevisse and Goosse 2016: §615), but with various pragmatic effects. Today, it does not seem to enjoy the same status as it did a century or two ago, since in widely consulted dictionaries such as Robert (Rey 2013) and Larousse (*Le Petit Larousse illustré* 2022) the locution is labelled *vieilli* (‘dated’).

C14S8 14.1.4.3 Innovations

- C14P14 Finally, some regionalisms are neither borrowings from Gallo-Romance dialects or other regional languages spoken in the territory, nor relics of a state of French no longer part of RF. These regionalisms are called ‘innovations’ and they are created through various different creative processes.
- C14P15 Semantic innovations are innovations affecting the semantic reference of an existing word, that is, its meaning (see Chapter 9). This occurs when a word belonging to RF receives a particular meaning in a given region. This is what can be observed for the word *poche*, used in RF to designate a trouser pocket, but which can refer to a supermarket bag in the west of France (Avanzi 2020a: 22). The same observation can be made for the word *cornet*, used in RF to denominate a paper cone for serving chips, chestnuts, or sweets, but used in eastern France and some regions of Switzerland to name a supermarket bag (Avanzi 2020a: 18).
- C14P16 Antonomasia refers to semantic innovations which are not based on common nouns, but on proper nouns. For example, in north-western France, speakers do not call a laundry drying rack an *étendoir* as in RF, but instead use the word *tancarville*. *Tancarville* was the name of a company that made drying racks whose shape was inspired by the famous bridge in the village of Tancarville in Normandy. Until very recently, it was sold exclusively in all the north-western regions of France (Avanzi 2020a: 50).
- C14P17 Finally, as is normal in all natural languages, speakers from the regions of mainland France regularly produce internal formal innovations. In French, there is no specific single word used to designate a person who drives a taxi (one says *chauffeur de taxi* (‘taxi driver’)). However, in the area of Montpellier, French speakers created the word *taxiteur* (by adding the suffix *-teur* to the lexical base *taxi*), because most trade names in French have a proper designation (Avanzi 2020a: 19). Other formal innovations rely on alternative patterns of deriving new words. For example, in the Lyon region, *gribouiller* (‘to scribble’) and *étendre* (‘to hang out [laundry]’) did not produce deverbal nouns in *-on* (*gribouillon* (‘scribble’)) and *-oir* (*étendoir* (‘laundry drying rack’)) as in RF; there, they led to derivatives in *-age* (*gribouillage* and *étendage*, see Avanzi 2020a: 128; see also 2.2.4, this volume, on word-formation processes).
- C14P18 It is important to note that it is not always easy to differentiate between innovations, dialectalisms, and archaisms, especially in northern France due to the proximity between Oïl dialects and French. For instance, it is not possible to state which category applies best to the

word *drôle*, used nowadays in western France to designate a small child (in past centuries, the word seemed to be used in RF, but with a slightly different meaning, see Simoni-Aurembou 1999: 579). Elsewhere, the difference between dialectalisms and earlier innovations can also be problematic (see the discussion of the word *poulaille* ('rooster') in the Lyon region, once falsely considered a borrowing from Franco-Provençal, Chambon and Chauveau 2004). Such examples do not necessarily imply that the categories used above are inconsistent, but remind us that the historical trajectory of regionalisms is often complex and demands painstaking examination of sources.

C14S9

14.2 REFERENCE WORKS ON FRENCH REGIONALISMS

C14S10

14.2.1 Seventeenth to nineteenth centuries: the prevalence of *cacologies*

C14P19

The earliest written sources identifying regionalisms are almost as old as the earliest scholarly books on the French language (see Chapter 6, this volume). Such references are found in so-called *cacologies*, a textual genre which developed out of the work of authors such as Vaugelas (2018 [1647]) and consisting of lists of words, expressions, or phrases deemed 'faulty' or 'incorrect' (see 6.2.3; 6.2.4, this volume). Most of these books were written in Paris during the seventeenth and eighteenth centuries, which explains why they contain only a few regionalisms (see Ayres-Bennett and Sejjido 2011: 311–12). The ones written outside Paris, by authors from different provinces of the Realm (Gascony, Limousin, Lyon, etc.), are much richer. A comprehensive list of references can be found in Rézeau (2000: 159–61).

C14S11

14.2.2 1900–1950: early scholarly studies of regional French

C14P20

In the first half of the twentieth century, very few dialectologists showed interest in the properties of locally spoken French, which explains why there are only three significant monographs from this period. One focuses on the French spoken in Grand'Combe (a village in the department of Doubs, see Boillot 1929). The second concerns the French spoken in Marseille (Brun 1931). The third, by Séguy (1950), describes the French spoken in Toulouse. These three books are organized like grammars, where all aspects of regional French structures (lexical, grammatical, and phonetic/phonological) are taken into account. The authors drew examples from conversations they had participated in, from the newspapers they read, or from student essays they marked.

C14P21

During this period, as French started to gain competitive advantage over the *patois* in society as well as in academic research, two other important monographs were published. One is now considered to be the first scholarly dictionary of regional French, focusing on Swiss French, especially as spoken in the region of Neuchâtel (Pierrehumbert 1926). The other one was published right after World War II (Martinet 1945) and focuses on the pronunciation of

French: it is the first study aimed at describing the way that phonological contrasts operate in different regions of France. It is particularly innovative, since it is based on the examination of questionnaires filled out by more than 400 men. In addition, it presents maps that clearly isolate some regions from others within mainland France.¹⁵

C14S12 14.2.3 1950–2010: the growth of studies of regional French

C14P22 A steady stream of scholarly collaboration and advances in high-performance computing enabled a profusion of works in the second half of the twentieth century, but also contributed to more specialized descriptions. The focus here is on publications on French at the national level. Note that this overview concerns only lexicology/lexicography (14.2.3.1) and pronunciation/phonology (14.2.3.2); the description of regional grammatical variation remains relatively less developed (see 14.3.2).

C14S13 14.2.3.1 *Lexicology/lexicography*

C14P23 In the field of the lexicon, the real development of what is commonly called *lexicographie différentielle* ('differential lexicography', Bavoux 2008; Wissner 2010) started in the early 1980s with the publication of several dictionaries on the regionalisms of the French spoken in certain specific locations or regions of France (among others, Médélice 1981; Tuaillon 1983; Rézeau 1984). The research field reached its high point at the end of the millennium, with the publication of the first dictionary of regionalisms from all of mainland France (Rézeau 2001). In many respects, this dictionary is unique and important in the history of the field. First, it is not dedicated to any one region, but the entire country. Second, it comprises hundreds of entries illustrated by thousands of authentic first-hand examples drawn from a very extensive list of works. Finally, and most importantly, this dictionary is the only one featuring maps, which were generated after the examination of a 'vitality survey'. Questionnaires were sent to a small sample of people long established in the region where they were living. They were asked to indicate whether they used a given regionalism or not. The use of this methodology marks a change in the paradigm, since up until then participant observation had been the predominant approach.¹⁶

C14S14 14.2.3.2 *Pronunciation/phonology*

C14P24 In the field of pronunciation studies, the phonological questionnaire method *à la* Martinet (1945)¹⁷ was little by little abandoned in favour of field recordings thanks to the advent of portable tape recorders at the end of the 1970s. At the national level, only two major projects

¹⁵ As regards French pronunciation, the work of Martinet (1945) inspired Deyhime (1967), who conducted a survey gathering the judgements of hundreds of French speakers from different regions of France using questionnaires.

¹⁶ The technique of participant observation involves direct contact between the interviewer (usually the researcher) and the interviewees. Note that the mail survey technique was used earlier to document morphosyntactic variation of dialects spoken in the Lyon region (Clédat 1887, 1888), but also in other social science fields, such as ethnography (Febvre 1961).

¹⁷ Martinet (1945) used a technique consisting of questionnaires where participants were asked to state if they pronounced particular pairs of words in the same way or differently.

have been conducted. The first led to the publication of a monograph (Walter 1982) which describes in detail the phonological systems of a certain number of idiolects. It is based on the examination of a field survey involving thirty-five localities representative of the main dialectal areas of France, Belgium, and Switzerland, involving approximately one hundred native speakers. The second study was initiated at the beginning of the 2000s and is still ongoing. Called *Phonologie du français contemporain* (PFC, ‘Phonology of Contemporary French’), it consists of a large corpus including recordings of speakers from all over the world. Within France, a dozen groups of roughly twelve non-mobile speakers of different ages and socio-economic backgrounds were recorded following a specific protocol (text and word list reading, structured and spontaneous interviews). The project has led to numerous publications (Durand et al. 2009; Detey et al. 2010b, 2016b, 2016c; Simon 2012;) in which many varieties of French (from mainland France and beyond) are described. A comprehensive summary of the main results of these studies and others focused on local survey points can be found in Armstrong and Pooley (2010).

