# New speaker language and identity: Practices and perceptions around Breton as a regional language of France

Davies-Deacon, Merryn

*Award date:*
2020

*Awarding institution:*
Queen's University Belfast

# New speaker language and identity: Practices and perceptions around Breton as a regional language of France

Submitted for the degree of Doctor of Philosophy by

Merryn Davies-Deacon

*B.A., M.A.*

February, 2020

School of Arts, English, and Languages

# Abstract

This thesis focuses on the lexicon of Breton, a minoritised Celtic language traditionally spoken in western Brittany, in north-west France. For the past thirty years, much work on Breton has highlighted various apparent differences between two groups of speakers, roughly equivalent to the categories of new speakers and traditional speakers that have emerged more universally in more recent work. In the case of Breton, the conceptualisation of these two categories entails a number of linguistic and non-linguistic stereotypes; one of the most salient concerns the lexicon, specifically issues around newer and more technical vocabulary. Traditional speakers are said to use French borrowings in such cases, influenced by the dominance of French in the wider environment, while new speakers are portrayed as eschewing these in favour of a "purer" form of Breton, which instead more closely reflects the language's membership of the Celtic family, involving in particular the use of neologisms based on existing Breton roots. This thesis interrogates this stereotypical divide by focusing on the language of new speakers in particular, examining language used in the media, a context where new speakers are likely to be highly represented. The bulk of the analysis presented in this work refers to a corpus of Breton gathered from media sources, comprising radio broadcasts, social media and print publications. Insights are also provided from interviews that were carried out with employees of Breton-language media to gain additional information on speakers' individual practices and beliefs and complement the linguistic data. The findings show that in these contexts, speakers often avoid some of the more extreme features associated with the new speaker stereotype, and that medium and register can be relevant factors in what sort of lexicon is used; also, notably, the traditional dialects of Breton to a certain extent serve as target varieties for many new speakers, showing that the two groups are not as separate from each other has some research has implied,

and that the standard new speaker variety is not necessarily the most prestigious. New speakers, and others with new speaker characteristics, do not form a homogeneous community, but instead have diverse ways of speaking and engaging with Breton and the rest of its speaker population.

# Acknowledgements

I have been fortunate to be able to undertake my PhD as part of a large multidisciplinary project, and my first thanks must go to the MEITS team at QUB: especially to Janice Carruthers. I couldn't have asked for a better supervisor (and choirmate), who has been constantly approachable, extremely easy to work with, and of course instrumental in a number of major decisions made regarding the structure and direction of this thesis. Particular thanks also to my second supervisor, Mícheál Ó Mainnín, for his constant encouragement and especially for the insightful comments as I approached submission. It's been a pleasure to work and share an office with the other members of the team, Deirdre, Daniel, and Robbie, as well as the other researchers who have joined us in the "cave" over the past three years. More generally, the insights into other linguistic disciplines and the additional opportunities I have been able to benefit from as part of MEITS have been invaluable, notably the chance to discuss my research with Doug Kibbee towards the end of the PhD, and the ability to undertake a placement at HarperCollins in the final stages. I must thank Wendy Ayres-Bennett for her support as PI of MEITS, and without whom many of these opportunities would have been unavailable.

Mersi bras/trugarez vras d'ar re am eus komzet ganto diwar-benn ar brezhoneg hag ar mediaoù; I owe a special debt of gratitude to my interview participants, who graciously and enthusiastically gave up their time and invited me into their workplaces. The "minor" part of my research became much more significant (and enjoyable!) than I anticipated: Chapter 5 could have been at least twice as long as it ended up. Trugarez vras ivez da Jérémy Ar Floc'h en deus sikouret gant an treuzskrivañ.

Over the past three years I have had the pleasure of discussing my work with other

scholars, who have invariably provided crucial information and important alternative perspectives. Many such interactions have taken place at conferences, but these connections have also in no small part been facilitated by the internet and the growing ability to share research online. I must therefore especially thank those who have been kind enough to share information by email, including Jean Le Dû, Mélanie Jouitteau, and Holly Kennard; and special thanks to a Twitter contact "on the ground". At conferences, it was especially rewarding to be part of a recurrent panel on minoritised languages in social media, and particular thanks go to the other members of the original group, Guillem Belmar Viernes, Maggie Bonsey, and Piotr Szczepankiewicz, and to our chair, Elin Haf Gruffydd Jones. I am also grateful to Fañch Bihan-Gallic for a remark that was influential during the final editing process, and to Anthony Harvey for his ongoing support and enthusiasm for my work, particularly those aspects outside the area of this thesis.

Undertaking my PhD at QUB has been a delight; the school of AEL, and French Studies in particular, has been the friendliest university department I've yet encountered. Specific thanks must go to the approachable and efficient administrative staff, especially Lynne Robinson, Tracy Duffy, and Linda Drain—and to those I got to know as fellow PhD students* and friends; in particular, Catherine, Daniel, Lauren, Mathew, Mike, Sarah, and Tom.

My mother and grandmother have been constantly supportive, interested in my work, and accommodating of my brief visits home to Redruth. Tas vy, che a wrug desky dhebm fatel ajon 'gan tavajow bian dhe voas pojek; na via gwres gena vy tra veth a'n par ma hep hedna. Joan and David: thank you for welcoming me into the family. Andrew—the most unexpected finding of my PhD—thanks for being my best friend.

*. Mais pas que [Quigley, 2019]

# Contents

# List of Figures

# List of Tables

# Chapter 1

# Introduction: Research context and questions

## 1.1 Overview of the thesis and structure of the chapter

This thesis focuses on Breton as an example of a minoritised language spoken in France undergoing revitalisation. It interrogates the contemporary use of Breton in three media contexts—radio, print journalism and Facebook—in order to investigate the extent to which the language conforms to the claims about new speakers typically found in existing academic work, with particular focus on the lexicon. Two methodologies for investigating these claims are used: firstly, a corpus-based investigation of Breton in the contexts listed above, including a detailed quantitative study of the lexicon found therein; secondly, interviews with a number of professionals working in Breton-language media, in order to gain a broader sociolinguistic understanding of the various considerations at play in Breton speakers' use of the language and its lexicon.

This first chapter sets out the research context, beginning with the place of Breton as a regional language of France with regard to both language policy and the status of different linguistic varieties; it then discusses the concepts of the new speaker and "néo-bretonnant", and the particular stereotypes associated with this latter group. The end of the chapter lists the research questions tackled by this thesis, with an indication of

how subsequent chapters will contribute to responding to these.

## 1.2 French policy on regional languages

France has long been known for its particular hostility towards regional languages, to an extent that marks it out from other democratic nation-states and has at times prompted comparisons with authoritarian regimes (Hornsby 2010:171). Some scholars have identified a desire to suppress regional languages stretching as far back as sixteenth-century legislation such as the *Ordonnances de Villers-Cotterêts*, although the apparent purpose of these edicts was to put an end to the use of Latin in legal situations, and the precise meaning of "langage maternel françois", the phrase found therein, is ambiguous (Lodge 1993:126). An attitude opposed to regional languages is more evident in post-revolutionary France: soon after the 1789 revolution a report asserted that "le fédéralisme et la superstition parlent bas-breton, l'émigration et la haine de la République parlent allemand ..." (Barère [1794], quoted in Perrot 1997:162). While this suppression of regional languages was carried out in the spirit of encouraging universal access to French as part of the promotion of equal rights for all French citizens, it equally engendered an attitude that viewed regional languages as inherently negative.

The effects of this stance on the part of the French state can be identified in the pejorative connotations that came to be associated with the word *patois*, as well as in the methods used to enforce the use of French in formal education, including the use of the *symbole*[1] in Brittany at the turn of the twentieth century (Broudic 2007) and in other parts of France and French colonies at various points in history (Milin 2018). This historic suppression of regional languages is undoubtedly still pertinent to the contemporary construction of language attitudes and ideologies in France, and has typically occupied a prominent position in regionalist political discourse. The ongoing struggle for linguistic recognition and the subjugation of regional languages in favour of French have formed an essential factor in arguments in favour of self-determination, being used as an example of how the French state has mistreated any regions that may

1. "Les enfants se surveillent mutuellement, celui qui serait surpris à proférer un mot de breton par inadvertance se voyant attribuer une pièce de bois et ne pouvant prendre part aux jeux de la récréation que lorsqu'il aurait lui-même entendu un de ses camarades parler breton" (Broudic 2007).

wish to preserve a distinct identity. In Brittany, the history of the political regionalist movement and the various expressions of Breton political identity throughout the modern era have been closely linked to linguistic matters: from the early Romantic nationalism motivated by the literary and linguistic work of the *Académie celtique* (see section 1.3 below) in the nineteenth century, to the concessions made to Breton, such as the instatement of radio broadcasting in the language, by occupying forces during the Second World War (Wmffre 2008:96), to the the frequent use of Breton in regionalist political graffiti around Brittany in the twenty-first century.[2] Today, regionalist political parties in Brittany stress their belief in the importance of the language: one of the most active, the centre-left *Union démocratique bretonne*, affirms for example its commitment to a "co-officialité des langues".[3] In the case of Brittany, the Breton term *emsav* ('uprising', or more literally 'getting oneself up') is often used to characterise the regionalist movement even in French (in e.g. Nicolas 2007), again demonstrating the important role of the language in political activism. This strong connection between language and politics contributes to the image of the "militant" speaker of a minoritised language, a concept that will be visited several times during this thesis, firstly in section 1.4.2.

Despite this state of affairs having been motivated by the historically neglectful attitude of the French state towards regional languages, it is nonetheless undeniable that various changes relating to language policy have occurred in recent years. During the second half of the twentieth century, in particular, various opportunities increased scope for the potential support of regional languages within French territory. To understand the position of Breton in contemporary society, it is useful to examine how language policy has developed more recently.

### 1.2.1 State language policy at a national level

It would be incorrect to assert that the development of French national language policy over the last sixty years has consistently offered an increasingly tolerant and support-

2. For examples, see http://stourmomp.blogspot.com, accessed 22 Jan 2020.
3. https://www.udb.bzh/nos-objectifs/, accessed 22 Jan 2020. It can also be noted that the *UDB* offers a full version of its website in Breton as well as in French.

ive atmosphere towards regional languages; instead, steps taken in this direction have been marred by simultaneous reaffirmations of the superiority of the French language, often—if sometimes unintentionally—to the detriment of the status of regional languages. The *loi Deixonne*, passed in 1951, was the first major addition to French legislation that promoted the use of regional languages; its provisions, though, were minimal, specifying a group of four regional languages, those not considered dialects of any other country's national language,[4] that were henceforth permitted to be taught in schools, if both teacher and pupils indicated their desire.[5] Taking advantage of this opportunity was made difficult by the lack of teacher training, and of textbooks (Jung and Urvoas 2012:34); moreover, the examinations in these languages that could be taken as part of the *baccalauréat* did not count towards the calculation of the average mark (Gardin 1975:32), meaning their utility as school subjects was difficult to defend, and their attraction to pupils surely limited. The *loi Deixonne* is thus one of several pieces of legislation that, while ostensibly favourable towards regional languages, offered them extremely limited practical support. Still, this symbolic concession must be noted, despite its limited effect, as the beginning of a gradual change in the French state's attitudes towards regional languages.

In 1975, the provisions of the *loi Deixonne* were reinforced in the *loi Haby*, which allowed "un enseignement des langues et cultures régionales ... tout au long de la scolarité".[6] However, in the same year, the *loi Bas-Lauriol* was passed, strengthening the hegemony of French, particularly in written form, by stipulating that its use was required in various contexts including advertising, product warranties, and inscriptions on public buildings, in an attempt to combat the increasing presence of English in these domains. This objective means that in the text of the *loi Bas-Lauriol*, French is placed in opposition to "langues étrangères"; regional languages are not mentioned, whether accidentally or through deliberate neglect.[7] The latter certainly seems more likely given the information that a ban on regional languages was initially to be included in the text of the law (Woehrling 2013:79), suggesting that the government eventually decided it

4. i.e. Basque, Breton, Catalan and Occitan.
5. Loi 51-46 du 11 janvier 1951, article 3 (available at https://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000000886638, accessed 14 Oct 16).
6. Loi 75-620 du 11 juillet 1975, article 12 (available at https://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000000334174, accessed 14 Oct 16).
7. Loi 75-1349 du 31 décembre 1975 (available at https://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000000521788, accessed 14 Oct 16).

would be best to ignore their existence.

This contradictory legislative attitude towards regional languages, whereby laws promoting their use are passed in tandem with other laws that restrict them, has persisted to the present day, and is echoed in other governmental responses to the situation of regional languages in France; the two most recently elected Socialist presidents, François Mitterrand and François Hollande, both promised more overt promotion of regional languages in their initial electoral campaigns, but both came to disappoint regional language activists (Laroussi and Marcellesi 1993; Le Nigen et al. 2014:146). Since the 1980s, various reports into the state of regional languages, including the Giordan, Poignant and Cerquiglini reports in 1982, 1998 and 1999 respectively (Laroussi and Marcellesi 1993; Ager 1999; Beacco and Cherkaoui Messin 2010), have been commissioned with a view to improving relevant policy provisions, but little action has ensued. As in the text of the *loi Bas-Lauriol*, many legal texts posit an opposition between French and foreign, without considering the different cases of regional and migrant languages, whose being spoken as native languages on French territory is in these cases apparently not sufficient to prompt a more thorough response to their status.

Both supportive and restrictive legislation have continued to appear in the years since the Haby and Bas-Lauriol laws. In 2008 the French constitution was modified to state that "les langues régionales appartiennent au patrimoine de la France".[8] However, in 1992, a clause had been inserted stating that "la langue de la République est le français",[9] which symbolically undermined support for regional languages. The insertion of this clause was part of France's response to its signing of the Maastricht Treaty and its new position within an increasingly integrated Europe, where French was believed to need to reassert its dominance on French territory in opposition to foreign languages, particularly English. Its adoption was particularly pertinent for the situation of Breton, however, as an attempt to integrate the Diwan schools, teaching through the medium of the language, into the state education system was legally prevented as a result of the clause (Woehrling 2013:85).

8. Constitution du 4 octobre 1958, article 75-1 (available at https://www.legifrance.gouv.fr/affichTexte.do?cidTexte=LEGITEXT000006071194, accessed 14 Oct 16).
9. ibid.

### 1.2.2 Language policy at the supranational level

France's position within Europe also brought its own directives regarding regional languages in the form of the European Charter for Regional or Minority Languages (ECRML),[10] drawn up by multiple member states of the Council of Europe in 1992, and signed by the French state in 1999. This charter aimed explicitly to protect regional languages across Europe in multiple domains including law, education, and media; to suit the varied situations present across the continent, 68 provisions were laid out in part III of the charter, from which nations ratifying it were permitted to choose a selection of at least 35 to enforce (CoE 1992:2). This freedom of choice, however, did not incite France to ratify the charter, and it remains one of only eight Council of Europe member states to have signed but not ratified it, as opposed to 25 ratifications.[11]

The French state did originally intend to ratify the ECRML, but a ruling from the *Conseil d'État* in 1996 judged it contradictory to the French constitution (Jung and Urvoas 2012:45), particularly to the clause stating that "la langue de la République est le français". Since then, attempts have been made to have the charter ratified (Le Nigen et al. 2014:146), including the commissioning of the Cerquiglini report in 1999 to investigate how this could be achieved (Beacco and Cherkaoui Messin 2010:105), but such attempts have had no success. The issue returned most prominently in 2012 when François Hollande, then presidential candidate, made a series of promises including the ratification of the charter—after his election, this remained unfulfilled. For language activists, this is a sign of the French state's continued indifference, if not outright hostility, towards regional languages; however, it should be noted that ratification of the ECRML is not in itself a meaningful act in their support. States that have ratified it are under no particular pressure to implement its provisions: the UK, for example, includes Scottish Gaelic and Cornish under the terms of the charter, but its government has recently withdrawn funding for Gaelic-language broadcasting (Devlin 2015) and for the work of the Cornish language office (Cornwall Council 2016), thereby effectively reneging on those terms of the charter that encourage it to provide financial

10. Abbreviations such as this will be spelt out at first mention, and specialist terms explained, but a list of both can be found in Appendix A to ease reading as the thesis progresses.

11. See http://www.coe.int/en/web/conventions/full-list/-/conventions/treaty/148/signatures, accessed 14 Oct 16. The other states that have signed but not ratified the charter are Azerbaijan, Iceland, Italy, Macedonia, Malta, Moldova, and Russia.

support for the promotion of regional languages. Nonetheless, the French state's refusal to ratify the charter can certainly be perceived as a symbolic affront, if not one that is meaningful in practice.

### 1.2.3 Language policy at the regional level

Academic studies of language planning and policy typically divide it into three categories: status planning, which relates to a language's position as an official national language or otherwise; corpus planning, which deals with the actual linguistic makeup of the language; and acquisition planning, which denotes the way in which the language is diffused through education (Wright 2007:165). Because French state language policy is minimal, and supranational policy often has little effect, much of this language planning for regional languages must be coordinated by local authorities: regional or departmental governments, or territorial collectivities, depending on the area in which the language is spoken. For example, acquisition planning is the subject of limited national legislation, which specifies that certain regional languages may be examined as a "langue vivante" as part of the *baccalauréat*;[12] the same legislation states that the details surrounding their teaching are to be established in each relevant *académie* by a *Conseil académique des langues régionales*, consisting of various representatives of educational institutions and associations in the area.[13] Any decisions about the actual content and provision of regional language teaching must therefore be made at this devolved level; likewise, initiatives relating to the use of regional languages in media and in the linguistic landscape tend to require the instigation of local authorities. As a result, different regional languages are likely to benefit from different levels of institutional support.

This situation can be investigated by examining the language policy that operates on some of the different regional languages of France, considering each one briefly before

12. Note de service de l'Éducation nationale 2012-162 du 18 octobre 2012 (available at http://www.education.gouv.fr/pid25535/bulletin_officiel.html?cid_bo=65827, accessed 14 Oct 16). Regional languages do now count towards the overall mark obtained, although their influence is minimal: see https://www.lexpress.fr/culture/reforme-du-bac-les-langues-regionales-a-la-sauce-blanquer_2061411.html, accessed 7 Aug 19.

13. Article D312-37 du code de l'éducation (available at https://www.legifrance.gouv.fr/affichCodeArticle.do?idArticle=LEGIARTI000025164747&cidTexte=LEGITEXT000006071191, accessed 14 Oct 16).

turning to Breton in more detail.

Corsican is an example of a language with strong support from local authorities. While not included under the terms of the *loi Deixonne* due to its linguistic similarities to standard Italian, which for a long time caused it to be frequently perceived as a dialect rather than a language (Blackwood 2004:309), it benefitted from the creation of the territorial collectivity of Corsica in 1982. The greater level of devolution to Corsica that this brought about enabled more explicitly favourable measures to be implemented. Such measures were also encouraged by a generally high awareness of a distinct regional identity due to Corsica's physical separation from mainland France and its having come under French control as recently as the eighteenth century (ibid.:307). Although substantial language shift began in the twentieth century (ibid.:308), as it did elsewhere in France, Corsican is still spoken by around half the population (Blackwood 2011:113). The increased powers of the territorial collectivity compared with a regional government mean that Corsican is prominently displayed in the linguistic landscape (Blackwood 2014) and compulsory Corsican language lessons have formed part of the education system since 1999 (Blackwood 2004:311), with the result that every child is raised with at least a basic understanding of the language.

Francoprovençal provides a contrasting example. Like Corsican, it was not recognised as a distinct language until recently (Martin 2011:2), and its recognition is still hindered by the fact that many speakers are not even aware of the term "Francoprovençal" (Vurpas 1993:184). The fact that the language's territory is not restricted to France but also covers parts of Switzerland and Italy, more overtly multilingual societies, gives the French authorities, on whom the onus to support Francoprovençal only partially falls, an excuse to engage less in its favour. No immersion or bilingual education in Francoprovençal is available, and it is not listed among the regional languages that may form part of the *baccalauréat*. For support, Francoprovençal relies on small voluntary organisations (Bert 2011:5); the regional government of the Rhône-Alpes formally declared that it would offer more institutional support for the language in 2009, but this declaration contained few concrete engagements other than pledging financial aid to academic research (Région Rhône-Alpes 2015). In 2015, the region signed an agreement with the neighbouring Aosta Valley region of Italy (ibid.), which could be a po-

tential source of greater support in the near future.

Occitan, on the other hand, has a well-established revitalisation movement, dating back to the establishment of the Félibrige in 1854 (IEO 2015:11), as well as a prestigious medieval literary tradition. As a language included in the terms of the *loi Deixonne*, the development of Occitan in education began early, and a network of immersion schools called Calandreta is now in operation all over southern France, first established in 1979[14] and now operating 52 primary schools and two secondary schools.[15] However, this covers a large territory in which multiple distinct dialects and several norms coexist, and while the number of Occitan speakers is "regularly reported" as around three million (Costa 2014:199), only 2514 children were educated in Calendreta schools in the school year 2008–9.[16]

For Breton, a tangible revitalisation movement first became active in the early twentieth century. As a non-Romance language, Breton was always clearly understood as distinct from French, allowing this movement to begin relatively early, particularly compared with Corsican and Francoprovençal. Action surrounding Breton has traditionally been the responsibility of Breton speakers and language activists themselves: formerly, decisions about Breton linguistic structure and education were mostly taken by voluntary associations, such as the *Entente des écrivains bretons*, which codified the first supradialectal orthography in 1907 (Wmffre 2008:24); the association *Ar Falz*, which first fought for Breton to be taught in state schools from 1933, and later trained teachers of Breton (ibid.:41); and the Diwan school system, a network of private (although non-fee-paying) Breton immersion schools first set up in 1977 (Vetter 2013:156), which now operates 46 nursery and primary schools, six *collèges* and one *lycée*.[17]

In recent years, however, the movement has acquired an increasingly official status. In 1985, a CAPES examination was created for new Breton teachers (Jung and Urvoas 2012:40), allowing them to gain the same qualification as teachers of any other subject, and thereby permitting education in Breton to fit better into the structures provided by the state. Over the past decades, state education in Breton, whether as a "langue

14. See http://calandreta.org/fr/historique-fr, accessed 25 Oct 16.
15. See http://calandreta.org/archives/-Mapa-.html, accessed 24 Oct 16.
16. See http://calandreta.org/fr/historique-fr, accessed 25 Oct 16.
17. See http://www.diwan.bzh/sections.php4?op=viewarticle&artid=25, accessed 25 Oct 16.

vivante" or through partial immersion in schools with an "option bilingue", has increased dramatically, seeing constant growth over the last few years (Région Bretagne 2012:23). The total number of pupils educated in some form of Breton bilingual or immersion setting in the school year 2018–19 was over 18000,[18] this being around 2.6% of the total number of pupils in education in the *Académie de Rennes*;[19] the language is estimated to have just under 200000 speakers in total (Broudic 2009:10). The *Office de la langue bretonne*, originally set up in 1999, was re-established as the *Office public de la langue bretonne* (*OPLB*) in 2004, since when it has been financed by the regional government of Brittany, as well as departmental councils,[20] and run "as a public service institution" (Ó hIfearnáin 2013:122).

Breton's increasingly official position has meant that the relevant regional and departmental governments have become increasingly involved in producing language policy. While bilingual street signage has been present in parts of western Brittany since the 1970s (Vigers 2013:178) at the instigation of individual municipalities, the regional government did not formally recognise Breton as a language of Brittany (along with Gallo) until 2004, when it produced a report on the situation of the two regional languages and listed certain undertakings that it would pursue with the aim of supporting them (Ó hIfearnáin 2013:117). Several documents are now available detailing the regional government's policy on Breton;[21] these deal in particular with acquisition planning, focusing chiefly on education as a way of sustaining the Breton-speaking population, as well as the encouragement of related cultural practices including Breton-language publishing (Région Bretagne 2012). In these documents, it is stated that work on the linguistic structure of Breton itself, i.e. corpus planning, is the responsibility of the *OPLB*.

Corpus planning thus forms one of two major strands of the work of the *OPLB*, the other being the provision of information on Breton to individuals and businesses and the promotion of its public use. Its corpus planning is heavily focused on expanding the Breton lexicon, particularly through the division known as TermBret, which focuses

18. See http://www.brezhoneg.bzh/56-sifrou-pennan.htm, accessed 15 Apr 19.
19. Around 690000 pupils; see http://www.ac-rennes.fr/cid102458/l-academie-en-chiffres.html, accessed 15 Apr 19.
20. The relevant departments are shown on a map in Appendix B.
21. See http://www.bretagne.bzh/jcms/JB080117_6202/fr/langues-de-bretagne, accessed 14 Oct 16.

on creating Breton equivalents for French vocabulary.[22]

### 1.2.4 Language policy, ideology, and approach to the standard

Where regional languages are officialised in the form of policy, several issues must be considered, often including that of the presence of multiple varieties, where the language's minoritised status may mean that no standard or norm has yet been accepted by the majority of speakers. Approaches to this issue may vary widely due to the coexistence of conflicting language ideologies, a concept that will recur in this thesis. Ideology is particularly pertinent in cases of language revitalisation, where the decision to use a language must often be made consciously and is thus intimately connected to individual and collective beliefs about that language's function, and, therefore, to language ideologies, which are "concepts or notions [that] are viewed as derived from, rooted in, reflective of, or responsive to the experience and interests of a particular social position" (Woolard 1992:237). In cases of language revitalisation, the dominant ideologies, those most reflected in the way in which the language is used and the form that it takes, will thus be derived from the social position of those speakers who are most central to the revitalisation movement.

An especially important concept in this domain is "the ideology of the standard language" (Milroy 2007:133), proponents of which advocate language standardisation with the aim of minimising variation: this can be seen as beneficial for minoritised languages in that the already small number of speakers is not further fractioned by dialectal splits. Activists and policymakers taking this approach would concentrate on standardising the language and subsequently supporting this standard form above all others. This results in a consensus on which form of the language to use in official contexts such as signage and education, and can seem a particularly logical path to take when the official state languages used in the same communities have also been strongly influenced by this ideology, as is the case for French, which has historically seen strict legislation in favour of standardisation. Notions of prestige are also relevant here, as such ideologies would result in the attribution of prestige to the standard variety, to

22. See http://www.fr.brezhoneg.bzh/146-les-services-de-l-office-de-la-langue-bretonne.htm, accessed 14 Oct 16.

the detriment of non-standard dialects. Thus, in such contexts, these dialects are at risk of becoming negatively valued in comparison.

However, it is equally possible to find opposing ideologies, resulting in a wealth of different perspectives and diverse linguistic practices among a language's speakers. An ideological approach favouring diversity over uniformity, and traditionalism over innovation, would be likely to reject standardisation in favour of existing dialects, resulting in policies such as the *polynomie* found on Corsica (Colonna 2015:45) that promotes the prestige of all regional dialects, a rare exception to the standard language ideology that typically prevails in contexts relating to policy for the minoritised languages of France. For Occitan, competing ideologies have led to two opposing practices: Blanchet (2012:20) argues that an approach that maintains the separateness of the regional varieties of southern France, such as Béarnais, Gascon, and Provençal, is more relevant to most speakers and "beaucoup plus acceptée socialement" than the "Occitanist" and standard language ideology-aligned attitude of promoting the notion of Occitan as a single language.

For Breton too, standard language ideology is attested in the materials produced by the *OPLB* and the regional government. The *OPLB*'s only concession to variation is in a historical context: in its online historical corpus of Breton, containing examples from the fifteenth century to the present, it gives attested written forms of words in their original orthographies (along with a transcription in modern standard Breton).[23] Elsewhere in the materials it supplies, it advocates standard Breton[24] only, such as in information sheets that provide Breton vocabulary relating to specific domains: for example, transport-related vocabulary, feeding directly into the display of Breton in the linguistic landscape (OPLB 2016). On its website,[25] the *OPLB* contrasts the traditional dialects, "aujourd'hui en très grande difficulté", with the "langue moderne", "fixé par des grammairiens et des lexicographes". Like French, this form of Breton is able to "servir de véhicule à l'expression de la pensée humaine", with the consequence that it can be considered suited to contemporary use. Moreover, while the differences between the dialects are often claimed to be significant in academic work (e.g.

23. See http://meurgorf.brezhoneg.bzh, accessed 14 Oct 16.
24. Standard and dialectal Breton will be outlined more fully in section 1.3.
25. See http://www.fr.brezhoneg.bzh/4-histoire.htm, accessed 14 Oct 16.

Ternes 1992:378), the *OPLB* instead downplays them, claiming that "les différences dialectales sont peu marquées ... le vocabulaire et la grammaire varient peu. En général l'intercompréhension est bonne".[26] The overall picture of modern Breton that the *OPLB* strives to promote is one of uniformity, after the French model.

The Breton regional government also promotes a standardised form of Breton, stating in its language policy documents that the *OPLB*'s chief objectives are "l'adaptation du lexique et de la norme écrite" (Région Bretagne 2012:18). Rather than promoting Breton in its existing form, the *OPLB* is charged by the regional government with developing it so as to ensure that it functions as an appropriate language for modern society. This includes settling "orthographe (y compris veiller à sa bonne utilisation), normes, corpus lexicale, termes techniques, toponymie, anthroponymie, etc." (ibid.); many of these items involve the expansion of the lexicon, allowing Breton to be used in increasingly specific domains. Moreover, the term "bonne utilisation" evokes "bon usage", recalling the idea that that Breton is being placed in a French-style paradigm where a single explicitly codified standard should prevail. The *OPLB* therefore demonstrates its adherence to standard language ideology.

### 1.2.5 Language policy: conclusion

French state language policy strongly promotes the supremacy of standard French, and Breton language policy thus seeks to counter the hegemony of French with an equally unified Breton, regardless of how appropriate this may be for a minoritised language. The work of the regional government and the *OPLB* suggests that Breton is required to be a language appropriate for any setting, including science and technology, hence the need for continued lexical expansion involving increasingly specialised vocabulary. As section 1.4 will show, this may be at odds with the role of Breton for certain speakers. Breton language policy is thus in danger of excluding some parts of the already small speaker community, not only in its insistence on the use of norms and standards, but also in the domains in which it promotes the use of the language. These domains, mostly relating to the provision of public services, are evidently nonetheless the easiest

26. See http://www.fr.brezhoneg.bzh/4-histoire.htm, accessed 14 Oct 16.

to reach from the *OPLB*'s position as an institution backed by regional government.

While the French government appears to have reversed its former position on regional languages to a certain extent, as shown in the 2008 modification to the constitution stating that "les langues régionales appartiennent au patrimoine de la France", and gradually increasing support for bilingual education, the potential effects of this are undermined by continued symbolic affronts such as the presence of "la langue de la République est le français" in the constitution and the non-fulfilment of Hollande's promise to ratify the ECRML. Recently, more tangible damage has occurred as a result of the French state's level of involvement in the education system, hindering efforts in minority language acquisition planning: at the beginning of the 2016 academic year, two Breton-speaking teachers were, apparently accidentally, transferred to the *académie* of Versailles, leaving pupils in the "filière bilingue" without the teachers they required for history lessons and necessitating significant amounts of administrative work in order to rectify the situation, which was eventually resolved in mid-October.

Regional language policy has expanded significantly in recent years; in the case of Breton, this was triggered by the regional government's overt decision to support the language in 2004. Yet the *OPLB*'s ideological alignment with standardisation risks marginalising speakers who use varieties of Breton closer to the local dialects. Where the authorities do take a more major role in language planning, conflict between them and voluntary associations is still evident in the existence of multiple school systems, resulting in an uneven distribution of Breton-language education provision (Ó hIfearnáin 2013:126). As language policy continues to expand, it is important to ensure that it does so in a way that does not further marginalise certain speakers by excluding their linguistic repertoires and practices, and that it acknowledges the potentially contrasting role of Breton rather than necessarily viewing it as a homogeneous counterpart to French.

## 1.3 Previous research on Breton and its varieties

As a Celtic language, Breton has been the subject of academic work in various literary and linguistic fields going back several hundred years. Academic interest in Brittany and Breton can be identified as long ago as the aftermath of the 1789 French revolution, when scholars began to study the Celtic origins of France in the belief that these were more fundamental to the nation than the structures of monarchy and aristocracy, now out of favour, that they judged to be the result of the Germanic invasions of France in the fourth and fifth centuries (Guiomar 1992:63). The *Académie celtique*, established in the early nineteenth century, organised scholarly activities relating to the Celtic history of France, including studies of Brittany and its language; in this climate, the works of Jean-François Le Gonidec, including a grammar and dictionary of Breton, were eagerly received (ibid.:69).

During the nineteenth century, a tradition of ballad collecting took hold in Brittany, with the resulting anthologies serving as important linguistic and literary sources for the study of Breton and its dialects. In the twentieth century, as the discipline of linguistics grew, theoretical descriptions of Breton and its dialects became more prevalent, with notable works including the *Atlas linguistique de la Basse-Bretagne* (Le Roux [1924] 1977), Kenneth Jackson's *A historical phonology of Breton* (Jackson 1967), and the four-volume *Trésor du breton parlé* (Gros 1966–1989), the last of which focused on the Trégorrois dialect. The subsequent growth of sociolinguistics and interest in minoritised languages motivated work on the status of Breton as a regional language undergoing revitalisation, including critical perspectives on the revitalisation movement and its linguistic decisions (e.g. Wmffre 2008), and eventually work on new speakers of Breton, as section 1.4 will relate in more detail.

Work on Breton, including both academic studies and learning materials, typically divides the language into four geographically delineated dialects, coinciding with the territories of the four bishoprics within which Breton survived into the twentieth century: Cornouaillais, Léonais, Trégorrois, and Vannetais.[27] Beyond this broad division, detailed studies of Breton acknowledge further differentiation at a more local level.

27. A map of these dialect areas is shown in Appendix B (p. 307).

Jackson (1967), for example, refers to the varieties of specific communes on multiple occasions, and in an appendix on "the chief features of the dialects" (ibid.:834), deals with the phonology of twentieth-century Breton by dividing the language firstly into the four dialect groups, and then into more specific local varieties, such as north-east Léon, east Léon, north Léon, and so on. Other work has gone further in reducing reliance on the concept of a discrete set of dialects: some researchers from the Breton sociolinguistic tradition also use the term "badume", coined by Breton linguist Jean Le Dû from the Breton expression *ba du-mañ* ('in these parts'; Le Dû 2009). This denotes local varieties of Breton, sometimes specific to a single parish. By invoking both badume and dialect, linguists are able to represent the geographical division of Breton, and its perceptual repercussions for speakers, more accurately.

With the caveat that this further subdivision must be borne in mind, so as not to risk falsely characterising the dialects as homogeneous, the use of major dialect groupings seems appropriate for the current study. Basing the dialect areas on the bishoprics, while at times criticised for its lack of linguistic grounding (Timm 2010:723), in fact illustrates the bishoprics' significance in that their cathedrals would have been important centres of power in Brittany, certainly until the institution of *laïcité* in 1905. Just as Paris emerged as the source of standard French due to its being the locus of administrative and religious power in France, the episcopal seats in western Brittany, which had no single major administrative centre, were obvious locations of power within the region, particularly given that Brittany is traditionally a strongly religious area. Indeed, until the twentieth century, written Breton was mostly produced by religious bodies or for religious purposes (Wmffre 2008:16; ibid.:22). The lack of widespread influence from a supradialectal standard is shown in the existence of three dialectally distinct versions of the religious journal *Lihereu Brediah er Fé/Lizeriou Breuriez ar Feiz/Lizero Breuriez ar Fe* in the mid-nineteenth century, each intended for a different *département* (ibid.:17). However, a literary standard, based mostly on the Léonais dialect, had emerged in major religious publications dating back as far as 1659 (ibid.:8).

Some writers have advocated different ways of dividing Breton into dialects, suggesting instead that the south-western and north-eastern areas have coalesced and that a division into three dialects is more appropriate. Jackson (1967:18), while maintaining

the four-way distinction, admits that a "transitional dialect" is spoken in an area comprising the far east of Cornouaille and far west of Trégor; more recently, some scholars have advocated more strongly for the three-way split (e.g. Wmffre 1998:2). Like the division into four dialects, this division into three can be related to the geographical distribution of power in Brittany. Industrial advances at the turn of the twentieth century increased the importance of inland areas as a centre of large-scale farming activity (Cucarull 2002:62–65). After the arrival of the railways in the 1860s (ibid.:83), the town of Carhaix, on the eastern edge of the former Cornouaille bishopric, became the site of an "étoile ferroviaire" (Le Guyader 2002:328) with five railway lines emanating from the town, and hence a centre of traffic from all over Brittany. As a result of this, the dialect boundary slightly to the east of this area would have become significantly less clear-cut. This thesis will therefore refer to five major dialects: those of the four bishoprics along with this central Breton variety.

As the previous section indicated, the dialects of Breton now coexist with a standardised variety whose roots lie in the older literary standard. This was codified more explicitly by literary organisations in the early twentieth century, and is now typically written using the orthography known as *peurunvan* ('fully unified'), one of the three major supradialectal orthographies to have been developed since 1900. This variety is now associated with a fully codified grammar as elaborated most fully and most frequently in the works of the writer Roparz Hemon (Timm 2010:725), and used in the most popular textbooks and teaching grammars (e.g. Chalm 2009), as well as being subject to continued development in the form of corpus planning by the *OPLB*. Academic work on Breton typically claims that this standard variety is the preserve of new speakers, or of "néo-bretonnants", categories that the next section will explore.

## 1.4 New speakers and "néo-bretonnants"

### 1.4.1 The new speaker in sociolinguistic theory

Academic literature on the revitalisation of minoritised languages has grown significantly in recent years, with earlier works in the field (e.g. Fishman 1991) still exerting an

important methodological influence over today's research. However, the concept of the new speaker as a universal, cross-linguistic category has only recently been introduced into this discourse. In a typology of speakers of endangered languages, Grinevald and Bert (2011:51) propose the category of "neo-speaker", characterising it as composed of speakers who learn the language as a result of revitalisation programmes, thereby acquiring it as a second language. This definition does not entail fluency: these speakers' command of the language may be limited, but they have a positive attitude towards the language and are motivated to learn it.