C14S15 14.2.4 2010 to the present: the contribution of crowdsourcing

C14P25 Given the material and technical constraints, the surveys mentioned above rely on data samples that are rather limited, both in terms of regions and the social characteristics of participants (a dozen participants for some isolated locations representing some bigger regions, in the best-case scenario). Taking advantage of the democratization of social media and smartphones, many scholars have started to launch online surveys to document regional and dialectal variation in their countries (see Avanzi and Thibault 2021 for a systematic review; see also Chapter 13, this volume).

C14P26 For French, Scherrer et al. (2015) were the first to use the research paradigm called ‘crowdsourcing’ or ‘citizen science’. They gathered the answers from just over one thousand speakers from across Europe who were asked to listen to a word pronounced in two different ways and to choose which one was closer to their own pronunciation. Using this method, they were able to map the pronunciation variants of just over thirty words.

C14P27 As part of the *Français de nos Régions* project, a dozen online questionnaires were launched within the French-speaking world (Avanzi et al. 2016; Avanzi and Thibault 2018, 2021; Avanzi 2019a, 2020b;).¹⁸ Participants were asked to provide information on their socio-economic and sociolinguistic background (country and postcode of the place where they spent most of their youth, geographical moves, age, gender, occupation, and highest education level obtained). In a second step, the participants had to answer around forty questions aiming to assess their use of a particular word or phrase. Each question was accompanied by a definition or usage context, an image and a list of answers.¹⁹ The participants were asked to select from a list of possible choices the answer(s) that best corresponded to their use. If necessary, they could select an ‘other’ box where they could freely enter text or continue to the next screen if they did not understand the question. An additional field also allowed

¹⁸ <https://francaisdenosregions.com/> accessed 17 Jul. 2022.

¹⁹ For example, to solicit the different names for a mop, a picture of a mop was accompanied by a brief definition worded as a question: ‘What do you call this piece of fabric used to wash the floor?’ This was followed by a list of items (*serpillière, panosse, wassingue, pièce, torchon*, etc.) with boxes participants could tick.

participants to leave comments after each question. In total, each questionnaire included between forty and sixty questions, which means that over the years hundreds of questions were asked. For each of them, the answers of 10,000 speakers were collected. Using this information, it was possible to create hundreds of maps (Avanzi 2017, 2019b, 2020a).²⁰

C14S16

14.3 A TYPOLOGY OF FRENCH REGIONALISMS

C14P28

Several of the maps from the *Français de nos Régions* project have been selected here to show how local features can affect different levels of French linguistic structures. From a linguistic point of view, the lexical level (14.3.1) is the best represented in these surveys, although a good portion of the questions aimed to document grammatical variation (14.3.2) and phonological variation (14.3.3).²¹

C14S17

14.3.1 Lexicon

C14P29

It is not possible to give a full overview here of the way the French lexicon varies regionally. I will focus on three items, chosen because they illustrate the three specific cases described in 14.1.4. The first one addresses the notion of ‘dialectalism’ by comparing two maps: one made from material gathered at the end of the nineteenth century (Gilliéron and Edmont 1902–10), when people still spoke Gallo-Romance dialects, the other from material gathered in the twenty-first century, as part of the *Français de nos Régions* project. The aim is to show that there is no one-to-one correspondence between both systems (14.3.1.1). The second attempts to map how an archaism is born. Using the example of the word *dîner* (‘dinner’), we see how the vitality of a regionalism varies across space but also across generations (14.3.1.2). The third traces the history of a recent innovation that has triggered a passionate debate that rages continuously on social media: the battle between *chocolatine* and *pain au chocolat* (14.3.1.3).

C14S18

14.3.1.1 Different ways of naming the ‘village fair’

C14P30

In France as in many other countries of the former Roman Empire (where Christianity has had a strong influence), the ‘village fair’ is an annual celebration that usually takes place

²⁰ The maps presented in this chapter were all generated using these materials. On these maps, the dots represent the administrative regions called *arrondissements* in France and Belgium, and *districts* in Switzerland. Due to lack of space, reference must be made to Avanzi (2019a, 2020c) for information on data selection and processing. On the epistemological aspects of the use of crowdsourcing in dialectology, see Avanzi and Thibault (2018, 2021).

²¹ The categories of regionalisms presented in 14.1.4 are used to structure the lexicon section. As for grammatical features, a different type of classification was chosen, since it can be tricky to provide the exact origin of some phenomena. The category of regionalism—archaism, innovation, or borrowing—is nevertheless mentioned when possible. As for pronunciation, since we are only dealing with the vowel system, none of the three types of regionalism applies.

on a weekend between May and October (during the hottest days of the year). It usually coincides with the celebration of the patron saint of the locality holding the fair. Nowadays, the religious origins of rural life associated with these fairs tend to have been forgotten, but the festivities remain roughly the same as a century ago: the celebrations are spread over one or more days, centred around a large meal eaten together, a ball, and various other attractions for young and old (e.g. raffles, *pétanque* competitions, merry-go-rounds or other rides, the sale of local products, etc.). Figures 14.1 and 14.2 respectively show the different ways of naming the ‘village fair’ in Gallo-Romance dialects at the end of the nineteenth century (when French was still a second language for most of the inhabitants of France) and in European French at the beginning of the twenty-first century. A comparison between the two maps highlights several features.

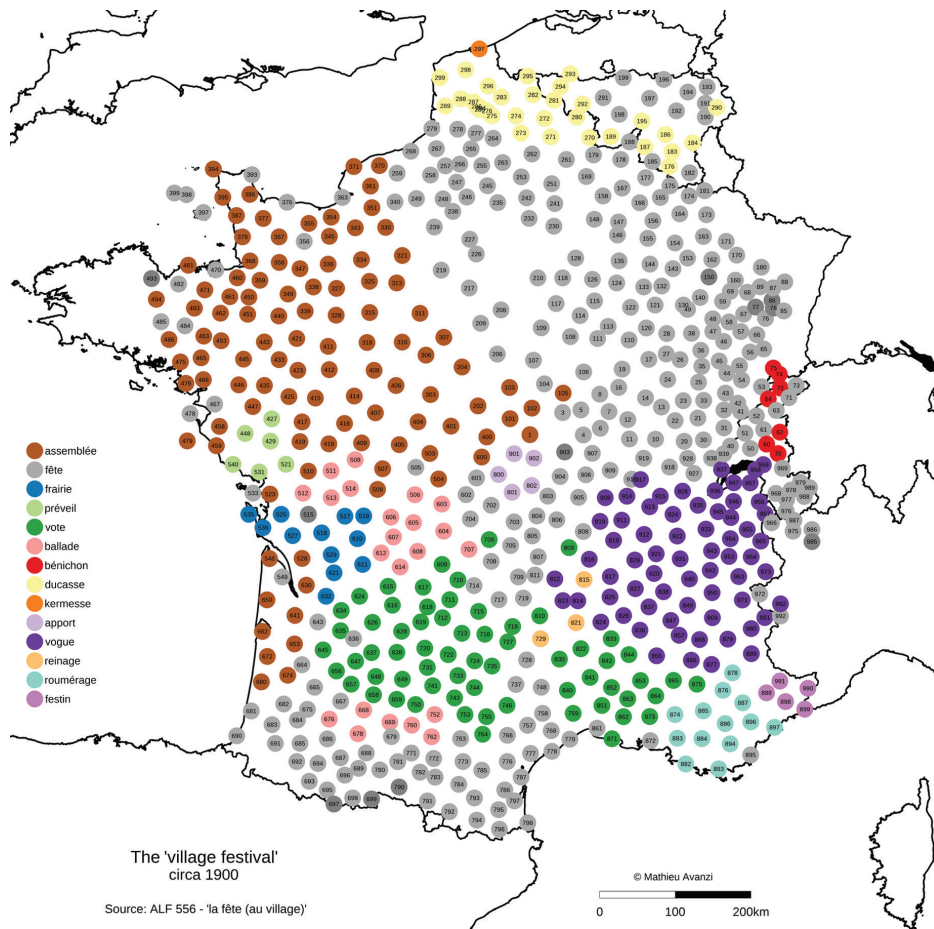
C14P31 First, we can see that at the end of the nineteenth century there were more terms to designate the village fair in dialects than there are in twenty-first-century French. According to the *Français de nos Régions* study, words such as *ballade*, *reinage*, and *roumèrage* are in fact no longer used (these variants were nevertheless included in the questionnaire).²² The second thing revealed by the comparison is the emergence of new words in regional French. In the south of France, French did not copy the Occitan word (*vota*, which would have given *vote*). Instead, a compound, *fête votive*, was created. Finally, it is worth commenting on the geographical area of some of the variants. In Belgium, *ducasse* has lost a lot of ground in the east, where it has been replaced by another variant (*kermesse* (‘fairground’)); on the other hand, it has been adopted by more people in southern Picardy. In the south, *fête votive* acted like a steamroller, replacing the more minority dialect variant *roumèrage*. In the west, we see that while the area of the word *frairie* has remained stable, the area of *assemblée* has been considerably reduced. Overall, these facts allow us to confirm that regional French is not just what remains of the patois when it died out, contrary to a widely held opinion (even among linguists). Although the two systems are closely connected, they each have their own unique dynamics, and it would be wrong to suggest that one is simply the residue of the other.