While Grinevald and Bert (ibid.) note that "this type of speaker has not been referenced in the literature yet", the term soon came to be more thoroughly explored and elaborated in the slightly altered form of "new speaker". Hornsby (2015:108) defines this category of speaker with reference to "transmission, attitude and origin": new speakers acquire the language in the context of formal education; they have a positive attitude towards it; and they need not originate from the same population as traditional speakers, i.e. speakers who have acquired the language as a first language through intergenerational transmission within the home and community. Definitions put forward in other academic work are broadly similar, but can involve minor differences. McLeod and O'Rourke (2015:154) similarly define new speakers of Scottish Gaelic by their acquisition of the language outside a home setting, but include the additional criteria that such speakers are able to speak the language competently and that they use it frequently. Walsh and Ní Dhúda (2015:173–174) also include these criteria in a study of new speakers of Irish, but their definition differs again from others in that rather than specifying that the language is acquired outside a home setting, they refer only to a community setting. By this definition, new speakers may acquire the language in the home through being the children of earlier generations of new speakers. These examples show that the definition of the new speaker is not entirely fixed: it appears to vary slightly from one language context to the next, but remains generally based on the central idea that the language is not acquired as a result of a continuous chain of intergenerational transmission dating back indefinitely.[28]

28. Work by the COST New Speaker Network has played an important role in examining the implications of the various definitions of the new speaker and the extent to which the category can be used as a meaningful starting point for linguistic and sociological analysis. Much of this work was carried out at the same time as the writing of the present thesis, and the final report (O'Rourke and Pujolar 2019) was

While the concept of the new speaker has emerged only very recently in general work on language revitalisation, this is not the case for literature that deals specifically with Breton, where the term "néo-bretonnant" has been present for over twenty years (first appearing in Jones 1995:428). Prior to the emergence of this category, the distinction that tended to be made was between written literary Breton, largely a product of the early twentieth-century revitalisation movement, and the spoken, usually dialectal, varieties used by traditional speakers (e.g. in Gros 1970; Hewitt 1977; Ternes 1992). With the emergence of the term "néo-bretonnant" came the recognition that this standard literary variety was becoming the source of a spoken standard, used particularly by those acquiring the language in formal education settings.

Jones' (1995) "néo-bretonnant" is in many respects based on the "militant" Breton speaker category identified by McDonald (1989). Rather than considering new speakers a homogeneous category that acts as a counterpart to traditional speakers, McDonald instead divides them into two groups: those based around the university of Brest in western Brittany (traditionally Breton-speaking territory) who place more value on the language of traditional speakers and are less inclined to become involved with the political movements associated with the revitalisation of Breton; and those based around the university of Rennes in the east, who, on the contrary, favour the standard language and are typically more politically active. This has parallels in other minoritised languages: academic work has reported that some new speakers of Irish subscribe to "an essentialist discourse in which the ideal of the Gaeltacht native speaker comes across very strongly" (O'Rourke and Walsh 2015:69), while others instead align themselves with "Dublin Irish" (ibid.:74); similarly, "the most active Manx speakers [fall] into two broad currents with regard to the nature and standard of the language" (Ó hIfearnáin 2015:113), referred to as "purists" and "authenticists". Returning to Breton, McDonald's account of her ethnographic fieldwork reveals that she spent more time among the militant group; its greater prominence in her work and its greater ideological and linguistic difference from the traditional speaker category presumably contribute to the fact that this is the group that becomes emblematic of the new speaker in Jones' work. The association of new speakers of Breton with a greater militant fervency

published shortly before the thesis was submitted. The findings summed up in this report therefore did not explicitly influence the methodology of this research; however, certain points made in this thesis do tie into the conclusions of the report, and references to it will be made in the relevant parts.

and certain linguistic characteristics (see below) creates a distinct opposition between this group and traditional speakers.

References to the "néo-bretonnants", largely based on the distinctions drawn by Jones (1995; 1998b), have continued in subsequent literature, and tend to refer to this category as a homogeneous group, making up the majority of new speakers. New speakers of Breton have existed since at least as long ago as the 1920s, and were already strongly motivated by political factors (Leach 2012:66). However, the modern concept of the "néo-bretonnant" can perhaps be dated back to the widespread social unrest in France in the late 1960s, which encouraged a heightened sense of Breton identity (Abalain 2000:66–67; Le Nevez 2013:92). The acquisition of significant competence in Breton in a formal education setting became possible in the late 1970s, when the Diwan immersion schools were established. If we are to take this as a starting point for the growth of a significant new speaker category in the Breton context, it must be noted that the original pupils of these schools may since have had children and raised them speaking Breton, meaning that not all new speakers are strictly L2 Breton speakers.

While not all new speakers are always classed as "néo-bretonnants" in prior research (Hornsby and Quentel 2013:78), all "néo-bretonnants" are considered to be new speakers (Timm 2003:34; Le Nevez 2013:92), and often to form the most numerous and visible part of the Breton new speaker category (Jones 1995; Jones 1998b; Hewitt 2014:1), meaning that there is scope for features assigned to "néo-bretonnants" to be assumed to characterise new speakers of Breton in general. Indeed, some work has used the term "néo-bretonnant" to refer to any new speaker of Breton;[29] in this thesis, it will be reserved for the militant category of new speaker, typically assigned various additional characteristics, as the next section will illustrate. For the term "new speaker", on the other hand, we shall use a deliberately broad working definition in order to account for the potential breadth of this category: a new speaker of Breton is a speaker who has acquired the language by means other than uninterrupted intergenerational transmission that dates back indefinitely, i.e. somebody who is not a traditional speaker.

29. For example, Hornsby (2009b:x) defines "néo-bretonnants" as "'new speakers' of Breton, i.e. L2 speakers".

### 1.4.2 Concepts and stereotypes around the "néo-bretonnant"

While the "néo-bretonnant" category has become the most prominent type of new speaker of Breton, this group is associated with a number of stereotypical traits that are not always so strictly linked to new speakers in other language contexts. Work on Breton and its speakers frequently makes a particular set of claims about both speakers and language varieties. The language of "néo-bretonnants", referred to as "néo-breton", is most frequently characterised as highly standardised and non-dialectal, resistant to variation (e.g. Jones 1995:428; German 2007:146; Rottet 2014:212); it matches the variety promoted by language planners, as discussed above, conforming to the ideology of the standard language. It is additionally claimed that "néo-breton" possesses several characteristics reflecting an ideology that favours purism and highlights Breton's position as a Celtic language by emphasising its typological distance from French: the avoidance of SVO word order (German 2007:170), the use of neologisms with neo-Celtic qualities (German 2007:153; Adkins 2013:58; Hewitt 2016:1), and general hypercorrection in the domains of syntax and phonology to avoid sounding like French (German 2007:153; Hewitt 2016:16). On the other hand, aspects of "néo-breton" that instead make it more similar to French than traditional Breton is include a French-like accent and intonation (Timm 2003:42; Madeg 2011:64; Hewitt 2016:16) and syntax calqued on French (Jones 1995:429; Madeg 2011:64; Hewitt 2016:16)—despite the conflicting claim that "néo-bretonnants" make a conscious effort to make their syntax different from that of French.[30] Additionally, the language of "néo-bretonnants" is said to be characterised by use of conservative morphological forms (Hewitt 2016:16) and a propensity towards grammatical mistakes or a general lack of fluency (Madeg 2011:64; Hewitt 2016:16), indicating structural influence from French, presumed to be the dominant and native language.

Just as numerous as these claims about the features of "néo-breton" are claims about the backgrounds and aims of its speakers. They are characterised as "militants and scholars" (German 2007:187), which entails links with political activism or a desire to express Breton identity (Jones 1995:428; Le Nevez 2013:92; Rottet 2014:213) as well as

30. Hewitt (2016:16) claims that syntax is "either French or hypercorrectly anti-French". It is perhaps the case that "néo-bretonnants" are thought to use French-influenced syntax in spontaneous speech, and "anti-French" syntax in contexts that allow more reflection before formulating sentences.

being middle-class and well-educated (Jones 1995:428; Timm 2001:454; Adkins 2013:58; Rottet 2014:213; Manchec-German 2018:15). In contrast with traditional speakers, they are said to be young (Jones 1998b:321; Hornsby 2005; German 2007:153; Adkins 2013:58; Rottet 2014:213) and located in urban areas (Jones 1995:428; Abalain 2000:76; Timm 2001:454; Timm 2003:34; German 2007:153; Hornsby and Quentel 2013:75), sometimes in eastern Brittany, outside the traditionally Breton-speaking area (Jones 1998b:321); moreover, they are literate in Breton (Jones 1998b:321; Hornsby and Quentel 2013:75), and desire to use it in as many domains as possible (German 2007:187; Hornsby and Quentel 2013:75–76), again relating to their strong sense of Breton identity. As with certain definitions of new speakers in other contexts, they usually acquire Breton as a second language (Timm 2003:34; Le Nevez 2013:92; Manchec-German 2018:15), in an educational setting (Jones 1995:428; Adkins 2013:59; Le Dû and Le Berre 2013:53; Rottet 2014:213).

With the "néo-bretonnant" category acting as the most prominent type of new speaker of Breton, these characteristics place the category of new speakers in opposition to that of traditional speakers, as Table 1.1 shows.

| **New speakers/"néo-bretonnants"** | **Traditional speakers** |
|---|---|
| younger | older (Jones 1998b:321; Timm 2003:34; Hornsby 2005:195; German 2007:153; Adkins 2013:59; Rottet 2014:212) |
| well-educated | less formal education (German 2007:153; Adkins 2013:58) |
| located in urban areas | rurally located (Jones 1998b:321; Timm 2001:454; German 2007:153; Adkins 2013:59; Hornsby and Quentel 2013:75; Rottet 2014:212) |
| middle-class | poor/working-class (Timm 2001:454; German 2007:150; Le Nevez 2013:92) |
| wish to use Breton in multiple contexts | use Breton in oral/intimate contexts only (German 2007:150; Hornsby and Quentel 2013:75; Le Nevez 2013:92) |
| literate in Breton | not literate in Breton (Jones 1998b:321; Adkins 2013:63; Hornsby and Quentel 2013:75; Rottet 2014:212) |
| speak Breton to emphasise Breton identity | ashamed of competence in Breton (McDonald 1989:285–286; Jones 1995:430; Adkins 2013:61) |

Table 1.1: Characteristics assigned to new and traditional speakers of Breton

The attribution of these various qualities divides the overall speaker population into two distinct categories, opposing "néo-bretonnants", seen as representative of new speakers, and their language variety to traditional speakers and theirs, with reference not only to how Breton is acquired but also to other aspects of their backgrounds and practices. The ideal new speaker, conforming to these definitions, is the stereotypical "néo-bretonnant" for whom Breton language policy is designed, as indicated by its focus on standardising the language, making it available in education, and extending its use to domains not associated with traditional speakers. This focus on the promotion of "néo-breton" feeds into several scholars' predictions that it is the only variety of the language that will survive into the future (Jones 1995:427; Jones 1998b:333; Timm 2005:37; German 2007:188). Despite this, the little quantitative research done on the demographics of Breton speakers indicates that new speakers make up only a very small proportion of the Breton-speaking population: estimates usually range from

around 5% to 10% (Hornsby 2009b:62–64), although due to the age profiles of the two groups, this is of course an increasing proportion. This prioritisation of a small subset of the speaker population, entailing the aim of making Breton a usable language for as many contexts as possible, has been cited in research as a factor contributing to traditional speakers' low sense of pride in their own linguistic competence in dialectal Breton (Adkins 2013). The same research often seeks to redress this balance, pointing out the artificial quality of standard Breton while viewing the traditional varieties as more authentic (e.g. Adkins 2013; German 2007; Hewitt 2014).

This, however, relies on a dichotomy between new and traditional speakers that appears somewhat inflexible, assigning criteria to both categories that need not necessarily correlate with the manner in which one acquires Breton. Indeed, despite the pervasiveness of the split between new and traditional speakers in academic literature, the same literature contains occasional indications that the line between the two groups can be blurred. Firstly, McDonald's (1989:156–157) division of new speakers into those based around Rennes, who conform more to the criteria now stereotypically associated with the "néo-bretonnants", and those based around Brest, whom she characterises as less politically active and more attentive to traditional Breton, recurs from time to time in subsequent work. While the latter group was largely ignored by Jones (1995; 1998b) in favour of illustrating the apparently stark divide between traditional speakers and the more extreme group of Rennes-based "néo-bretonnants", Hornsby and Quentel (2013:78) use the term "native authenticists" to refer to new speakers who insist on the superiority of traditional speakers' language, taking an ideological position that opposes the ideology of the standard; in the same way, Hewitt (2014:1) draws a distinction between "neologising activists", who conform more closely to the stereotypical description of the "néo-bretonnant", and "popularising activists", who place a greater value on traditional Breton. Rottet (2014:213), like McDonald, divides new speakers into "Celticists", who favour Celtic-based neologisms, and "traditionalists", who "see as authentic only the actual usages of traditional native speakers". Researchers emanating from the tradition of the university of Brest, who often favour the Breton of traditional speakers, may speak Breton natively but are set apart from the typical traditional speaker paradigm by virtue of being educated and able to read

and write both standard and dialectal Breton; inversely, they may have initially learnt Breton in a formal educational setting but made a deliberate effort to acquire a regionally marked variety. Likewise, the courses that provide most adult education in Breton, often criticised for their use of the standard language (Hornsby 2015:110), in fact show evidence of a degree of dialectal specificity: for example, certain courses taking place in the Vannetais area prioritise the Vannetais dialect (Adkins 2014:124). Similarly, even the Diwan immersion schools, seen elsewhere as a principal breeding ground for "néo-breton" (German 2007:152; Manchec-German 2018:1), use local dialects to a significant degree (Adkins 2014:124). Indeed, the Diwan schools' original policy was to prioritise "local" or "popular" forms of Breton (McDonald 1989:204–205), and still today, "many individual teachers do their best to teach a form of Breton that is in tune with that which is used in [the] local community" (German 2007:153). At a more official level, the departmental council of Finistère has run the *Quêteurs de mémoires* project, a master-apprentice style programme bringing younger learners into contact with traditional speakers; however, "its success has been limited" (Manchec-German 2018:23). As an example of a recent initiative, this nonetheless highlights the possibility that new speakers' attitudes, and the types of language they wish to use, could be changing.

Just as the "néo-bretonnant" category therefore appears not to be so easily mapped onto the new speaker, some traditional speakers (as defined by virtue of their having acquired Breton through uninterrupted intergenerational transmission) also exhibit characteristics that distance them from their stereotypical description. Le Coadic (2013:34) shows that despite the claim that many traditional speakers feel no attachment to Breton identity, many of them do in fact engage with such an identity in diverse ways; for example, Ó hIfearnáin (2013:120) surveys a sample of traditional speakers who are involved in Breton language teaching, and are themselves young, well-educated and literate in Breton, which contradicts the typical definition of traditional speakers. The existence of this group shows that despite a perceived gap between new and traditional speakers, and an apparent unwillingness on the part of each group to interact with the other, this is in many respects not true: indeed, younger new speakers can look to traditional speakers as mentors (Adkins 2013:64). Kennard (2014) shows that beliefs about the linguistic qualities of the Breton acquired by new

speakers in educational settings are not always borne out: school pupils acquiring Breton are shown, after some years of education, not to rely on SVO sentence structures, as criticised among new speakers by Hewitt (2016), but instead to use syntactic patterns that more closely resemble traditional Breton. Le Pipec (2013) points out the ambiguity of the term "néo-bretonnant" with reference to the fact that new speakers may acquire Breton from new speaker parents or grandparents rather than in an educational setting: as a replacement for the new/traditional speaker binary, he posits a continuum from "acquis" to "appris" and from "exogène" to "endogène" that shows the greater complexity of the situation, reflecting the fact that educated Breton speakers, and those who learn Breton in adulthood, are more explicitly able to make choices about their use of a specific language variety so as to correspond with their own identity, and enable communication with speakers from diverse backgrounds. This marks a recognition of the potential unsuitability of the usual dichotomy, replacing it with a continuum providing a more detailed representation of the actual state of affairs. However, such an approach is rare in the literature, and as shown above, many scholars continue to make use of a binary distinction between new and traditional speakers, associating each group with particular characteristics, and ascribing certain linguistic qualities to the variety of Breton thought to be used by new speakers that result in the devaluation of this variety. This identification of a discrete category of new speakers with seemingly homogeneous aims and practices encourages the development of language policy that reflects only a restricted vision of the functions that Breton has or should have in society, and may fail to take into account any recent societal developments in the Breton-speaking population.

The dichotomisation of Breton speakers into "new" and "traditional", and the equation of the former category with the stereotypical "néo-bretonnant", bringing with it an extensive collection of associated attributes as identified above, characterises the Breton-speaking population as two disparate homogeneous masses, and language policy certainly caters for this categorisation. However, any groups of speakers who do not match these stereotypes—for example, the 17% of speakers under 40 in Carhaix who acquired the language from parents or grandparents,[31] according to a 2013 survey (OPLB

31. While these may not be traditional speakers in all cases, this figure still goes against the stereotype that younger speakers of Breton learn the language through institutional means.

2014:12), or the sample of traditional speakers working in education identified by Ó hIfearnáin (2013:120)—risk being marginalised by this policy. Perceptions and practices of actual speakers need careful examination in order to ensure that provisions affecting Breton are put in place with an understanding of the diverse potential roles of the language and its importance in the construction of individual local identities, particularly as those identities evolve and are renegotiated.

## 1.5 Breton-language media

As a minoritised language, Breton does not have a particularly dominant position in the media. Its presence on television in particular is minimal: while in the UK the place of regional languages on television can be as prominent as an entire channel in the language (in the case of S4C in Wales), this is not the case in France. For Breton, regular television broadcasting is restricted to a daily 30-minute news programme on state regional television. In 2000, a privately-funded channel in Breton was launched; Moal (2004) provides an assessment of the first years of this channel's operation. However, Moal's warning that "the overall political climate doesn't seem too favourable" (ibid.:33) was to prove important, as by 2008, the channel no longer produced content in Breton (Broudic 2011). The internet has nonetheless enabled access to a wider range of Breton-language video content, particularly via the streaming website Brezhoweb; this service is provided by volunteers, however (ibid.), and produces little original content.

Radio broadcasting in Breton is more developed, with a choice of stations available. In France, the nationwide public radio service receives state funding and operates a number of stations broadcasting both nationally and regionally. Its regional station franchise, France Bleu, comprises 44 local radio stations,[32] roughly one between each two of the departments of France. Accordingly, the administrative region of Brittany, consisting of four departments, is served by two of these local stations, France Bleu Breizh Izel and France Bleu Armorique, which operate in the west and east of the region respectively. The area covered by France Bleu Breizh Izel therefore roughly cor-

32. See https://www.francebleu.fr/france-bleu-dossier-de-presse, accessed 19 Jan 2017.

responds to the territory in which Breton has traditionally been spoken, and the station pays homage to this in its use of the name *Breizh Izel*, the Breton term for this area.[33] It broadcasts three five-minute news bulletins in Breton each weekday morning, along with longer programmes that often focus on cultural subjects, coming to about three hours every day in total. Additionally, daily short segments in French aim to teach the listener a Breton word or phrase. France Bleu Armorique also broadcasts in Breton, in smaller quantities.

Associative radio stations, not affiliated with the state network, also exist. The principal associative network broadcasting in Breton is Radio Breizh, a group of four stations originally established in order to provide Breton-language programming before such provisions were made by Radio France. Each station is localised to a particular dialect area where it broadcasts on FM frequencies, although in practice programmes are shared among stations, meaning that distinctions among stations according to dialect are often not observed. All are additionally available online. Two of these stations broadcast programmes only in Breton, while the other two also produce and broadcast French programmes. Like France Bleu Breizh Izel, Radio Breizh broadcasts news bulletins, longer cultural features, and programmes that focus on specific Breton phrases.

Written media in Breton is also produced by a range of sources, although none has achieved particularly high circulation figures. Print media in Breton has existed since the nineteenth century, originally in the form of religious periodicals (Wmffre 2008:17). In the early twentieth century, these were joined by literary publications such as *Gwalarn*, contributors to which were instrumental in the initial phases of the revitalisation of Breton, including the development of the *peurunvan* orthography, now the most commonly used in both published books (Abalain 2000:85) and official documentation. Today, both literary periodicals (e.g. *Al Liamm*) and religious ones (e.g. *Kannadig Imbourc'h*) are still produced in Breton, and are equally joined by a number of more general publications, including the weekly newspaper *Ya*, which had nearly 1200 subscribers in 2011 (Broudic 2011). Occasional Breton columns are also included in local editions of the French newspapers *Le Télégramme* and *Ouest-France*, which have much larger circulation figures.

33. France Bleu Breizh Izel and France Bleu Elsass (Alsace) are the only local stations on the France Bleu network to be named in a regional language.

Going by the typical definitions of new and traditional Breton set out in the previous section, we are encouraged to perceive the media, particularly the specialist Breton-language press and dedicated radio stations, as a context that will be populated mostly by new speakers. If traditional speakers restrict their use of the language to the home and local community, as research has asserted, it follows that the language found in the media, which is produced for public platforms and accessed from across substantial geographical distances, should mostly be produced by new speakers. Other attributes assigned to traditional speakers equally suggest that their role in the production of such media should be limited: their supposed "illiteracy" in Breton precludes them from creating or consuming print journalism, while the belief that the two groups of speakers do not typically communicate with each other also suggests that the media will make no effort to reach this type of speaker.

This is even more acute in the case of the internet and social media: canonical traditional speakers will have restricted access to the latter as a result of their supposed greater age and associated lack of technological skills. Additionally, "the internet-user population consists primarily of technologically proficient, educated, white, middle-class, professional males" (Hewson 2003:291), many of these traits corresponding with stereotypical "néo-bretonnant" attributes.[34] Indeed, use of the internet among Breton speakers seems particularly low: "plus de 9 bretonnants sur 10 ignorent un outil aussi populaire qu'Internet" (Broudic 2009:189). The internet is thus also a pertinent context that we can expect to be a major source of new speaker language.

The media, including the internet, are thus important settings for the use of Breton by new speakers, and will serve as the core source of data for the current study; the specific sources used in this research are described more fully in section 2.2.1.

34. As Hewson points out, however, "this trend is becoming less prominent with the rapid growth of the internet" (Hewson 2003:291), and in the time since this statement was made, the situation will undoubtedly have evolved.

## 1.6 Research questions

In light of the discussion in this chapter, this thesis aims to investigate the stability of the binary divide between new and traditional speakers of Breton, with reference to both linguistic data and speakers' perceptions. As section 1.4.2 pointed out, these categories have been a particular focus of much previous research, which has frequently relied on particular stereotypes around the "néo-bretonnant", while at the same time indicating that the Breton-speaking population may in fact be more difficult to group into two discrete categories. Language policy for Breton tends to operate on the assumption that there is a homogeneous category of "néo-bretonnant" new speakers, using a specific standardised variety; accordingly, official corpus planning has often focused on the expansion of Breton vocabulary. The lexicon will hence form the basis of the linguistic enquiry undertaken here. As the previous section shows, the media provides an ideal context for investigating the language of new speakers, and chapter 2 will set out how this will be undertaken. This study will hence be of relevance to work on new speakers of other languages, as it focuses on one where this category has been present in the literature for some time, highlighting potential issues to be navigated as the field develops in additional contexts.

Emerging from the discussion in this chapter, the research questions examined in this thesis are as follows:

1. Do features attested in practice support or challenge the way in which academic literature depicts "néo-breton" in terms of its linguistic features, with specific reference to the lexicon?

Academic literature has characterised "néo-breton", the language variety thought to be used by "néo-bretonnants" and by extension the majority of new speakers of Breton, in terms of various attributes, listed earlier in this chapter. By focusing on Breton in a domain typically associated with new speakers, this thesis will seek to determine whether or not some of these features are indeed attested, and to what extent standard Breton is used in certain non-official contexts. In order to confine this study to an appropriate size, it will focus on the lexicon of Breton, a domain that is salient in lan-

guage policy and provides a number of opportunities for analysis. Chapters 3 and 4 in particular will contribute to answering this question, focusing on the analysis of a corpus of Breton in media contexts, and presenting broad statistical trends and more fine-grained observations respectively.

2. Is the depiction of "néo-bretonnants" in academic literature matched by how new speakers engage with Breton?

As with the stereotypical characteristics of "néo-breton", this chapter has enumerated some of the features typically ascribed to the "néo-bretonnant" in the academic literature. Investigating both linguistic aspects of the corpus gathered and relevant metadata regarding speakers' backgrounds and self-declared identities will help determine whether or not this rigid characterisation of the "néo-bretonnant" category is valid, and additionally reveal the attitudes and ideologies that motivate the use of Breton among new speakers, showing how they balance concerns of identity and communication. This will be the main focus of chapter 5, which presents findings from interviews carried out with professionals working in Breton-language media.

3. How does use of Breton vary among media and registers?

By analysing the use of Breton in a variety of media domains—radio, print journalism and the internet—this research will aim to point out any variation among different media sources, drawing conclusions about how context, medium and register may bring about variation even within a broadly standardised linguistic variety. Again, this will be covered in chapters 3 and 4, but particularly in the former chapter, which will use statistical methods to directly compare parts of the corpus according to various criteria. Relevant observations will also be made in chapter 5 based on interview participants' discussion of their use of Breton and the roles of the organisations that they work for.

4. What patterns may be emerging relating to changes in the way Breton is used and in its linguistic characteristics, and what is the role of the media in this?

Finally, this thesis will suggest how practices around Breton may be evolving among

contemporary speakers as the speaker population develops, based on this particular research context. Again, chapters 3, 4, and 5 will be of relevance: chapters 3 and 4 will highlight trends in language use that may indicate departures from the typical descriptions of "néo-breton", while chapter 5 will identify similar trends relating to speaker practices.

# Chapter 2

# Methodology

## 2.1 Introduction

This chapter describes the methodologies used for gathering and analysing the data investigated in this thesis. As identified in the research questions, the linguistic focus of this study is on lexicon. The analysis of the Breton lexicon is thus realised through the creation and annotation of a corpus. This forms the bulk of the investigation, and the methodologies involved are described in section 2.2. Additionally, in order to investigate some of the non-linguistic stereotypes surrounding "néo-breton" listed in section 1.4.2, a second phase of the data collection and analysis involved carrying out a number of interviews with speakers. This allowed the examination of speakers' perceptions and the relationship between language and identity to be probed in more detail. The methodologies employed for this part of the project are described in section 2.3. Appendix C contains a link to publicly accessible versions of the corpus and interview transcriptions.

## 2.2 Corpus methodology

Section 1.4.2 lists various linguistic criteria proposed in academic literature that distinguish "néo-breton" from the language of traditional speakers, which fall into a range

of areas, encompassing Breton's phonology, morphology, syntax and lexicon. This study focuses on the last of these four categories. As stated in chapter 1, the lexicon of Breton is overtly regulated by the *OPLB*, this forming part of a long history of codification resulting in multiple dictionaries over the course of the language's revitalisation. Moreover, the lexicon of "néo-breton" has attracted particular comment in academic literature (e.g. Jones 1995:428; German 2007:153; Adkins 2013:58; Hewitt 2016). It seems a particularly appropriate feature to consider in determining whether contemporary usage correlates with claims made about new speakers.

Existing research on the "néo-breton" lexicon tends to highlight specific lexemes in order to show that new speakers apparently use a highly standardised, Celtic-derived vocabulary; systematic studies of linguistic data are rare. Rottet (2014:237) examines a range of twentieth- and twenty-first-century Breton dictionaries, showing the variety of approaches taken to finding equivalents for French headwords such as *végétarien*, rendered in various ways among his sources: *vejetarian*, an orthographically adapted loan from French; *plant-debrer*, a compound literally meaning 'plant-eater', made from pre-existing Breton vocabulary, although *plant* clearly derives from French; and *debrerlinad*, *linad* being an uncommon Breton word evidently used to avoid the use of any French component. This exemplifies the various attitudes of the dictionaries' writers towards the role new Celtic- and French-derived vocabulary should play in Breton, and Rottet's findings certainly indicate that the Breton of speakers who use such dictionaries may be more diverse in nature than is typically suggested. Examining actual usage will thus be a crucial correlate to this work, particularly given that the language of new speakers has not previously been examined using a corpus-based approach. Compiling a corpus permits the quantitative analysis that forms a central component of this thesis, and will also provide data for future research on modern Breton. The rationale for the use of a corpus methodology is discussed in more detail in section 2.2.2.

### 2.2.1 Data sources

Section 1.5 sets out the rationale for collecting data from media contexts. While the sources concerned can be associated with new speakers, going by the typically de-

scribed characteristics of the "néo-bretonnant", no attempt is made to exclude data produced by speakers outside this category. As this project interrogates beliefs about the nature of "néo-breton", its users, and the contexts in which it occurs, it would be dangerous to assume the reliability of such beliefs as a basis for investigation. Instead, the data in themselves form the starting point for linguistic analysis. It is not the intention of this project to examine only stereotypical "néo-breton", but instead to investigate a context where "néo-breton" is thought to be the dominant variety.

Within the broad domain of enquiry of media Breton, three types of discourse are examined, one oral and two written: radio programmes, print journalism, and online social media. While none of these discourse types represents the most informal, conversational register, nor the most formal literary one, differences in register that are salient for this research exist among the three. Social media posts tend to be produced by individuals acting in a personal capacity and are not subject to editorial control, permitting the use of lower registers than in the other two contexts. The register of articles in printed media can vary according to the subject being discussed and the intended audience; radio programmes contain a mixture of scripted segments and interviews, the latter being more likely to be a source of lower-register language due to their conversational nature and greater spontaneity. As a result of these differences, register must be borne in mind when analysing the data, and it is likely that differences will consequently be identifiable across the three discourse contexts, giving a fuller picture of how the Breton of new speakers might vary according to register.

#### 2.2.1.1 Radio

Section 1.5 outlines the extent of radio broadcasting available in Breton. The data for the radio subcorpus are taken from both France Bleu Breizh Izel and Radio Breizh, and from both news programmes and longer feature programmes. Examining samples from both state and associative radio allows investigation of whether the independence of the latter from central control may allow it to include less standard varieties of Breton in its news programmes and features. The France Bleu data come from the news programme *Keleier Breizh*, broadcast every weekday morning, and *Breizh O Pluriel*, an

hour-long programme that focuses on Breton culture and includes interviews with local artists, broadcast every Saturday and Sunday. The Radio Breizh data are taken from the ten-minute news bulletin *Keleier ar vro*, broadcast every weekday across the network, and the hour-long cultural programme *Sevenadurioù*, broadcast monthly on two of the four stations. As the only news bulletins on their respective stations, *Keleier Breizh* and *Keleier ar vro* are obvious data sources. *Breizh O Pluriel* and *Sevenadurioù* were chosen for their similarly topical subject matter, and their lack of linguistic focus; in linguistically-themed programmes, such as those that teach Breton phrases, certain tokens would be over-represented, and it is also likely that speakers would be more aware of how they use language. Moreover, the programmes selected do not explicitly focus on traditional Breton and its speakers, unlike some of those broadcast on the Radio Breizh network: as the aim of this research is to examine language associated with new speakers, this type of programme must be excluded.

#### 2.2.1.2 Print journalism

Three print publications are used as sources for this research. The first such source is *Ya*, a newspaper published once a week—the only publication in Breton to be released so frequently. As the most widely read Breton-language periodical (Broudic 2011), it was an obvious source for this project. *Ya* covers a variety of subjects, including news about local events and sports, as well as features on the topic of culture and other minoritised languages. It is published by the association Keit Vimp Bev, based in a small town in inland western Brittany, the area where traditional Breton has endured most successfully.

The second print source is the magazine *Bremañ*, published every two months by Skol an Emsav, one of the largest organisations to provide Breton adult education classes, based in Rennes in eastern Brittany (and hence outside the traditionally Breton speaking area). This is thus a significant difference from *Ya*, and therefore the publication seems useful to include as a potential source of contrast in its use of language. Again, it is one of the most widely circulated Breton-language publications (ibid.), and like *Ya*, has a broadly generalist subject matter. In this case, articles with an overtly linguistic

focus tend to be restricted to Skol an Emsav's other magazine, *#brezhoneg*, aimed at learners; *Bremañ* is intended for those who read Breton competently.

The third print source is the magazine *Brud Nevez*, which, like *Bremañ*, is published every two months. This is the longest-established of the magazines in the sample, having existed in its current incarnation since 1977, and before that as *Brud* from 1957 (Wmffre 2008:746). *Brud Nevez* is produced by the publishing house Emgleo Breiz, a principal advocate of the *skolveurieg* orthography developed as an alternative to the de facto standard, *peurunvan*, during the 1950s, and through its use of this less popular orthography is situated more peripherally within the Breton revitalisation movement, meaning again that different linguistic characteristics may become apparent during analysis. Again, it publishes features relating to news and culture in Brittany and in the wider world, but it additionally includes poems and short stories. Due to the specificities of the literary register, these sections of *Brud Nevez* do not form part of this research.

#### 2.2.1.3 Social media

The third context to be examined is online social media. While still associated with new speakers of Breton, as discussed in section 1.4.2, this context is crucially different from the previous two in its lack of an editorial process and openness to any participant. Interactional turns are brief and tend to be expressed in informal language, characteristics that usually typify spoken rather than written communication: accordingly, computer-mediated communication of this type has been characterised as a new, hybridised, variety where, despite the written medium, "features of the spoken language varieties also form part of the new electronic register" (Kallweit 2018:400). Examining language of this type, not restricted to contributions made by those specifically employed or commissioned to produce content, also allows access to a larger, less professionalised body of language users, while maintaining the focus on domains typically associated with new speakers.

Social media offer a variety of ways to communicate, associated with varying degrees of public access. The most popular social media platform in France is Facebook, which

surpassed 30 million users in the country in mid-2015.[35] Twitter, typically a more public platform, is also widely used, and a number of its users tweet regularly in Breton; a list of such accounts is maintained by the website *Indigenous Tweets*.[36] However, several factors mean that Twitter is less conducive to sustained interactions than Facebook: many of the accounts using Breton are operated on behalf of organisations rather than individuals; tweets are often used to link to content on other websites rather than inciting a conversation within Twitter itself; and at the time of data collection each tweet was limited to 140 characters. Accordingly, looking at Breton-language tweets revealed that few lead to conversational interaction, and it consequently seemed a poor source for data.

While many interactions on Facebook take place through private messages or on users' profile pages, restricted from public view, communication also occurs through "groups", pages where members with a common interest can share information and seek opinions. A number of such groups specifically cater for users of Breton, the most populous and active among them being *Facebook e brezhoneg*, which as an "open group" allows any user logged into Facebook to view its members' posts. Originally created in order to develop a translation of the Facebook interface into Breton and lobby for its implementation, which was achieved in late 2014 (Guivarch-Tonnard 2014), it has since developed into a more general discussion group, with posts in both Breton and French, discussing Breton culture and identity as well as linguistic matters. Unlike some other similar groups, it is not overtly political in nature and is not intended as a lobbying group for the promotion of the increased use and visibility of Breton in public life, although inevitably this is frequently discussed. At the time of data collection, the group had over 10000 members,[37] whose interactions form the source of the third body of data under investigation here.

35. See http://www.ouest-france.fr/high-tech/facebook-un-francais-sur-deux-actif-sur-le-reseau-3653032, accessed 24 Jan 17.
36. See http://indigenoustweets.com/br, accessed 24 Jan 17.
37. The number of members given at https://www.facebook.com/groups/334727793245979 on 24 January 2017 was 10658.

### 2.2.2 Data collection methodology

This project uses a corpus-based methodology, where the data are gathered in three subcorpora corresponding to their three different sources. This entails the creation of a machine-readable file—the corpus itself—containing the data collected as well as metadata relating to the sources used. The data are annotated in order to enable the identification and analysis of relevant features. Baker (2010:8) stresses the potential interface of sociolinguistics and corpus linguistics in that both make use of quantitative techniques to analyse samples of data and draw conclusions about wider populations, showing the utility of corpora for sociolinguistic research in their capacity for the analysis of data. Corpus methodologies can be particularly useful for comparing differences across registers (Conrad 2015:310), a dimension of variation particularly relevant to this research, given that three different media contexts are involved. As the overall objective of this part of the study is to show whether evidence from linguistic data confirms or invalidates claims made in the literature about the nature of "néo-breton", it is especially appropriate to compile a corpus: this activity frequently "[brings] to light data that are not in accordance with ... what the specialised literature has assumed to be the case" (Valera-Hernández 2014:169).

Most existing corpora of Breton focus exclusively on the language of traditional speakers; various examples can be accessed online.[38] Two linguistic atlases also provide prominent sources of data for traditional Breton: the *Atlas linguistique de la Basse-Bretagne* (*ALBB*; Le Roux [1924] 1977), compiled during the 1920s, and the more recent *Nouvel atlas linguistique de la Basse-Bretagne* (*NALBB*; Le Dû 2001). Both follow the principles of traditional dialectology in eliciting certain lexemes, often farming-related, from older, rurally-located informants, and seeking principally to illustrate the range of dialectal variation in Breton phonology and its lexicon. All these sources thus deliberately exclude new speakers and their varieties as a result of adherence to the traditionalist ideologies that historically guided dialectological work. "Néo-breton" does feature, however, in Meurgorf, the *OPLB*'s online "dictionnaire historique"[39] based on a selection of texts ranging in publication date from the

38. e.g. *Banque sonore des dialectes bretons* (http://banque.sonore.breton.free.fr/index.html), and the sound files from which the *Nouvel atlas linguistique de la Basse-Bretagne* was compiled (Le Dû 2014)
39. See http://meurgorf.brezhoneg.bzh/bibliographie, accessed 23 Jan 17.

fifteenth century to today. This corpus is based on dictionaries, literary periodicals, and published fiction and non-fiction books, orientating it towards formal registers of language. Both the highest literary registers of language[40] and lower registers of traditional dialectal Breton are therefore available in corpora; there is scope for the collection and examination of data of intermediate registers, providing evidence that will help answer the question of whether a division between traditional Breton and "néo-breton" pertains in all contexts.