C14S19 14.3.1.2 *Dîner time*

C14P32 Today in Paris, the Île-de-France, and the surrounding departments, people have *petit-déjeuner* (‘breakfast’) shortly after getting up in the morning. They have *déjeuner* (‘lunch’) around noon, and finally *dîner* (‘dinner’) at the end of the day. Yet in some regions of French-speaking Europe, people eat *déjeuner* in the early morning, they have *dîner* at noon, and they finish with a *souper* (‘supper’) in the evening. Historically speaking, the coexistence of these two systems and their resulting hierarchy (the second being considered archaic) is quite new (Goosse 1989).

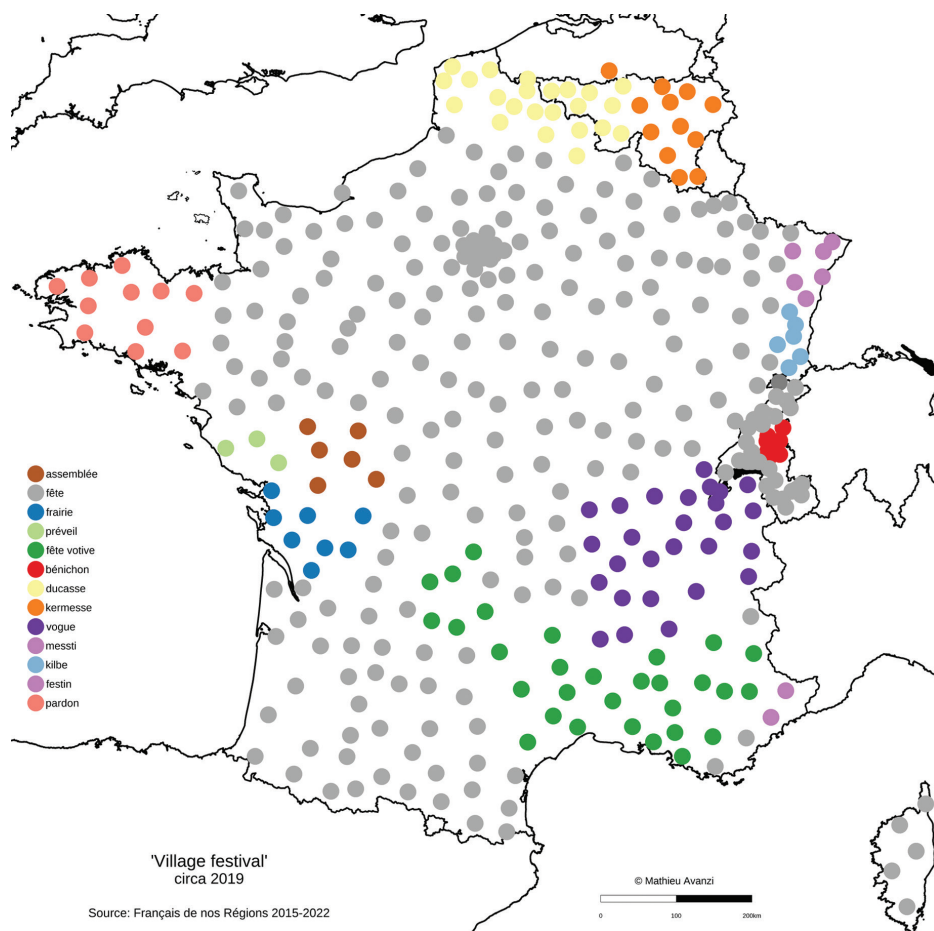
C14P33 In the Middle Ages, we know that there were different words to designate the different meals of the day. Two of them (*déjeuner* and *dîner*) originated from Latin *DISJEJUNARE (lit.

²² The *Atlas linguistique de la France* (ALF, Gilliéron and Edmont 1902–10) focused on Gallo-Romance dialects, which explains why no data are available in it for Corsica, Brittany, Alsace, Moselle, and the Basque regions. Some localities were included in the survey even though they did not belong to mainland France (Channel Islands, Wallonia, French-speaking Switzerland, and some valleys in Italy) because they belong to Gallo-Romania.



C14F1 FIGURE 14.1 Different ways of naming the ‘village fair’ in Gallo-Romance dialects (c. 1900)

‘to break the fast’). The other form, *souper*, is derived from the name of an emblematic dish taken at the end of the day, *soupe* (‘soup’). We do not really know when a distinction was made between *déjeuner* (to take the first meal of the day) and *dîner* (to take the second meal of the day), but we do know that the lexical triad became established in the sixteenth century. It has been estimated that towards the end of the eighteenth century in Paris, the *dîner* hour (i.e. the midday meal) began to shift towards the end of the afternoon, which had the consequence of delaying the time of the first meal of the day, actually triggering the emergence of two breakfasts, a first one taken very early and a second one taken a few hours later. For this reason, from the nineteenth century onwards, the need was felt to distinguish the first breakfast from the following bigger lunch. This is how the phrase *petit-déjeuner* entered dictionaries of RF, the word *déjeuner* designating in this variety the midday meal. The change did not happen overnight, as can be seen in phrases such as *déjeuner du matin* (‘morning breakfast’) as opposed to *déjeuner du midi* (‘midday lunch’), or *premier déjeuner* (‘first lunch’)



C14F2 FIGURE 14.2 Different ways of naming the ‘village fair’ in European French (c. 2019, N = 13,132)

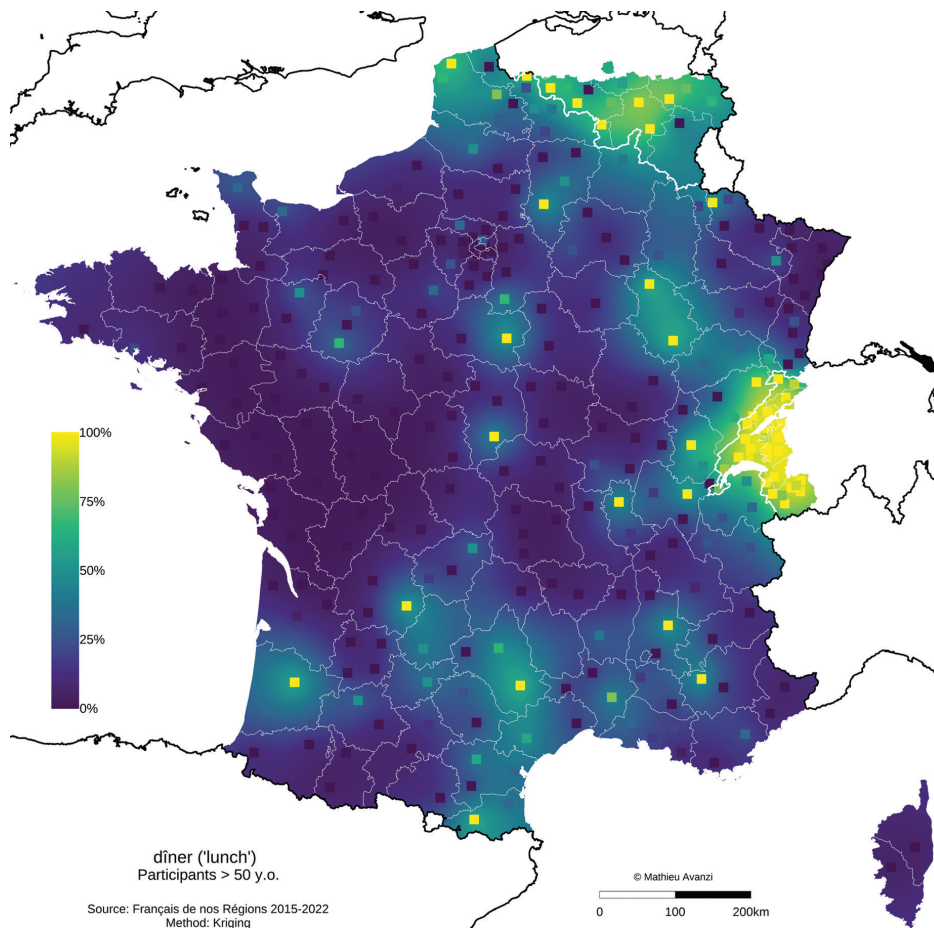
as opposed to *second déjeuner* (‘second lunch’), mostly attested at the beginning of the nineteenth century. The word *petit-déjeuner* did not appear for the first time in a dictionary until 1932 (in the dictionary of the French Academy), at first without a hyphen (which was only added very recently). In the nineteenth century in Paris, *dîner* time had moved back to the time when one usually has *souper*, so that, in this region, *dîner* finally started to designate the evening meal, with *déjeuner* being imposed at the same time to designate the ‘midday meal.’²³ Little by little, this Parisian innovation spread to all regions of France. A situation of homonymy then arose, between the generation of older and younger folks, the former using the old system (*déjeuner/dîner/souper*), the latter the new system (*petit-déjeuner/déjeuner/*

²³ On semantic change in general, see 9.2, this volume.

dîner). An awkward ambiguity may then have set in (if you are invited to *dîner*, is it at noon or night?). It was resolved with the disappearance of the archaic triad in RF. Peripheral areas of France did not follow the same trajectory.