Corpus creation entails the selection of conventions for transcribing and annotating the data. This project uses the Text Encoding Initiative (TEI), a set of conventions[41] that stores the annotated data and associated metadata, such as broadcast date or page number, in a single file. The TEI is based on the XML markup language, which is internationally recommended for interoperability and the long-term preservation of data, and is particularly well-suited to the creation of corpora of written texts, making it especially appropriate for this work, given that two of the three corpora are compiled from written sources. However, this does not mean it is unsuited to encoding speech, particularly as radio programmes tend to be at least partially scripted and contain few instances of overlapping conversational turns. Precedent has additionally been set for encoding oral data using the TEI, involving the development of specific conventions relating to the transcription of speech (Baker 2010:104).[42] Moreover, this encoding standard "is flexible in that it allows (indeed requires) individual corpus users to add to or adapt its guidelines where needed" (ibid.:15). This is a necessity for this research, given that Breton corpus linguistics is not a field with an associated established theoretical framework, and that the types of language this project focuses on are particularly new to systematic analysis of this type. Some elements of the transcription and annotation must therefore be designed for the purposes of this project.

40. Since the beginning of the revitalisation process, this tends to have been produced by new speakers, although literature predating the twentieth century would have evidently been the product of educated traditional speakers (see Le Pipec 2013:105). This literary register of traditional Breton largely ceased to be used as the language's decline accelerated in the early twentieth century.

41. These are specified at http://www.tei-c.org.

42. See also the "Transcriptions of Speech" section of the *TEI Guidelines* at http://www.tei-c.org/release/doc/tei-p5-doc/en/html/TS.html (accessed 3 May 17).

#### 2.2.2.1 Ethical considerations relating to online data

In the case of internet data, ethical issues around the protection of human subjects must be taken into account. While posts in public Facebook groups are openly available without any kind of access restrictions, most users are unlikely to consider this public accessibility when sharing content: indeed, "simply because a user has posted on a Facebook page does not mean they intend or expect for their comment to be reproduced" (Pihlaja 2017:223). Researchers therefore have a duty to respect Facebook users' privacy by implementing measures that protect their identity. In accordance with this, when collecting the Facebook data for analysis in this thesis, an initial step to ensure ethical soundness involved the removal of names and profile identifiers from the posts sampled.

This measure was not sufficient by itself, however. Zimmer (2010:7) recounts a research project, also based on data from Facebook, where

> "while the ... researchers might have felt simply removing or coding the subjects' names or other specific identifiers from the dataset was sufficient, had they followed the European Union's guidance, they would have recognised that many of the subjects' ''physical, physiological, mental, economic, cultural or social identity'' could also be used for re-identification. Even after removing the names of the subjects, since the dataset still includes race, ethnicity, and geographic data, re-identification of individual subjects remains a real possibility."

To this end, no demographic information was collected relating to the Facebook users, and no details were retrieved from their profiles: the data analysed came only from the *Facebook e brezhoneg* group and were purely linguistic.

Working with human subjects would have required ethical approval (as it did for the interview-based part of this research; see section 2.3.4 below), which would have been difficult to obtain in this case, as it would have required written consent from all 99 "speakers" in the Facebook sample. A further precaution was therefore taken to ensure that the obtention of the Facebook data could not be classified as human subjects

research. This comprised the avoidance of quoting passages from this part of the corpus, as quoted material would be searchable online and would enable the identification of the original user: "if a researcher publishes the text that a subject posted, without the subject's name or even with a pseudonym, the dynamic searchability of Internet-accessible archives creates the possibility that a third-hand reader of the research can trace the comments back to their original writer" (Walther 2002:206). Citation of the Facebook data is thus limited in the thesis to individual words, paraphrases, or translations.

Markham and Buchanan (2012:6) include "exact quoting" in their list of criteria that indicate that human subjects research is being undertaken; additionally, Walther (2002:207) claims that "the analysis of Internet archives is not human subjects research if a researcher does not record the identity of the message poster and if the researcher can legally and easily access such archives". Consequently, this part of the project need not be classified as human subjects research and thus avoids the need to obtain consent from all 99 users. As an additional precaution, the Facebook subcorpus has been removed from the publicly accessible version of the corpus (see Appendix C).

### 2.2.3 Identifying words to tag

After gathering the samples from each source, certain tokens were tagged to enable analysis. The majority of words used in Breton, particularly those occurring more frequently, are not considered emblematic of either "néo-" or traditional Breton, as they are used in all varieties of the language. Examining every single lexical token would therefore have been both impractical and unnecessary. Instead, the lexemes of primary interest to this study were those typically viewed as characteristic of either new or traditional speaker language, rather than being part of the common stock shared across varieties. These fall into three groups:

1. transparent borrowings from French, considered to be preferred by traditional speakers (Rottet 2014:213) as they reflect an ideology that acknowledges the recent dominance of French in the community;

2. words specific to particular dialects of Breton, also thought to be used by traditional speakers only (Jones 1998a:132);

3. words borrowed from other Celtic languages or older forms of Breton, and "neo-Celtic" words formed from Breton morphemes of Celtic origin, thought to be characteristic of "néo-breton" as they represent an ideology that favours purism (German 2007:185).

It was therefore necessary to devise a workflow that would extract these three types of lexeme from the bulk of the vocabulary encountered. Particular care was required in the case of type 1: not all words of French origin are relevant, as around two-fifths of basic vocabulary in Breton consists of well-assimilated borrowings from French (Ternes 1992:374). These common words are not perceived as French and therefore used unquestioningly by promoters of even the least French-influenced forms of Breton, indicating the importance of beliefs driven by superficial resemblances between Breton and modern French rather than the actual linguistic origins of the lexicon; the need for Breton to *appear* Celtic comes across as the most important factor for those favouring the purist ideologies with which "néo-breton" is stereotypically associated.[43] These highly-assimilated words are therefore not of interest to this study, not being perceived as emblematic of "néo-breton". As such, it was necessary to find a way of separating them from more salient loan words.

Subdividing this category makes it possible to pinpoint which specific types of loanword were relevant to the analysis. Words and morphemes of non-Celtic origin in Breton—the vast majority having entered via French, even if they ultimately derive from another language—can be divided into three types:

a words used in Breton since before the revitalisation process began, fully assimilated and used without question by promoters of "néo-breton" (*pont*, 'bridge');

b words used in Breton since before the revitalisation process began, not taken up by promoters of purist "néo-breton" (*mersi*, 'thank you'), and hence now considered

43. A similar focus on surface forms is observed in revived Cornish by Mills (2015), who notes that promoters of the least English-influenced variety of the language reject words that resemble surviving English cognates, but allow those derived from English words no longer in use.

the preserve of traditional speakers;

c words not present in pre-revitalised Breton, adopted into the language by certain speakers during the revitalisation period (*telefon*, 'telephone').

Only types b and c in this list correspond to type 1 in the above list: type a is not relevant to the analysis. The workflow developed therefore needed to identify words of types 1, 2, and 3 from the first list, while excluding words of type a in the second list. With these conditions in place, it acted as a filter for establishing which words required annotation and analysis, allowing this to concentrate on words typically thought to be used in the language of either new or traditional speakers exclusively.

**Workflow:**

If a multi-word expression is used where a single word could be used in French and a multi-word expression is not routinely used in French,[44] include the expression.

If a free morpheme[45] within the lexeme has an etymologically and semantically[46] related counterpart in modern French, follow Flowchart 1:[47]

44. e.g. *saver tiez* for 'architecte' (Rottet 2014:236).
45. This denotes any morpheme except bound affixes.
46. The inclusion of "and semantically" allows words such as *paotr* ('boy', from French *poutre* 'beam'; Ternes 1992:374) to be rejected.
47. This section of the workflow allows words of type 1, such as *telefon*, to be identified, whilst excluding those words of French origin that are not of interest to the study (type a).

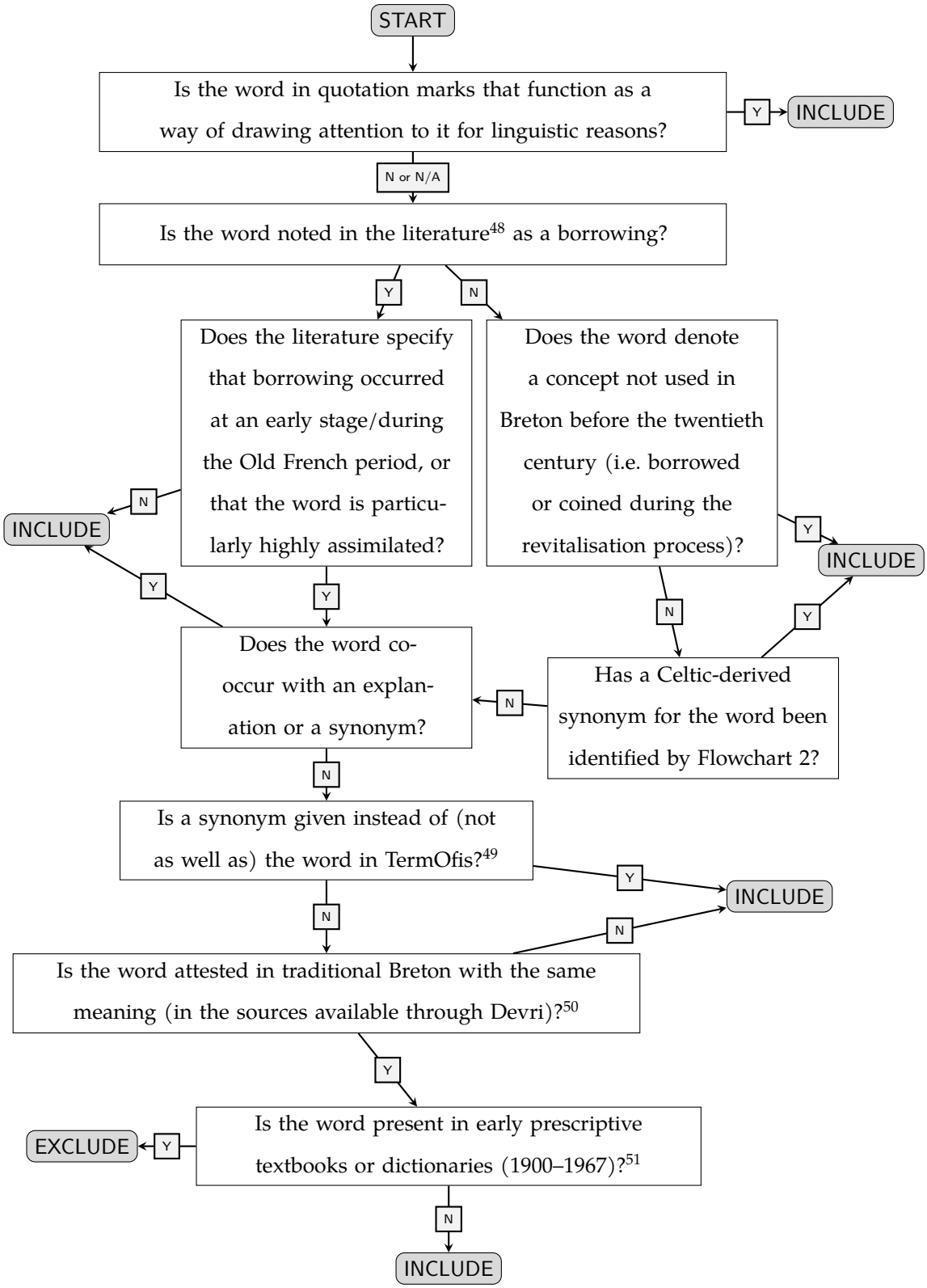


Figure 2.1: Flowchart 1, for identifying words of type 1

If the word is not relevant to Flowchart 1, or if a polymorphemic word[52] or multi-word expression is excluded by Flowchart 1, follow Flowchart 2:[53]

48. Including textbooks.

49. The *OPLB*'s glossary of officially recommended terms, available at http://www.brezhoneg.bzh/87-termofis.html, accessed 16 Apr 19.

50. A historical corpus of Breton compiled at Université Rennes 2, available at http://devri.bzh, accessed 16 Apr 19.

51. The fourth wave of the Breton revitalisation movement, or "quatrième *emsav*" (Kergoat 2012:4), taking a less nationalist and pro-standardisation approach than previous incarnations, was hastened by the political events of 1968 (Abalain 2000:66ff). If a French borrowing appears in revivalist dictionaries prior to this date, it cannot be considered stigmatised in even the most purist varieties of "néo-breton".

52. This denotes compounds or lexemes created through affixation.

53. This part of the workflow identifies words of types 2 and 3, and includes questions that deal specifically with dialect-specific words.

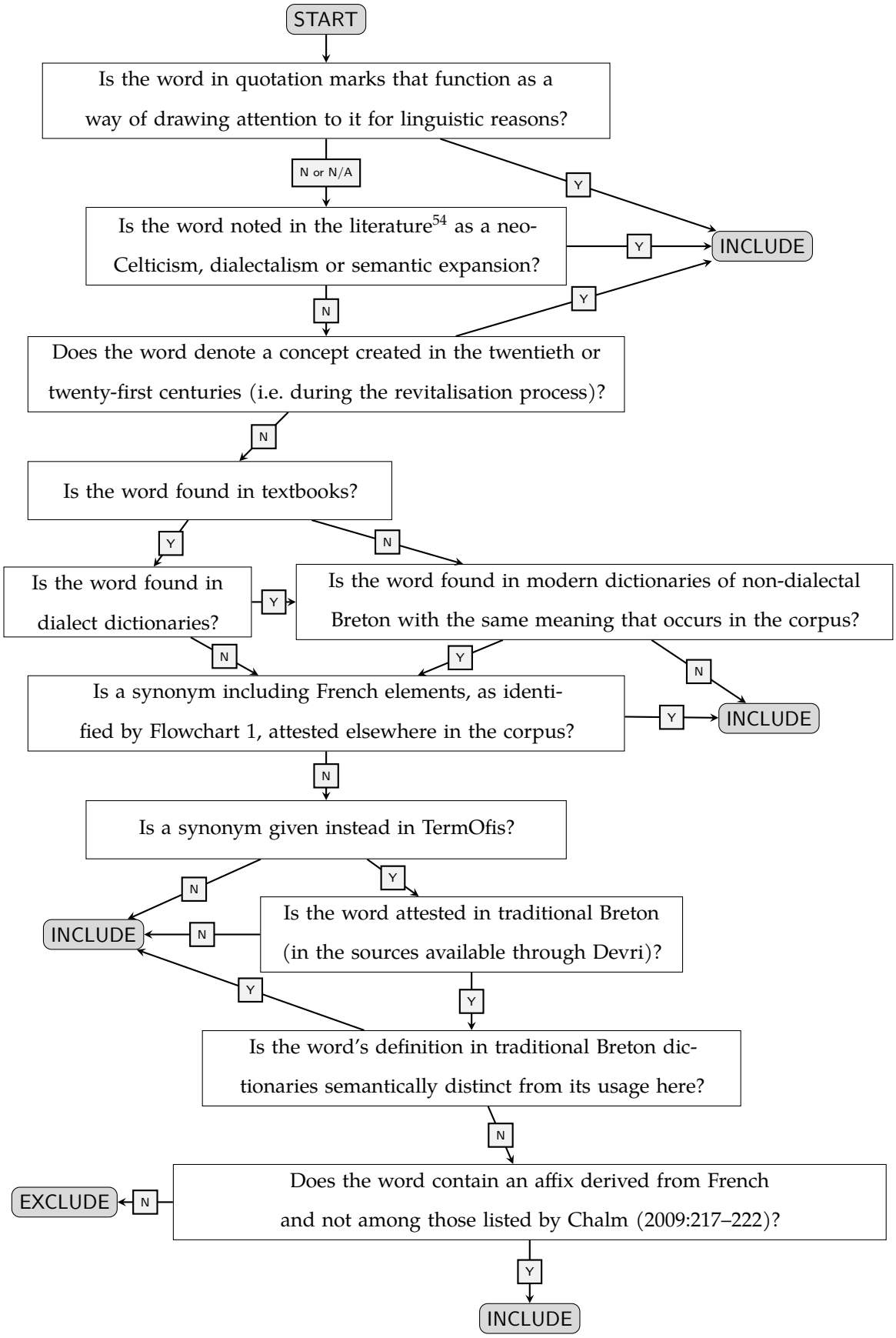


Figure 2.2: Flowchart 2, mostly for identifying words of types 2 and 3

### 2.2.4 Outcome of the workflow

The aim of the workflow was to isolate and tag any word that could be perceived as indicating an alignment with either traditional Breton or "néo-breton", without this necessarily meaning anything about the word's actual use or origin, given that it is the position of the "néo"-/traditional Breton dichotomy within academic and militant discourse that is under investigation. This resulted in a fairly generous approach to tagging, where some Celtic-derived[55] words that are in fact found in pre-1900[56] dictionaries with the attested meaning are tagged nonetheless. Both *abeg* ('reason') and *baleadenn* ('walk'), for example, are highlighted by Timm (2001:456) as "proposed lexical innovation[s]", but are in fact both found with these meanings as early as 1499 (Lagadeuc [1499] 1867:4), and continue to be found into the nineteenth century: see section 4.3.5 for more detailed discussion. Nonetheless, the association of these words with "néo-breton" by these scholars indicates that they are at least considered to be features of "néo-breton" in some quarters, and consequently that they are of interest here.

This same rationale was used for tagging words for which a word deriving from a different language, but with an equivalent meaning, had already been selected for annotation. While the Celtic-derived word *gouel* ('festival') would not otherwise have been selected, being present in both pre-1900 Breton and TermOfis, its French/international-derived synonym *festival* is also present in the corpus and tagged due to its absence from Roparz Hemon's 1964 dictionary. *Gouel* is therefore also tagged by the second flowchart, as it may be used—consciously or otherwise—in order to avoid the more French-sounding *festival*.

The words tagged therefore corresponded to those that could, to at least some extent, be assigned to either traditional Breton or "néo-breton" based on stereotypical descriptions of these varieties. Based on the outcome of annotation, the above list of types of words of interest (at the beginning of section 2.2.3) can be expanded to the following:

54. This is a potential source of problems as the literature does not always match with what is attested in traditional Breton. For an example, see section 4.2.3.
55. For clarification on the use of "derived" in this context, see section 3.2.3.
56. 1900 was used as an approximate cut-off point for the growth of the "néo-breton" movement: if words are found in pre-1900 sources, they are considered to be present in traditional Breton.

1. French-derived words found in Breton before 1900, not found in either modern or mid-twentieth-century dictionaries;
2. Words deriving from either French or Celtic, found in Breton before 1900, where a different word is instead advised in TermOfis as a translation of the French equivalent;
3. Celtic-derived words found in Breton prior to 1700 but not between 1700 and 1900;
4. Words not found in Breton before 1900, deriving from any language (in practice, mostly Celtic or French);
5. Words deriving from either French or Celtic, found in Breton before 1900, but only with a different meaning from that attested in the corpus;
6. Words attested as dialect-specific (mostly Celtic-derived but some French-derived);
7. Words deriving from any language, identified in the literature on "néo-" and traditional Breton as salient borrowings or coinages;
8. Multi-word expressions, containing words deriving from any language, used as equivalents of single French words;
9. Any French-derived equivalent of a Celtic-derived word already tagged;
10. Any Celtic-derived equivalent of a French-derived word already tagged.

Not tagged, therefore, were longstanding French-derived words that became sufficiently integrated into Breton not to be contested by modern dictionaries, and longstanding Celtic-derived words that have never been seen as contentious or believed to have been coined by users of "néo-breton". This body of words constitutes those that would not be associated with either "néo-" or traditional Breton specifically, i.e. those that can be considered part of the stock common to both varieties.

To minimise the risk of false positives for the third and fourth of the categories listed above, words were checked using Devri,[57] an online historical corpus of Breton created

57. http://www.devri.bzh

by the Université Rennes 2 containing words from nearly a thousand different sources, dating as far back as the ninth century. For this purpose 1900 was taken as a cut-off point for words that pre-date the growth of the "néo-breton" movement. Despite using this source, it is still possible that a small number of words used before 1900 would fail to be identified as such, but their evident infrequency among early written sources of Breton is perhaps grounds enough for including them among the salient tokens.

### 2.2.5 Determining validity among lexemes

Despite the use of the workflows, in some instances it was not immediately obvious whether to tag a word or leave it untagged, due to varying degrees of difference from the prescribed standard. Some cases of such uncertainty arose from the need to ensure consistency between audio and written sources. Word-initial <c'h>, for example, typically corresponding to oral /x/ or /ç/, is realised as [f] by speakers from the *pays bigouden* in south-western Brittany, meaning that the presence of this sound in words written with initial <c'h> can be identified as a dialectal feature, and by extension, according to the criteria set out in the literature, as a feature marking at least some degree of alignment with traditional Breton, implying that tagging should occur. While this feature can be identified in the radio data, it is impossible to do so in the written sources, as speakers would be far more likely to retain the standard spelling <c'h> than to use one that might indicate this dialectal pronunciation. In cases of dialectal variation such as this, therefore, it seemed most appropriate to take a conservative approach, considering salient only those tokens that would differ from the standard in written as well as spoken contexts, thereby ensuring a consistent rate of token selection across the different subcorpora. As a result, some tokens in the radio corpus that differ from the standard are untagged: however, as this is a lexical enquiry and such borderline cases tend to be both few in number and different from the standard purely in terms of phonology, this is not a significant omission. It is worth bearing in mind, however, the fact that as the radio is a spoken medium, its speakers are evidently able to display dialectal allegiance through phonological means, and therefore can do so with more subtlety than in writing.

Similar decisions had to be made in other situations. The most common infinitival ending in Breton is typically written <añ>, although <iñ> is also frequent. For most verbs with these endings, the modern Breton form has been standardised as either <añ> or <iñ>, but historical records of Breton tend to show significant amounts of variation, and in some cases this persists among modern dictionaries. In terms of pronunciation, <añ> and <iñ> are typically realised as [ã] and [ĩ] respectively, but as these are final unstressed syllables, which often tend to be reduced to schwa, it can be difficult to tell the difference between the two in an oral context. Moreover, in both the written and the spoken data many verbs are inflected and lack an infinitival ending, meaning the alternation is often irrelevant. These points, in addition to the fact that the alternation between <añ> and <iñ> or between [ã] and [ĩ] does not index any specific dialectal or new/traditional distinction, meant it seemed most reasonable not to tag any tokens that deviated from the standard only with respect to this feature.

Another case where a similar decision had to be made was that of the lexeme *surf* and its derivations (hereafter "*surf**"). TermOfis gives a different spelling, *seurf**, as the Breton translation for the French/international word *surf* and derived words. However, all attestations in the corpus except two are of *surf**. The word is attested only among the data from *Brud Nevez*, making the matter more complicated as this publication uses a different orthography from that used by TermOfis. However, an <eu><ẽu> alternation of this type between the two orthographies is not found elsewhere: the word *neuial* ('to swim'), for example, retains <eu> in both,[58] and tokens of *seurf** are also present in the *Brud Nevez* sample. The alternation between *surf** and *seurf** therefore appears not to be driven by orthography, but by a conscious decision on the part of the author. This alternation thus provokes different tags in the corpus: *surf** is spelt in the same way as the French word, but *seurf** is spelt differently, and different tags are used depending on whether the orthography of the tagged token changes or remains the same as in the source language. This alternation can therefore be considered a salient distinction in that *seurf** is constructed so as to be further from the French/international source word, and therefore more typical of canonical "néo-breton". As a result, it does seem worth tagging *seurf** and *surf** separately rather than conflating them as a single lexeme.

58. One of the most visible features of the *skolveurieg* orthography used in *Brud Nevez* is in fact in a sense the inverse of this distinction: the indefinite articles, spelt *ul*, *un*, *ur* in the usual *peurunvan* orthography, are spelt *eul*, *eun*, *eur* using *skolveurieg*. This pattern does not extend beyond these contexts, however.

Identifying non-standard tokens therefore relied on their matching two criteria. Firstly, consistency across written and oral contexts where the token appears in multiple subcorpora; secondly, potential ideological distance from standard or prescribed vocabulary. This could involve either a different degree of adaptation from French, reflected in the way the token is tagged, as in the case of *seurf*/surf**; or the use of a variant that can be attributed to dialectal usage, again annotated specifically to note the dialect in question. In addition to this, it was borne in mind while annotating the corpus that the focus of this enquiry is primarily lexical: annotation of purely morphological or phonetic deviations from the standard was kept to a minimum.

This reasoning likewise contributed to decisions about whether or not to tag other borderline cases. The Breton word for *sorcier*, for example, is *sorser*: clearly a borrowing from the French, but a long-established one that would not in itself trigger annotation. In the corpus, *sorser* is attested, but so is the variant *sorc'hser*. While this is further in appearance from the source—and, indeed, contains the typically Breton digraph <c'h>—this difference from *sorser* would not be reflected in the annotation, as both these tokens would require the tag used to mark a different orthography from the source language. Additionally, this use of <c'h> does not appear to correspond with any dialectal usage. On the other hand, a token of *bourc'h* (= French *bourg*) occurs in the corpus, contrasting with the prescribed spelling *bourk*. Unlike *sorc'hser*, this variant can be seen to correspond with a dialectal pronunciation, attested in the *NALBB*; it therefore reflects a particular dialectal usage, and has accordingly been tagged as a specific word.

### 2.2.6 Categorising lexemes

Relevant tokens were assigned a number of tags relating to particular categories that were thought likely to present potential starting points for analysis. Each token was assigned to as many categories as necessary, with no restriction on the number of tags.

The categories are as follows:

#### 2.2.6.1 Word origin

Lexemes with components originating in French, English or Welsh[59] are marked accordingly, with a subcategory in the first case for those forming part of learned vocabulary in French, with cognates in other non-Romance languages. Morphemes present in earlier stages of Breton that had fallen out of use by the nineteenth century and were reintroduced during the revitalisation process are also noted. Other words, consisting solely of morphemes present in traditional Breton, form an additional category.

#### 2.2.6.2 Semantic relationship with source word

Where a word is semantically different from its source lexeme, whether this is traditional Breton or another language, it is marked as either different or related through semantic expansion. Where a new word has been created on the basis of a word in a different grammatical category, this is also marked.

#### 2.2.6.3 Polymorphemic words and multi-word expressions

Words containing multiple derivational morphemes and multi-word expressions are marked as such, and the first of these two types is divided into compound words and words created through affixation. Polymorphemic words and multi-word expressions are also classified according to whether the morphemes or words they contain are found elsewhere in Breton, or whether they are taken from French in order to coin the token.

#### 2.2.6.4 Parts of speech

Words are grouped into the following categories: nouns, verbs, adjectives, adverbs, discourse markers, prepositions, determiners and pronouns, interjections, verbal particles, and conjunctions.

59. As a closely related language with a comparatively well-stocked lexicon, Welsh has acted as a source for new Breton lexemes since the earliest efforts at lexical expansion: see Rottet (2014:225–226) for the example of *alarc'h* ('swan'; from Welsh *alarch*).

#### 2.2.6.5 Attestation

The attestation of a word in any of a number of sources, including several dictionaries and the *OPLB*'s online terminology database, TermOfis, is noted. Section 3.3.1 gives more information on the sources consulted.

Words for which an attestation is given but the meaning is different are marked distinctly. Words for which an alternative is given in TermOfis instead of what is attested are also marked.

#### 2.2.6.6 Code-switching

The use of other languages, often French, in the form of longer phrases, and the use of French to specify the meaning of an uncommon Breton word (as evident from the surrounding context) are marked separately.

#### 2.2.6.7 Dialectal vocabulary

Dialectal terms, and the broad dialect area[60] from which they originate, are marked. However, some former dialectalisms, particularly from Léonais, are now part of standard Breton: the effects of this are considered in section 3.4.

#### 2.2.6.8 Orthographic modification (magazine and internet data only)

In the case of morphemes and words not originally from Breton, it is noted whether or not their orthography has been brought in line with Breton conventions.

#### 2.2.6.9 Phonological modification (radio data only)

In the case of morphemes and words not originally from Breton, it is noted whether or not their phonology has been altered from the source lexeme.

60. Words that could be used in multiple dialects are marked as such.

#### 2.2.6.10 Markers of explicitness of borrowing or coinage

Cases where the speaker or an interlocutor remarks upon the word or draws attention to it for linguistic reasons are also noted.

#### 2.2.6.11 Other morphological information

Words modified through inflection or initial consonant mutation are marked as such. Clippings, such as *bak* (< French 'bac', from 'baccalauréat') are also marked, as are tokens with feminine or diminutive suffixes.

### 2.2.7 Background information on the three subcorpora

The content of the corpus is worth discussing, as this would undoubtedly have effects on the analysis. As section 2.2.1.1 sets out, the radio data come from both news programmes and longer feature programmes. Both deal almost exclusively with news affecting Brittany. This is not to say that the range of subjects covered is not varied; topics include politics, education, festivals, transportation strikes, terrorism, and theatre, among others. Every news programme concludes with a segment about sport, and every such section in the sample mentions both football and the Vendée Globe, a quadrennial round-the-world sailing race in its final stages at the time the programmes were broadcast (December 2016–January 2017), in which many of the principal competitors, including the eventual winner, were of Breton origin.

The news programmes were deliberately selected to ensure that each was presented by a different person and equally that the two Radio Breizh programmes were from the same broadcast dates as two of the three France Bleu programmes, so as to maximise the opportunity of generating comparable vocabularies. In both cases, therefore, the first subject covered was the same: both programmes from 18 January 2017 began with a report on a protest against mining in central Brittany during a visit from Emmanuel Macron in his capacity as presidential candidate, while both from 19 December 2016 began with a report on the distribution of money raised by the Redadeg, a biennial

sponsored run in support of Breton-language education and other enterprises that use or promote Breton. While both pairs of broadcasts equally finished with reports on football and the Vendée Globe, the other subjects they covered were different, although all maintained a focus on Brittany.

The other radio programmes share with each other a similar broad theme of culture, but focus on different aspects of it. As both broadcasts were an hour long, the samples were cut to the first twenty minutes (in the case of the France Bleu sample) and the first fifteen minutes (for the Radio Breizh sample), with musical interludes excluded. The France Bleu programme discusses two main topics: a travelling theatre production created by a Breton company, and a film writing workshop taking place on an island off the west coast of Brittany. The Radio Breizh programme, broadcast from a conference in Brest, has only one main subject: the addition of the *fest-noz*[61] to UNESCO's "intangible cultural heritage" list. As well as discussion of this status, and of the *fest-noz* and its future, the excerpt includes analysis of a film about the *fest-noz* tradition, providing a thematic link with the France Bleu programme. As both longer programmes discuss cultural matters, the lexical fields used in the two programmes overlap. However, as the Radio Breizh programme deals specifically with a topic rooted in traditional Breton culture, less of the specialist vocabulary related to this topic will be considered contentious in regard to the matter under examination: such vocabulary will already be of long standing in Breton, minimising the need for either borrowings or coinages.

The three printed sources have some differences in subject matter. *Ya*, the only newspaper of the three, with weekly issues, covers local news and so has some overlap with the radio news programmes, again including articles on the subject of sport as well as news about local culture, including the Redadeg. However, it also includes news from outside the Breton context. In the sample, there is an article about one of the last traditional speakers of Manx, as well as one about a creator of fictional languages for popular fantasy films and television series. These articles evidently relate to the theme of minoritised and non-national languages, one that is of evident interest to *Ya*'s readership.

61. 'night festival': an evening of music and dancing, often accompanied by crêpes and cider. Music is typically provided by a *binioù–bombard* duo—Breton bagpipes and a shawm-like instrument—or by a *bagad* (a band) consisting of several *bombardoù*, a few *binioù* and often percussion. Some *festoù-noz* mix these traditional instruments with others such as electric guitars.

*Bremañ* also branches out from the local context, but focuses on different areas. While some of the articles in the issue sampled from relate to Brittany, most of those in the sample itself do not. However, for the most part they have a unifying theme of concern for other minority populations and for movements in favour of regional self-determination. The 4000-word sample contains articles of varying lengths relating to the Kanak people of New Caledonia, the people of French Guiana, the ETA separatist movement in the Basque country (which had announced a permanent ceasefire shortly before the publication of this issue of *Bremañ*), and the republican movement in Northern Ireland, the last in the form of a report on the recent death of Martin McGuinness. These articles indicate a more political, perhaps more overtly partisan, outlook on the part of *Bremañ*: while *Ya* restricts itself to more cultural topics, and in the case of news from outside Brittany, to subjects that focus on language, *Bremañ* is more likely to take an interest in other minority populations around the world and their struggles for official recognition. In its coverage of Brittany, too, it more openly asserts an ideological stance, elements of which it assumes are shared with its readership: one of the articles in the sample states "ken anat eo evit lennerien Bremañ emañ Naoned e Breizh hag ez eus 5 departamant e Breizh": 'it is most obvious to readers of *Bremañ* that Nantes is in Brittany and that there are five *départements* in Brittany.' At this point, it might be anticipated that this clear position with regard to non-linguistic ideologies could similarly be reflected in linguistic ones: as "néo-breton" is typically associated with ideologies of political militantism (see section 1.4.2), it could be expected that the presence of similar ideologies in *Bremañ* would indicate that this publication is more ready to use stereotypical "néo-breton". This will be explored in subsequent chapters.

*Brud Nevez* is different from *Bremañ* and *Ya* in that the sample is restricted, like the radio programmes, to matters of Breton interest. Issues tend to focus on a particular theme: in the sample, that theme is the sea, evidently a subject known well in Brittany. News items are restricted to a single page at the end of the magazine (not included in the sample) and are focused solely on Brittany itself; the articles in the sample are more personal features relating to the authors' own experiences with the sea. This difference from the other written sources is compounded, at least cosmetically, by the use of the less common *skolveurieg* orthography.

Several of the subjects covered in the radio and printed data also arise in the Facebook data, including the recently installed cable car in Brest and a Breton-language theatre production. Politics is also a frequent subject on Facebook: without the moderation of an editor, users are free to contribute their personal opinions, which in many cases unsurprisingly reveal a distrust for the French state. One of the most common subjects of all is questions and comments about the language itself; posts where this was the main subject from the outset were excluded from the corpus, so as to avoid sampling posts with large proportions of use of words that the users might not otherwise employ. However, given the frequency with which comments turned to linguistic matters, it was impractical to exclude all posts with any degree of linguistic content.

On the whole, there is a wide variety of subjects covered in the Facebook data: as well as providing political opinions, users advertise future events or appraise past ones; they ask for advice on where to learn Breton, or whether certain facilities are available in the language; they share news about Breton publications and occasionally send general messages of goodwill, such as at the new year. The Facebook data thus involve the widest variety of topics and registers; the analysis in chapters 3 and 4 will reveal whether the language used is similarly heterogeneous in terms of its inclusion of features that align with traditional Breton or "néo-breton".

## 2.3 Interview methodology

Section 2.1 explained the need for an element of additional research beyond the creation and analysis of the corpus: in order to gather information on speakers' perceptions and identities, it was necessary to move away from strictly linguistic analysis and speak to language users directly. For this project, this was done in the form of a number of interviews in order to find out, firstly, whether users of Breton in the media conform to the characteristics of the stereotypical "néo-bretonnant" (see section 1.4.2), and secondly, what strategies these speakers employ in cases of new or contested vocabulary, and how concerns of communication and identity are reconciled. The analysis of these interviews forms a smaller portion of this research than the analysis of the corpus, and is presented in chapter 5.

### 2.3.1 Determining participants

As the corpus focuses on Breton in the media, it was decided that interviewees should also be drawn from these contexts. This was not restricted to those who produced the samples in the corpus, however; some of these speakers were interviewed, but some interviewees were not part of the corpus. This was done for various reasons: partly pragmatic ones, as it might have been difficult to recruit enough interviewees if this restriction had been placed on participant eligibility; and partly in order to elicit data from a wider range of participants. As a result, the distribution of speakers in the corpus and interviewees can be conceptualised according to the following Venn diagram:

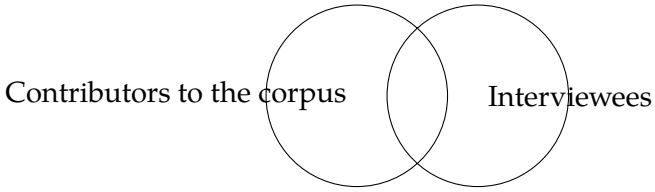


Figure 2.3: Contributors to the corpus versus interviewees

The initial aim was to speak to people involved with all three media contexts examined, i.e. radio, print publications, and Facebook. However, it was subsequently decided that interviews would be restricted to employees of the first two media types, for both practical and ethical reasons. Deciding on potential interviewees to contact from among the users of *Facebook e brezhoneg* would have required the identification of some of the more active users in the sample taken for the corpus, which in turn would have necessitated a more in-depth analysis of who was producing the data gathered. As explained above, profile information of this sort was not stored along with the data from the group, and so determining whom to contact as a potential interviewee would have been difficult without breaching this ethical stipulation. Additionally, online identities are often more complex and fragmented than their offline counterparts (Kennedy 2006). As a result, it was possible that Facebook users could be adopting a virtual persona far removed from their offline self-presentation, which meant that conducting in-person interviews with people whose language use was being examined solely in an online context could have been inappropriate. Employees of the radio stations and

print publications, on the other hand, were sought out through their employers and were able to give informed consent in all cases; most of those interviewed also used Breton on Facebook, often both personally and professionally. As a result, information from Facebook users *was* gathered, but with the caveat that those users interviewed were also employees of the radio or print organisations under examination, and would of course not be representative of the entire population of Facebook users.

### 2.3.2 Recruiting participants

Participants were sought with the initial aim of interviewing at least one employee from the different types of media as categorised in the research, i.e. the three print publications, *Bremañ*, *Brud Nevez*, and *Ya*, and the two types of radio station, public and associative. Further participants in excess of these numbers were not turned down, however, and as Radio Breizh consists of four stations that operate independently, all of which were contacted separately, it was anticipated that responses from multiple sources of the same type would be received.