C14P34

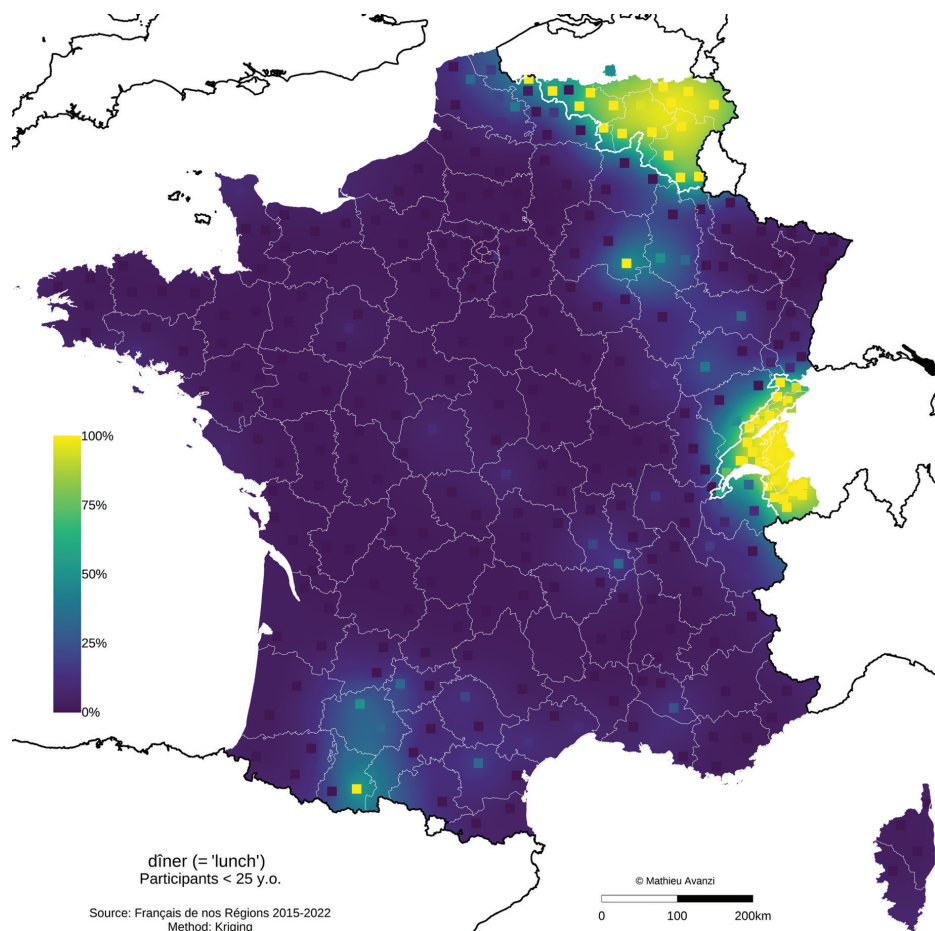
Nowadays, as we can see in Figures 14.3 and 14.4, when comparing the map generated from answers of participants younger than twenty-five with the map generated from answers of participants over fifty, the word *dîner* ('midday meal') is disappearing within mainland France. It survives in some peripheral areas but is slowly disappearing. In Belgium and Switzerland, it remains more common, because these regions are more sheltered from Parisian innovations. The existence of independent political, education, and media systems in Belgium and Switzerland can explain why the older system survives and linguistic change did not take place in these areas.²⁴ This is a very vivid example of social change driving linguistic change (see also Chapter 10, this volume).



C14F3

FIGURE 14.3 Percentage of use of the word *dîner* to designate lunch, data from participants over fifty years old (N = 1,855)

²⁴ For more information on French in Belgium and French in Switzerland, see 16.2; 16.3, this volume.



C14F4 **FIGURE 14.4** Percentage of use of the word *dîner* to designate lunch, data from participants under twenty-five years old (N = 2,862)

C14S20 *14.3.1.3 Pain au chocolat or chocolatine?*

C14P35 Throughout south-western France, the pastry known elsewhere as *pain au chocolat* or *petit pain (au chocolat)* ('chocolate croissant') is sold under the name *chocolatine*. To explain the origin of this last term—emblematic of the linguistic debates that have raged across the Internet since the popularization of social networks (see 6.2.7, this volume)—a whole array of more or less eccentric hypotheses has been formulated, either on the origin of the pastry or on the etymology of the word *chocolatine*. An examination of attestations of each of these terms through history (in online archives such as Gallica²⁵ or Google Livres²⁶) makes it possible to offer a new account of the issues.

²⁵ <https://gallica.bnf.fr/edit/und/rechercher-dans-gallica> accessed 5 Feb. 2023.

²⁶ <https://books.google.fr> accessed 5 Feb. 2023.

- C14P36** • The first mentions of recipes for *pain au chocolat* date back to the nineteenth century. At first, they do not refer to the pastry now known by the term, but to chocolate preparations made from almond powder or to a specific type of dessert (a bread made with a chocolatey dough).
- C14P37** • When the phrase *petit pain au chocolat* appeared in the early 1920s, it is impossible to know whether the *pain* ('bun') in question was a piece of bread into which a chocolate bar was inserted, or a pastry baked with a bar of chocolate in it. From the 1970s, the adjective *petit* ('small') becomes increasingly rare as the occurrences of *pain au chocolat* increase.
- C14P38** • The word *chocolatine* first appeared in a newspaper article from 1853.²⁷ It then designated a chocolate sweet in the shape of a *dragée* (a coated almond). In the following decades, we find the word used with different meanings: 'chocolate-based liquor', 'cocoa-based chocolate and grape sugar', 'quinine dragée', and then finally in the early 1950s as a synonym of 'chocolate pastry', in a magazine printed in south-western France.
- C14P39** Taken together, this information helps us to recreate in Figure 14.5 a diachronic scenario to explain what can be observed.
- C14P40** At first, almost everywhere in France, this chocolate pastry must have spread under the name *petit pain au chocolat*. The phrase disappeared from RF to make way for a shorter version: *pain au chocolat*. As an archaism, the longer form has persisted in peripheral north-eastern areas. In the south-west, the word *chocolatine*, which is a local innovation (the suffix *-ine* being commonly used in French to derive words related to food), spread in an area whose centres are Bordeaux and Toulouse.²⁸

C14S21 **14.3.2 Grammar**

- C14P41** Because of their rareness in French corpora (Blanche-Benveniste 1991; Avanzi et al. 2016; Avanzi and Thibault 2020), grammatical regionalisms are still the poor relation in regional variation studies. Anyone taking a close look at the books on regionalisms published in the 1980s and 1990s (see 14.2.3.1) can nevertheless find sporadic information on the gender of certain words (14.3.2.1), the selection and order of pronouns (14.3.2.2), variation in verb constructions (14.3.2.3), or the use of certain verbs to express specific tense and modal meanings (14.3.2.4).²⁹

²⁷ *La Presse*, 2 Jul. 1853 (<https://gallica.bnf.fr/ark:/12148/bpt6k476280h/f4.item.r=chocolatine.zoom> accessed 5 Feb. 2023).

²⁸ In Belgium, the word *couque*, which is used in certain areas, is a borrowing from Flemish (Francard et al. 2021 [2010]).

²⁹ Because they are very numerous and their description could fall under the study of lexicon, no mention is made here of numerous grammatical morphemes used in regionally marked syntactic structures, such as adverbials expressing different temporal or modal values or the use of certain prepositions (see Avanzi and Thibault 2020 for an overview).



C14F5 **FIGURE 14.5** Different ways of naming a ‘chocolate croissant’ in European French (N = 11,010)

C14S22 **14.3.2.1** *Word gender*

C14P42 From a diachronic point of view, noun gender can manifest variation in French.³⁰ Many dictionaries of regionalisms printed in the 1980s/1990s document gender variation for words that are feminine in RF but masculine in specific regions (such as *couleuvre* (‘snake’), *éclair* (‘lightning’), *horloge* (‘clock’), *vis* (‘screw’)), or that are masculine in RF but feminine in certain areas (such as *lièvre* (‘hare’), *ongle* (‘nail’)). However, the data gathered as part of the *Français de nos Régions* project did not deliver any significant traces of such variation at the regional level.