As all the organisations maintain websites that display phone and email contact details, or offer the possibility of getting in touch through an online form, making this initial enquiry was straightforward. In most cases, emails were sent to the organisation's general contact address with a brief explanation of the project and an invitation to employees to respond if they were willing to be involved. Using a snowball sampling technique, these initial contacts were also asked if they could provide contact details for colleagues they believed might wish to participate, and a few additional interviewees were recruited through this method. Further communication with intended participants was mostly via email, although some phone calls were made once contact had already been established. Fittingly, social media also played a role in the recruitment process, as initial contact was made with one participant through Twitter direct message, and with another, the interview was arranged via Facebook messenger.

Responses were received from all five media types apart from *Ya*, two employees of which were contacted by email and phone but without success. As another participant suggested, the lack of response in this case could have been due to the fact that *Ya* em-

ploys very few permanent staff, instead mostly relying on freelance journalists. This is a typical organisational model for Breton print publications, as other interviews discussed.

In total, eight interviews were carried out with employees of the media organisations, comprising two from public radio, four from associative radio, one from *Brud Nevez* and one from *Bremañ*. In all cases, participants were involved with writing content for their respective organisation and, in the case of the radio employees, presenting programmes. A representative of the *Office public de la langue bretonne* was also interviewed in order to gather information about the workings of TermOfis, the *OPLB*'s online database of recommended terminology. While the specific discussion of TermOfis could not be included in the analysis for reasons of space, quotations from this participant are nonetheless included where relevant, and his demographic data are equally taken into consideration.

As the interviews were not the principal focus of this thesis, and only a small number were carried out, qualitative analysis of these data was more appropriate than quantitative analysis would have been. Obtaining a sample of interviewees that was balanced for demographic attributes such as age or gender was felt to be unnecessary and overcomplicated with this small sample size. Information of this type was collected from interviewees, however (see section 2.3.4 below), and is detailed in chapter 5.

### 2.3.3 Theme of the interviews and language choice

The interviews were not intended to gather linguistic data, as the corpus focused on this. Instead they were carried out for the purpose of gathering data relating to speakers' backgrounds and to their perceptions of certain aspects of the use of Breton, both in relation to their own language and the use of Breton within their organisations. This information was gathered in order to probe some of the stereotypical characteristics of "néo-bretonnants" (see section 1.4.2), given that the corpus data responds to those on the linguistic side.

Interviews took place in a variety of locations, with some occurring in public spaces

such as cafés and bars, and others at the offices of the participants; interviewees were invited to suggest locations themselves in order to make the process more convenient for them. Interviews were carried out in French. Two principal reasons motivated this decision: firstly, it was decided that interviewees would be in a better position to critically assess their own use of Breton if they were not speaking the language at the time. Secondly, my own competence in spoken Breton is considerably less than in French, and so using French was more appropriate from a practical point of view since both interviewee and interviewees had a high level of fluency. Indeed, had I decided in advance of the interviews to carry them out in Breton and made an effort to improve my command of the language accordingly, my Breton would have remained largely non-dialectal, by virtue of having mostly learnt the language from books and having no particular allegiance to any dialect area. This might have led participants to assume that I favoured a more standardised variety myself and was carrying out the interview from an ideological standpoint in this regard. As a consequence, they might have given the answers they believed I wished to hear, or modified their own language, especially given that "a complete stranger arouses the level of attention paid to speech" (Tagliamonte 2012:106). Additionally, this would again have been likely to influence any self-reporting of their language attitudes. Conducting the interviews in French allowed this to be avoided.

### 2.3.4 Questions asked

Interviews were semi-structured, in that a list of questions was devised beforehand, tailored to the type of media and organisation depending on the interview, and participants were typically asked the questions from this list, supplanted by any further questions that seemed appropriate based on the answers already provided. Similarly, the questions were not always asked in a set order, but were instead posed at what was judged to be the most suitable time based on the direction of the discussion. This type of interview is the "commonest in qualitative work" (Arksey and Knight 1999:7), allowing both interviewer and interviewee to choose how much detail to provide on certain topics, by asking more or fewer questions or by providing longer or shorter responses respectively. The semi-structured approach also "anticipates analysis, and

facilitates the organisation of that final stage" (Gillham 2005:75). As the interviews were carried out in the later stage of the project, when the analysis of the corpus data had already been carried out and written up, it was beginning to become clear which topics would be the most salient in the interview context; again, therefore, the semi-structured format seemed especially appropriate.

My position as a researcher required navigating at this point in the project. Arksey and Knight (1999:13) point out that "many feminist social scientists claim that for interviews to be successful there needs to be a shared culture between the interviewer and the interviewee". With this in mind, I made reference to my own background as a Cornish speaker where appropriate, as well as to my own experiences learning Breton and to previous periods of time spent in Brittany, including one stay of several months. While it was clear that I was an outsider from my name, accent, and affiliation to a UK university, I thus hoped to mitigate this by showing that my own cultural and linguistic background was similar to that of the participants and that I was able to empathise with many of the struggles of minoritised language speakers from a personal rather than merely academic perspective.

In accordance with university ethics regulations, interviewees were provided with an information sheet stating in broad terms the purpose of the study, and asked to sign a consent form,[62] copies of which were stored separately from the transcripts of the interviews, in which, again following university regulations, participants remain anonymous. Operating in this manner meant that progress through the institutional ethics process was possible without any problems, and permission from the school ethics committee was obtained prior to conducting the interviews.

The interview questions were grouped into four categories. The first contained simple demographic questions in conformance with the work on "speaker variables" carried out by the MEITS project, under which this research was conducted. Collaborative work within the MEITS team had resulted in the decision to store the following key variables for all speakers interviewed across the project: age, gender (with the options male, female, and other), current place of residence and length of time lived there, previous places of residence and length of time lived there, and language usage and expos-

62. This form is reproduced as Appendix D.

ure. Additionally, cross-strand work within the MEITS project resulted in the adoption of a methodology whereby identity would be measured on a continuum (Carruthers and Fisher, in preparation): for this project, the evaluation of French and Breton identity was affected by this. Rather than using the typical speaker variables, then, which are common in work situated in a more traditional variationist paradigm (Milroy and Milroy 1997), this thesis measures identity in a way more similar to the Likert scale, "commonly used to measure attitude" (Jamieson 2004:1217). However, while a Likert scale is divided into categories, returning a discrete measurement, the scales used by the MEITS project are continuous, providing more versatile data that can be processed by the various multidisciplinary strands of the project in appropriate ways. By collecting data according to this cross-project methodology, data can be compared across individual outputs and can be included in an overall databank allowing larger-scale analysis.[63] In the case of this thesis, this different way of measuring identity also acknowledges that this is a complex variable that is not always easy to class according to pre-defined categories. The scales presented to participants are reproduced in Appendix E, and the results of this methodology are discussed in section 5.2.

The second part of the interview asked participants about their linguistic background and acquisition of Breton; the third covered their use of traditional and online media in both Breton and French, while the fourth asked about their role in the organisation for which they work and about how lexical issues are dealt with in the organisation, including issues around dialectal words, borrowings, and neologisms. However, questions of a more linguistic nature, relating to specific lexemes and entailing comparative elements across interviews, were avoided, as responses would have been self-reported, potentially giving rise to inaccuracies, which would have had significant negative effects given the small sample size (Milroy and Gordon 2003:54).

The interview questions are included as Appendix F.

63. The decision to refer to "minoritised" rather than "minority" languages also stems from this collaborative work.

### 2.3.5 Analysis of the interviews

Interviews were transcribed in full and encoded using TEI, in the same way as the corpus data, although the transcriptions were left unannotated. Full transcription of the interviews was done in order to ensure that themes were easily identified across the nine interviews. However, as no linguistic analysis was undertaken, not every word of the interviews was transcribed: elements such as false starts, semantically empty discourse markers, and paralinguistic features were ignored.

The interviews were analysed according to the established principles of thematic analysis, entailing "the generation of themes via coding", "arguably the most common analytic approach taken by qualitative researchers using interviews" (Roulston 2013:305). No in-depth analysis of the language used in the interviews was being sought; as the data emerging from this part of the research form only a secondary correlate to the main focus of the study, i.e. the corpus data, there was no need for a detailed investigation of interviewees' discursive choices. Thematic analysis, focusing on the themes and subjects arising during the interviews and moving from descriptive coding to the identification of overarching themes (King and Horrocks 2010:153), was an appropriate level of detail for this part of the research.

An iterative approach was taken to the analysis, consisting of an alternation "between emic, or emergent, readings of the data and an etic use of existing models, explanations, and theories" (Tracy 2012:184). The structure of the interviews themselves was based broadly on the research questions and on the need to collect certain data about speaker profiles and other non-linguistic information. On investigating the interview transcripts for the purposes of determining themes, however, this was not done with any use of these pre-existing models as a guide, but instead, the themes were gathered directly from the data, and then grouped into broader categories. When these categories were grouped together for the purposes of writing up findings from the interviews, on the other hand, these were ordered according to the structure of the interviews, and consequently informed by the research context and questions. This iterative model allowed the background and findings from earlier parts of this research to be borne in mind while constructing the relevant chapter, but also permitted important insights

from the interviews to emerge, even in cases where they did not fit with previous findings or preconceptions. Iterativeness is crucial in thematic analysis as it facilitates a "dynamic and interactive" methodology where "the researcher is constantly moving back and forth between design and implementation" (Tobin and Begley 2004:390), allowing the research not to be unduly grounded in preexisting beliefs.

Findings from this part of the research are presented in chapter 5, along with illustrative quotations from the interviews.

# Chapter 3

# Contemporary Breton usage 1: Quantitative analysis of the corpus data

## 3.1 Introduction

This is the first of two chapters dealing with the corpus of Breton from media sources. It considers the corpus data in bulk, presenting results from a number of quantitative analyses based on the annotation of the corpus. Examining the data from various perspectives will reveal whether the Breton found in these media contexts conforms to the stereotypical description of "néo-breton", and particularly the extent to which this varies under different conditions, responding particularly to research questions 1 and 3, as well as 4. The chapter contains four sections investigating different aspects of the tagged tokens and their distribution: parts of speech and word origins (section 3.2), attestation in dictionaries and semantic expansion (section 3.3), dialectal features (section 3.4), and various methods for measuring tokens' morphological integration into the Breton lexicon (section 3.5).

These aspects of the data can provide a way of determining the extent to which different parts of the corpus fit the "néo-breton" model, in terms of their use of borrowings,

coinages and dialectal vocabulary. In order to do so, the frequency of tags added to the corpus during the annotation process will be examined in depth. Their occurrence will be compared across the three subcorpora (radio, print, and Facebook), as well as among different publications and radio programmes, in order to identify how different types of Breton may be used in different contexts as a result of medium, register, and other factors.

## 3.2 Words and their origins

### 3.2.1 Makeup of the corpus and statistical information

The corpus was gathered with the aim of obtaining linguistic data in Breton. However, several small portions of the data sampled were in French, and these were transcribed in order to provide context for the Breton data. In the radio subcorpus, French can be found in links that start and end programmes; in the printed sources, it is restricted to a small amount of front matter in *Bremañ*, accompanied by the notice that its inclusion is only "evit kaout an niv. CPPAP"—in order to be entitled to a publication number from the *Commission paritaire des publications et agences de presse*, which evidently stipulates that each issue must provide such information in French. French was by far the most prevalent among the Facebook data: there, entire conversational exchanges took place in the language, reflecting its unmoderated environment; despite the instruction in both French and Breton on the group's header image to "rédiger vos messages en breton", the sample shows no evidence of the use of French being viewed negatively in practice. A prolonged exchange in English was also found among the Facebook data. While all this was included in the corpus, it was not analysed or included in word count figures.

The radio subcorpus was compiled first, obtaining a little over an hour's worth of data, and following this the other two subcorpora were constructed so as to be roughly equivalent to the first in terms of the number of words collected. Excluding segments in French and English, therefore, the corpus contains a total of 34344 words,[64] comprising

64. As multi-word expressions were counted as single tokens for annotation purposes, similar multi-

10843 from the radio sources, 11566 from Facebook, and 11936 from the print publications. This third subcorpus was further split as evenly as possible according to the three different publications, with 4050 words coming from *Bremañ*, 3945 from *Brud Nevez*, and 3941 from *Ya*.

Of this total of 34344 words, 1653 were proper names of some kind, comprising 394 names of brands or organisations, 652 place names, and 607 other proper names, this last category mostly composed of personal names and titles of books and plays. These categories were excluded from the analysis for reasons of space.

This leaves a total of 32691 words. Following the workflow for determining salient tokens, 3904 of them were selected for annotation, this being a rate of 11.9%,[65] corresponding to 1344 individual lexemes, each appearing between 1 and 124 times. Over half the lexemes—789 out of the 1344—appear only once, and this reflects an overall skew towards a small number of occurrences: likewise, 227 lexemes appear twice and 92 three times. At the other end of the scale, *yezh*[66] ('language') is the most frequent tagged token with 124 occurrences; the next most frequent, *film* ('film'), is found noticeably less frequently with 79 occurrences, and the next most frequent after that, *holl* ('all') and *raktres* ('plan'), each occur 49 times.

This figure of 11.9% indicates the approximate percentage of words in the corpus that could potentially indicate an alignment with either traditional Breton or "néo-breton", as determined by the workflow in section 2.2.3. The figure is fairly consistent among the individual subcorpora, with 12.3% of tokens annotated in the radio subcorpus, 11.7% in the Facebook sample and 11.8% among the print data.

Breaking down the subcorpora themselves, the percentages of annotated tokens stray a little further from the 11.9% average in some cases, but still not hugely. The radio data can be split in two ways, according either to the station type (public or associative) or to the programme format (news or features). 12.8% of the tokens from the France

word expressions that went untagged were counted as single words for this figure. Additionally, non-language specific discourse markers, such as *hm*, *ah*, and the sound rendered as *hein* according to French orthographic conventions, were transcribed but not included in word counts. This is discussed more in section 4.3.1.

65. Percentages and other figures are given to 1 decimal place.

66. *Yezh* is an interesting case, not only because of its high frequency as a tagged token. This word will be examined in more detail in section 4.2.3 below.

Bleu data are annotated, slightly more than the Radio Breizh figure of 11.9%. The gap is wider when the data are split in the second way: in the news programmes, 13.7% of the tokens are annotated, compared with 11.1% for the longer feature programmes. This perhaps reflects the news programmes' need to deal with topics that do not relate to the traditions of Breton life, including politics and sport, which require more specialised vocabulary and more words not found in Breton prior to 1900,[67] both categories of words that are likely to contain contested vocabulary as outlined in section 2.2.3. However, while the feature programmes deal more often with cultural customs that are traditional in the Breton context, they do equally have to speak of newer inventions and make use of technical vocabulary, particularly in the domains of film and cinema.

The largest amount of variation in this respect is found among the printed sources. The *Brud Nevez* sample contains the smallest proportion of annotated tokens at 7.7%; *Ya* has 13.4%, and *Bremañ* has the highest, at 14.5%. Again, subject matter is likely to account for this discrepancy, at least partially. Given that the issue of *Brud Nevez* sampled from is on the theme of the sea, a traditionally important subject in Breton life, it is likely that less need to choose between borrowings and coinages would have been required in compiling this publication than in the cases of *Bremañ* and *Ya*, both of which relate a thematically wider range of local and international news stories. However, a lower rate of use of vocabulary that indexes support for either traditional Breton or "néo-breton" may also point to a particular ideological stance that seeks to avoid such conflict or to be more inclusive of readers from diverse backgrounds. By examining the tokens that *are* annotated, and whether they may indicate an overall alignment with typical "néo-breton" or with something else, we will be able to determine the extent to which this is the case.

### 3.2.2 Parts of speech: influence of medium and register

While analysis that focuses only on parts of speech cannot be interpreted as meaningful with regard to the "néo-"/traditional Breton divide, it can reveal how differences in register have linguistic consequences. The 3904 annotated tokens covered a range of

67. For the use of 1900 as a cut-off point, see section 2.2.4.

parts of speech, with their frequencies listed below.

| | **Tokens** | **Types** |
|---|---|---|
| Nouns | 2265 | 833 |
| Verbs | 562 | 206 |
| Adjectives | 439 | 194 |
| Adverbs | 274 | 70 |
| Discourse markers | 152 | 21 |
| Prepositions | 94 | 15 |
| Determiners and pronouns | 94 | 7 |
| Interjections | 17 | 6 |
| Verbal particles | 4 | 2 |
| Conjunctions | 3 | 2 |
| **Total** | **3904** | **1356**[68] |

Table 3.1: Parts of speech of tagged tokens

68. This is slightly greater than the overall number of types because some types functioned as different parts of speech depending on their syntactic context. Breton adjectives, for example, can be used adverbially, and verbs can be used as nouns.

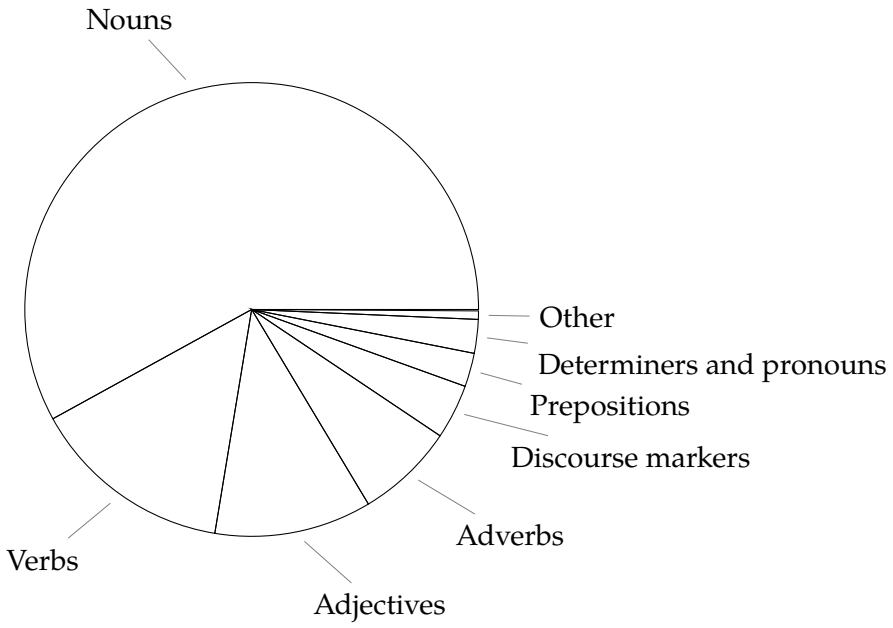


Figure 3.1: Graph of Table 3.1 (tokens)

Nouns are thus by far the most frequent of tagged items, constituting over half (58.0%) of tokens and 61.4% of types. Indeed, open-class lexical items (i.e. nouns, verbs, adjectives, and adverbs) are generally more numerous among the tagged tokens than closed-class items. Where closed-class items are tagged, it tends to be because they are indicative of a particular dialect: this is the case for 13 out of the 15 types in the case of prepositions, for example. As non-dialectal tagged tokens tend to be recent coinages or otherwise more technical vocabulary, this is unsurprising.

Breaking down these data in proportion to the three subcorpora, the following figures emerge.

| | | | Raw | | | Normalised[69] | | |
|---|---|---|---|---|---|---|---|---|
| | **Tokens:** | **Total** | **Radio** | **Facebook** | **Print** | **Radio** | **Facebook** | **Print** |
| *[70] | Nouns | **2265** | 664 | 742 | 859 | 673.6 | 705.7 | 791.6 |
| | Verbs | **562** | 176 | 201 | 185 | 178.5 | 191.2 | 170.5 |
| * | Adjectives | **439** | 94 | 199 | 146 | 95.4 | 189.3 | 134.6 |
| * | Adverbs | **274** | 118 | 70 | 86 | 119.7 | 66.6 | 79.3 |
| * | Discourse markers | **152** | 126 | 25 | 1 | 127.8 | 23.8 | 0.9 |
| * | Prepositions | **94** | 51 | 23 | 20 | 51.7 | 21.9 | 18.4 |
| | Determiners/pronouns | **94** | 29 | 35 | 30 | 29.4 | 33.3 | 27.6 |
| * | Interjections | **17** | 10 | 7 | 0 | 10.1 | 6.7 | 0 |
| | Verbal particles | **4** | 2 | 2 | 0 | 2.0 | 1.9 | 0 |
| | Conjunctions | **3** | 0 | 2 | 1 | 0 | 1.9 | 0.9 |
| | **Total** | **3904** | **1270** | **1306** | **1328** | | | |

Table 3.2: Tagged tokens by subcorpus

69. As the total size of each subcorpus was slightly different, the raw numbers of tokens were divided by the total number of words in the relevant subcorpus and then multiplied by 11000 to obtain a normalised figure. The figures in the first part of the table should thus be interpreted as 'number of tokens per 11000 words'. A similar process was applied to other divisions of the corpus and subcorpora in order to ensure that populations being compared were of the same size.

70. In this chapter, an asterisk is used to mark table rows in which statistical significance is attested.

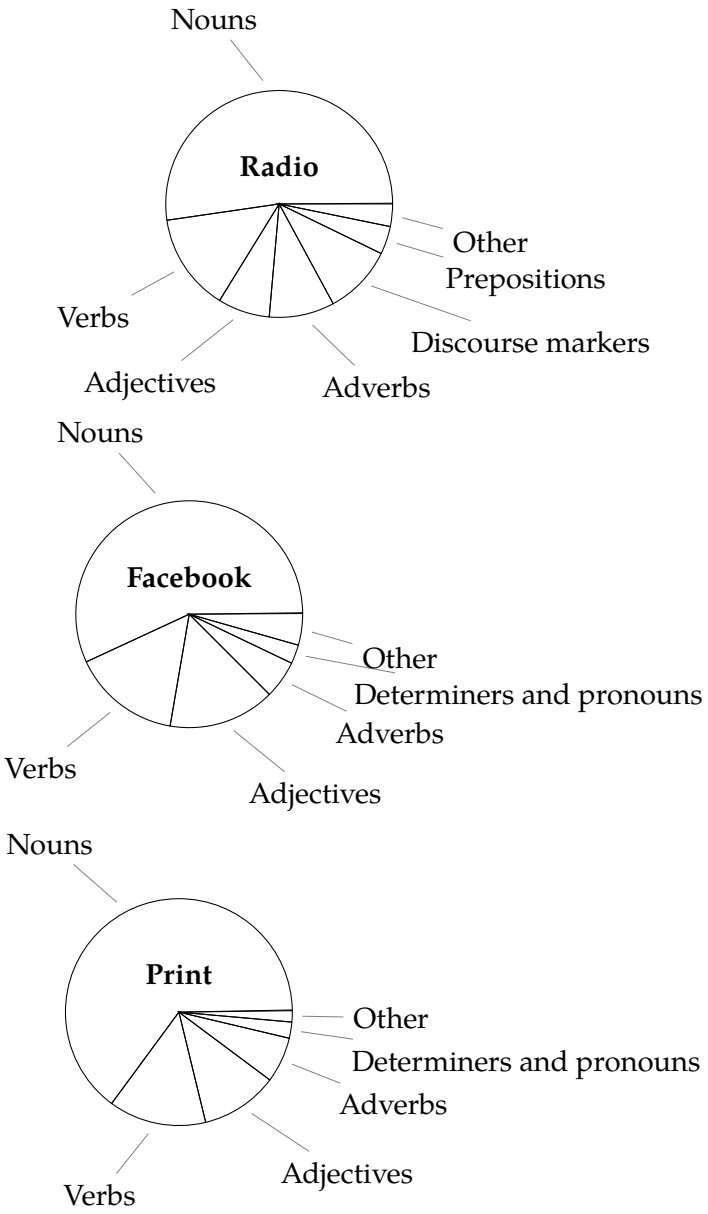


Figure 3.2: Graph of Table 3.2

These figures show some discrepancies among the subcorpora, reflecting the differences in register among the three different media types. Discounting the figures for verbal particles and conjunctions, where the sample size is too low to draw any meaningful conclusions, a few points of interest can be noted, and in fact the difference among subcorpora is statistically significant[71] for nouns, adjectives, adverbs, discourse markers, prepositions, and interjections. In the case of interjections, while the radio

71. The data were tested using a chi-square test at a 0.05% significance level: this is the case for all mentions of statistical significance in this chapter. Information on how this test was applied can be found in Appendix G.

and Facebook data display rates of occurrence that are fairly consistent with each other, there are no occurrences at all in the printed media, providing evidence that this is a more formal register. This difference of register is still more obvious when discourse markers are considered: here, there is a clear difference among all three of the subcorpora. The radio data show the highest rate of discourse markers, befitting the oral nature of this medium; the print journalism again has a very low rate, showing its more formal written style. The data from Facebook are in between with respect to this measure, showing that while computer-mediated communication shows characteristics of oral discourse (Ferrara, Brunner, and Whittemore 1991:10), its written nature means that the eventual register of language found on social media can be situated in an intermediate position.

Some of the less stark differences across corpora may also point to differences in register: the higher proportion of tagged adjectives in the Facebook data may reflect an overall higher proportion of adjectives, perhaps indicating a greater presence of subjective opinions and feelings in this unmoderated environment where users write on their own behalf.

A final point to consider from these figures is the high proportion of prepositions found in the radio data compared with the other two subcorpora. This is a word class whose occurrence would not be expected to vary across registers. As only 15 preposition types are tagged across the corpus, it is possible to examine their distribution in more detail.

| | Normalised | | |
|---|---|---|---|
| **Preposition** | **Radio** | **Facebook** | **Print** |
| *a-fet* 'concerning' | 1.0 | 1.0 | |
| *àr* 'on' | | 3.8 | 4.6 |
| *àr-bouez* 'towards' | | | 0.9 |
| *àr-lerc'h* 'after' | | | 0.9 |
| *aveit* 'for' | 4.1 | 1.0 | |
| **di*[72] 'to' | | | 0.9 |
| *diàr* 'from' | | | 3.7 |
| *diàr-benn* 'about' | | | 0.9 |
| *dibaoe* 'since' | 2.0 | | 1.8 |
| *diouzh* 'at, against' | 1.0 | 1.0 | |
| *doc'h* 'from' | 17.2 | 1.0 | |
| *e-kerzh* 'during' | 1.0 | | 4.6 |
| *èl* 'like' | 1.0 | 4.8 | |
| *endan* 'under' | | 1.0 | |
| *get* 'with' | 24.3 | 8.6 | |

Table 3.3: Prepositions by subcorpus

As this table shows, the higher proportion of tagged prepositions in the radio data is caused by a far greater presence of two particular types: *doc'h* and *get*. Both these words are markers of the Vannetais dialect, and their increased appearance among the radio data may at first appear to indicate a larger number of Vannetais speakers in this subcorpus than in the others. However, many of the other prepositions tagged are also markers of Vannetais: indeed, this includes all those that contain <àr>[73] (standard Breton <war>), a number of tokens of which feature in the two written subcorpora. Given that it is unlikely that the rate of occurrence of prepositions meaning 'from' or 'with'

72. The sole attestation of this word in the corpus is in the inflected form *dimp* ('to us'), hence the preceding asterisk to mark a reconstruction of the dialectal lemma.

73. In the radio subcorpus, *àr* would have been used in the transcription where prepositions written with <war> in the standard orthography lacked /w/ or /v/ at the beginning of the relevant syllable. However, this did not occur.

would vary among registers or media, the fact that *doc'h* and *get* are this prevalent only in the radio subcorpus perhaps suggests that different words play the role of dialect markers depending on whether the medium is oral or written. With this in mind, the distribution of dialectal features will be examined in more detail in section 3.4 below.

### 3.2.3 Origins of word elements

Tokens were tagged in order to mark the language via which their constituent elements would have entered modern Breton. This is more salient for the study than marking their language of ultimate origin, given that words entering Breton via French would still be considered French borrowings even if they had in turn been borrowed into French from another language. For example, *week-end* was tagged as French, despite ultimately deriving from English.[74] To set such words apart from vocabulary that originated in French, they were also given an extra tag to mark their status as "international".[75]

Likewise, words containing elements found in English but not in French, even if not originally from English, were tagged as English. Any words remaining whose elements were all attested in Breton prior to 1900 were tagged as such, with a different tag used in the case of words for which no attestation is indicated in Breton later than the early modern period (i.e. prior to around 1700), in order to identify any cases of the type pointed out by Calvez (2012:49), who notes that in the case of *lu* ('army'), "un terme vieux-breton dont on est sans attestation écrite depuis plusieurs siècles ... a été réintroduit". Similarly, words containing elements derived from Welsh were also given a specific tag, given that Welsh has been identified as a source of "neo-breton" coinages (German 2007:185).

Following this system, the vast majority of the tokens selected (98.4%) were tagged as

74. Accordingly, use of the terms "derived" and "origin" here refers to the language via which words came into Breton, rather than their ultimate origin.

75. In the context of this thesis, this term does not necessarily refer to learned words with Latin or Greek elements, although many of the "international" words in the corpus do fit into this category. The definition here is slightly wider, referring to any word used in both French as well as in major non-Romance languages (in practice, words found in French and English). "International" here therefore does not refer to the category of "learned internationalisms" evoked by Dowling (2014:9), but instead is used in the sense employed by Lerat (1988), who includes both these Latin- or Greek-derived words and "le mot allogène" (ibid.:487).

either French or Breton; the remaining 61 tokens contained neither French elements nor Celtic-derived elements attested in Breton between 1700 and 1900.[76]

The overall distribution of word origins is as follows:

| | |
|---|---|
| Breton (contains only elements found in Breton before 1900 and not in French) | 2356 |
| French (contains elements used in French—may also be found in Breton before 1900) | 1470 |
| Welsh (contains Welsh elements not found in Breton before 1900) | 30 |
| English (contains elements used in English and not in French) | 24 |
| Older Breton (contains elements found in Breton with the same meaning prior to around 1700 but no more recent attestations until after 1900) | 19 |
| Other languages (not part of French vocabulary) | 5 |
| **Total** | **3904** |

Table 3.4: Word origins among tagged tokens

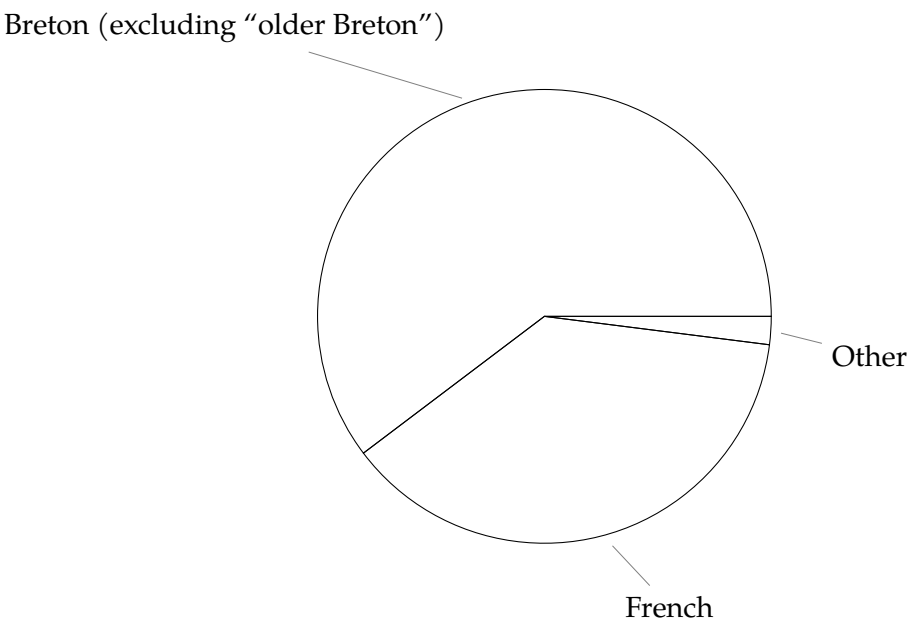


Figure 3.3: Graph of Table 3.4

The "other languages" category consisted of Basque (3 tokens), Antillean Creole and Portuguese. The last two of these were used where the token was specific to these lan-

76. See section 2.2.4 for the significance of this time period.

guages (*gwoka* and *samba de roda* respectively, both dances), and would presumably also have been used in French, although no attestations of this were available. However, the use of Basque (two tokens of *euskarat* [adjectival 'Basque'] and one of *euskarad* [nominal 'Basque']) is a different case: a form of the French/international word *basque* would have been available. Instead, Breton (in this instance) uses a term directly taken from the Basque language itself. This use of vocabulary from other minoritised languages for terms relating to the language in question recurs in the corpus, particularly in place names. While this research does not analyse this category, we can note *Euskadi* for the Basque Country, *Tolosa* ('Toulouse', from Occitan), *Doire* ('Derry', from Irish), and Baiona ('Bayonne', from Basque). Further examples of place names that follow this pattern can be found among the place names prescribed by the *OPLB* in its place name database, KerOfis,[77] which advises a number of place name translations for towns in the Occitan-speaking area that are closer to their Occitan names than their French names: for example, *Nisa*, *Montpelhièr*, *Marsilha* and *Ais de Provença*. This is also the case where other minoritised languages are concerned: *Elzas* is prescribed for Alsace. However, these recommendations are somewhat inconsistent: KerOfis gives *Strasbourg* rather than a name closer to the Alsacien *Strossburi*. Nonetheless, the general preference for minoritised-language place names indicates a political alignment with speakers of these other minoritised languages, implying that the *OPLB* wishes to emphasise the connection between speakers of Breton and users of other minoritised languages across the world.

While nineteen tokens come into the "older Breton" category, containing elements used in ways attested before 1700 but not again until after 1900, seventeen of them are in fact tokens of the same word, *dazont*.[78] The process of resurrecting older words invoked by Calvez (2012) therefore does not appear to occur particularly frequently. However, related processes can be observed elsewhere in the corpus. Some words attested between 1700 and 1900, which must be tagged as "Breton" rather than "older Breton", are flagged in various sources as being typically associated with an earlier stage of the language. *Koun* ('memory'), which occurs four times in the corpus, is identified by

77. The following examples were checked at http://www.fr.brezhoneg.bzh/40-kerofis.htm, 19 October 2017.

78. In this case, the elements themselves have occurred continuously in Breton: it is their composition as the noun *dazont* that is considered specific to the pre-modern language. See German (2007:186) and section 4.2.1.

Calvez (2012:649) as "attesté jusqu'au XVIIe siècle, puis tomb[é] en désuétude avant d'être introduit de nouveau dans la composition de néologismes". Two related compounds, *ankounac'haat* ('forget') and *maen-koun* ('commemorative stone'), are found in the corpus, along with *koun* itself. While *maen-koun* is a neologism, *ankounac'haat* is attested several times in nineteenth-century Breton, including in Le Gonidec's dictionary of 1850, where *koun* is also found—hence the exclusion of *koun* itself from the "older Breton" category during the annotation process, despite Calvez's remark. In the dictionary, however, *koun* includes a note that it is "peu usité aujourd'hui par d'autres que par des vieillards ; mais on le retrouve dans ses composés" (Le Gonidec 1850:212). The use of *koun* in the corpus is therefore notable, as its presence suggests that speakers are willing to include it in their language despite the fact that in the mid-nineteenth century it was already considered obsolescent. The appearance of *ankounac'haat* is less remarkable, given that it was included in Le Gonidec's dictionary without such comment, and that compounds of *koun* seem to have been common in the nineteenth century.

Out of the 1470 words tagged as French, 362 (24.6%) are internationally used words, also found in English and other languages, and often deriving from English (*match*) or Greek (*ideogram*), as well as other sources (*spaghetti*).

Splitting the corpus into its three constituent parts reveals further information on how tokens of different origin are used. The following table shows the number of tagged tokens in each subcorpus by origin, normalised so that each subcorpus is equivalent to 11000 words, and their proportion as a percentage of tagged tokens for each subcorpus. These percentage figures provide one way of measuring the tendency of the discourse, as the tokens under examination are those that could in some way indicate levels of allegiance with stereotypical "néo-breton". A higher proportion of French-derived words in a particular subcorpus could therefore indicate the use of a type of Breton further from this stereotype in that context.

| | | # tokens (normalised) | | | % of tagged tokens in the subcorpus | | |
|---|---|---|---|---|---|---|---|
| | **Origin** | **Radio** | **Facebook** | **Print** | **Radio** | **Facebook** | **Print** |
| * | **Breton** | 713.2 | 801.7 | 746.5 | 55.4% | 64.5% | 61.0% |
| * | **French** | 567.1 | 420.4 | 432.2 | 44.0% | 33.8% | 35.3% |
| * | *International* | *118.7* | *95.1* | *133.6* | *20.9%*[79] | *22.6%* | *30.9%* |
| * | *Non-international* | *448.4* | *325.3* | *298.6* | *79.1%* | *77.4%* | *69.1%* |
| * | **Welsh** | 0.0 | 2.9 | 24.9 | 0 | 0.2% | 2.0% |
| * | **English** | 1.0 | 9.5 | 12.0 | 0.1% | 0.8% | 1.0% |
| | **Older Breton** | 5.1 | 7.6 | 5.5 | 0.4% | 0.6% | 0.5% |
| | **Other** | 2.0 | 0.0 | 2.8 | 0.2% | 0 | 0.2% |

Table 3.5: Word origins by subcorpus

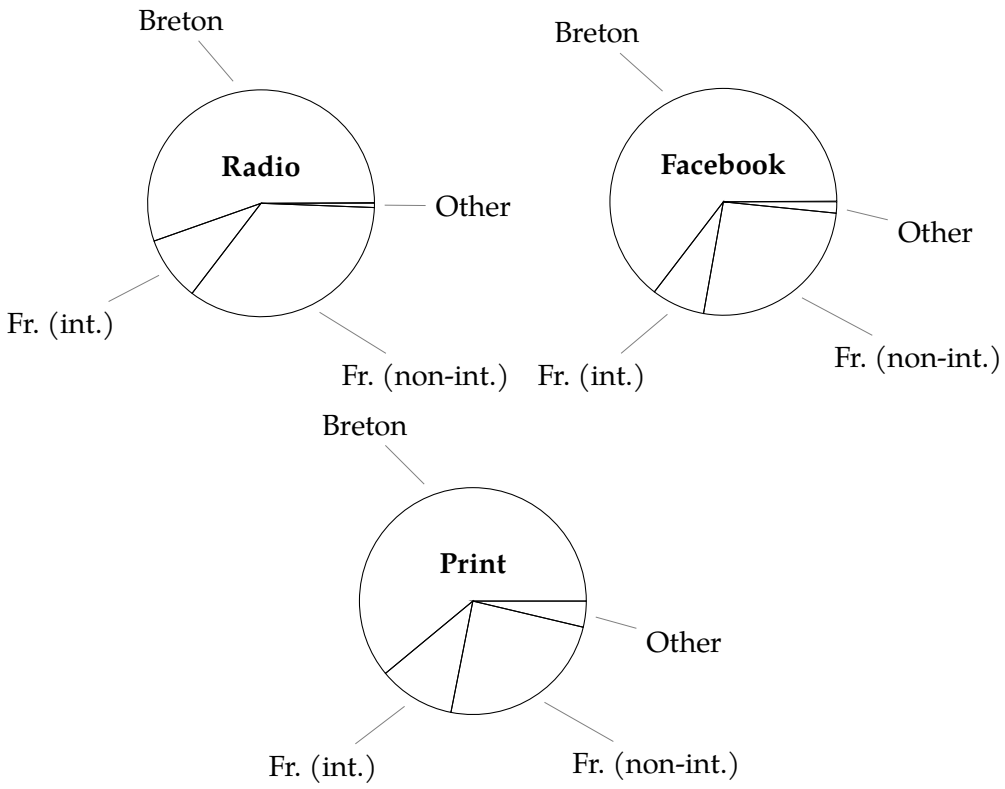


Figure 3.4: Graph of Table 3.5

79. For the "international" and "non-international" rows, the percentage figures are as a proportion of French-derived words per subcorpus; this is the case for the following three tables also.

For each language of origin, a chi-square test (see Appendix G) was carried out to compare the three subcorpora. This measure of statistical significance accounts for any difference in the number of tagged tokens per subcorpus: where fewer tokens are tagged, a large percentage difference may not be as significant as it appears. The chi-square test takes this into account.

The differences among subcorpora are striking: for all word origins except "other" and "older Breton", a statistically significant difference can be observed when comparison is carried out across the three subcorpora. Breton-derived words outnumber French-derived words in all cases, although the radio corpus has the smallest number of tokens tagged as Breton and the largest number tagged as French. The Facebook subcorpus is at the other extreme, with Breton-derived tokens twice as numerous as French-derived ones. This could perhaps indicate that the radio medium is more readily associated with traditional Breton, and that speakers on the radio may either deliberately or subconsciously choose to use more French-derived words in an effort to appeal to traditional speakers; the Facebook subcorpus seems in this respect to be more aligned with "néo-breton".

When the radio data are split according to whether or not the speaker is an employee of the station, the figures vary. Among employees, 36.8% of tokens come from French, a proportion closer to those found in the other two subcorpora and similar to the figure for the corpus as a whole (37.7%). Among guests, however, 51.5% come from French. As a large number of the words spoken by employees are scripted—for example, in news bulletins—this suggests that a higher proportion of Breton-derived words are used in scripted speech, while spontaneous speech is more likely to contain French-derived words.

On Facebook, in contrast, the written medium means that users are able, to some extent, to plan their vocabulary choice in advance, and this context provides the largest proportion of tokens derived from Breton and the smallest derived from French. Unless the populations of Facebook users and radio guests are significantly different, it can be extrapolated that Breton speakers take different approaches to their lexicon depending on whether they are speaking or writing. This could be related to a difference

in register, although as noted above, the register of the Facebook subcorpus is not so far removed from oral communication as that of the print subcorpus. It appears instead that Facebook is a more stereotypically "néo-breton" space than the radio in this regard, where users make the choice to use a greater rate of Breton-derived words.

Another difference across the subcorpora can be observed in the case of international words used in French. Among French-derived words, the print sources use a higher proportion of international words than the other media, suggesting that they are more tolerant of French-derived words if they are also used in other languages. This also demonstrates how the higher register of print media has effects on the vocabulary used: these international words often refer to newer or more technical concepts. Conversely, the number of tokens from French that are not used in other languages is the lowest of all three subcorpora in this case, implying that these French-specific borrowings may be considered less acceptable in this medium than elsewhere, and suggesting a more stereotypically "néo-breton" outlook.

The print subcorpus also contains the greatest number of tokens whose origins were tagged as neither Breton nor French. Use of Welsh-derived words is associated with alignment with "néo-breton", and their greater frequency here, along with the apparent lesser tolerance of non-internationally used French-derived words, indicates again that the strongest associations with stereotypical "néo-breton" may be found among the printed sources. The radio subcorpus appears to contain more words from older Breton, but in fact all five occurrences are tokens of the same word, and all occur in the same section of one programme, with four of the five tokens uttered by a single speaker. When the borderline cases *koun* and *maen-koun* are also considered, as discussed above, these are found only in the print sources (two tokens of *koun* and one of *maen-koun* in *Bremañ*; two of *koun* in *Ya*). Regardless of whether *koun* and *maen-koun* should be counted as coming from older Breton, it is clear that the use of words derived from Celtic sources other than contemporary traditional Breton (i.e. from Welsh and older Breton) occurs at the highest rate in print, and at the lowest on the radio. This less frequent use of words from Celtic sources matches the higher rate of use of words derived from French on the radio, demonstrating that the type of language is further from stereotypical "néo-breton".

Returning to the print subcorpus, the more frequent occurrence of English-derived tokens suggests a higher level of receptivity to words from languages other than French. Combined with the higher proportion of tokens that are used in French but also in other languages (the "international" category), this indicates that the printed sources are more tolerant of using words that exist in other languages, but less tolerant of using words that are used only in French. This again conforms more stereotypically to the supposed "néo-breton" pattern, but cannot be deemed excessively purist, as borrowings from languages other than French appear tolerable. This apparent contradiction will be examined further by considering the three publications separately.

Before doing this, the radio subcorpus can be examined in more detail. We have already seen the difference between radio employees and guests; below is the radio subcorpus split by station type (public versus associative).

| | **# tokens (normalised)** | | **% of tagged tokens** | |
|---|---|---|---|---|
| **Origin** | **Public** | **Associative** | **Public** | **Associative** |
| **Breton** | 374.6 | 340.1 | 55.9% | 54.9% |
| **French** | 295.0 | 273.1 | 44.0% | 44.0% |
| *International* | *52.0* | *66.1* | *17.6%* | *24.2%* |
| *Non-international* | *243.0* | *207.0* | *82.4%* | *75.8%* |
| **Welsh** | 0.0 | 0.0 | 0 | 0 |
| **English** | 1.1 | 0.0 | 0.2% | 0 |
| * **Older Breton** | 0.0 | 4.9 | 0 | 0.8% |
| **Other** | 0.0 | 1.9 | 0 | 0.3% |

Table 3.6: Word origins in the radio subcorpus by station type

There is little variation between the two categories here: the only statistical significance to be observed is in the case of the older Breton tokens, where, as we have seen, all are tokens of the same word (*dazont*) and all occur within a short space of time as part of the same programme. The associative radio sample does contain a higher

rate of internationally-used French-derived tokens, which may comparatively indicate a slightly lower tolerance of French-derived words in general.

Breaking these data down by speaker type, as above, there is again little difference between public and associative radio in general. Both show the higher rate of French-derived tokens for guests than for employees of the stations. The associative radio seems slightly more extreme in that 33.0% of tokens are derived from French among employees (40.3% for public radio) and 54.4% are derived from French among guests (48.3% for public radio)—so the gap is larger in this case. Among guest speakers on associative radio, the number of French-derived tokens is therefore higher than the number of Breton-derived tokens. Speakers who deviate from the supposed "néo-breton" norm of using coinages in preference over borrowings perhaps feel more welcome on associative than public radio. However, the use of dialect may also be relevant here, and will be returned to in section 3.4.

The radio data can also be split by type of programme, as below.

| | **# tokens (normalised)** | | **% of tagged tokens** | |
|---|---|---|---|---|
| **Origin** | **News** | **Features** | **News** | **Features** |
| **Breton** | 405.5 | 310.0 | 57.7% | 52.7% |
| **French** | 296.3 | 271.4 | 42.2% | 46.1% |
| *International* | *61.3* | *57.4* | *20.7%* | *21.2%* |
| *Non-international* | *235.0* | *213.9* | *79.3%* | *78.8%* |
| **Welsh** | 0.0 | 0.0 | 0 | 0 |
| **English** | 1.0 | 0.0 | 0.1% | 0 |
| * **Older Breton** | 0.0 | 5.0 | 0 | 0.8% |
| **Other** | 0.0 | 2.0 | 0 | 0.3% |

Table 3.7: Word origins in the radio subcorpus by programme type

In this instance, there is again a small amount of variation, but none is statistically significant (other than, again, the "older Breton" category): the figures for the proportion

of internationally-used words in particular are extremely consistent. Split by speaker type, too, there is little variation. It appears that the programme format does not have an effect on the rate of use of words by origin.

The print subcorpus can also be split according to the three publications.

| | | # tokens (normalised) | | | % of tagged tokens | | |
|---|---|---|---|---|---|---|---|
| | **Origin** | ***Bremañ*** | ***Brud Nevez*** | ***Ya*** | ***Bremañ*** | ***Brud Nevez*** | ***Ya*** |
| | **Breton** | 300.6 | 146.9 | 297.7 | 61.7% | 53.6% | 64.6% |
| * | **French** | 172.0 | 120.8 | 138.6 | 35.3% | 44.1% | 30.1% |
| | *International* | *43.5* | *30.7* | *59.5* | *25.3%* | *25.4%* | *43.0%* |
| | *Non-international* | *128.6* | *90.2* | *79.1* | *74.7%* | *74.6%* | *57.0%* |
| * | **Welsh** | 7.2 | 0.9 | 16.7 | 1.5% | 0.3% | 3.6% |
| | **English** | 4.5 | 3.7 | 3.7 | 0.9% | 1.4% | 0.8% |
| | **Older Breton** | 0.0 | 1.9 | 3.7 | 0 | 0.7% | 0.8% |
| | **Other** | 2.7 | 0.0 | 0.0 | 0.6% | 0 | 0 |

Table 3.8: Word origins in the print subcorpus

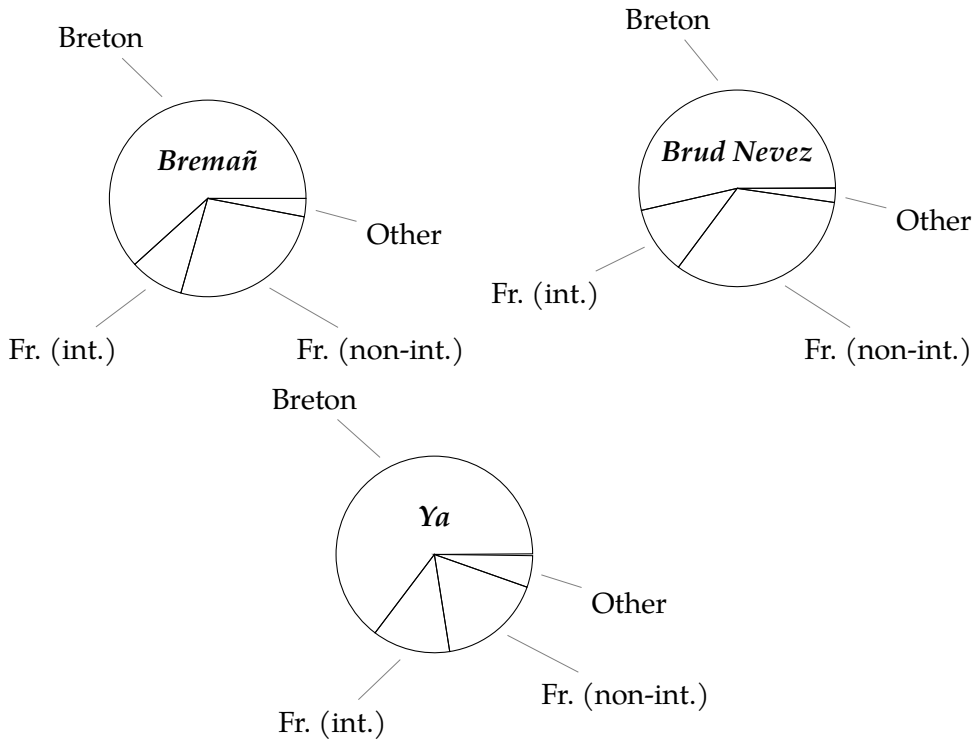


Figure 3.5: Graph of Table 3.8

In this case, there is some striking variation, particularly when *Brud Nevez* is compared with the other two sources. We have already seen that the total number of annotated tokens is lower in the *Brud Nevez* sample, but with the data broken down in this way it is possible to see that this is particularly the case for Breton-derived words, and, compared with the other print sources, also for Welsh-derived words. Consequently, while *Brud Nevez* also contains the smallest number of French-derived tokens, the number of Breton-derived tokens in this source is so small that proportionally, the French-derived tokens appear significantly more numerous than in *Bremañ* and *Ya*, and statistical significance is apparent in the cases of French and Welsh.

*Ya* contains the highest number of Breton-derived tokens, in both absolute[80] and percentage terms. It also exhibits an unusually high proportion of internationally used words in relation to the total number of French-derived tokens. In this instance, subject matter may play a role: part of the sample is an article about constructed languages, which includes a number of names of languages used in fantasy films and television

80. The figures in the table are normalised, but this also holds true for the raw data.

series, such as *dothraki*, *irathient*, and *shiväisith*. However, the fact remains that *Ya* contains the smallest number of non-internationally-used French-derived tokens, and this, together with its higher rate of Breton-derived, older Breton-derived and Welsh-derived tokens, suggests at this stage that *Ya* may be the most conformant to the "néo-breton" stereotype. *Bremañ* is broadly similar, while *Brud Nevez* is the most unorthodox in its far lower rate of use of Breton-derived tokens. The one figure that remains similar across all three publications is the occurrence of English-derived tokens, which is higher than in the other two subcorpora, perhaps reflecting a more international outlook on the part of the printed publications.

### 3.2.4 Summary

This section focused initially on parts of speech, revealing differences in register among the subcorpora, and then on word origins, which provided a preliminary measure of conformance to stereotypical "néo-breton" for the different parts of the corpus. These initial enquiries have thus provided some clear indications about how the type of language used varies among the different parts of the corpus, and how this may correlate with different attitudes towards the form "néo-breton" should take.

Nouns were seen to be the most frequently tagged items, contributing to a trend of there being more open-class words among the tagged tokens in general. This reflects the fact not only that nouns are typically the most frequently-used word class, but also that in this particular study, many of the tokens of interest denote newer concepts not used in Breton vocabulary prior to 1900. The paucity of interjections and discourse markers in the print subcorpus provides an indication that the data taken from print publications are of a higher register, while the fact that some are still present among the Facebook data shows that computer-mediated communication of this type displays characteristics of oral discourse despite the fact that it is a written medium. The resulting continuum of register, from the print subcorpus (highest) to the radio data (lowest), provides a backdrop for subsequent investigations of the data.

Turning to word origins, it was noted that the majority of tagged tokens do not involve borrowing from other languages, but that tokens involving borrowing from French

form a substantial minority. There are certain areas across and within the subcorpora where significant variation can be observed in terms of the use of words of different origins. In general, the radio seems most disposed towards stereotypical traditional Breton in its use of a higher rate of French borrowings. This occurs more frequently among guests than among employees, however, which may result from the fact that guests' contributions are less often scripted. Writing on Facebook is not so inclined towards French borrowings, instead tending more towards the use of Celtic neologisms, suggesting an outlook that broadly maintains stereotypical "néo-breton" characteristics and indicating a greater tendency towards purist ideologies in this online context. The print sources, in general, also conform to this stereotype in terms of word origins, and appear more tolerant of internationally-used French-derived words than of other French-derived words, as well as using more words from other languages and from older Breton, which may be a consequence of the higher register employed. *Brud Nevez* noticeably goes against this trend, however, suggesting that there may be room for a range of approaches in the domain of Breton print journalism. While *Ya* and *Bremañ* are similar, *Ya* uses more internationally-used and English-derived words, suggesting particularly low tolerance of borrowings that come specifically from French.

## 3.3 Dictionaries and other sources

### 3.3.1 Dictionary attestations

Examining a range of dictionaries, including some published before the decline and revitalisation of Breton are considered to have begun and others from throughout the revitalisation period, can reveal how attitudes to words of different origin in Breton have changed through time, and comparing this with the corpus provides a way of measuring the conformance of the language in question with prescribed standards. To this end, two sets of tags were used, based on a specific range of dictionaries and similar sources. For each source, one tag marked the occurrence of the token in the source with the same meaning, while a second tag indicated that it occurred only with a different meaning.

In addition to tagging words that appeared in Breton prior to 1900 based on attestation in Devri,[81] two dictionaries from this period were also specifically included in this selection. The first, *Le Catholicon*, dates from 1499 and is the first published dictionary of Breton, giving equivalent words in both French and Latin. The second dictionary is that of Jean-François Le Gonidec, edited by Théodore Hersart de la Villemarqué, published in 1850. This selection therefore spans the chronological range of pre-1900 Breton dictionaries.

Attestation in dictionaries is of course not proof of use in the vernacular at the time of the dictionary's publication. Even dictionaries of traditional Breton will reflect the ideological motivations of their creators; Le Gonidec and Villemarqué in particular have been noted for their inclusion of words based on Welsh and unattested otherwise in the language of traditional Breton speakers (Trépos 1960:394; Morvannou 2004). Nonetheless, they also include French-derived words that do not appear in some later prescriptive dictionaries, and their definitions of certain words are indications of cases of semantic expansion that have occurred during the revitalisation of Breton. Their reliance on borrowings from Welsh also indicates that tendencies towards neo-Celtic purism are not restricted to the twentieth century, and are therefore not an exclusively "néo-breton" phenomenon.

Tokens were also checked in five post-1900 sources, comprising two dictionaries, two online databases, and one textbook. The dictionaries were selected so as to use one published prior to the political change in direction that occurred as a result of the events of 1968, and one published after it. This does not mean that every dictionary published after 1968 will be more tolerant of French-derived words: Rottet (2014:227) examines ten dictionaries, eight of which were published after this date, and asserts that they "are fairly evenly split in terms of alignment with 'traditionalist' or 'purist' ideologies".

The pre-1968 dictionary used here is the *Dictionnaire breton-français* of Roparz Hemon, one of the most prolific grammarians of Breton and editor of the literary journal *Gwalarn*, which played a major role in the "néo-breton" movement during the early twentieth century. Hemon's dictionary went through several editions, later versions of which were named the *Nouveau dictionnaire breton-français* to mark their significant

81. As noted above, this is not certain to have captured every token.

expansion from their predecessors. Hemon was one of a number of high-profile Breton speakers to be condemned after the second world war for suspected collaboration with the occupying forces (Wmffre 2008:175); his linguistic contributions to the Breton movement therefore remain contentious, and his support for what is now the de facto standard Breton orthography, *peurunvan*, was a factor in speakers' refusal to adopt this system, as they believed the orthography itself was a direct result of collaboration (ibid.:179). However, Hemon was a prolific linguist whose work on Breton cannot be disregarded.

The post-1968 dictionary is a 1994 publication from the publishing house Mouladurioù Hor Yezh. This is the second edition of a pocket-sized dictionary that has become one of the most popular Breton-French dictionaries in use today, with translations of the French definitions and headwords into English and German also available.[82] This dictionary is in effect the successor to earlier classical works such as Hemon's publications: it was produced by scholars of Breton coming from the same tradition, and is in fact based on another of Hemon's works, *Brezhoneg eeun*. Using this dictionary as a source reflects the fact that the intent of this research is not to compare dictionaries with major ideological differences,[83] but instead to track any changes in the mainstream "néo-breton" current.

Given the small size of the Mouladurioù Hor Yezh dictionary, and its age, it is unsurprising that it does not contain particularly technical vocabulary or words for more recently invented concepts (for example, 'broadband'). Two online terminology databases were therefore also consulted. The first, TermOfis, a general database of mostly technical terms, is regulated by the *Office public de la langue bretonne*, and therefore is accorded the legitimacy of being produced by the body officially responsible for Breton language planning. The second, Brezhoneg21, was chosen for its comprehensive coverage of scientific terminology, as a more moderate alternative to the works of Preder, a terminology commission notorious for its overuse of Celtic-derived vocabulary (Rottet 2014:223).[84] As a development of terminology recommendations made prior

82. See https://www.brezhoneg.org/br/catalogues?search_api_views_fulltext=geriadur+divyezhek (accessed 26 Oct 17).

83. This is done by Rottet (2014).

84. Indeed, Brezhoneg21 cites Preder as having taken the tendency towards neo-Celtic coinages too far, asserting that by them, "kaset e vo neuze ar pleg 'puraat' kalz pelloc'h ... Ur mikrolekt brezhonek

to the establishment of the *OPLB* and TermOfis for use in the Diwan education system, it can likewise be considered of semi-official status.

While these two sources tend to contain more specialised vocabulary, the final source is the beginners' textbook *Brezhoneg ... buan hag aes*, written by Per Denez, another important figure in the early decades of the revitalisation of Breton. This textbook contains only the most basic Breton vocabulary, but this is nonetheless a context where choices can often be made between French-derived or Celtic-derived terms.[85] As in the case of the 1994 dictionary, the existence of translations of *Brezhoneg ... buan hag aes* into several other languages, including English, Welsh, German and Esperanto, indicates its popularity. While it postdates 1968 by a few years, it can be considered part of the earlier "néo-breton" tradition, as Per Denez was active at the same time as Roparz Hemon.

The dictionaries can therefore be classed as belonging to a sequence of traditions.[86] The two pre-1900 dictionaries date from an era before the growth of the "néo-breton" movement began. Hemon's dictionary and *Brezhoneg ... buan hag aes* form part of the earlier "néo-breton" tradition, while the 1994 dictionary and the two terminology databases are more recent. Analysing how the corpus contains words from this range of sources provides another measure for determining the form taken by "néo-breton" today.

### 3.3.2 Dictionary attestation in the corpus

Using the first of the two sets of tags mentioned at the beginning of section 3.3.1, marking tokens' attestation in the specified sources with the same meaning as in the sample, we can investigate the extent to which the different dictionaries are represented in the corpus, splitting it first according to the three subcorpora. For TermOfis, a further row indicates the number of tokens used in the corpus in ways that contradict the recommendations in TermOfis: these are classed as "proscribed by TermOfis".

nevez a savo e-giz-se, peuzdigomprenus evit Yann vrezhoneger" ('the instinct for "purification" [was] thus taken further ... in this way, they created a new Breton microlect, scarcely comprehensible to the ordinary Breton speaker'; http://brezhoneg21.com/pivomp.php; accessed 13 Nov 17).

85. This can be seen in the fact that Denez chooses to give *demat* as a greeting rather than French-derived *boñjour* or *salud*, or the more idiomatic Breton phrase *mont a ra mat?* (roughly 'how's it going?').

86. They are accordingly listed chronologically in the tables.

| | | # tokens (normalised) | | | % of tagged tokens in the subcorpus | | |
|---|---|---|---|---|---|---|---|
| | **Source** | **Radio** | **Facebook** | **Print** | **Radio** | **Facebook** | **Print** |
| | Breton before 1900 | 420.0 | 380.4 | 351.1 | 32.6% | 30.6% | 28.7% |
| * | *Le Catholicon* (1499) | 77.1 | 137.9 | 121.6 | 6.0% | 11.1% | 9.9% |
| | Le Gonidec/ Villemarqué (1850) | 129.9 | 122.7 | 125.3 | 10.1% | 9.9% | 10.2% |
| * | Hemon (1964) | 555.9 | 579.2 | 652.5 | 43.1% | 46.6% | 53.3% |
| * | *Brezhoneg* ... (1972) | 67.0 | 77.0 | 94.9 | 5.2% | 6.2% | 7.8% |
| | Mouladurioù Hor Yezh (1994) | 758.8 | 683.8 | 728.0 | 58.9% | 55.1% | 59.5% |
| | brezhoneg21 (online) | 365.2 | 379.5 | 400.9 | 28.3% | 30.6% | 32.8% |
| * | TermOfis (online) | 687.8 | 702.8 | 832.2 | 53.4% | 56.6% | 68.0% |
| | *Proscribed by TermOfis* | *217.1* | *230.2* | *174.2* | *16.9%* | *18.5%* | *14.2%* |

Table 3.9: Dictionary attestations by subcorpus

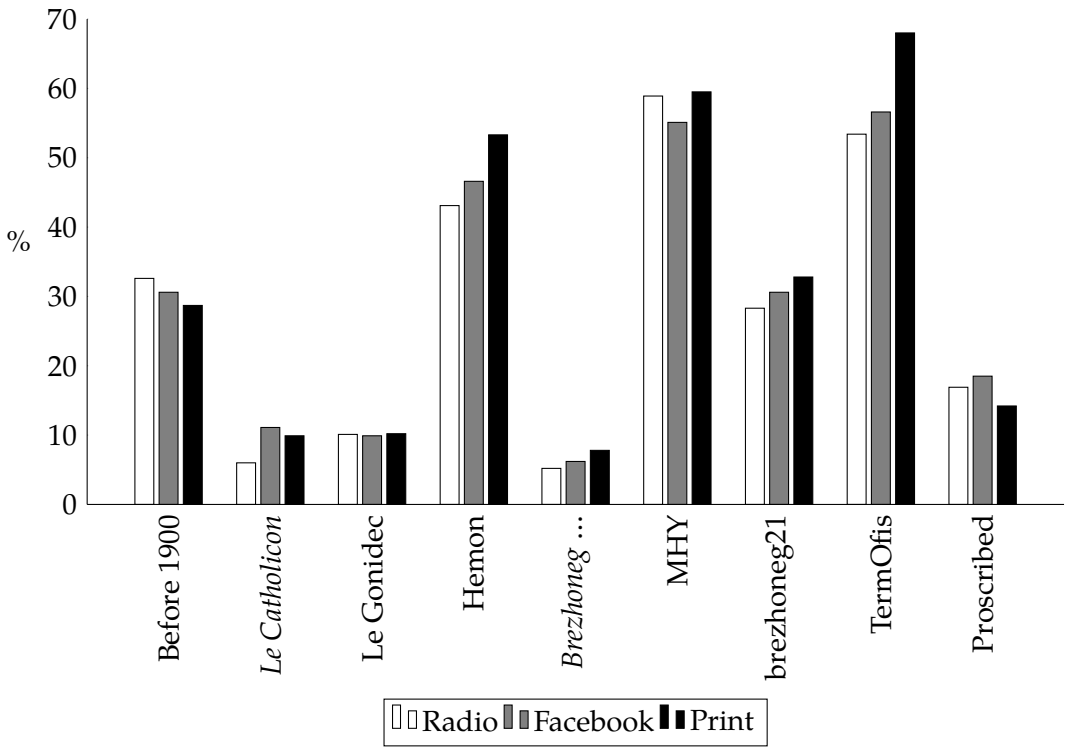


Figure 3.6: Graph of Table 3.9

Statistical significance can be observed in certain cases, although this is sometimes skewed by particular words. It initially appears that words found in *Le Catholicon* are far less frequent among the radio data than elsewhere, but much of this is due to the much lower rate of occurrence of *yezh* ('language'), the most frequent tagged token in the corpus overall. In the radio data *yezh* appears only four times, while it appears 71 times in the Facebook data and 49 times in the printed sources. This is due to the nature of the topics discussed: much of the discussion on Facebook focuses on the Breton language itself, while a number of the articles in *Ya* cover other languages. This noticeable difference is therefore unlikely to be meaningful in the sense of revealing any ideological stance.

Significance can however also be observed for Hemon's dictionary, *Brezhoneg ... buan hag aes*, and TermOfis: for all three, the number of tokens found in them that occur in the print subcorpus is noticeably higher than in the other two subcorpora, and the radio subcorpus contains the smallest number in each case. This suggests that the printed sources follow the supposedly typical current of "néo-breton" most closely, their vocab-

ulary resembling not only that of the standard "néo-breton" reference works from the conservative side of the movement, but also that of TermOfis, the most official terminology database available. Conversely, the printed sources use the smallest number of tokens that are proscribed by TermOfis. However, in this case, it is not the radio sources with the highest number of tokens here, but instead the Facebook data. While it has already been seen that Facebook users tend to conform to the typical current of "néo-breton", and these figures generally add weight to this, it is also evident that Facebook users employ a high number of terms in ways that are not officially recommended. This suggests that the Facebook users display some creativity with language, but not in a way that necessarily correlates with an attitude that is opposed to typical "néo-breton": they remain close to typical "néo-breton" in the sense of preferring Celtic-derived words, but this does not mean that they use standard language.

Again, the radio data can be split by station.

| | | **# tokens (normalised)** | | **% of tagged tokens** | |
|---|---|---|---|---|---|
| | **Source** | **Public** | **Associative** | **Public** | **Associative** |
| | Breton before 1900 | 218.6 | 202.1 | 32.6% | 32.6% |
| | *Le Catholicon* (1499) | 39.3 | 37.9 | 5.9% | 6.1% |
| * | Le Gonidec/ Villemarqué (1850) | 81.7 | 49.6 | 12.2% | 8.0% |
| | Hemon (1964) | 308.8 | 249.7 | 46.0% | 40.3% |
| | *Brezhoneg* ... (1972) | 40.3 | 27.2 | 6.0% | 4.4% |
| | Mouladurioù Hor Yezh (1994) | 417.0 | 345.0 | 62.2% | 55.6% |
| | brezhoneg21 (online) | 206.9 | 160.3 | 30.8% | 25.9% |
| * | TermOfis (online) | 391.6 | 300.3 | 58.4% | 48.4% |
| * | *Proscribed by TermOfis* | *93.4* | *122.4* | *13.9%* | *19.7%* |

Table 3.10: Dictionary attestations in the radio subcorpus by station type

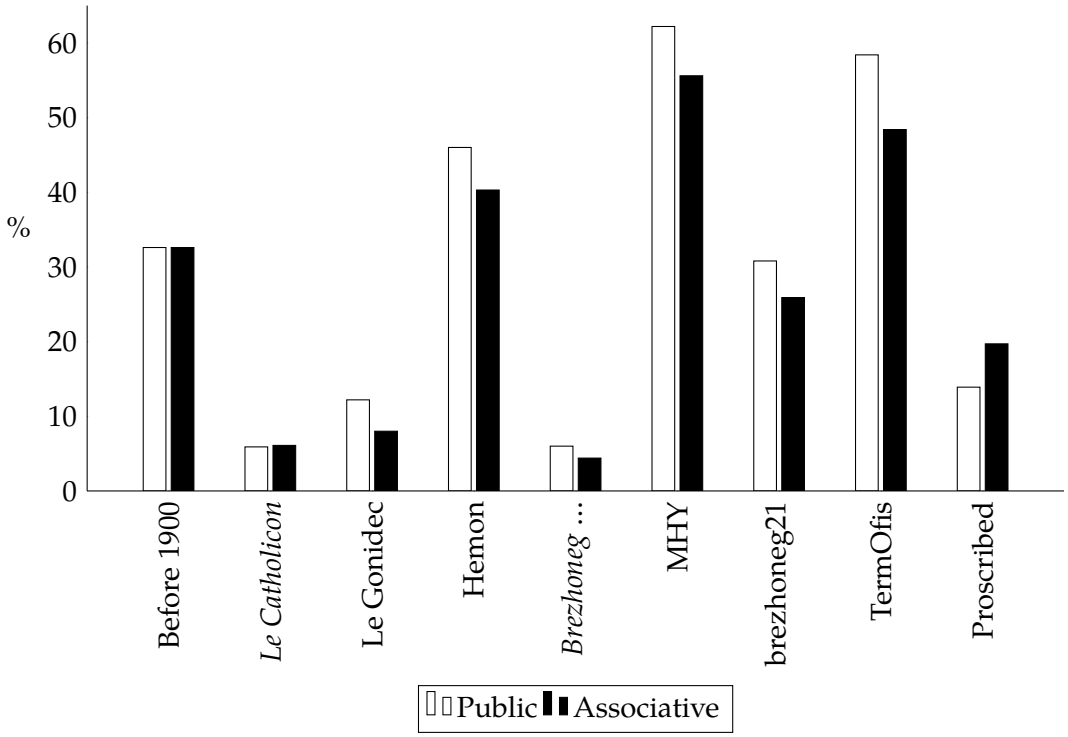


Figure 3.7: Graph of Table 3.10

Across the two contexts, the rate of use of words found in the different dictionaries is mostly very consistent. Statistical significance can however be observed for Le Gonidec and Villemarqué's dictionary, where public radio uses a higher rate of tokens than associative radio, and for TermOfis and words proscribed by TermOfis, where public radio uses more words in TermOfis and fewer words proscribed by it. Taken as a whole, these three significant differences may suggest that public radio conforms more to standardised versions of Breton than associative radio does—which may be expected through the former's belonging to the state-sponsored Radio France network. Similar differences, while not large enough to be statistically significant, can be observed in the case of all the other sources other than *Le Catholicon*, which, in its use of a higher rate of French-derived words, is not part of the trend towards neologisms followed to a greater or lesser degree by the other sources. This certainly suggests that public radio tends to use more standard forms.

The radio data can also be split according to programme type, once again.

| | | # tokens (normalised) | | % of tagged tokens | |
|---|---|---|---|---|---|
| | **Source** | **News** | **Features** | **News** | **Features** |
| | Breton before 1900 | 235.0 | 186.2 | 33.4% | 31.6% |
| | *Le Catholicon* (1499) | 37.4 | 39.6 | 5.3% | 6.7% |
| | Le Gonidec/ Villemarqué (1850) | 74.9 | 55.5 | 10.7% | 9.4% |
| * | Hemon (1964) | 271.4 | 284.3 | 38.6% | 48.3% |
| | *Brezhoneg* ... (1972) | 40.5 | 26.7 | 5.8% | 4.5% |
| | Mouladurioù Hor Yezh (1994) | 393.0 | 366.5 | 55.9% | 62.3% |
| * | brezhoneg21 (online) | 174.7 | 190.2 | 24.9% | 32.3% |
| | TermOfis (online) | 357.7 | 330.8 | 50.9% | 56.2% |
| * | *Proscribed by TermOfis* | *139.3* | *79.2* | *19.8%* | *13.5%* |

Table 3.11: Dictionary attestations in the radio subcorpus by programme type

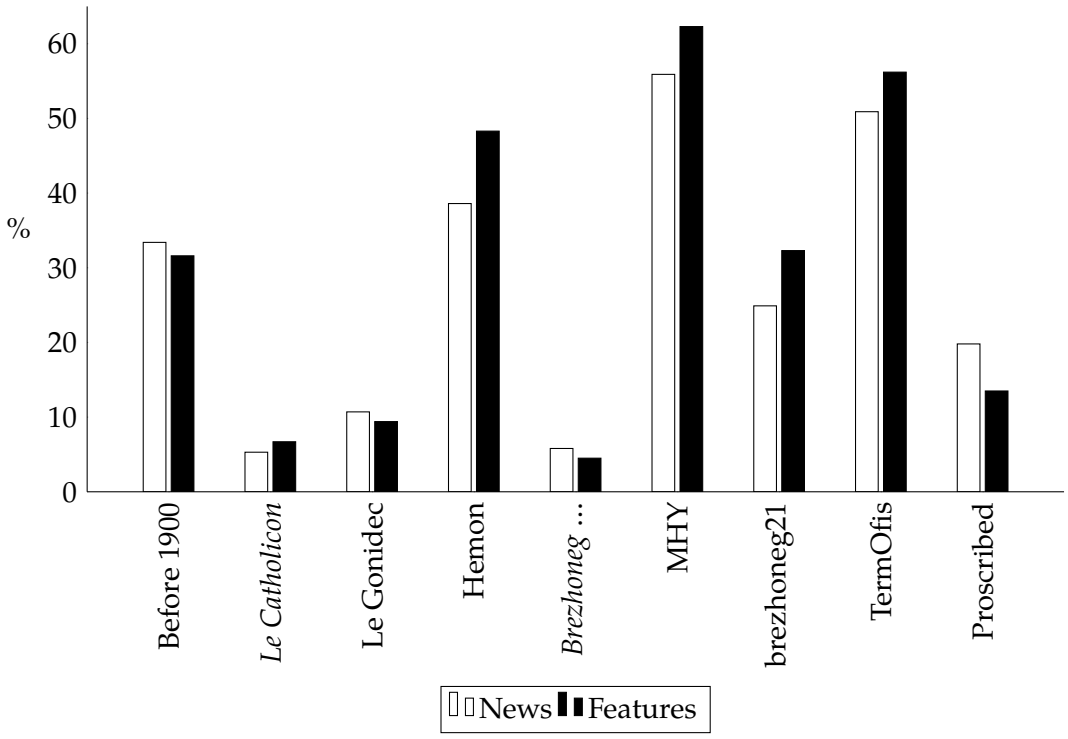


Figure 3.8: Graph of Table 3.11

Some examples of statistical significance can be found among these data. More of the tokens used in the feature programmes are found in Hemon's dictionary than in the news programmes, and similarly, more are found in Brezhoneg21. This latter finding suggests that the feature programmes deal with more technical scientific vocabulary, most probably in their discussion of cinema.

The news programmes also use more tokens proscribed by TermOfis, which may indicate a higher tolerance of non-standard lexicon in these contexts. Taken in conjunction with the fact observed above that associative radio uses more terms proscribed by TermOfis, this indicates that the news programmes on associative radio may be the locus of the most non-standard language. Again, this may reflect a potentially broader range of guests interviewed on these programmes, with more diverse linguistic repertoires.