³⁰ On nominal gender in general, see 2.2.2, this volume.

- (5) a. il le mange
 he.NOM:3.SG:MASC it.ACC:3.SG:MASC eat.PRES:3.SG
 ‘he eats it’
- b. il y mange
 he.NOM:3.SG:MASC it.ACC:3.SG:MASC eat.PRES:3.SG
 ‘he eats it’

C14P46 As shown by Avanzi (2018), in a large area around Lyon, speakers have a neutral-value accusative pronoun that has the same form as the locative pronoun of RF, *y* (‘there’).³² The influence of the same dialects also explains the fronting of the pronoun *personne* (‘nobody, no one’) that has been observed in the same region, as in (6b):

- (6) a. j’ ai vu personne
 I.NOM:1.SG have.PRES:1.SG seen.PRES:3.SG nobody
 ‘I’ve not seen anyone’
- b. j’ ai personne vu
 I.NOM:1.SG have.PRES:1.SG nobody seen.PAST.PART
 ‘I’ve not seen anyone’

C14P47 Avanzi (2020c) shows that this usage is much less frequent today than the use of *y* as an accusative pronoun.

C14S24 14.3.2.3 *Variation in verb valency and verb structure*

C14P48 Another domain of variation concerns the alternation between accusative and dative complements of some verbs such as *aider* (‘to help’), *indifférer* (‘to leave indifferent’), *informer* (‘to notify’), *empêcher* (‘to hinder’), and *ressembler* (‘to look like’).³³ For most of these verbs, Grevisse and Goosse (2016: §284–5) mention regional variation, which is confirmed by the presence of such items in many dictionaries of regionalisms. Data gathered by the *Français de nos Régions* project show that for certain verbs, the use of non-standard variants might be more frequent in some areas of the French-speaking world (e.g. *aider*, see Avanzi 2019a), but for the most part, this variation is not regional.

C14P49 In previous stages of the language, certain verbs such as *tomber* (‘to fall’), *échapper* (‘to escape’), and *marier* (‘to marry’) took a direct object when used transitively. Speakers would say *tomber/échapper quelque chose* (lit. ‘to fall/escape something’) or *marier quelqu’un* (‘to marry somebody’). Nowadays, in RF, such verbs are used in factitive constructions: *faire/laisser tomber quelque chose* (‘to let something fall’, i.e. ‘to drop’), *laisser échapper quelque chose* (‘to let something escape’), or in a pronominal construction: *se marier avec quelqu’un* (lit. ‘to get married with someone’). Our data show that the old way of using the verb *marier* is no longer attested anywhere in mainland France. In contrast, the older transitive

³² See also 16.4.4.3, this volume.

³³ The concept of verbal valency is introduced in 2.3.3, this volume.

constructions of the verbs *tomber* and *échapper* remain common in many regions of France (Avanzi 2020a: 206–7).

C14S25 14.3.2.4 *Verb tenses*

C14P50 Verb tenses and modals also constitute a domain that is sensitive to regional variation (Ayres-Bennett and Carruthers 1992).³⁴ Speakers from southern France as well as from the Lyon region and French-speaking Switzerland have a double compound past called the *passé surcomposé* which involves the use of both auxiliaries, *avoir* ('to have') and/or *être* ('to be'):

- (7) a. j' ai eu fumé
 I.NOM:1.SG have.PRES:1.SG have.PAST.PART smoked
 'I have/had smoked'
- b. je suis eu parti
 I.NOM:1.SG be.PRES:1.SG have.PAST.PART left
 'I have/had left'

C14P51 In RF, grammarians do accept the use of *passé surcomposé* but only when it occurs in a subordinate clause (embedded in a phrase initiated by a conjunction such as *quand* or *lorsque*, 'when'). The examples in (7), where the *passé surcomposé* is linked to a verb which is the nucleus of a main clause, are therefore regional uses. Type (7a) and type (7b) are not found in the same areas. Within French-speaking Europe, (7a) is much more widespread than (7b). The first is used in the entire southern area of France, as well as in the region of Lyon and in Brittany (Avanzi 2020a: 201–11). The second is hardly used outside Switzerland (Borel 2020).

C14P52 When used as an auxiliary in RF, the verb *aller* ('to go') is normally followed by an infinitive to express near future (its value is close to the English progressive and 'going-to' future, see also 3.4.4.5, this volume). In Old and Classical French, the verb *vouloir* ('to want') was used in the same way. Rézeau (2001: 1042 ff.) states that it used to be attested in many French regions when referring to the weather, as in (8a). In other contexts, such as in (8b), it is restricted to the French spoken in Jura (Avanzi 2020a: 118).

- (8) a. il veut pleuvoir
 it.NOM:3.SG want.PRES:3.SG rain.INF.
 'it is going to rain'
- b. tu veux tomber
 you.NOM:2.SG want.PRES:2.SG fall.INF.
 'you are going to fall'

C14P53 Finally, another grammatical regionalism stems from the use of the verb *savoir* ('to know') to express physical capacity, as in, *je ne savais pas dormir à cause de mon lumbago* ('I could not sleep because of my back pain').³⁵ This is typical of the Nord-Pas-de-Calais area, but is even more

³⁴ For an overview of French verb tenses and modals, see 2.2.3, this volume. For a semantic account, see 3.4, this volume.

³⁵ In RF, this modality can be expressed only with the verb *pouvoir* ('can').

common in Belgium, where it is used with fewer syntactic and semantic restrictions (Avanzi and Barbet 2021).

C14S26 14.3.3 Phonology

C14P54 French vowel inventories traditionally identify around sixteen vowels, depending on whether the criterion of quantity (opposition between short and long) is taken into account (Martinet and Walter 1982; Armstrong and Pooley 2010; Lyche 2010). Because many vowels feature poor functional efficiencies, many oppositions are sensitive to regional variation.³⁶ Among the weakest points of the system are the opposition between the two low vowels (/a/~ɑ/) and the opposition between the two front nasals (/ɛ/~œ/; 14.3.3.1; see also 1.2.1.2, this volume). Mid vowels represent another weak point (14.3.3.2). Vowel quality represents a final weak point (14.3.3.3).

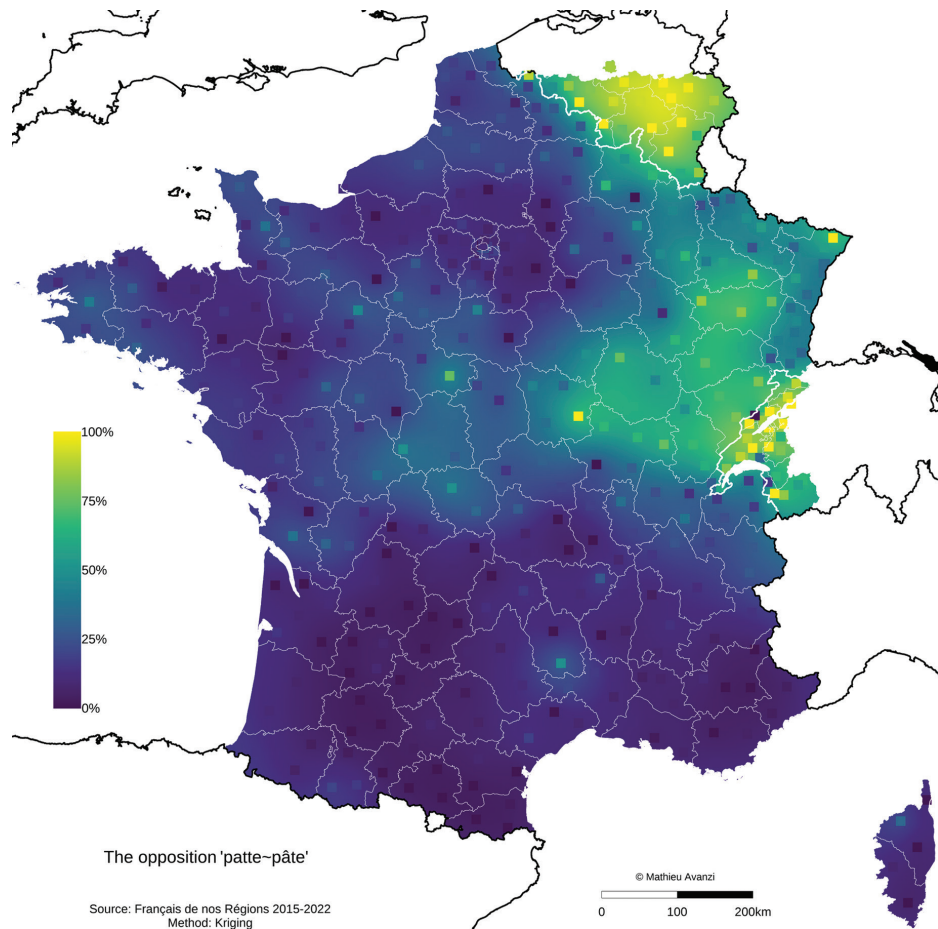
C14S27 14.3.3.1 The /a/~ɑ/ and /ɛ/~œ/ oppositions

C14P55 In studies dealing with the pronunciation of RF, there is a debate over the necessity of acknowledging the existence of two phonemes for the low vowels of French, /a/~ɑ/, and the two front nasal vowels, /ɛ/~œ/ (see 1.2.1.2, this volume). Indeed, in many regions, speakers do not distinguish between the two members of these phoneme pairs, and phonetically realize the resulting phonemes respectively [a] and [ɛ], in pairs like *patte~pâte* ('paw~'pasta'), *rat~ras* ('rat'~'cropped'), *malle~mâle* ('trunk'~'male'), etc. (see Delattre 1957) or *brin~brun* ('strand'~'brown'), *empreinte~emprunte* ('mark'~'borrow'), *Alain~alun* ('Alan'~'alum'), etc. (see Nève de Méverginies 1984). The data recently gathered within the *Français de nos Régions* project shows that the distribution of both oppositions is indeed undergoing a change, but that it does not affect all regions of French-speaking Europe at the same pace (see Avanzi 2020b for details).

C14P56 The map in Figure 14.6 shows that there are indeed considerable differences within France itself. As we can see, the *patte~pâte* opposition is especially vibrant in the north-eastern quarter of the country, within an area that includes the fifteen departments between Belgium and French-speaking Switzerland. This shows that the political boundaries that separate French-speaking Belgium and French-speaking Switzerland from France are not always linguistic ones.

C14P57 The map in Figure 14.7 shows that the opposition between *brin* and *brun* is much more robust in the south than in the north. In this landscape, Corsica—where only a third of the participants declared that they express this phonological opposition (34.7%)—is closer to the northern part of France. Otherwise, we see that the opposition has only sporadically survived in the north.

³⁶ The schwa is one such example because its regional variation depends on multiple factors (see 1.2.2, this volume). Outside phonology, the variation of final consonant realization in words such as *sourcil* ('eyebrow'), pronounced [sursi] or [sursil], *alphabet* ('alphabet'), pronounced [alfabɛ] or [alfabet], etc. is another major domain of regional variation in pronunciation (see Avanzi and Thibault 2021). Many random words present pronunciation variants. They often consist of borrowings, e.g. the word *Spritz* found in the expression *Aperol Spritz* can be pronounced with an initial [s] or [ʃ], depending on the region of speakers (see Avanzi 2020: 108).



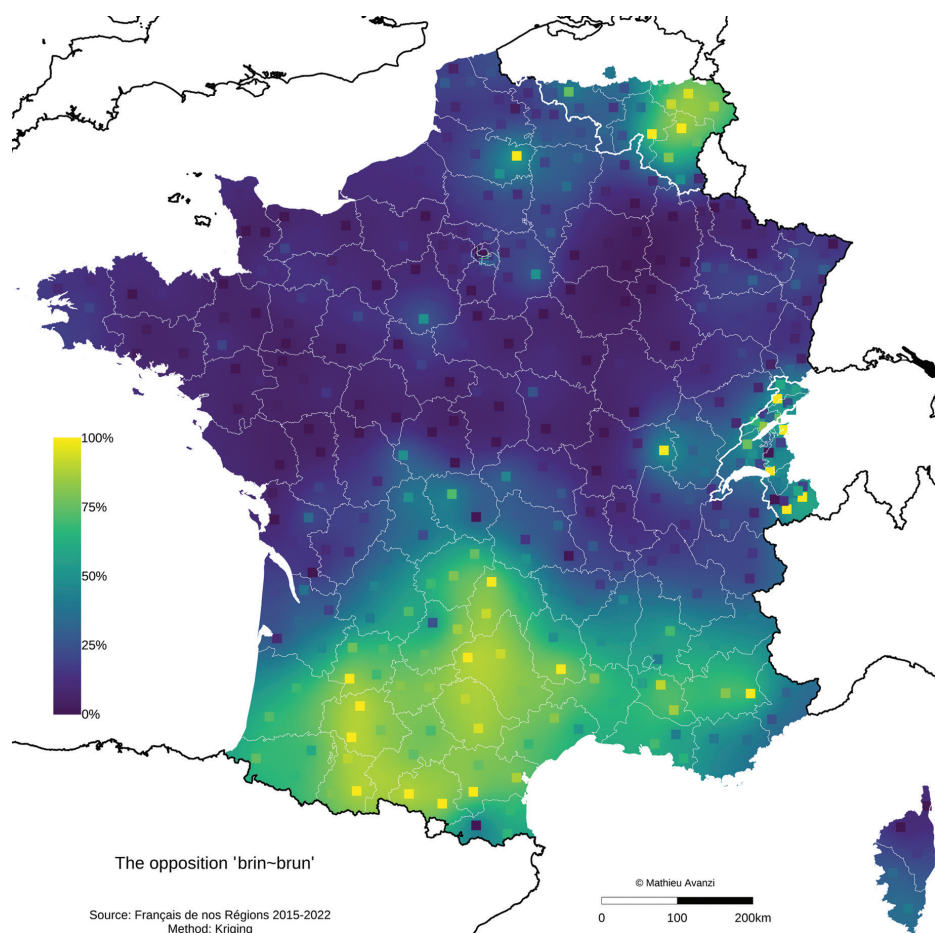
C14F6 FIGURE 14.6 Percentage of participants for whom there is a phonological contrast between *patte* /a/ and *pâte* /ɑ/ (N = 11,249)

C14S28 14.3.3.2 Mid vowels: /e/~ /ɛ/, /o/~ /ɔ/, and /ø/~ /œ/

C14P58 In RF, /e/ and /ɛ/ are different phonemes, as can be demonstrated by the existence of minimal pairs such as *piqué~piquet* ('stung'~'peg') (see also 1.2.1.2, this volume), but this is not the case in many regions of mainland France, as can be seen on the map in Figure 14.8.

C14P59 We can see that in France, the entire southern half and some peripheral regions in the northern part (Normandy and Nord-Pas-de-Calais) do not maintain the phonological contrast between *piqué* and *piquet* (both are pronounced with the close-mid vowel /e/). In RF, as well as in the other parts of France, Belgium, and Switzerland, the opposition remains quite stable.

C14P60 To understand the variation affecting the opposition /o/~ /ɔ/, it is necessary to distinguish between open and closed syllables. In closed syllables, such as in the pair *saute~sotte*

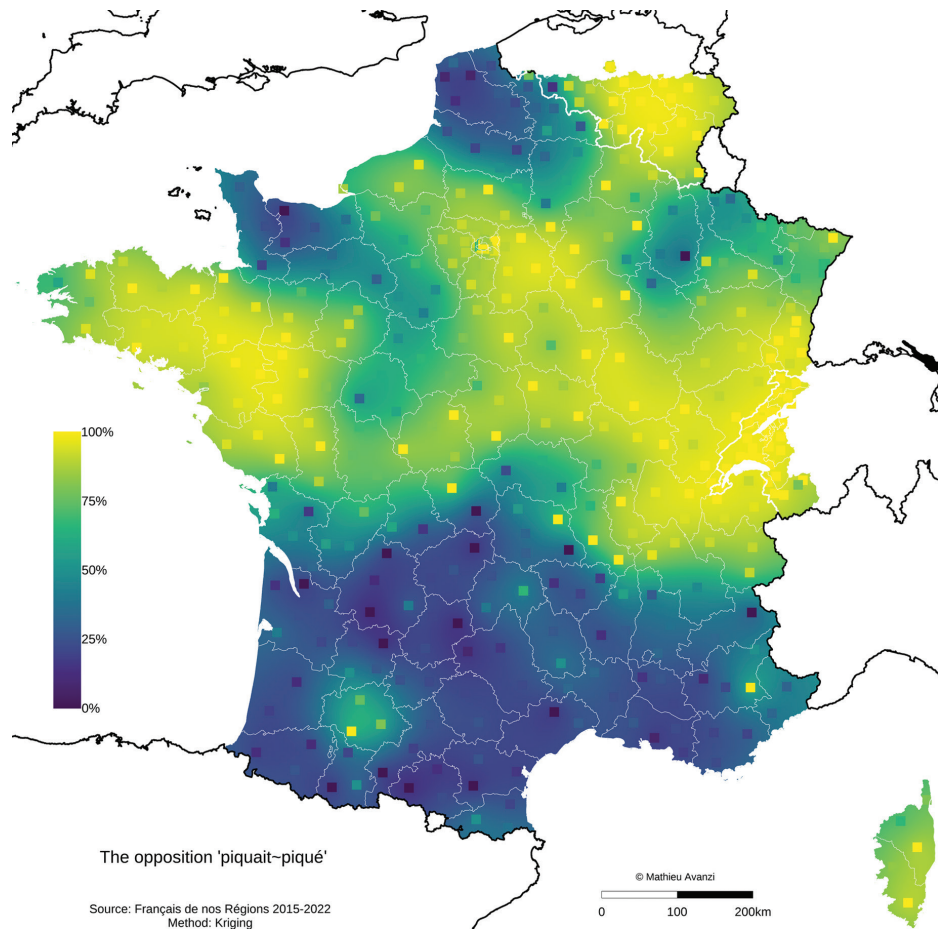


C14F7 **FIGURE 14.7** Percentage of participants for whom there is a phonological contrast between *brin* /*ɛ̃*/ and *brun* /*œ̃*/ (N = 9,058)

(‘jump’~‘fool f.’), the opposition is still phonological for many speakers in northern France, as can be seen in Figure 14.9.