Finally, the data from the printed sources can be split according to the three publications.

| | | **# tokens (normalised)** | | | **% of tagged tokens** | | |
|---|---|---|---|---|---|---|---|
| | **Source** | ***Bremañ*** | ***Brud Nevez*** | ***Ya*** | ***Bremañ*** | ***Brud Nevez*** | ***Ya*** |
| * | Breton before 1900 | 90.5 | 112.5 | 148.9 | 18.6% | 41.0% | 32.3% |
| * | *Le Catholicon* (1499) | 24.4 | 27.0 | 70.7 | 5.0% | 9.8% | 15.4% |
| * | Le Gonidec/ Villemarqué (1850) | 38.0 | 47.4 | 40.0 | 7.8% | 17.3% | 8.7% |
| | Hemon (1964) | 252.6 | 133.8 | 265.2 | 51.9% | 48.8% | 57.6% |
| * | *Brezhoneg* ... (1972) | 21.7 | 36.2 | 37.2 | 4.5% | 13.2% | 8.1% |
| | Mouladurioù Hor Yezh (1994) | 292.4 | 154.3 | 280.0 | 60.0% | 56.3% | 60.8% |
| * | brezhoneg21 (online) | 172.9 | 99.5 | 127.5 | 35.5% | 36.3% | 27.7% |
| * | TermOfis (online) | 378.4 | 140.3 | 310.8 | 77.7% | 51.2% | 67.5% |
| * | *Proscribed by TermOfis* | *48.9* | *58.6* | *67.0* | *10.0%* | *21.4%* | *14.5%* |

Table 3.12: Dictionary attestations in the print subcorpus

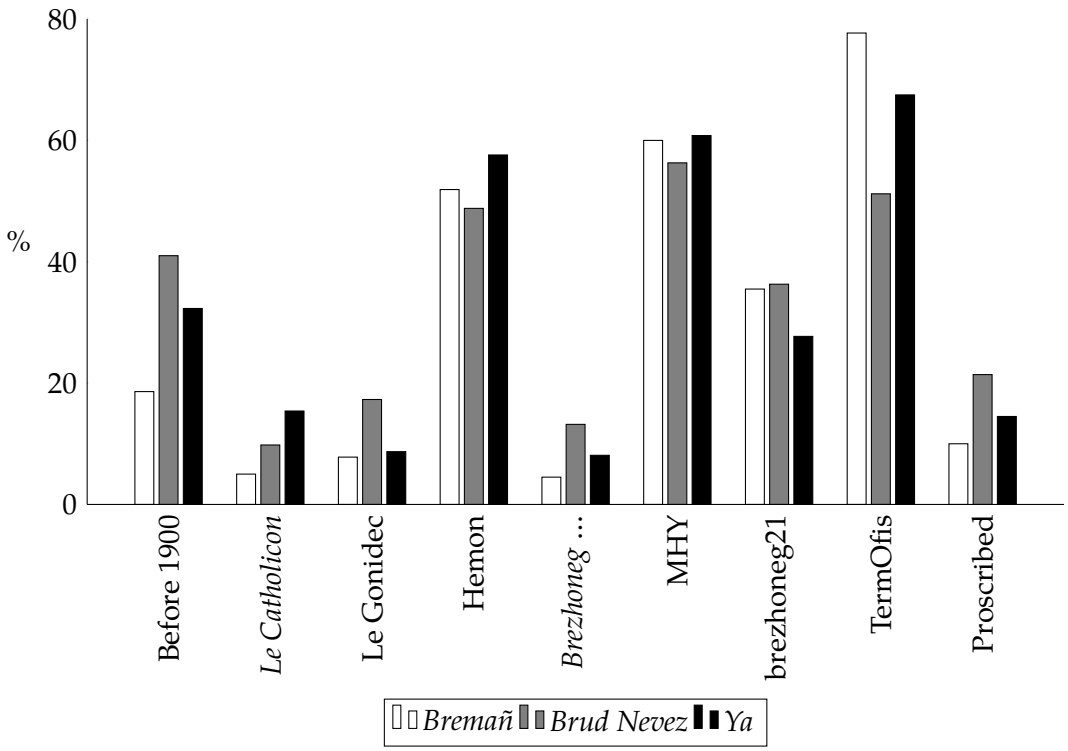


Figure 3.9: Graph of Table 3.12

Across the three sources, the number of tokens is statistically significant in most categories: all except the two twentieth-century dictionaries (Hemon and Mouladurioù Hor Yezh). *Bremañ* has a noticeably low rate of use of tokens of words found in Breton before 1900, suggesting it is more inclined towards the use of more recently created vocabulary. It also has the highest rate of use of tokens found in TermOfis, and, in numerical terms, of tokens in Brezhoneg21, suggesting that it makes the most reference to slightly more technical terminology, which again entails a higher rate of use of newer vocabulary. *Ya*, on the other hand, contains an unusually large number of tokens found in *Le Catholicon*, although again, this can be attributed to a high rate of occurrence (42 instances) of the word *yezh* in the sample, which contains multiple articles on the subject of language.

*Brud Nevez* was previously identified as different in its attitude towards word origins, and this again plays out in terms of dictionary attestation. In proportional terms, it includes a higher rate of use of words found in *Brezhoneg ... buan hag aes*, suggesting that it makes more frequent use of basic vocabulary; this is backed up by its noticeably lower

rate of use, in absolute terms, of tokens found in TermOfis and Brezhoneg21. The trend towards lower use of contested vocabulary in *Brud Nevez* observed earlier is reflected in this less frequent use of technical terms. *Brud Nevez* also contains a higher percentage of tokens proscribed by TermOfis, indicating that it is less likely to use standard Breton than the other printed sources. Again, this contributes to the impression that *Brud Nevez* does not subscribe to typical "néo-breton" ideologies; despite its high register, its language is more non-standard,and perhaps more accessible through its comparative avoidance of specialised and contested vocabulary.

Before moving on from this discussion of neologisms, it is worth briefly discussing some further statistics relating to word origins and proscription. Tabulating dictionary attestations with word origins shows that in the category of words proscribed by TermOfis, French-derived words are greater in number, making up 58.8% of the sample. This suggests that those who regulate Breton today continue to prioritise Celtic- over French-derived words in the majority of cases. This may be at odds with speaker practices involving the creation of new vocabulary: 78.3% of the tagged tokens in the sample with French-derived elements are not found in Breton prior to 1900, indicating that many of these contested items of vocabulary with French-derived elements are more recently created. Both Breton and French appear to have been extensively used as sources for coining new vocabulary since the beginning of the revitalisation period, despite attempts to prioritise those of Breton origin.

### 3.3.3 Semantic expansion

The second set of tags relating to dictionaries, marking tokens used with meanings different from those found in each source, can be used to measure semantic expansion. Specific cases of this will be considered as part of the more detailed analysis in the next chapter; this section deals with general statistical trends.

Semantic expansion is a cross-linguistically common way (McMahon 1994:174) of increasing the function of a language's lexicon without requiring the creation of new lexical forms. It can thus function as a way of circumventing the issue of whether to coin new words or borrow from other languages, which in the case of Breton can avoid

alignment with a "new" or "traditional" position. Despite this, semantic expansion has nonetheless been stigmatised by opponents of "néo-breton". One apparent case is *yezh* ('language'), which is the most frequent tagged token in the corpus, occurring 124 times. Given the frequency with which linguistic topics are discussed by the members of *Facebook e brezhoneg*, this word owes its high rate of occurrence to an abundance of tokens in the Facebook subcorpus. It is included among the tagged tokens because of its citation as a "néo-breton" word by German (2007:186), who asserts that "in varieties where it is actually used, *yezh* means the sound made by running water", and that traditional speakers would instead use *langaj* (< French *langage*), indicating that *yezh* is a lexeme that has undergone semantic expansion. However, *yezh* is found with the meaning 'language' as early as the fifteenth-century *Catholicon*, as well as in other texts from before 1900. While it may indeed have fallen out of use in less formal linguistic registers of the kinds used by German's informants, it cannot be claimed that use of *yezh* with the meaning 'language' is a practice that originated with "néo-breton" speakers. However, the fact that German insinuates that this is the case suggests that it may be generally believed among speakers that *yezh* is a "néo-breton" word, and hence that use of *yezh* aligns a speaker with a "néo-breton" attitude, regardless of the word's history. Speakers who wish to distance themselves from the "néo-breton" stereotype may therefore avoid using *yezh*. It is because of the possibility of this perception, as evidenced in the literature, that *yezh* was included among the tagged tokens. However, it does not contribute to the following statistics on semantic expansion, as the tags for semantic expansion were based on the occurrence of a word in the dictionaries and other sources examined. As *yezh* does not occur in any of these sources without the meaning 'language', it was not tagged as a word affected by semantic expansion.

Tokens used with meanings that differed from the meanings attested for them in pre-1900 sources or in any of the dictionaries or terminology databases examined were counted as cases of semantic expansion. There are 901 such tokens in the corpus. 84.2% of these are Celtic-derived (i.e. from Breton or Welsh). Semantic expansion thus appears to be a process that tends to be applied to Celtic-derived words, indicating that these words are favoured for the expansion of the lexicon, fitting the supposed preference in "néo-breton" for the use of words of Celtic origin. It therefore perhaps func-

tions as a way of favouring Celtic-derived words while not going so far as to require a large amount of neologism creation, which has been viewed as excessively purist. This allows speakers to avoid hastening the spread of French-derived vocabulary (either through borrowing or through semantic expansion, which, as we have seen, is far less common for French-derived words), while also decreasing the chance of being criticised for a purist outlook—although we have seen in the case of *yezh* that such criticism may still occur.

We can visualise the distribution of tokens affected by semantic expansion in the usual way:

| | | **# tokens (normalised for the subcorpora)** | | | | **% of tagged tokens in the subcorpus** | | | |
|---|---|---|---|---|---|---|---|---|---|
| | **Word used with different meaning from attestation in:** | **Total** | **Radio** | **Facebook** | **Print** | **Total** | **Radio** | **Facebook** | **Print** |
| * | Breton before 1900 | 401 | 124.8 | 125.5 | 134.6 | 10.3% | 9.7% | 10.1% | 11.0% |
| * | *Le Catholicon* (1499) | 148 | 58.8 | 51.4 | 33.2 | 3.8% | 4.6% | 4.1% | 2.7% |
| * | Le Gonidec/ Villemarqué (1850) | 398 | 90.3 | 173.1 | 117.0 | 10.2% | 7.0% | 13.9% | 9.6% |
| * | Hemon (1964) | 222 | 97.4 | 70.4 | 47.9 | 5.7% | 7.6% | 5.7% | 3.9% |
| * | *Brezhoneg* ... (1972) | 34 | 27.4 | 5.7 | 0.9 | 0.9% | 2.1% | 0.5% | 0.1% |
| * | Mouladurioù Hor Yezh (1994) | 54 | 10.1 | 24.7 | 16.6 | 1.4% | 0.8% | 2.0% | 1.4% |
| | brezhoneg21 (online) | 225 | 81.2 | 60.9 | 74.6 | 5.8% | 6.3% | 4.9% | 6.1% |
| * | TermOfis (online) | 72 | 33.5 | 23.8 | 12.9 | 1.8% | 2.6% | 1.9% | 1.1% |
| | **Total** | **901** | **271.9** | **311.9** | **281.1** | **23.1%** | **21.1%** | **25.1%** | **23.0%** |

Table 3.13: Semantic expansion among the tagged tokens

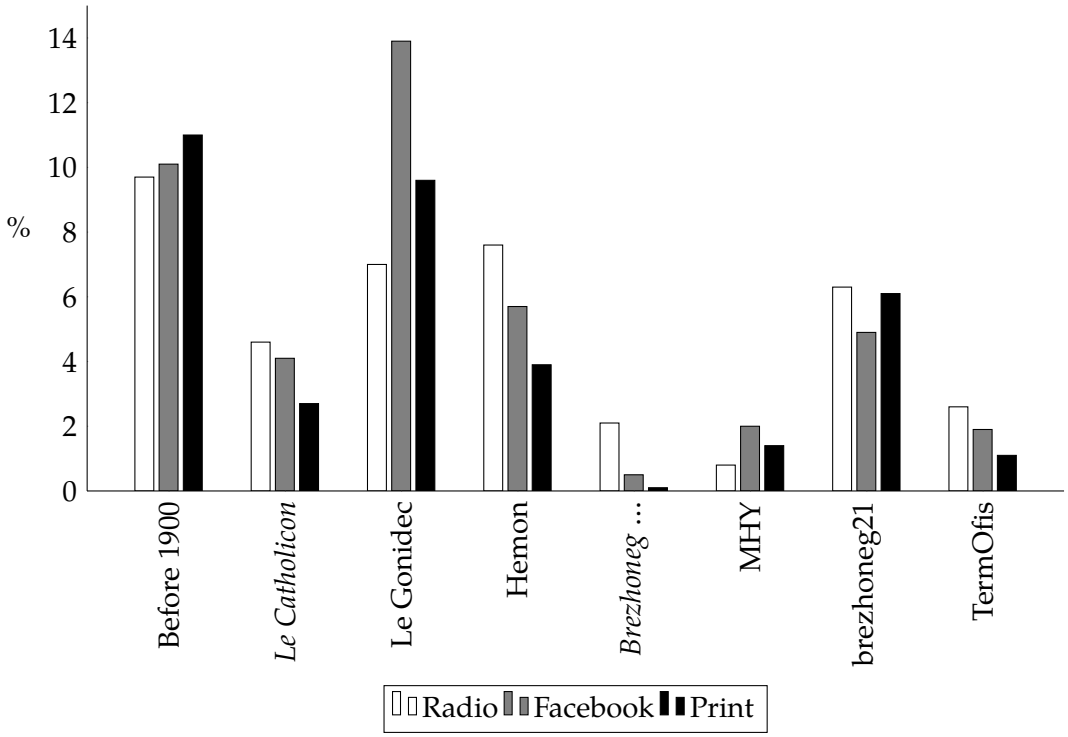


Figure 3.10: Graph of Table 3.13 (comparison of the subcorpora)

It may appear from the fact that 23.1% of tagged tokens involve some kind of semantic difference from the attestations and definitions examined that semantic expansion is extremely common in the corpus, affecting nearly a quarter of all tagged tokens. However, this apparently high rate reflects the fact that dictionaries often include only a small selection of possible definitions: *Brezhoneg ... buan hag aes*, for example, being a beginners' textbook, may provide only one of multiple glosses for a word if this one meaning is the only one that is required for the purposes of the example dialogues in the textbook; brezhoneg21 is a scientific database, and so may not include non-scientific glosses. It is more meaningful on this broad level to concentrate on the figures in the first row: that is, those tokens with a different meaning in the corpus from those found for the same word before 1900. These words are therefore those where semantic expansion appears to have occurred if the use of the word in contemporary Breton (i.e. in the corpus) is compared with the meaning(s) attested for it before the revitalisation of the language began. 10.3% of tokens, then, fit this criterion, showing that semantic expansion still appears to be fairly widespread, if not quite as prevalent as the earlier figure implied. This occurs at a fairly constant rate among the three subcorpora, and indeed

there is no statistically significant variation to be observed when the corpus is divided further in the usual ways, i.e. by subcorpus, radio station type, radio programme type, and publication.

When the figures for individual dictionaries are considered, however, statistically significant variation can in fact be observed in all cases across the three subcorpora, with the exception of brezhoneg21. This follows different patterns depending on the dictionary. On the radio, for example, there is a higher rate of semantic difference from *Brezhoneg ... buan hag aes*, but this can be accounted for with reference to the word *setu*. This makes up the vast majority of the tokens with this tag overall: 29 out of the total of 34. In *Brezhoneg ...*, *setu* is glossed as 'here is'; in spoken Breton, it has come to be used as a discourse marker, similar to French *voilà*, as in this excerpt from the radio subcorpus:

> *kregiñ a reoc'h setu goût a rez pet litrad hoc'h graet ar bloaz-mañ da skouer*
> 'do you start *setu* do you already know how many litres you have made this year for example'

In this excerpt, *setu* marks the point at which the interviewer decides to restart his conversational turn, effectively marking *kregin a reoc'h* as a false start and *goût a rez* as the true beginning of the sentence. It therefore does not correspond with the way it is used in *Brezhoneg ... buan hag aes*, where it is introduced in the first unit—in fact, the book's very first example sentence—as part of the example *Setu un den ha setu un ti* ('Here is a man and here is a house'). The use of *setu* in the way seen in the excerpt above is particularly associated with oral styles, given that it is a discourse marker, and hence occurs disproportionately in the radio subcorpus. This skews the observed figures for *Brezhoneg ... buan hag aes* as a whole.

Elsewhere, the significant variation among the subcorpora again indicates that different attitudes towards the use of certain types of language may be present. TermOfis and Hemon's dictionary, both flagship sources of standard "néo-breton", albeit from different periods of the language's revitalisation, follow the same pattern in that the radio data show the most cases of semantic difference from their glosses and the print

subcorpus conforms to them the most closely. This echoes the general findings from the previous chapter, where, in overall terms, the radio was in many ways found to be the furthest from stereotypical "néo-breton" and the printed sources the closest to it.

When rates of semantic expansion in the different subcorpora compared with Le Gonidec and Villemarqué's dictionary and the Mouladurioù Hor Yezh dictionary are investigated, a different pattern can be found, linking these sources despite their having been published over a century apart. This time, the radio data are now the closest followers of these dictionaries, while the Facebook subcorpus deviates the most from their glosses. Rather than being seen as purveyors of highly standardised Breton in the same way as Hemon's dictionary and TermOfis, these two publications can be considered more neutral. They represent a fairly conservative type of Breton—we saw above that Le Gonidec and Villemarqué's dictionary already had some purist tendencies despite being published in the mid-nineteenth century—but one that certainly has a more relaxed attitude towards borrowings than that advocated by the early "néo-breton" movement. This is perhaps a better fit for the variety used on the radio, which, as we have already seen, tends to include more borrowed vocabulary than the other sources, and hence the lower rate of semantic distance between the radio subcorpus and the lexicon in these dictionaries can be explained. The fact that the Facebook data are the least conformant to this more neutral variety again suggests that this subcorpus displays a particular type of nonstandardness that nonetheless involves a fairly high level of "néo-breton" purism.

Statistical significance can also be found in the case of *Le Catholicon*, where the print subcorpus is less semantically distant from its lexicon than the other two subcorpora. However, in this case, it is difficult to establish a reason for this, and it seems reasonable to conclude that this difference is merely an effect of which lexemes happen to occur in the different subcorpora. Breaking the corpus down into its constituent parts, and ignoring the discrepancies in the data caused by the discursive use of *setu* (see above), the only division that constantly produces statistically significant results is that of the print corpus into the three publications, tabulated below.

| | | # tokens (normalised) | | | % of tagged tokens in the sample) | | |
|---|---|---|---|---|---|---|---|
| | **Word used with different meaning from attestation in:** | ***Bremañ*** | ***Brud Nevez*** | ***Ya*** | ***Bremañ*** | ***Brud Nevez*** | ***Ya*** |
| | Breton before 1900 | 47.1 | 28.8 | 58.6 | 9.7% | 10.5% | 12.7% |
| * | *Le Catholicon* (1499) | 6.3 | 13.0 | 14.0 | 1.3% | 4.7% | 3.0% |
| * | Le Gonidec/ Villemarqué (1850) | 36.2 | 20.4 | 60.5 | 7.4% | 7.5% | 13.1% |
| | Hemon (1964) | 11.8 | 10.2 | 26.1 | 2.4% | 3.7% | 5.7% |
| | *Brezhoneg* ... (1972) | 0.0 | 0.0 | 0.0 | 0 | 0 | 0 |
| | Mouladurioù Hor Yezh (1994) | 4.5 | 28.8 | 9.3 | 0.9% | 1.0% | 2.0% |
| | brezhoneg21 (online) | 21.7 | 19.5 | 33.5 | 4.5% | 7.1% | 7.3% |
| * | TermOfis (online) | 0.9 | 4.6 | 7.4 | 0.2% | 1.7% | 1.6% |
| * | **Total** | **86.0** | **55.8** | **139.6** | **17.7%** | **20.3%** | **30.3%** |

Table 3.14: Semantic expansion in the print subcorpus

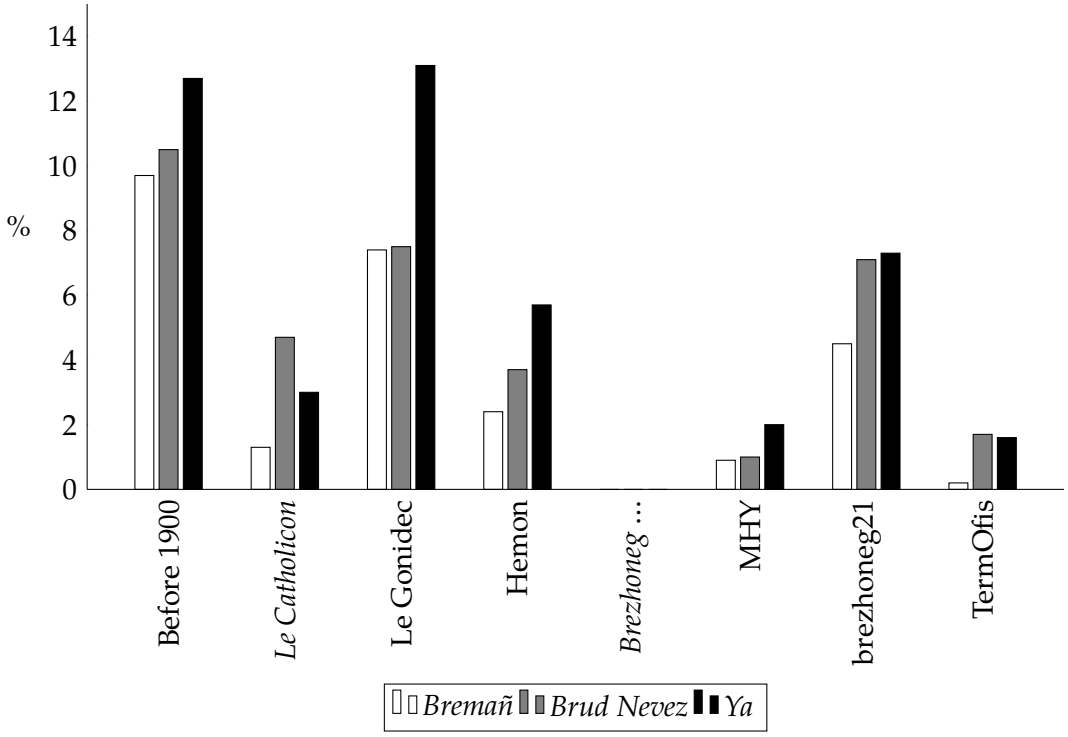


Figure 3.11: Graph of Table 3.14

In this case, statistical significance can be observed for *Le Catholicon*, Le Gonidec and Villemarqué's dictionary, Hemon's dictionary, and TermOfis, as well as for the overall figure. With the exception of *Le Catholicon* and TermOfis, this represents a trend (also seen for the other sources, despite the lack of statistical significance) of greatest conformance for *Bremañ*, with the highest level of deviance from glosses and attestations occurring in *Ya*, while *Brud Nevez* occupies an intermediate position. This is an interesting finding when it is considered that the statistics in the previous chapter often grouped *Bremañ* and *Ya* together as vehicles of a more stereotypical "néo-breton", making *Brud Nevez* the outlier; here, in terms of the overall figures, it is *Ya* that is the outlier, with a much higher rate of semantic expansion than the other two publications. *Ya* therefore appears to use "néo-breton" more innovatively than *Bremañ*, making use of vocabulary in less well-established ways. This may also correspond with *Ya*'s greater tolerance for international borrowings, indicating that in general terms, it more readily borrows or adapts existing vocabulary, from Breton or from other languages, while *Bremañ* is more likely to prefer new coinages. This adds nuance to the previous section's preliminary conclusion that *Ya* may be the most stereotypically "néo-breton" publication: it now appears instead that both *Bremañ* and *Ya* conform to "néo-breton" trends, but in different ways.

### 3.3.4 Summary

In general, the trends identified in section 3.2 were found to continue with reference to dictionary attestations and semantic expansion. Section 3.3.2 showed that going by the measure of dictionary attestations, the print subcorpus again fits best with typical "néo-breton" in its stronger correlation with Hemon's dictionary than in the case of the other two subcorpora. Conversely, the radio data are least correlated with attestation in Hemon's dictionary and can thus be said, again, to be the furthest from the typical "néo-breton" current. Unsurprisingly, the print subcorpus also contains the smallest number of tokens proscribed by TermOfis. It is the Facebook data, however, that show the most frequent use of such tokens, showing that while Facebook users appear in general to use a more purist form of Breton, they are also unconstrained by the prescriptive limits of TermOfis, using language in a less standard manner.

Among the print media, *Bremañ* tends to use vocabulary that is not present in Breton before 1900, indicating a preference for recent coinages and more technical vocabulary, while *Brud Nevez* on the other hand uses more basic vocabulary (shown by the fact that more of the tagged tokens in *Brud Nevez* are found in *Brezhoneg ... buan hag aes*). This publication also uses more words proscribed by TermOfis, indicating, as in the previous section, that it conforms less with standardised "néo-breton". In the case of semantic expansion, covered in section 3.3.3, various observations again lend weight to the trends already pointed out. The radio subcorpus has again been shown to avoid excessive purism, while the Facebook data do conform to this trait, but remain noticeably non-standard. Examining semantic expansion has added depth to the conclusions previously drawn about *Bremañ* and *Ya*, indicating that while both can be considered to conform to the characteristics of "néo-breton" in general, the latter does so in a way that draws on a wider range of sources, including internationally used borrowings and more frequent use of semantic expansion rather than outright coinages.

## 3.4 Dialectal features

While the analysis undertaken so far has relied on the association of higher proportions of Breton-derived vocabulary with a closer alignment with typical "néo-breton", this requires fine-tuning given that one of the beliefs around "néo-bretonnants" is that they subscribe to standard language ideology. Any dialectal features, therefore, most of which derive from Breton, can be counted as features that avoid use of the standard, thus contradicting this ideology and going against typical "néo-breton". This section will thus examine the occurrence of dialectal features. These were identified using dialect atlases, signposting in modern dictionaries or textbooks, and attestation in dialect-specific dictionaries (many of which are indexed in Devri).

Five tags were used to mark different dialects. As stated in section 1.3, Breton is typically divided into four dialects, but scholars have also suggested that the central area of Brittany can also be considered a specific dialect area. Tags were therefore used for Léonais, Trégorrois, Cornouaillais, Vannetais and central Breton. Out of these, Vannetais is linguistically the most distinct, and has the ideological backing of a longstand-

ing written tradition "promoted and strengthened" at the turn of the twentieth century in the creation of a distinct orthography (Wmffre 2008:22). For these reasons, we might expect Vannetais to be especially prominent among the dialects used in the corpus. Léonais might also be expected to be prevalent, given that out of the dialects it was the most drawn upon as a source for standard literary Breton.

Some of the dialectal words attested in the corpus could have belonged to a range of dialects, and were tagged accordingly. Out of the 72 types tagged as dialectal, four could have marked more than one dialect: *dijoloiñ* ('to move') can be used in Vannetais or central Breton, *endan* ('inside') and *kerzhet* ('to walk') in Vannetais or Trégorrois (i.e. eastern Basse-Bretagne), and *bale* ('to walk') in Léonais or Cornouaillais (i.e. the westernmost area).[87] *Bale* is additionally a notable case in that it has become incorporated into standard Breton, but was nonetheless tagged as dialectal in opposition with the synonyms *pourmen* and *kerzhet*, used in other dialects, which also occur in the corpus (see section 4.3.5 for further discussion of these words). Léonais words such as this one, therefore, may mark an allegiance with standard "néo-breton" rather than with dialectal Léonais, and the results will show further examples of this.

The distribution of apparently dialectal words was as follows:

| **Dialect** | **Tokens** | **Types** |
|---|---|---|
| Vannetais | 189 | 49 |
| Léonais | 54 | 9 |
| Trégorrois | 15 | 8 |
| Central Breton | 17 | 6 |
| Cornouaillais | 25 | 2 |
| **Total** | **275** | **70** |

Table 3.15: Dialect occurrence among tagged tokens

It can be noted from this that Vannetais is indeed the most prevalent of the dialects,

87. This is why the total figures in Table 3.15 are lower than the sums of the previous rows.

with the largest number of tokens as well as a number of types noticeably larger than for the others; Léonais is, again as suspected, the next most prevalent, certainly in terms of the number of tokens. The remaining dialectal tokens appear to be distributed fairly evenly, although 21 of the 25 tokens in the case of Cornouaillais are of *bale*, which is unlikely to indicate a dialectal leaning as much as a "néo-breton" one. The situation of the dialects therefore appears somewhat polarised with respect to the lexicon, with the only significant degree of dialectal variation in the contexts under investigation involving Vannetais. Out of the nine Léonais words that occur, four have become features of standard Breton, reducing the number of Léonais-specific tokens to twelve and bringing it into line with the other dialects (apart from Vannetais) in terms of how often it is found.

### 3.4.1 Dialectal lexicon across the corpus

The following table shows the distribution of dialectal words across the three subcorpora.

| | | **# tokens (normalised)** | | | **% of dialectal tokens in the subcorpus** | | |
|---|---|---|---|---|---|---|---|
| | **Dialect** | **Radio** | **Facebook** | **Print** | **Radio** | **Facebook** | **Print** |
| * | **Vannetais** | 100.4 | 42.8 | 41.5 | 78.0% | 63.4% | 58.4% |
| * | **Léonais** | 8.1 | 17.1 | 25.8 | 6.3% | 25.4% | 36.4% |
| * | **Trégorrois** | 5.1 | 7.6 | 1.8 | 3.9% | 11.3% | 2.6% |
| * | **Central Breton** | 15.2 | 1.0 | 0.9 | 11.8% | 1.4% | 1.3% |
| * | **Cornouaillais** | 5.1 | 1.9 | 16.6 | 3.9% | 2.8% | 23.4% |
| * | **Total** | **133.9** | **70.4** | **86.6** | | | |

Table 3.16: Dialect by subcorpus

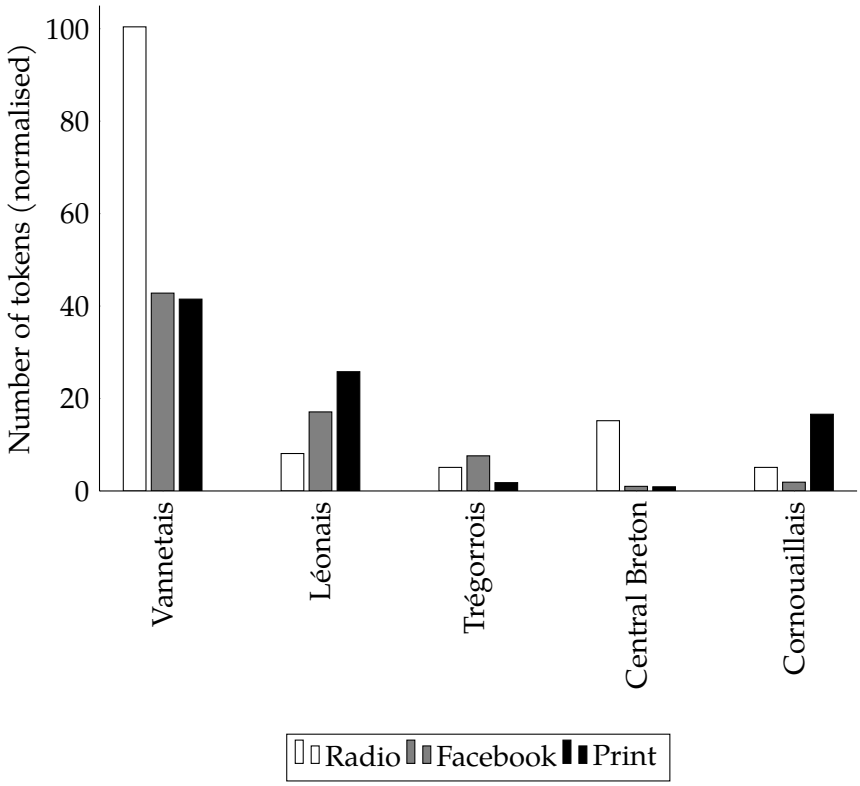


Figure 3.12: Graph of Table 3.16

There is clear variation among the subcorpora, with Vannetais and central Breton used far more frequently on the radio than elsewhere. It also seems conversely that Léonais and Cornouaillais are more present in print. However, again, examining the actual tokens found reveals that the figures for Léonais and Cornouaillais are mostly contributed to by those that could be associated with standard Breton. Recalling the fact that only four of the Cornouaillais tokens are of a word that is associated exclusively with Cornouaillais (*dibaoe*—'since'), the distribution of these tokens would perhaps be a better indicator of where Cornouaillais is found. *Dibaoe* is in fact found twice in the radio data and twice in *Brud Nevez*. Its occurrence on the radio correlates with the higher rate of use of words from other dialects in this medium, lending support to the supposition that the radio is more frequently a locus for the use of dialectal features. The use in *Brud Nevez*, too, tallies with the earlier findings that this magazine appears to defy the trend towards standard "néo-breton" found in the other print sources examined. In general, then, the majority of dialectal variation can be observed to occur on the radio.

The radio subcorpus can be split, as above.

| | | # tokens (normalised) | | % of dialectal tokens in the section | |
|---|---|---|---|---|---|
| | **Dialect** | **Public** | **Associative** | **Public** | **Associative** |
| * | **Vannetais** | 34.0 | 65.1 | 62.7% | 88.2% |
| * | **Léonais** | 8.5 | 0.0 | 15.7% | 0 |
| | **Trégorrois** | 3.2 | 1.9 | 5.9% | 2.6% |
| | **Central Breton** | 8.5 | 6.8 | 15.7% | 9.2% |
| * | **Cornouaillais** | 5.3 | 0.0 | 9.8% | 0 |
| * | **Total** | **59.5** | **73.8** | | |

Table 3.17: Dialect in the radio subcorpus by station type

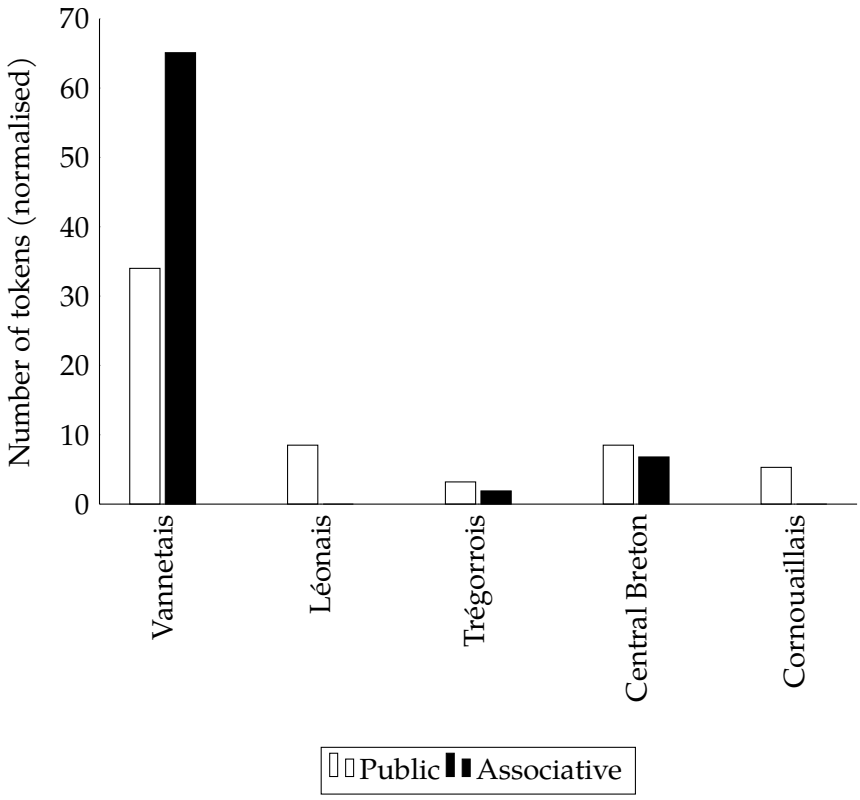


Figure 3.13: Graph of Table 3.17

Again, Léonais and Cornouaillais appear to skew in the same direction, this time towards public rather than associative radio. This may indicate that public radio is more likely to use forms associated with the standard language. Both radio attestations of

*dibaoe* occur on public radio, however, and Trégorrois and central Breton forms also occur more frequently on public radio, suggesting instead that there is no dialectal difference between the two radio types. For Vannetais, on the other hand, where there is a much larger sample size, there is a significantly higher rate of occurrence on associative radio. The details of this distribution will be examined in more depth below.

These data can be compared with those obtained when the radio subcorpus is split by programme type.

| | | **# tokens (normalised)** | | **% of dialectal tokens in the section** | |
|---|---|---|---|---|---|
| | **Dialect** | **News** | **Features** | **News** | **Features** |
| | **Vannetais** | 79.0 | 22.8 | 83.5% | 63.9% |
| | **Léonais** | 4.2 | 4.0 | 4.4% | 11.1% |
| | **Trégorrois** | 3.1 | 2.0 | 3.3% | 5.6% |
| | **Central Breton** | 10.4 | 5.0 | 11.0% | 13.9% |
| * | **Cornouaillais** | 0.0 | 5.0 | 0 | 13.9% |
| * | **Total** | **96.7** | **38.8** | | |

Table 3.18: Dialect in the radio subcorpus by programme type

There is less variation to observe between the two programme types than between the two radio stations. Vannetais is noticeably more represented in news programmes than feature programmes, but this is not statistically significant as a proportion of dialectal tokens. The only specific dialect for which statistical significance can be observed is Cornouaillais: as noted above, this is likely to mark allegiance to Léonais or standard Breton rather than to the Cornouaillais dialect. Again, the two radio attestations of *dibaoe* occur in a feature programme: as they are both produced by the same speaker, it is difficult to make any generalisations about the two contexts based on this. Like the previous split, therefore, programme type appears to have little effect on dialect use. Overall, dialectal tokens appear far more frequently in news programmes, and this is due to a high rate of Vannetais tokens from one speaker: this will be discussed below.