C14P61 In Nord-Pas-de-Calais and southern mainland France, the *loi de position* (which stipulates that a vowel is closed in an open syllable and open in a closed syllable; see 1.2.1.2, this volume) applies: the contrast is neutralized. In open syllables, such as in the pair *seau*~*sot* (‘bucket’~‘fool m.’), there is no contrast for most speakers in mainland France. The opposition survives only in the eastern fringe of mainland France, as well as in Belgium and Switzerland (see Figure 14.10).

C14P62 Finally, the regions where the opposition between /*ø*/ and /*œ*/ is maintained can be seen in Figure 14.11. The map shows the regions where speakers make a distinction



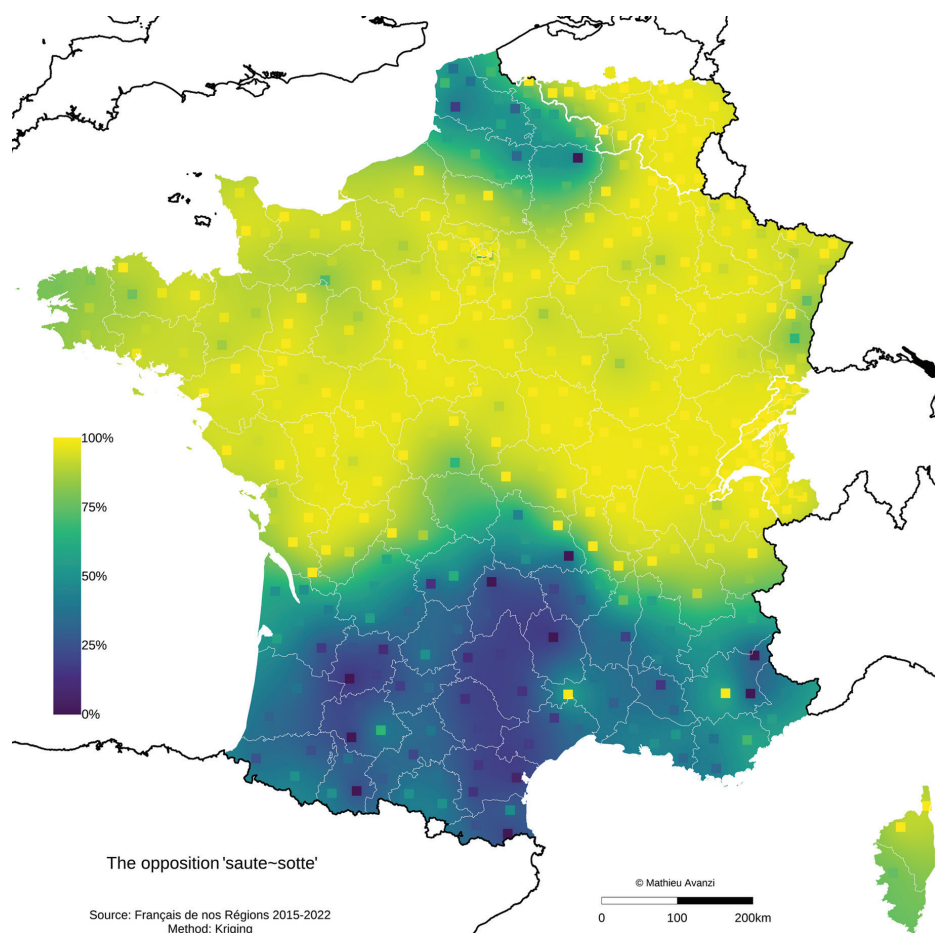
C14F8 **FIGURE 14.8** Percentage of participants for whom there is a phonological contrast between *piquait* /ɛ/ and *piqué* /e/ (N = 11,249)

between the words *jeûne* ('fast' (n.), pronounced with /ø/) and *jeune* ('young', pronounced with /œ/).

C14P63 Note that for this specific case, it is not possible to determine whether the distinction relies on vowel quality or vowel quantity, because of the presence of a circumflex accent, which in the north-eastern regions triggers a long vowel (see 14.3.3.3).

C14S29 *14.3.3.3 Short and long vowels*

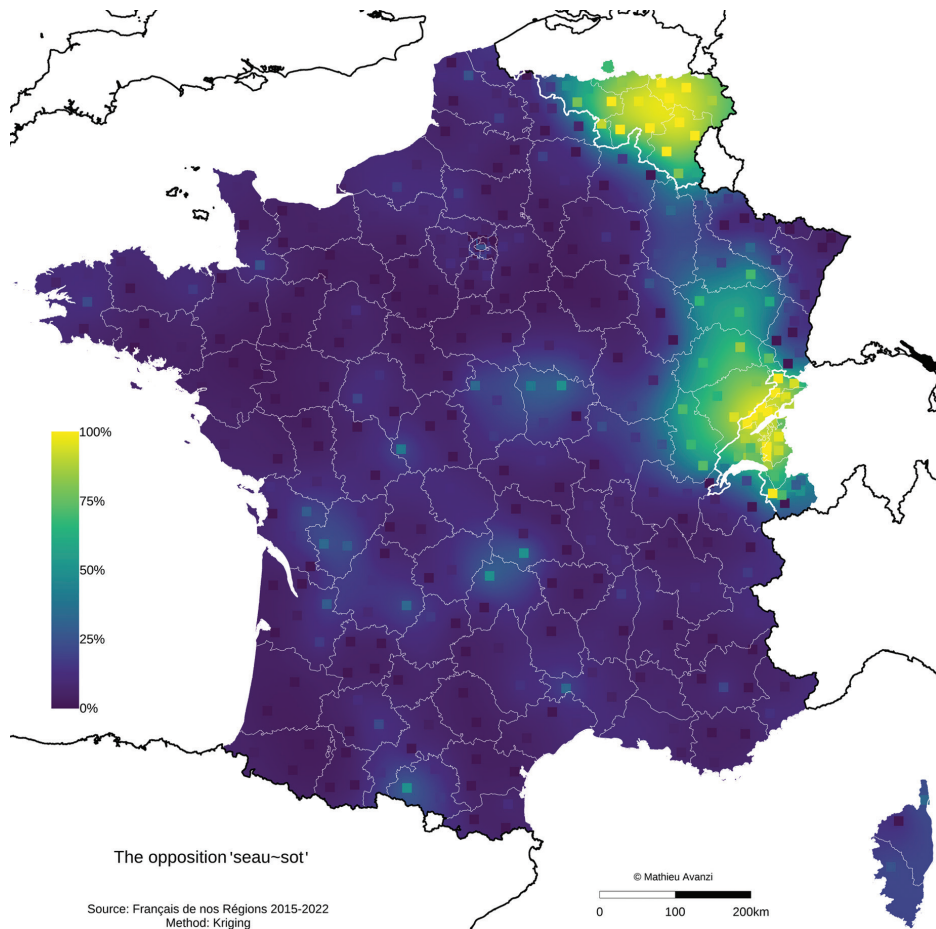
C14P64 Until around the mid-twentieth century (Martinet 1945; Lyche 2010; see also 1.2.1.2; 7.3.3, this volume), vowel quantity in RF used to express a phonological contrast between short and long vowels and, as a consequence, distinguished masculine and feminine forms of a



C14F9 FIGURE 14.9 Percentage of participants for whom there is a phonological contrast between *saute* /o/ and *sotte* /ɔ/ (N = 8,862)

given word (*nu~nue* ('naked m.'~'naked f.)); *collé~collée* ('stuck m.'~'stuck f.)) or potential homophones, whether without (*bout~boue* ('boil'~'mud'); *lit~lie* ('bed'~'dregs')) or with a circumflex accent (*mettre~mâitre* ('put'~'male primary school teacher'); *pêche~pèche* ('sin'~'peach')).

C14P65 The analysis of Figure 14.12 showing the percentage of participants who claimed to make a contrast between *faites* ('do.PRES:2.PL, pronounced with /ɛ/) and *fête* ('festival, pronounced with /ɛ:/) makes it possible to distinguish three cases. First, in most areas, the opposition between long and short is not observed. Second, there are areas where the opposition between long and short is still produced by some speakers, but mostly older ones: this concerns all the departments of the eastern fringe of France, close to the Belgian and Swiss borders. Third, there are regions where the opposition is still phonological for everyone. This is the case in



C14F10 **FIGURE 14.10** Percentage of participants for whom there is a phonological contrast between *seau* /o/ and *sot* /ɔ/ (N = 8,862)

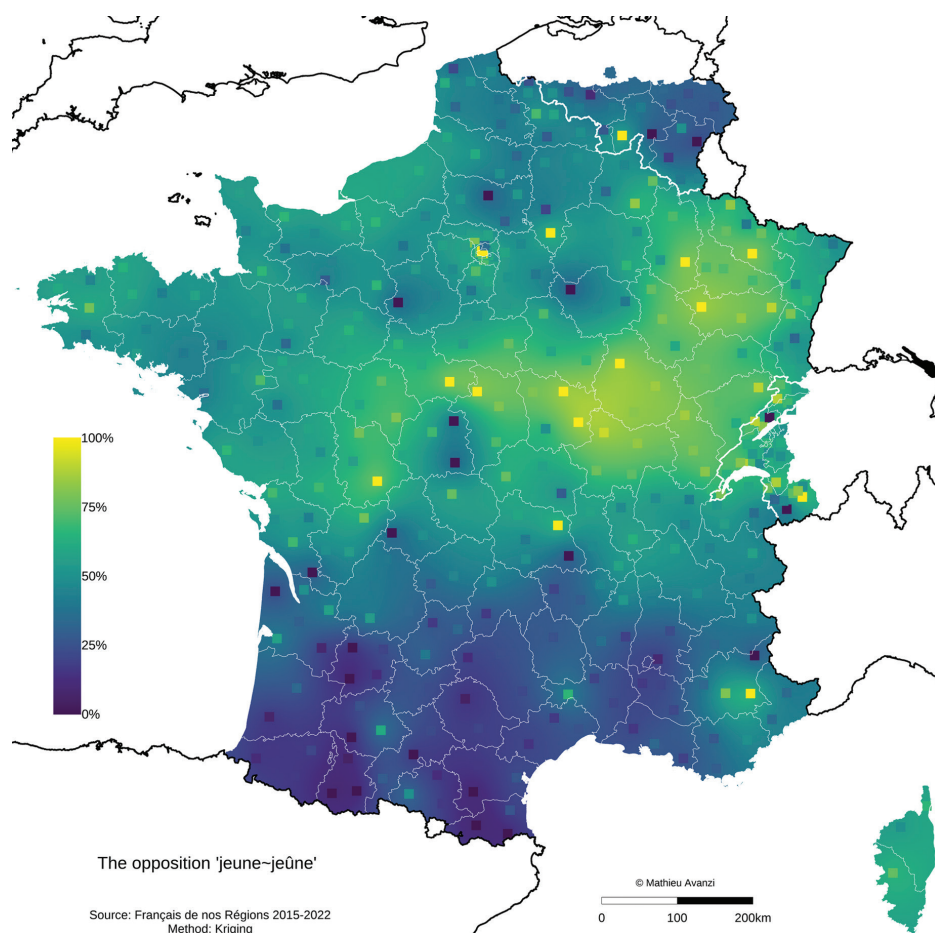
the French-speaking regions of Belgium and the majority of French-speaking Switzerland (with the exception of Valais).

C14S30

14.4 CONCLUSION

C14P66

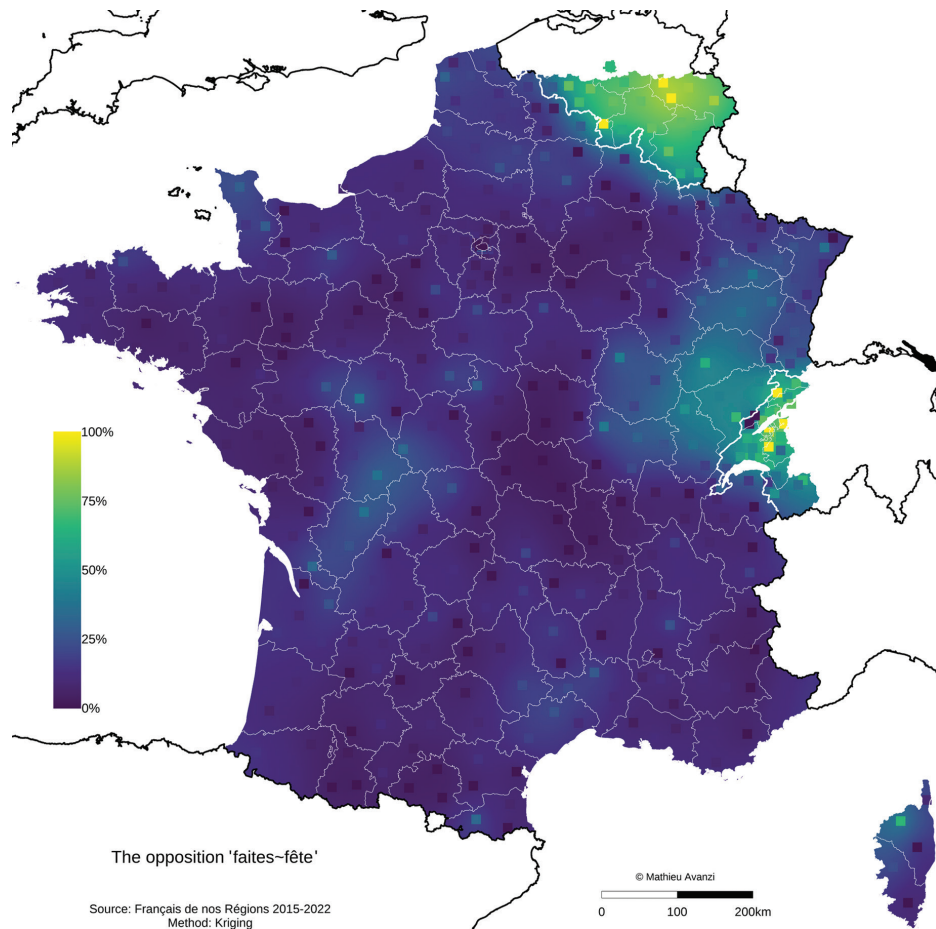
This chapter has aimed to give an overview of the ways French varies regionally in mainland France. It started by introducing important notions such as Reference French and by defining regionalisms, both generally and in the particular French context. It then presented a systematic overview of studies of regional French. Such studies are almost as old as the earliest description of central French and emerged when the Gallo-Romance dialects and regional languages of France started to lose ground. The final section demonstrated the



C14F11 FIGURE 14.11 Percentage of participants for whom there is a phonological contrast between *jeune* /œ/ and *jeûne* /ø/ (N = 8,862)

potential that crowdsourcing offers for mapping lexical, grammatical, and phonological variation across space.

C14P67 Crowdsourcing has allowed for a fresh look at certain phenomena that are disappearing primarily due to dialect levelling, with some interesting reorganization across political boundaries (see the example of *dîner* to refer to lunch, for example, 14.3.1.2; or the fate of phonological opposition such as in /a/~/ɑ/, /ɛ/~/œ/, and short vs. long vowels, 14.3.3.1 and 14.3.3.3). This method has also provided data making it possible to re-examine the relationship between Gallo-Romance dialects and regional French, with some types of interference between the two systems not always as regular as expected (see the example of the ‘village fair’, 14.3.1.1; and the examples of pronouns *y* and *personne*, 14.3.2.2). The limits of crowdsourcing are well known (see Avanzi and Thibault 2019). Among other things, it



C14F12 FIGURE 14.12 Percentage of participants for whom there is a phonological contrast between *faites* /ɛ/ and *fête* /ɛ:/ (N = 11,249)

should be emphasized that crowdsourcing provides information on what speakers think they do, rather than on what they actually do. At present, we cannot rely on such data to assess the frequency of phenomena or specify their context of use in everyday life by taking diaphasic variation into account, for example. Another issue involves the question of diastratic variation (see Chapter 10, this volume). Diastratic factors, too, need to be taken into account when studying regional variation. While age was shown to have an impact on the geography of some regionalisms (such as with *dîner* to refer to lunch, 14.3.1.2), this type of analysis still needs to be performed for all other features. It is also necessary to take into account the social profile of informants (defined in terms of education or occupation), as such factors have been shown to clearly influence the use of regional features.³⁷

³⁷ See Armstrong and Pooley (2002) for a comprehensive list of references, as well as Chapter 10, this volume.