Finally, the print sources can again be split.

| | | # tokens (normalised) | | | % of dialectal tokens in the subcorpus | | |
|---|---|---|---|---|---|---|---|
| | **Dialect** | ***Bremañ*** | ***Brud Nevez*** | ***Ya*** | ***Bremañ*** | ***Brud Nevez*** | ***Ya*** |
| * | **Vannetais** | 7.2 | 5.6 | 29.8 | 88.9% | 19.4% | 83.8% |
| * | **Léonais** | 0.0 | 20.4 | 5.6 | 0 | 71.0% | 16.2% |
| | **Trégorrois** | 0.9 | 0.9 | 0.0 | 11.1% | 3.2% | 0 |
| | **Central Breton** | 0.0 | 0.9 | 0.0 | 0 | 3.2% | 0 |
| * | **Cornouaillais** | 0.0 | 13.9 | 2.8 | 0 | 48.4% | 8.1% |
| * | **Total** | **8.1** | **41.7** | **38.2** | | | |

Table 3.19: Dialect in the print subcorpus

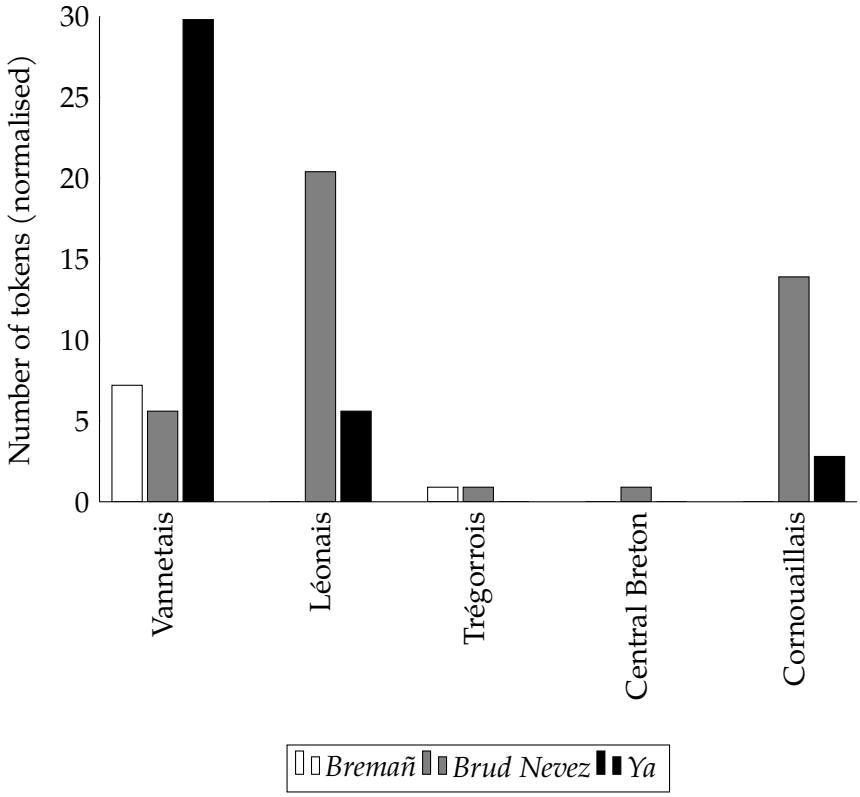


Figure 3.14: Graph of Table 3.19

This appears to conflict with the apparent trends observed in the previous sections. While the evidence found so far indicates that *Brud Nevez* was the least standard of the

three publications, here it is *Ya* that uses the highest number of indisputably dialectal tokens, again coming mostly from Vannetais. This suggests that *Ya* may have a more inclusive approach to non-standard varieties than *Bremañ* does.

Going by earlier evidence, we may expect *Brud Nevez* to show an equally favourable attitude towards dialect: in the event, it shows unusually high proportions of Léonais and Cornouaillais tokens compared with the other publications. As above, this could indicate a tendency towards either these dialects or standard literary Breton. It is worth noting that a number of the articles in the *Brud Nevez* sample cover locations around Brest—that is, in the Léonais area. The use of more Léonais in this publication therefore may indeed indicate that the writers are Léonais users. However, the position of *Brud Nevez* as a literary magazine could also play a role here, leading writers to use more literary language, which is typically drawn from Léonais. The only publication showing indisputable tolerance of dialectal vocabulary is *Ya*, again suggesting that this source may favour a slightly different type of "néo-breton" from that attested in previous studies.

In general, therefore, dialectal features are observed most often on the radio, and for the most part this is scarcely affected by the type of radio station or the type of programme. The print sources vary in a different way from what was previously observed, with *Ya* appearing to be the most likely to tolerate dialectal forms, but still remaining less frequent a locus of dialect use than the radio.

### 3.4.2 Individual speakers and dialect use

Using individual dialectal features is not in itself sufficient to imply that a speaker is proficient in any specific dialect: the features used must be consistent with a single dialect area rather than drawn from multiple locations. "Néo-bretonnants", supposedly subscribing to anti-French and pro-Celtic ideologies, are said to hypercorrect their syntax to emphasise Breton's grammatical distance from French (Hewitt 2016:16); it is equally possible that speakers of Breton who similarly wish to emphasise the Celtic nature of the language, but additionally take an ideological position that favours traditional linguistic features, could use features drawn from multiple dialects rather than

from a single geographical area. This, however, would mark such speakers out as distinct from traditional speakers and, like the avoidance of dialectal features altogether, seem unacceptable to "native authenticists" (see section 1.4.2), for whom tradition and support for the dialects would constitute a more important ideological position than the need to demonstrate Breton's distance from French. With this in mind, the data were examined broken down to the level of individual speakers.

In the radio and Facebook sources, 42 speakers (out of a total of 139) used dialectal vocabulary of some kind: nineteen of these speakers were in the radio subcorpus and 23 on Facebook. It is harder to tell how many different voices are present in the print sources, due to the presence of anonymous and editorial texts: seventeen sections use dialectal vocabulary in the print sources, out of a total of 39 sections. It is unclear whether there are as many as seventeen different speakers represented here, although the number of users of dialectal features in the printed sources would be no fewer than eleven.

Most contributors make use of only very small numbers of dialectal tokens, with the median number used being three, despite some of these speakers contributing large amounts of speech or long passages to the corpus. Three of the writers in the print sources use over ten dialectal tokens, the highest count being 26; none do so on Facebook, probably a result of the typically shorter post and comment length; and three speakers on the radio use more than ten tokens, the highest using 40 tokens, all from Vannetais. These speakers evidently have a large effect on the statistics examined above, although it can be noted that even when the most prolific user of Vannetais is discounted, this remains the most frequent of the dialects.

Seventeen of the speakers use words from multiple dialects in a way that could give a conflicting impression of their own dialectal allegiance, apparently supporting the belief that "néo-breton" speakers use features drawn from multiple dialects in an attempt to convey authenticity. However, in many cases one of the tokens in question is a Léonais word that has become part of standard Breton, meaning that the speaker is not mixing two dialects but rather using standard Breton with some dialectal features from a particular dialect (but perhaps resulting in speech just as objectionable to the

"native authenticist"). One Facebook user, for example, uses Vannetais *mes* ('but'), and *parlant* ('dialect'), but also Léonais *va* ('my'). However, *va* is effectively integrated into standard Breton through its use in a grammar book by the influential Roparz Hemon (Hemon 1984), despite the fact that pan-dialectal *ma*, found in most textbooks, would be expected. Where some speakers may see a contradiction here, the sanctioning of *va* in a well-known grammar book contributes to the eradication of such a contradiction where new speakers relying on such textbooks are concerned.

Beyond such cases, however, some speakers do indeed appear to use words drawn from multiple dialects that are not part of the standard. One Facebook user uses Léonais *edo* (an imperfect form of the verb 'to be') but also Vannetais *aveit* ('for') and *endan* ('under'), and uses *diouzh* rather than *ouzh* in the phrase *selaou diouzh*, 'listen to', a feature of central Breton. Use of words from two neighbouring dialects could indicate that the speaker is from a border area, but this explanation evidently cannot be used in this case, as the Léonais and Vannetais areas are not contiguous (see Appendix B). This does therefore appear to be a case where dialect words are not used to indicate allegiance with one particular dialect, but rather perhaps to show alignment with this less typically "néo-breton" practice in general. However, this Facebook user is nonetheless unable to give the impression of being a traditional speaker.

It has been noted that Vannetais is the most prevalent dialect in the corpus, and that this is likely to be influenced by its linguistic distance from the other dialects, which has allowed it to maintain a more distinct tradition. This linguistic distance does however raise the question of how comprehensible the use of dialect vocabulary is to speakers who are unfamiliar with the dialect in question. As all the data in the corpus are from media intended for consumption by the whole Breton-speaking population, not restricted to particular dialect areas, dialect use has to be balanced with the need to be understood by speakers who know little of the dialect's specificities. This theme will recur in this section, as well as in chapter 5.

As a result, if speakers are to employ strategies that signal their allegiance to dialectal Breton, they must ensure that their output remains comprehensible. An effect of this can be seen in how the dialectal tokens identified fall into part of speech categories, as

the following section will discuss.

### 3.4.3 Dialectal tokens and parts of speech

Section 3.2.2 already noted the high frequency of dialectal tokens among prepositions, and this is mirrored when the figures are examined from the opposite angle. The following table shows the dialectal tagged tokens by their part of speech, visualised in the graph.

| **Part of speech** | **Number of dialectal tokens** |
|---|---|
| Prepositions | 86 |
| Verbs | 68 |
| Nouns | 49 |
| Determiners | 32 |
| Adverbs | 25 |
| Adjectives | 7 |
| Particles | 4 |
| Interjections | 2 |
| Conjunctions | 2 |
| **Total** | **275** |

Table 3.20: Part of speech of dialectal tagged tokens

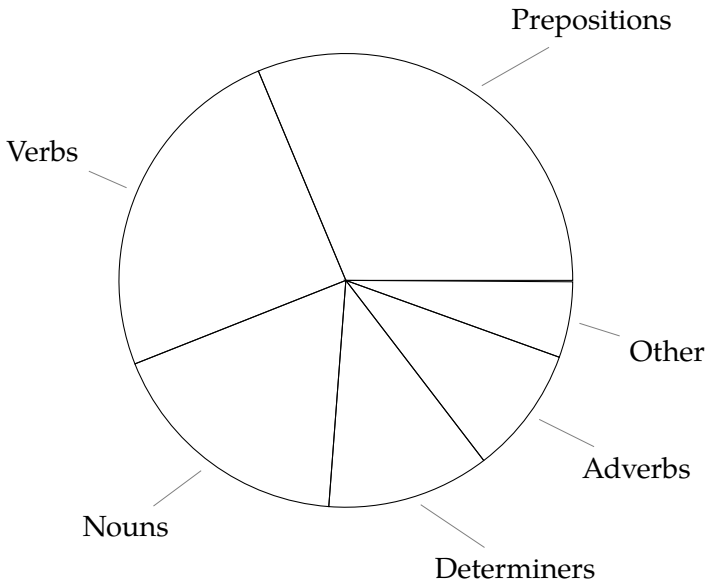


Figure 3.15: Graph of Table 3.20

In the overall figures for tagged tokens, open-class words are most prevalent, with nouns by far the most frequently tagged part of speech at 58.0% of all tagged tokens, followed by verbs, as section 3.2.2 showed. However, among dialectal tokens, prepositions are the most common, and determiners are also notably more prevalent; closed-class words thus seem more frequent in general.

The distribution of part of speech types can be compared with that for the tagged tokens overall, and with the corpus as a whole.

| | Number of each POS as a percentage of whole sample | | |
|---|---|---|---|
| **Part of speech** | **Tokens tagged as dialectal** | **Tagged tokens** | **Whole corpus (estimated from sample)**[88] |
| Nouns | 17.8% | 58.0% | 21.2% |
| Verbs | 24.7% | 14.4% | 19.9% |
| Determiners/pronouns | 11.6% | 2.4% | 15.6% |
| Prepositions | 31.3% | 2.4% | 13.1% |
| Verbal particles | 1.5% | 0.1% | 6.6% |
| Adjectives | 2.5% | 11.2% | 6.0% |
| Adverbs | 9.1% | 7.0% | 5.6% |
| Conjunctions | 0.7% | 0.1% | 5.3% |
| Discourse markers | 0.0% | 3.9% | 0.6% |
| Interjections | 0.7% | 0.4% | 0.2% |

Table 3.21: Part of speech among dialectal tokens, all tagged tokens, and the whole corpus (estimated)

The increased proportion of prepositions tagged as dialectal can therefore be seen not to reflect the distribution of tagged tokens, or of words in the corpus as a whole. Among tagged tokens, the proportions of prepositions, determiners, pronouns, conjunctions, and verbal particles are notably lower than in the whole corpus; those that *are* tagged tend to be dialectal, as these parts of speech are of course less subject to the need for borrowing or neologism. Nouns, on the other hand, are particularly highly affected, as previously discussed. The low proportion of nouns tagged as dialectal is therefore even more striking: this will be discussed below.

Most dialectal word types occurred only a very small number of times, with the median frequency being 2. Seven words occurred ten or more times, and their distribution was as follows:

88. This is based on a sample taken from each of the three subcorpora, with a total word count of 853. The percentages do not sum to 100%, as personal, brand and place names, being excluded from the analysis, were not included.

| | **Total** | **Radio (normalised)** | **Facebook (normalised)** | **Print (normalised)** |
|---|---|---|---|---|
| *get* 'with' (Gw)[89] | **33** | 24.3 | 8.6 | |
| *bale* 'to walk' (L/standard) | **21** | 3.0 | 1.9 | 14.7 |
| *doc'h* 'from' (Gw) | **18** | 17.2 | 1.0 | |
| *-senn* 'that' (Gw) | **17** | 17.2 | | |
| *boud* 'to be' (Gw) | **16** | 8.1 | 1.0 | 6.5 |
| *va* 'my' (L/standard) | **12** | | 6.7 | 4.6 |
| *hiriv* 'today' (KB) | **10** | 9.1 | | 0.9 |

Table 3.22: Most frequent dialectal words

It can immediately be seen that the majority of these words are from Vannetais, with Léonais the next most prevalent. Of the two Léonais words, *bale* has indisputably become part of standard Breton, meaning it cannot be considered to reflect dialectal usage. *Va*, as discussed above, can also be considered part of standard Breton, and moreover part of the hyper-Celticised early twentieth-century variety promoted by stereotypical "néo-bretonnants". Its high rate of appearance among the Facebook data therefore suggests that this is a context in which users of the most stereotypically "néo-breton" variety might be found.

On the other hand, the Facebook data also include a fairly high rate of use of Vannetais *get* (standard/KLT[90] *gant*), indicating that a number of Facebook users wish to display alignment with Vannetais. *Get* is far more frequent, however, on the radio (as noted in section 3.2.2), along with *doc'h* and *-senn*, also from Vannetais. This contributes to and exemplifies the greater use of Vannetais, and of dialectal tokens overall, in the radio subcorpus. Many of these tokens come from a single radio presenter, who contributes 40 dialectal tokens, the largest number for a single speaker; all are from Vannetais. The

89. In this table, the following abbreviations are used: Gw = Vannetais; L = Léonais; KB = central Breton.

90. KLT collectively refers to the three dialects Cornouaillais, Léonais, and Trégorrois, and can be used to denote features of Breton used in varieties of Breton other than Vannetais.

second most prolific dialect user by this measure is, however, a writer in *Ya*, who produces 26 dialectal tokens, again, all from Vannetais, indicating that she is a consistent user of this dialect. Those who use the greatest number of dialect features therefore appear consistent in this regard. It must also be noted that in the radio subcorpus, despite the large number of dialectal tokens from a single speaker, this has no skewing effect on the overall data: even without taking these particular examples into consideration, the radio data still present the largest number of dialectal tokens by a significant margin.

Despite the fact that a number of writers in the print subcorpus use Vannetais features, some of those that are most common in the other sources, particularly on the radio, are entirely absent in this context: *get*, *doc'h* and *-senn* do not appear at all in the sample. This is likely to be because of their phonological proximity to the standard versions of these words: *gant*, *deus* and *-se* respectively. *Get* is typically pronounced /gət/ or /dʒət/, which is still recognisable from the orthographic form <gant>; /dɔx/ and /sən/ can likewise be considered fairly close to the orthographic representations of their standard equivalents. Writers or editors seem therefore to have chosen to use these standard versions in writing, despite their use of other dialectal features indicating that they would pronounce these words in a Vannetais way. On the other hand, as the Vannetais versions of the words do have conventional orthographic representations even within the most common supradialectal orthography (i.e. <get>, <doc'h> and <senn>), these spellings *were* used in transcribing the radio data, meaning that these particular Vannetais tokens appear more frequently in the written representation of this subcorpus than in the other two subcorpora. Conversely, this explains why *boud* is comparatively more frequent in the printed sources: in this case, the pronunciation /but/ is perhaps too different from the orthographic representation of the standard, <bezañ> (pronounced /beã/ or /bɛzã/), for this standard orthography to represent it.

Writers on Facebook, unlike those in the print subcorpus, seem not to have chosen to represent their Vannetais pronunciation with standard orthography, or certainly not to the same extent: while *-senn* does not appear in the Facebook subcorpus, *doc'h* appears once and *get* is particularly frequent. Facebook users therefore appear more willing to exhibit their dialectal allegiance through orthographic choice, perhaps made more

possible by the lack of an editor.

It was observed above that the majority of the most frequent dialectal tokens are closed-class words, and particularly prepositions. It seems reasonable to suggest that this high use of dialectal prepositions allows speakers to signpost their contributions as dialectal without relying on their audience's knowledge of large amounts of dialectal vocabulary. With the knowledge that *get*, *doc'h*, and *-senn* are Vannetais, speakers of other dialects (including the standard) can associate an utterance with this dialect without having to know the meaning of less commonly occurring Vannetais nouns and verbs. The conscious use of dialect words evidently goes beyond a merely communicative function to contribute to the construction of the speaker's identity as a dialect user, a highly salient consideration for new speakers for whom language and identity are particularly intertwined. However, by restricting dialect use to these well-known words, this also has a communicative purpose, meaning the utterance can still be understood by speakers of other dialects. In a diverse, geographically diffuse community, this allows concerns of identity and communication to be balanced. This notion will recur in chapter 5, particularly in section 5.6.1.

Cases of the frequent occurrence of dialectal prepositions (compared with other tokens) do not seem to have been previously identified in the literature on minoritised languages; the effect of this pragmatic attitude that balances communication with identity will also be explored in section 5.6. While the current study is not carried out within a strictly variationist paradigm, this observation from the data nonetheless highlights "the merits of exploring identity construction and motivation in L2 users [in this case, new speakers, who are perceived to belong to this category] when examining their patterns of variation" (Nance et al. 2016:185); it shows how identity construction is intimately bound up with specific details of language use. While dialect users writing in the print sources examined evidently also express their own dialects through the use of words such as Vannetais *boud*, their need to convey personal identity appears less great, given that they avoid more dialect-specific spellings. This makes sense given that they are writing on behalf of the publication rather than in a personal capacity: considerations of identity are less important.

Finally, in the Facebook subcorpus, a few examples of more unequivocal use of dialect can be noted. In such cases, engagement from users of the standard or of other dialects is low. One post, a rare example of its kind, uses four Vannetais-specific words (out of a total of 21 words), two of which are nouns; the only reply it has attracted has a similarly high rate of use of Vannetais, with four Vannetais words out of thirteen, and further Vannetais spelling conventions.[91] Highly dialectal posts are certainly tolerated within the context of the Facebook group, but thus appear to be ignored by speakers who are not proficient in the dialect in question.

### 3.4.4 Summary

From the analysis it is apparent that dialectal features are visible in Breton-language media, but not to a particularly large extent. However, their appearance indicates that at least some of those who use Breton in these contexts wish to convey alignment with particular dialect areas, and in doing so position themselves at odds with "neo-breton". Some formerly dialectal words, however, particularly those from Léonais, such as *bale*, have become part of modern standard Breton, as seen by their higher and more universal rate of occurrence in the corpus.

The radio is the context where the highest use of unequivocally dialectal vocabulary is used, and this could be a result of two factors: firstly, that dialects are less likely to be represented when the medium is written (particularly in the case of the print sources, as we have seen); secondly, the possibility that radio could be more deliberately inclusive of traditional speakers, who may be excluded from written media as a result of an (assumed) inability to read Breton. Including a restricted number of dialectal features on the radio thus does not fit the stereotype of a wholly standardised "néo-breton", seeming more welcoming to traditional speakers, but equally ensures minimal loss of comprehension from those unfamiliar with the dialects.

In the print subcorpus, *Ya* is notably more tolerant of dialectal features, once those that have become part of the standard have been discounted; it is perhaps surprising

91. For example, use of the definite article *en* instead of standard *an*. This was not tagged because the phonetic proximity of *an* and *en* (particularly given that the article is unstressed and often realised using schwa) would have made it impossible to distinguish these consistently in the radio subcorpus.

that *Brud Nevez* does not take this role, as it has been seen to include larger numbers of non-standard and French-derived words. Its comparative lack of dialectal features could reflect the fact that it is more literary in its content than are the other sources. On the other hand, *Ya*'s inclusion of dialect again signifies that its particular brand of "néo-breton" draws more heavily from the existing resources of Breton rather than relying on neologism, and can perhaps be considered less conservative in this respect. Meanwhile, the Facebook subcorpus shows how users negotiate the trade-off between communication and dialectal identity.

A small number of speakers use a lot of dialectal tokens, and are generally consistent in that these tokens tend to be taken from one dialect in particular. In most cases, this is Vannetais, which is by far the most prevalent of the dialects in the corpus, presumably aided by its greater linguistic distinctiveness and history of codification. Additionally, dialect use tends to be restricted to a small number of markers, many of which are closed-class words such as prepositions, easing communication while allowing the speaker to display dialectal identity: the Facebook subcorpus is a particular example. This indicates that dialects continue to be valued by new speakers.

## 3.5 Integration of tagged tokens into the lexicon

Finally, the tagged tokens were examined to determine the extent to which they were integrated into Breton according to a variety of measures. Due to the nature of this investigation, much of the analysis below applies only to words containing borrowed elements, although the statistics observed must often be compared with those observed for Breton-derived words in order to show the relative measures of integration.

### 3.5.1 Phonological and orthographic integration of borrowings

Three measures were used for determining the extent to which the tagged tokens have been integrated into Breton linguistic structure. The first applies specifically to the 1529 tagged tokens involving borrowing from another language (French, English, Welsh, or "other"). These tokens were tagged to mark whether the phonology or orthography

(depending on the tokens' presence in the spoken or written medium) of the borrowed elements was different from in the source language.

Phonology and orthography were kept separate in these tags as there is no direct correspondence between orthographic and phonological modification. A word could retain its original spelling but be pronounced differently (e.g. utterance-final *tout*, /tut/ in Breton); it could be pronounced in the same way but be orthographically different (e.g. *lig* < French *ligue*); it could be both spelt and pronounced differently (e.g. *kampionad* < French *championnat*); or it could be spelt and pronounced in the same way (e.g. *euro*). As a result, there is no means of direct comparison between the radio subcorpus and the two written subcorpora here.

The figures are shown below.

| | **Tokens** | **Percentage** |
|---|---|---|
| Retained phonology | 414 | 73.7% |
| Changed phonology | 148 | 26.3% |
| Retained orthography | 332 | 33.2% |
| Changed orthography | 635 | 66.8% |
| **Total** | **1529** | |

Table 3.23: Phonological and orthographic modification among tagged tokens

While nearly three quarters of the spoken tokens in question retained the phonology of the original language in borrowed elements, this tendency is thus reversed for orthography, where two thirds of tokens have had their orthography altered. The greater readiness to alter orthography than phonology points towards the fact that phonology is considered a more intrinsic property of a word, while orthography can be changed more easily; it also highlights the importance of orthography in the Breton revitalisation movement, and the need for a word to appear Breton in a visual sense for it to be accepted as part of the lexicon.

There is little variation in this matter at subcorpus level. As a spoken medium, radio

cannot be compared with the other two contexts here; the Facebook and print subcorpus figures are very similar to each other. Public and associative radio also show only a negligible difference from each other. The difference between news and feature programmes is slightly more pronounced, with 68.9% of relevant tokens keeping the source phonology in news programmes, compared with 78.6% in feature programmes. This difference is not statistically significant, however.

Likewise, there is little difference in this regard among the three print sources. Given that *Brud Nevez* has shown characteristics of being further from the "néo-breton" stereotype in terms of some of the features examined above, we might expect it to be notably different again from the other two publications here. *Brud Nevez* does indeed have a slightly higher proportion of tagged tokens where the orthography remains unchanged (40.0%, compared with 34.5% in *Bremañ* and 32.7% in *Ya*); however, again, this is not of statistical significance.

It therefore appears that divisions within the corpus have little effect on the proportions of borrowed tokens with changed phonology and orthography compared with those where the phonology and orthography are unaltered. The main point to be made from investigating these figures is that while two thirds of borrowed tokens have their orthography altered from the source language—this being French in 96% of cases—only just over a quarter have altered phonology. Borrowing therefore generally seems able to accommodate French phonology with ease, while the orthography of a word is more likely to be altered to conform to Breton conventions. As well as the factors mentioned above, this also reflects the fact that all speakers of Breton also speak French and are already fully competent users of French phonology; borrowings that do not occur in a bilingual context of this type tend instead to maintain the orthography of the source language but potentially undergo phonological change: for example, *croissant* in English or *shampooing* in French. Other factors of this community bilingualism will be explored in section 4.4.

### 3.5.2 Derivational morphology and multi-word expressions

#### 3.5.2.1 Occurrence of derivational morphology and multi-word expressions

Tokens were tagged to note the presence of any derivational morphology as a second measure of integration into the linguistic system of Breton. The presence of words combining native and borrowed elements could be interpreted as an indicator of greater tolerance of the latter, whereas keeping the two types of element morphologically separate could indicate a desire not to mix borrowed with native elements. Multi-word expressions are also included in this analysis for the same reason.

Before investigating this in detail, it is appropriate to examine the overall proportions, among the tagged tokens, of words involving derivational morphology, i.e. affixation or compounding resulting in the creation of a distinct lexeme, and forming multi-word expressions. These are as follows (with monomorphemic tokens also included for comparison):

| | **Tokens** | **Percentage** |
|---|---|---|
| Monomorphemic[92] | 1842 | 47.2% |
| Affixation | 1357 | 34.8% |
| Compounding | 522 | 13.4% |
| Multi-word expressions | 199 | 5.1% |
| **Total** | **3904**[93] | |

Table 3.24: Derivational morphology among tagged tokens

As we can expect, the proportions of tokens involving derivational morphology vary across the three subcorpora, no doubt a result of the different registers of language employed.

92. In this section, this category also includes tokens involving only inflectional affixes.
93. The figures in the table sum to slightly more than this as a few tokens involved both affixation and compounding.

| | | # Tokens (normalised) | | | Percentage of tokens | | |
|---|---|---|---|---|---|---|---|
| | | **Radio** | **Facebook** | **Print** | **Radio** | **Facebook** | **Print** |
| * | Monomorphemic | 711.2 | 539.3 | 529.0 | 55.2% | 43.4% | 43.2% |
| * | Affixation | 319.6 | 506.9 | 469.1 | 24.8% | 40.8% | 38.3% |
| * | Compounding | 194.8 | 141.7 | 166.8 | 15.1% | 11.4% | 13.6% |
| | Multi-word expressions | 69.0 | 59.0 | 63.6 | 5.4% | 4.7% | 5.2% |

Table 3.25: Derivational morphology by subcorpus

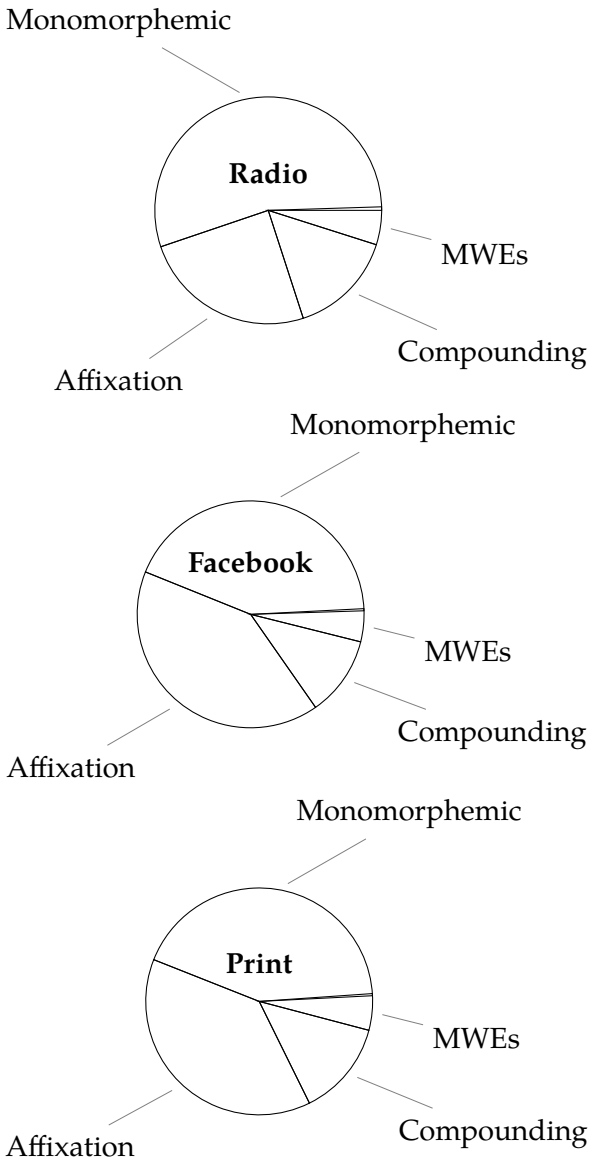


Figure 3.16: Graph of Table 3.25

Statistical significance can be observed for all categories except multi-word expressions. The radio subcorpus shows the highest proportion of monomorphemic words among the tagged tokens, reflecting the spoken register's tendency towards a morphologically simpler style. Conversely, there is a higher rate of affixation in the two written subcorpora. Compounding, on the other hand, defies this overall pattern, being most frequent on the radio and least frequent on Facebook. Its high rate of occurrence in the radio subcorpus can be traced to the data from associative radio, where statistical significance is again indicated compared with the figure for public radio. In this case, it appears that the data are skewed by their subject matter: much of the discussion in the associative radio sample is of theatre and includes several instances of the word *pezh-c'hoari* ('play'). In general, therefore, the rate of compounding therefore appears not to vary by register, unlike affixation.

#### 3.5.2.2 Derivation and multi-word expressions by word origin

In order to investigate these morphosyntactic features in a way that might reveal information about the treatment of borrowed elements, tokens in these categories were tagged in a way that noted the language from which the elements were derived. To permit a more detailed level of analysis, indicating whether borrowed elements in these tokens were borrowed for use specifically as part of the morphologically complex token or could be found independently of this, the tagging went into more detail than the general tagging for word origins (see section 3.2.3). Instead of tagging any polymorphemic word with a French-derived element as such, those containing French-derived elements found in Breton in a standalone capacity were tagged as "Breton compounding". This permitted a more detailed system for categorising tokens, particularly where they contained borrowed elements. For example, for compounds, this operated as follows:

| | **Tag combination** | **Example** |
|---|---|---|
| Breton elements only | Breton origin, Breton compounding | *mell-droad* 'football' |
| At least one French-derived element, all of which occur standalone in Breton | French origin, Breton compounding | *teulfilm* 'documentary' |
| One French-derived element not occurring standalone in Breton and one other element | French origin, Breton/ French compounding | *ti-bank* 'bank' |
| French-derived elements not occurring standalone in Breton only | French origin, French compounding | *kilometr* 'kilometre' |

Table 3.26: Categorisation of word origins in polymorphemic tokens

Words with elements derived from French can in this way be placed on a scale of integration: those most integrated into Breton will be used standalone as well as with derivational morphology, and co-occur with Breton elements in morphologically complex words. Those less integrated will be present in a more restricted range of situations; that is, only as part of a token where they co-occur with other French-derived elements, without the freedom to stand alone or to be combined with Breton elements.

With this in mind, we can examine the distribution of the relevant tokens, starting with those containing affixation and looking at its occurrence across the three subcorpora.

| | # Tokens (normalised for the subcorpora) | | | | % of tokens with affixation | | | |
|---|---|---|---|---|---|---|---|---|
| | **Total** | **Radio** | **Facebook** | **Print** | **Total** | **Radio** | **Facebook** | **Print** |
| Celtic root,[94] Breton affix | **964** | 226.2 | 368.1 | 326.2 | **71.0%** | 70.8% | 72.6% | 69.5% |
| French root found standalone in Breton, Breton affix | **84** | 9.1 | 31.4 | 38.7 | **6.2%** | 2.9% | 6.2% | 8.3% |
| Breton root, French affix | **5** | 2.0 | 1.9 | 0.9 | **0.4%** | 0.6% | 0.4% | 0.2% |
| French root found standalone in Breton, French affix | **35** | 11.2 | 7.6 | 14.7 | **2.6%** | 3.5% | 1.5% | 3.1% |
| * French root not found standalone in Breton, Breton affix | **224** | 56.8 | 87.5 | 70.0 | **16.5%** | 17.8% | 17.3% | 14.9% |
| Root of another origin, Breton affix | **3** | 0.0 | 0.0 | 2.8 | **0.2%** | 0 | 0 | 0.6% |
| French root not found standalone in Breton, French affix | **42** | 14.2 | 10.5 | 15.7 | **3.1%** | 4.4% | 2.1% | 3.3% |
| **Total** | **1357** | | | | | | | |

Table 3.27: Tokens with affixes by subcorpus

First considering the overall counts of the different types, it can be noted that some occur much more frequently than others. French affixes are rarely appended to Celtic roots, for example, and among French roots, Breton affixes are a lot more common

94. This category includes roots from Breton, Welsh, and older Breton, which were put in the same category for this analysis due to the low occurrence of Welsh and older Breton roots and their overall alignment with stereotypical "néo-breton".

than French ones, even where the root does not form a standalone borrowed word in Breton. Among those categories of tokens using French affixes, the most numerous is that where the root is also from French and is not used standalone in Breton. This indicates that French-derived affixes tend to be restricted to use with roots that both also come from French and are not particularly integrated into Breton. Breton affixes, on the other hand, can be seen to be favoured over French ones, even where the root is French. This makes the borrowed element more integrated, perhaps rendering it more acceptable.

We can now turn to the distribution of these categories across the three subcorpora. Statistical significance can be found in only one of the seven categories, this being the category of words where a French root, found standalone in Breton, occurs with a Breton affix. In this case, the number of tokens on the radio is much lower than in the written sources. It is unclear why this is, particularly when the proportion is not compensated for by an increased number of tokens in any of the other categories. A possible reason is the fact that many such words are abstract nouns and as such may belong to a higher register than that typically found on the radio. This may also be the case for words with Celtic roots and Breton affixes, which in absolute terms are also fewer on the radio than elsewhere, but do not show up as statistically significant using this method.

Breaking down the radio data by programme type also reveals some significant discrepancies:

| | | # Tokens (normalised) | | % of tokens with affixation | |
|---|---|---|---|---|---|
| | | **News** | **Features** | **News** | **Features** |
| * | Celtic root, Breton affix | 94.6 | 130.7 | 58.0% | 83.5% |
| | French root found standalone in Breton, Breton affix | 6.2 | 3.0 | 3.8% | 1.9% |
| | Breton root, French affix | 1.0 | 1.0 | 0.6% | 0.6% |
| * | French root found standalone in Breton, French affix | 11.4 | 0.0 | 7.0% | 0 |
| * | French root not found standalone in Breton, Breton affix | 45.7 | 11.9 | 28.0% | 7.6% |
| | Root of another origin, Breton affix | 0.0 | 0.0 | 0 | 0 |
| | French root not found standalone in Breton, French affix | 4.2 | 9.9 | 2.5% | 6.3% |

Table 3.28: Tokens with affixes in the radio subcorpus by programme type

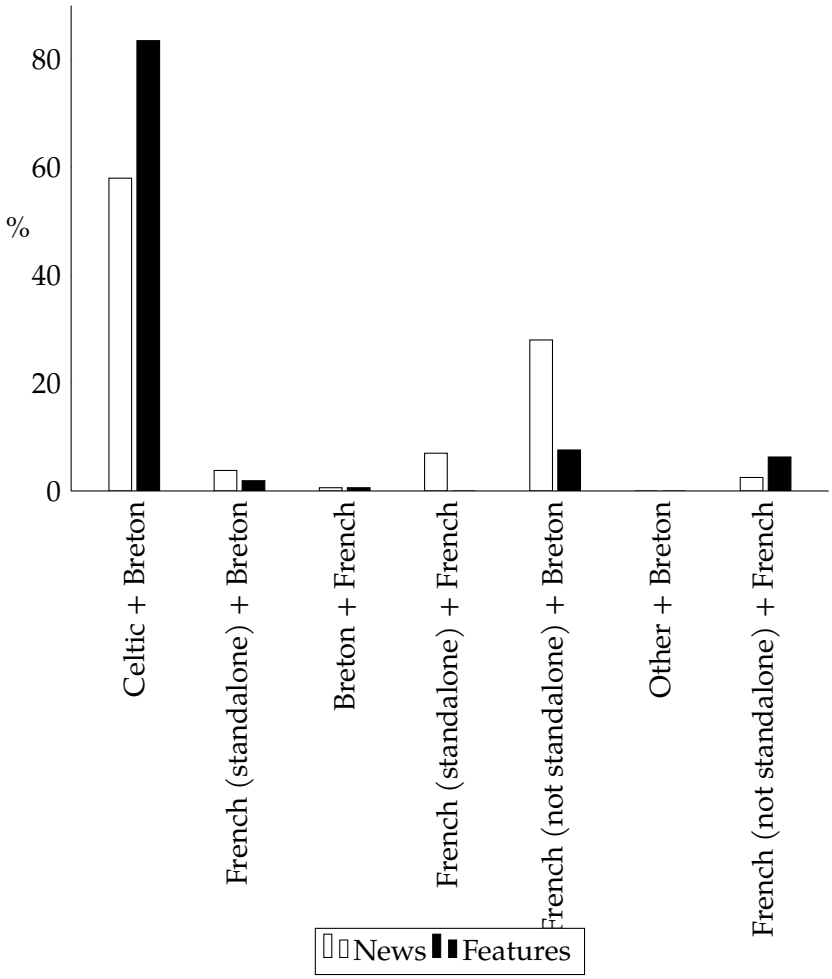


Figure 3.17: Graph of Table 3.28

Statistical significance can be observed in three categories: tokens with Celtic roots are more common in the feature programmes, while two categories are more common in news programmes: tokens with French roots found standalone in Breton and French affixes, and tokens with French roots *not* found standalone in Breton and Breton affixes. Evidently there is no way of meaningfully grouping these categories. The news programmes appear from these figures to have a greater tolerance of words containing French elements; the only category for which this is not the case is the final one, of French roots not found standalone in Breton with French affixes. This, however, conforms with the implication that French elements are less tolerated in the feature programmes: where they do occur, they are more likely not to be integrated into Breton, as more of those that do not occur standalone make use of French rather than Breton affixation. Thus, if a French word is not already integrated into Breton, it is less likely

to become integrated through affixation. The news programmes perhaps use a more casual style of speech as they are discussing topics in less depth: this could be a reason for their greater use of French-derived roots. However, it can be noted from the figures earlier in this chapter that among the tagged tokens in general, the trend goes the other way, with slightly more use of Breton-derived words in the news programmes, albeit not to a statistically significant extent. This difference may reflect a difference in vocabulary or register.

It is also worth breaking down the figures for the print subcorpus.

| | | **# Tokens (normalised)** | | | **% of tokens with affixation** | | |
|---|---|---|---|---|---|---|---|
| | | ***Bremañ*** | ***Brud Nevez*** | ***Ya*** | ***Bremañ*** | ***Brud Nevez*** | ***Ya*** |
| | Celtic root, Breton affix | 170.2 | 40.9 | 113.5 | 68.6% | 57.1% | 77.2% |
| | French root found standalone in Breton, Breton affix | 21.7 | 8.4 | 8.4 | 8.8% | 11.7% | 5.7% |
| | Breton root, French affix | 0.0 | 0.0 | 0.9 | 0 | 0 | 0.6% |
| * | French root found standalone in Breton, French affix | 2.7 | 4.6 | 7.4 | 1.1% | 6.5% | 5.1% |
| | French root not found standalone in Breton, Breton affix | 38.0 | 15.8 | 15.8 | 15.3% | 22.1% | 10.8% |
| | Root of another origin, Breton affix | 2.7 | 0.0 | 0.0 | 1.1% | 0 | 0 |
| * | French root not found standalone in Breton, French affix | 12.7 | 1.9 | 0.9 | 5.1% | 2.6% | 0.6% |

Table 3.29: Tokens with affixes in the print subcorpus

In this case, statistical significance can be observed in two categories. There is an unusually low proportion of tokens with French roots found standalone in Breton and French affixes in *Bremañ*, and, conversely, an unusually *high* proportion of tokens with French roots not found standalone in Breton and French affixes in the same publica-

tion. *Bremañ*, which earlier evidence has shown to represent a more conservative "néo-breton" tradition, therefore appears to take different attitudes to French borrowings depending on whether the borrowed element is already present and integrated into Breton vocabulary. Those that are already in use in Breton are treated as Breton and given Breton affixes; those that are not are treated as French and are more likely to receive French affixation, which tends to be restricted to use with these unintegrated elements. This has the effect of retaining the lack of integration, and this unwillingness to integrate these newer French borrowings can be interpreted as a sign of alignment with a more stereotypical form of "néo-breton". Again, *Ya* differs from this, in this case showing figures similar to *Brud Nevez*. This could again be a reflection of the fact that *Bremañ* uses a more formal linguistic register, but also indicates, again, that in certain respects *Ya* is more tolerant of French-derived vocabulary.

Moving on to compounding, there is less variation among the data. The overall figures are as follows, along with a breakdown by subcorpora.

| | # Tokens (normalised for the subcorpora) | | | | % of tokens involving compounding | | | |
|---|---|---|---|---|---|---|---|---|
| | **Total** | **Radio** | **Facebook** | **Print** | **Total** | **Radio** | **Facebook** | **Print** |
| Celtic elements[95] only | **404** | 138.0 | 112.2 | 138.2 | **77.4%** | 70.8% | 79.2% | 82.9% |
| * At least one French-derived element, all of which occur standalone in Breton | **89** | 45.7 | 23.8 | 17.5 | **17.0%** | 23.4% | 16.8% | 10.5% |
| One French-derived element not occurring standalone in Breton and one other element | **11** | 3.0 | 2.9 | 4.6 | **2.1%** | 1.6% | 2.0% | 2.8% |
| French elements not occurring standalone in Breton only | **16** | 8.1 | 1.9 | 5.5 | **3.1%** | 4.2% | 1.3% | 3.3% |
| English elements not occurring in French only | **2** | 0.0 | 1.0 | 0.9 | **0.4%** | 0 | 0.7% | 0.6% |
| **Total** | **522** | | | | | | | |

Table 3.30: Compound words among the subcorpora

French-derived elements not occurring standalone in Breton seem slightly more likely to occur in compounds with other French-derived elements than with Breton elements, again suggesting that there is resistance to morphologically integrating words that are not already considered part of the Breton lexicon. Additionally, they are rare in compounds overall, compared with other types. It therefore seems rare that a French ele-

95. Again, roots from Breton, Welsh, and older Breton were grouped here.

ment is borrowed into Breton for the sole purpose of use in compounds.

When the data are split according to subcorpus, the only statistical significance to be observed is in the category of tokens where there is at least one French-derived element and all of these occur standalone in Breton. Here, the print media show the lowest proportion, reflecting their tendency to make less use of French-derived words in general—the higher proportion of Breton-only compounds in this subcorpus also indicates this. The radio data have the highest proportion of tokens where there is at least one such French element, conversely highlighting the radio's greater use of French-derived words, but this does not reveal any new information. Similarly, when the subcorpora are broken down in the usual ways, there is only one case of statistical significance: within the radio subcorpus, feature programmes contain three tokens in the third category (i.e. one French-derived element not occurring standalone in Breton and one other element), while news programmes contain none. As all these tokens are of the same word, it is unlikely that this is a meaningful difference.

Similarly, the proportion of different types of multi-word expressions shows little variation when the corpus is broken down. Here are the overall figures:[96]

96. Apparent gaps in the list of categories are due to not all categories being attested in the corpus.

| | # Tokens (normalised for the subcorpora) | | | | % of total multi-word expressions | | | |
|---|---|---|---|---|---|---|---|---|
| | **Total** | **Radio** | **Facebook** | **Print** | **Total** | **Radio** | **Facebook** | **Print** |
| Breton words only | **146** | 50.7 | 46.6 | 43.3 | **73.4%** | 73.5% | 79.0% | 68.1% |
| At least one French-derived word, all of which occur standalone in Breton | **36** | 11.2 | 9.5 | 13.8 | **18.1%** | 16.2% | 16.1% | 21.7% |
| French-derived words not occurring standalone in Breton only | **9** | 6.1 | 1.9 | 0.9 | **4.5%** | 8.8% | 3.2% | 1.4% |
| * English-derived words not occurring in French only | **7** | 0.0 | 1.0 | 5.5 | **3.5%** | 0 | 1.6% | 8.7% |
| Words from other languages only | **1** | 1.0 | 0.0 | 0.0 | **0.5%** | 1.5 | 0 | 0 |
| **Total** | **199** | | | | | | | |

Table 3.31: Multi-word expressions in the subcorpora

Here, the avoidance of mixing words from different languages is still more apparent, with no instances at all of words from languages other than Breton occurring in multi-word expressions with Breton-derived words, apart from cases where a French-derived word already occurs standalone in Breton (i.e. the second category in the table). Again, Breton-derived tokens make up the majority, and again, as stated above, there is little variation among the subcorpora, with the only statistical significance to be observed in the category of English-derived words not occurring in French, all but one of which

occur in the print subcorpus.

Having investigated some morphosyntactic properties of the tagged tokens, we are able to give more support to some of the earlier findings. In its use of more monomorphemic words and less affixation, the radio subcorpus employs morphologically simpler language, which is unsurprising given that it is the only spoken subcorpus. By the same measure, the highest linguistic register can be found in *Bremañ*, which uses the most affixation. It also shows characteristics of stereotypical "néo-breton" in its less frequent integration of French elements, which tend throughout the entire corpus to be morphologically or syntactically paired with other French elements more often than Breton elements are, but particularly so in *Bremañ*. For compounding and multi-word expressions, however, there is in most respects little difference among the different divisions of the corpus, and those differences that are found may be due to subject matter rather than register or "néo-breton" tendencies, as in the case of the higher rate of compounding on associative radio, which is likely to be due to discussion of theatre and the high frequency of the word *pezh-c'hoari*. Investigating these properties therefore does not reveal any particular difference among or within subcorpora, but does show in the overall numbers of tokens that mixing elements with origins in different languages is rare, perhaps indicating a reluctance to integrate borrowings into Breton by morphological or syntactic means if they are not of long standing in the language already.

### 3.5.3 Mutation and inflectional morphology

#### 3.5.3.1 Occurrence of mutations and inflectional morphology

As a third measure of integration into Breton, inflectional morphology was also examined. Feminising suffixes were also tagged separately[97] and are included in this section of the analysis. Mutation, a type of sandhi that characterises the Celtic languages (Fife 2010:7) involving alteration of an initial consonant in certain environments,[98] was

97. These could be argued to be derivational. However, as words with feminising suffixes were not counted as separate lexemes for this research, they are considered along with inflection rather than derivation.

98. For example, feminine singular nouns mutate after articles: *taol* ('table') becomes *an daol* ('the table'); verbs mutate after certain particles, e.g. the particle that marks a continuous action: *gwelout* ('to

also examined and is equally included here: as it is a characteristic of Celtic languages, it might be expected to occur less often among words considered "non-Celtic". Inflectional morphology itself was divided into Breton and non-Breton. This latter category includes any inflections that do not normally occur in Breton, such as *-s* to mark plurality, found on the token *audiobooks*.

The distribution of inflection and mutation across the corpus was as follows.

| | **# Tokens (normalised for the subcorpora)** | | | | **Percentage of tokens** | | | |
|---|---|---|---|---|---|---|---|---|
| | **Total** | **Radio** | **Facebook** | **Print** | **Total** | **Radio** | **Facebook** | **Print** |
| Breton inflection | **862** | 270.9 | 272.0 | 284.8 | **22.1%** | 21.0% | 21.9% | 23.3% |
| Non-Breton inflection | **6** | 0.0 | 2.9 | 2.8 | **0.2%** | 0 | 0.2% | 0.2% |
| Mutation | **325** | 95.4 | 95.1 | 120.7 | **8.3%** | 7.4% | 7.7% | 9.9% |
| Feminising suffix | **17** | 9.1 | 1.9 | 5.5 | **0.4%** | 0.7 | 0.2% | 0.5% |
| Unmutated and uninflected | **2789** | 937.4 | 906.4 | 840.5 | **71.4%** | 72.8 | 73.0% | 68.7% |
| **Total** | **3904** | | | | | | | |

Table 3.32: Inflection and mutation among tagged tokens

Looking first at the overall figures, it can be observed that inflection is drawn from Breton paradigms in the vast majority of cases, i.e. over 99% of inflected tokens. The characteristics of the six tokens with non-Breton inflection will be considered below. Feminising suffixes are also fairly rare among the tagged tokens: they tend to be used only for professions in Breton.

In this case, there is not much difference among the subcorpora. This is unsurprising, as unlike derivational morphology, which creates more complex vocabulary, inflectional morphology does not have this property, and so differences in register are less

see') becomes *o welout* ('seeing'); etc.

likely to affect proportions of inflectional morphology. A lower rate of mutation might be expected among the radio data given that speech is less careful and harder to correct than written production, and that mutation does not occur in French, which will be the first language of most speakers, so speech errors may be more likely. However, at first glance, while the number of mutated tokens is slightly lower in the radio subcorpus, this is not statistically significant. When the radio data are divided into news and feature programmes, though, statistical significance can be observed, with 8.6% of tagged tokens in news programmes mutated as opposed to 6.1% in feature programmes. This will be explored further below.

This is one of only two instances of statistical significance that can be observed across the various ways of splitting the corpus. The second is among the data from print journalism, in the case of tokens with Breton inflection. We have seen that these make up over 99% of cases of inflection in total, and therefore it does not seem unreasonable to generalise here and assume that these figures represent inflection overall. The proportion of tokens in *Bremañ* that are inflected is unusually high, with 28.3% of tokens inflected, compared with 21.7% in *Brud Nevez* and 18.8% in *Ya*. Breaking these data down by parts of speech, the following figures emerge:

| | | **# Tokens (normalised)** | | |
|---|---|---|---|---|
| | | ***Bremañ*** | ***Brud Nevez*** | ***Ya*** |
| * | **Nouns** | 105.0 | 35.3 | 55.8 |
| | **Verbs** | 30.8 | 22.3 | 30.7 |
| | **Other** | 1.8 | 1.9 | 0.0 |
| * | **Total** | **152** | **64** | **93** |

Table 3.33: Inflected tokens in the print subcorpus

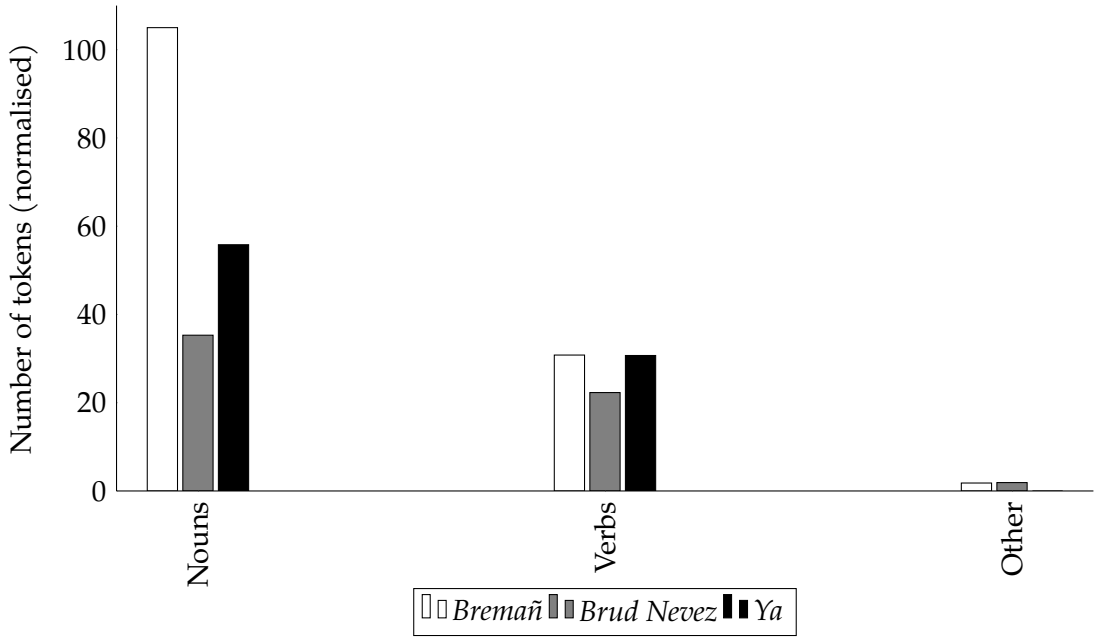


Figure 3.18: Graph of Table 3.33

While the three publications are mostly fairly consistent in the number of tagged tokens that are inflected, the one category where this is not the case is that of nouns, where *Bremañ* has a noticeably higher number, and some discrepancy can also be observed between *Brud Nevez* and *Ya*. It has already been established that differences in register are unlikely to cause differences in inflectional morphology of this type, and given that this relates specifically to plural markers, the only type of morphology tagged in this way, this is really nothing more than a measure of the frequency of plural nouns. Therefore, it is apparent that there are more plural nouns in *Bremañ* than in the other two print sources, at least among the tagged tokens. This does not appear to be attributable to register or have any significance in relation to "néo-breton", but perhaps merely reflects the subject matter of the three publications, with the sample from *Bremañ* happening to deal with topics that address groups of things or people.[99]

99. We have seen that *Bremañ* has a more political focus, concentrating particularly on linguistic and ethnic minorities, so it does not seem unreasonable to assume that some of the plural nouns in question are the names of minority peoples.

#### 3.5.3.2 Mutation and inflectional morphology by word origin

Examining the distribution of inflection and mutation in this general way is therefore not particularly revealing. However, it is also worth breaking down this distribution by word origin, given that mutation is specific to the Celtic languages. For this, word origins were grouped into two categories: Celtic (tokens with Welsh elements, tokens with elements found only in Breton prior to 1700, or tokens consisting only of elements found in Breton prior to 1900) and non-Celtic (tokens with French, English or other non-Celtic elements). Their distribution across the corpus is as follows:

| | # Tokens (normalised for the subcorpora) | | | | Percentage of tokens | | | |
|---|---|---|---|---|---|---|---|---|
| | **Total** | **Radio** | **Facebook** | **Print** | **Total** | **Radio** | **Facebook** | **Print** |
| Celtic; Breton inflection | **592** | 162.3 | 188.3 | 215.7 | **15.2%** | 12.6% | 15.2% | 17.6% |
| Non-Celtic; Breton inflection | **270** | 108.5 | 83.7 | 69.1 | **6.9%** | 8.4% | 6.7% | 5.6% |
| Celtic; non-Breton inflection | **0** | 0 | 0 | 0 | **0** | 0 | 0 | 0 |
| Non-Celtic; non-Breton inflection | **6** | 0.0 | 2.9 | 2.8 | **0.2%** | 0 | 0.2% | 0.2% |
| Celtic; mutation | **220** | 63.9 | 70.4 | 76.5 | **5.6%** | 5.0% | 5.7% | 6.3% |
| * Non-Celtic; mutation | **105** | 31.4 | 24.7 | 44.2 | **2.7%** | 2.4% | 2.0% | 3.6% |
| Celtic; feminine suffix | **12** | 5.1 | 1.0 | 5.5 | **0.3%** | 0.4% | 0.1% | 0.5% |
| Non-Celtic; feminine suffix | **5** | 4.1 | 1.0 | 0.0 | **0.1%** | 0.3% | 0.1% | 0 |
| Celtic; unmutated and uninflected | **1648** | 501.2 | 580.1 | 501.3 | **42.2%** | 38.9% | 46.7% | 41.0% |
| Non-Celtic; unmutated and uninflected | **1141** | 436.2 | 326.2 | 339.1 | **29.2%** | 33.9% | 26.3% | 27.7% |
| **Total** | **3904** | | | | | | | |

Table 3.34: Inflection and mutation by word origin

These data show a pattern in inflection across words of particular origins. While the greater number of Celtic words with Breton inflection could be assumed to reflect the greater number of Celtic tagged tokens overall (2405, versus 1499 non-Celtic tagged tokens), there is a statistically significant difference even when this difference in population size is taken into account. This can also be observed for tokens involving muta-

tion and those not affected by inflection or mutation, as well as, evidently, for tokens with non-Breton inflection, where all six tokens are of non-Celtic origin. Tokens of Celtic origin are therefore significantly more likely to be inflected or mutated, while those of non-Celtic origin are conversely significantly more likely not to be affected by either of these phenomena: 76.1% of them are uninflected and unmutated, as opposed to 68.5% of Celtic tokens. In this respect, non-Celtic words appear on the whole to be less integrated into Breton. This may indicate that speakers are choosing not to use words of non-Celtic origin when a mutation or inflection is required, or that they are not mutating or inflecting these words as often when they do use them.

Taking into account the overall distribution of Celtic and non-Celtic tagged tokens, a few examples of statistical significance can be observed when the corpus is divided into parts. One of these examples is found when the three subcorpora are compared: the number of non-Celtic tokens that are mutated is significantly higher in the print subcorpus, and the lowest in the Facebook subcorpus. The fact that more of them are mutated indicates that more of these tokens are integrated into Breton, and hence, again, that those non-Celtic tokens used in the print subcorpus are likely to be longer-established borrowings. On Facebook, conversely, the fact that more non-Celtic tokens are less integrated implies that more of them are recent or one-off borrowings.

It was shown above that significantly more tokens were mutated in news programmes than in feature programmes, and the suggestion was made that this could be due to a greater proportion of scripted content in news programmes. When this is broken down to the level of Celtic versus non-Celtic tokens, it becomes evident that the difference occurs most noticeably among non-Celtic tokens: there are 25.0 mutated non-Celtic tagged tokens[100] on news programmes and only 6.9 on feature programmes (for Celtic tokens, 35.3 on news programmes and 28.7 on feature programmes). As with the print subcorpus data, this suggests that the non-Celtic vocabulary used in news programmes is more integrated into Breton and perhaps primarily made up of longer-standing borrowings. The feature programmes therefore appear comparatively more tolerant of less integrated, more recently borrowed vocabulary, perhaps again a result of their greater discussion of the arts and consequent higher use of technical terms in

100. These are normalised figures based on equal-sized samples for the two types of programme.

this field.

Two instances of statistical significance can be found when the print subcorpus is broken down by publication. The first is in the category of non-Celtic words that occur with mutation, where the rate of occurrence in *Bremañ* (41.6 tokens)[101] is noticeably higher than in the other two publications (15.8 in *Brud Nevez* and 11.2 in *Ya*). This suggests that *Bremañ* is more likely to use mutation with borrowed words, indicating that the borrowings it uses may be longer-established, or that it takes a stricter approach to mutation overall, using it even with those words that other sources may choose not to mutate because they are less established borrowings. This indicates a willingness to ensure that the text is as grammatically Breton as possible, and fits with *Bremañ*'s apparent use of a higher linguistic register.

The second instance of statistical significance within the print subcorpus is in the category of words with non-Celtic inflection. It has already been seen that there are only six such tokens in the entire corpus. Out of these six, half occur in the print subcorpus, all three of which can be found in *Ya*. The three tokens are of different words, all English terms not typically used in French: *terminal speakers*, *dark elves* and *play-offs*. *Ya* therefore appears more tolerant of the use of borrowings of this type, although it does not go as far as grammatically integrating them into Breton, instead using their original inflectional morphology where necessary. While these three tokens all tend not to be used in French, with the possible exception of *play-offs*,[102] this is not the case for all six of the tokens in this category that occur in the corpus. The other three occur in the Facebook subcorpus and are all used in French. One, *audiobooks*, comes originally from English, but the other two originate in French: *bretonnants* (Breton speakers) and *bretonnismes* (words or figures of speech used in local French but usually derived from Breton). Facebook users therefore appear more comfortable using French words that they may consider untranslatable into Breton and not integrating them with respect to inflectional morphology, while the printed publications may feel more pressure to carry out such integration or use a Celtic equivalent word. Again, the unmoderated nature of the Facebook context allows its users to take different approaches to how

101. Again, these figures have been normalised.
102. *Play-off* is occasionally found in French, but the native equivalent (*match de*) *barrage* is more common.

loanwords are adapted or left unchanged. In *Ya*, on the other hand, the use of non-Celtic inflectional morphology is restricted to words that are more alien to the Breton context.

### 3.5.4 Summary

This section began by observing orthographic and phonological change, noting that the former was far more frequent in borrowed words. This is likely to be influenced by the significant level of contact with French, a theme that will recur in future sections. It then examined morphology, noting that both derivation and inflection tend to make use of morphemes that are native to Breton: for inflection, even in the case of words that contain components borrowed from French. The small number of words that retain French-style inflection are particularly recent borrowings, hence the least integrated into Breton. In this domain, little variation can be observed among the subcorpora, showing that register is not of much significance here.

For mutation, slightly more significant variation can be found, but still not a particularly large amount. Mutation occurs at a high rate in print journalism, and particularly so, comparatively speaking, in the case of non-Celtic tokens, implying that those non-Celtic tokens used in print journalism tend to be the most integrated into Breton. This is especially the case in *Bremañ*.

This section shows that while a large number of borrowings are integrated in at least one of the ways examined, some instead retain orthography, phonology and/or morphology from the source language and can be dealt with differently depending on the speaker's attitudes. This is particularly apparent from the differences between *Bremañ* and *Ya*, and variation and change of this type is hardly specific to Breton: in the same way that anglicisms in French may begin by retaining their original morphology and eventually become more morphologically integrated,[103] it is possible that these words in Breton will become more integrated into the language in the future, just as generations of borrowings will have done in the past. The analysis above certainly shows

103. For example, Saugera (2012:140) hypothesises that "*post-it* will eventually inflect when it is no longer felt to be a trademark, just like other integrated nominalised trademarks (e.g. *des scotchs*, *des tupperwares*)."

that even in more typical "néo-breton" contexts, borrowing is an important device for enriching the Breton lexicon.

## 3.6 Conclusion

Overall, the corpus exhibits a moderate degree of conformance to "néo-breton": Breton-derived vocabulary is more common among tagged tokens than borrowings, dialect is fairly restricted, and Breton affixes are favoured over French ones. In general, this additionally correlates with register: typically, the print sources, which use the highest register out of the three subcorpora, use a type of language that conforms far more to scholars' claims about "néo-breton". However, exceptions to this general tendency can also be identified. While the Facebook data are in an intermediate position in terms of register, they include the most Breton-derived tagged tokens, including some instances of purist vocabulary that are in excess of that prescribed by TermOfis. The heteroglossic characteristics of the Facebook subcorpus mean that some of these instances of extreme purism may in effect be "cancelled out" by other users' different approaches, meaning that according to the macro-level statistical analysis applied in this chapter this subcorpus appears not to tend towards either pole in particular. The more fine-grained qualitative analysis in the next chapter will attempt to show more clearly some of the nuances among the Facebook data.

Notable variation also occurs among the print sources. Despite being a printed publication and containing some extremely high-register contents in the form of literary contributions (although these are not in the sample), *Brud Nevez* can be seen to defy the "néo-breton" trend followed by the other publications in various ways, showing its ideological incompatibility with stereotypical "néo-breton". This indicates how, guided by these ideological differences, conflicting attitudes and practices can coexist in the media context: the choice of whether or not to use stereotypical "néo-breton" thus does not necessarily depend on register. Additionally, the Facebook subcorpus shows that more "néo-breton" contexts are not always those that conform the most to the standard, demonstrating that while ideologies favouring standardisation and those favouring purism often overlap, the extent to which this occurs can differ.

Unlike *Brud Nevez*, both *Bremañ* and *Ya* follow more typical "néo-breton" tendencies, but they take notably different approaches to this. *Bremañ*, the inheritor of the tradition of Breton periodicals, is more conservative, being more likely to use neologisms and words found in Hemon's 1964 dictionary; *Ya*, far more recently established, makes more use of certain borrowings: specifically, internationally-used words and words derived from languages other than French, particularly English. It is also more open to the use of dialectal features. While it broadly maintains the usual "néo-breton" qualities, then, it does appear to do this in a less conservative way than *Bremañ*, showing instead a greater tolerance of both external influence and internal variation. This perhaps mirrors the attitudes of the two publications more generally: while *Bremañ* is a successor of the "néo-breton" tradition of periodicals with a clear political interest, *Ya* is a newer publication of a different type that does not come from the same line, allowing it to take a more innovative approach to language. This perhaps indicates a change in progress: a new type of "néo-breton" may be emerging, with different ideological motivations, an observation that is relevant to research question 4.

Less variation can be identified when the radio subcorpus is split: there is some indication that public radio uses more standard vocabulary than associative radio, but not all the data bear this out. Employees use significantly less French-derived vocabulary than guests do, but given the trends in variation and register, it is most likely that this is due to the fact that more of the employees' speech is scripted and therefore belongs in a sense to a written, and therefore higher, register. Both types of radio station appear accommodating of a more dialectal style, as well as showing indicators of being further from stereotypical "néo-breton" in general. As well as this befitting the lower register of the radio data compared with the other data under examination, it corresponds with the fact that the radio, as an oral medium, is the most accessible to traditional speakers, who may struggle to read Breton or lack internet access. While almost all those who participate in radio programmes are new speakers, they may be aware of their potentially large traditional speaker audience: this will be discussed further in chapter 5.

At this stage in the analysis, another salient point relates to the use of dialects: it was found that this is typically done in a way that balances concerns of communication

and identity, allowing speakers to communicate with others of different backgrounds while expressing their own allegiance to a certain dialect. Concerns of identity will likewise be relevant in the next chapter, which examines more specific features of the corpus.

# Chapter 4

# Contemporary Breton usage 2: Qualitative analysis of the corpus data

## 4.1 Introduction

This chapter continues the focus on the corpus with a move away from quantitative analysis, focusing on particular aspects of the data in more detail. This will reveal further nuances to the trends already identified in chapter 3. Section 4.2 investigates a number of word-formation processes that have led to the creation of words appearing in the corpus, examining how these fit into stereotypical perceptions of "néo-breton". Section 4.3 focuses on a selection of specific words that are found in the corpus, showing how their use intersects with these stereotypes. Finally, section 4.4 focuses on the fact that French is a language that speakers in the corpus will have in common, examining how this may influence their Breton. This chapter will also make reference to the practices used by dictionaries and terminology databases, which equally reveal information about how "néo-breton" is constructed and framed, and provide data for answering research question 1 in particular, as well as 3 and 4.

## 4.2 Word-formation processes

This section follows on from the focus on morphology in the last section of the previous chapter with a more detailed look at some word-formation processes that can be observed from the corpus data. Evidently, the lexicon of Breton has required significant expansion since 1900 in order to contain words for technical and newer concepts. In cases other than direct borrowing from other languages—almost always French, as the previous chapter established—and any ensuing orthographic and/or phonological modification, this requires new words to be coined, or existing vocabulary to undergo semantic change: this section examines a few ways in which this occurs in the corpus. All four of the word-formation processes discussed in this section[104]—calques, clipping, semantic expansion, and affixation—are common in other languages: they are all discussed by Wise (1997) in relation to French, for example. In minoritised languages such as Breton, however, they may attract particular comment, as this section will note.

### 4.2.1 Lexical calques

Scholars have criticised "néo-breton" on the grounds that it apparently involves calquing from French, believing this to be a sign that new speakers are unable to master Breton syntax (Hewitt 2016:16). We might therefore expect to find this feature in abundance in the corpus; in fact, however, there are only a small number of examples at the lexical level.

For the purposes of this research, lexical calques were defined according to a specific criterion: terms where every constituent word had the same meaning within the phrase in question as when they stand alone were not included. This meant tagged tokens such as *prezidant a enor* (< *président d'honneur*) and *domani foran* (< *domaine public*) were not investigated here, as the specific syntactic arrangement of elements is not itself important for the interpretation of the phrase, and additionally it is impossible to claim that such structures are borrowed from French rather than being the logical way of

104. The ludic use of language also forms part of this section, but is not strictly a word-formation process—rather, it is a technique that may involve particular word-formation processes.

arranging the elements according to Breton grammar.

In the same way that some of the tagged tokens were characterised as internationally used, some of the lexical calques in the corpus come into this category. *Lakaat e pleustr*, for example, attested six times in the corpus, can be related to the French equivalent *mettre en pratique*, but also to English *put into practice*. Similarly, *rouedad sokial*, found twice in the corpus, derives from French *réseau social*, which in turn is calqued on English *social network*. Both these examples involve semantic expansion through the use of metaphor, done in the same way in Breton, French, and English. Both can additionally be considered to be part of standard Breton due to their inclusion in terminology databases: *rouedad sokial* is found in TermOfis and *lakaat e pleustr* in brezhoneg21. Again, their "international" quality may render them more acceptable due to their presence in languages other than French. It is interesting to note, however, that the definition given for *lakaat e pleustr* in brezhoneg21, which gives both French and English equivalents, does not actually use the terms *mettre en pratique* or *put into practice*, instead giving the synonyms "appliquer ... mettre en œuvre, utiliser ... apply, carry out, implement".[105] The fact that this phrase is a calque is therefore downplayed,[106] perhaps to emphasise its legitimacy as part of the Breton language.

Like the two examples in the previous paragraph, another internationally used calque found in the corpus is *republik bananezel* (< *république bananière*), which occurs among the Facebook data. Unlike the previous cases, however, this term is not recommended in dictionaries or terminology databases: TermOfis advises the use of *republik bananez*. In this case, *bananezel* is a direct translation of *bananière*, *-el* being a frequently used adjectival suffix, while *bananez* literally means 'bananas'.[107] The term found in the corpus is thus more of an exact calque from the French, while that prescribed in TermOfis more closely fits with the English version, and uses a noun adjectivally, a feature of Breton that is far less common in French. This again suggests that "international" calques are, at an official level, more tolerated than French-derived ones. As section 3.2.3 showed,

105. http://www.brezhoneg21.com/geriadurBG.php, accessed 23 Apr 18.

106. *Mettre en œuvre* is of course syntactically identical, but this at least avoids a direct translation of *pleustr* (as *pratique*).

107. Despite being semantically plural, this is the unmarked form of the noun, being one of a number of Breton words which in morphological terms follow a collective-singulative rather than singular-plural paradigm. To express the concept of a single banana it would be necessary to add a singulative suffix, giving *bananezenn*.

this appears to fit the attitude displayed by "néo-breton" sources.

As noted, calqued structures found in French alone are less frequent in the corpus, and presumably less tolerated. Two of those that do appear (*demat*, 'hello', literally 'good day', < *deiz* 'day' and *mat* 'good', cf. *bonjour*; *dazont*, 'future', literally 'to come', < *da* 'to' and *dont* 'come', cf. *avenir*) occur numerous times (five and 17 times respectively). According to Devri, they are also attested multiple times in Breton prior to 1900, appearing therefore to be highly integrated into the language, but both have been stigmatised as "néo-breton" usages in the literature (German 2007:186), indicating a negative attitude to this kind of calque among scholars. Another French-only calque that occurs fairly frequently (three times) in the corpus is *tro-spered* (< *tour d'esprit*). This term is listed in Hemon's dictionary and the Mouladurioù Hor Yezh dictionary; an online search also reveals that it appears to be widely used.[108] Unlike the two previous examples, however, *tro-spered* seems not to be attested prior to 1900: it does not appear at all in Devri, and its earliest entry in Meurgorf is from 1980. *Tro-spered* may therefore be an exception to the usual trend whereby only those French-only calques with a long history of use in Breton are tolerated in "néo-breton" usage; indeed, unlike *demat* and *dazont*, it is not even contested in the literature—despite being an apparent example of the calquing from French that is thought commonplace in "néo-breton". Other than this example, this stereotypical feature of this variety appears not to occur in the corpus, suggesting that users of Breton in these contexts are more aware of syntactic subtleties than they are typically considered.

Calquing from French syntactic structures at the sentential level—for example, SVO word order for a reason other than the topicalisation of the subject—may occur more frequently in contemporary Breton than in traditional descriptions of the language (Timm 1989:377), but this sort of syntactic investigation is outside the scope of the current lexicon-based research. We can nonetheless conclude that lexical calques, at least, are infrequent, especially those involving non-internationally-used structures, thus, in this respect, contradicting the claim that "néo-breton" relies heavily on French linguistic structure.

108. The presence of searchable Breton online is minimal, so the fact that over a thousand results are returned from a Google search for "tro-spered" (conducted 23 April 2018) indicates that its use is quite widespread.

### 4.2.2 Clippings

Clipping is a common process in French (Fridrichová 2013:1), although it does not often form new lexemes with new meanings, in many cases instead shortening an existing lexeme to provide an alternative that often becomes more widely used than the original form, particularly in more familiar registers, such as *la fac* or *les infos*. Nonetheless, clippings are examined here in order to determine whether the frequency with which they are found in French might mean that they are equally widespread in Breton. Again, if clipping is found to be a frequently used process, this may imply that French linguistic structure is of particular influence on ("néo"-) Breton.

A number of clippings can be observed in the corpus, although the majority of them are found in French and borrowed into Breton, e.g. *prepa*, *maths*, *impro*, *tele*, *kom* (< *com*), *mikro* (< *micro*). In some cases the clippings found in the corpus have established and normally uncontested Breton equivalents: *aktu* (< *actualités*), for example, is used in place of the Breton *keleier*; *diko* (< *dictionnaire*) is used rather than *geriadur*. *Aktu* and *diko* can certainly be considered non-standard, given their absence from dictionaries and terminology databases. Other French-derived clippings in the corpus are usually similarly absent from these sources, although a few (*eko* [< *éco*], *bio*, *metro*) are found in TermOfis, and *oto* (< *auto*) and *foot* also appear in TermOfis accompanied by their respective alternatives *karr* and *mell-droad*. In the case of *aktu* and *diko*, using the clipping is perhaps considered more acceptable than using the original full French word; it would be highly unusual to see *actualités* and *dictionnaire* (notwithstanding any adaptation to Breton orthography), given that *keleier* and *geriadur* are uncontested elements of the Breton lexicon. However, it is surprising to see these French-derived words at all, for the same reason; in these cases, the user was perhaps influenced by the presence of these clippings in French, and wished to use a similar register in Breton, needing to borrow from French in order to do so. These therefore appear to be examples of French influence.

Clipped versions of words that do not derive from French are rarer in the corpus: three words in this category can be identified. The first two, *jedo* ('maths') and *rouedad* ('internet'), like the French-derived clippings listed above, do not form new lexemes but

instead form shortened versions of the existing words *jedoniezh* and *kenrouedad*[109] respectively. It can be noted that in both French and English, the equivalents of both these words often occur in clipped form (*maths*, *Net*). As with the lexical calques examined above, the occurrence of equivalent clippings in other languages may make them more acceptable in Breton, although in this case it could be claimed that the fact that these words in particular are clipped is a result of influence from French linguistic structure. Clipping the Breton words, however, provides a shorter word for the concepts in question without resorting to borrowing the French clippings as in the previous examples (*maths* is attested elsewhere in the corpus). As this vocabulary is Celtic-derived, but its use in a clipped form echoes the equivalent process in French, it appears to fit the "néo-breton" stereotype of avoiding French lexical borrowings but relying on borrowed structural features.

The third clipping in this category is *egor* ('space'). Unlike the others, this is a true case of word-formation, in that the clipping has a distinct meaning not related to the source lexeme. *Egor* appears to be absent from Breton prior to 1900: it does not appear in Devri at all, nor is it present in the available sources of dialectal Breton. Given that *egor* has no cognate in modern Welsh, it appears that *egor* is derived from *egorant* ('immense') by means of clipping, perhaps involving the analysis of *-ant* as an adjectival suffix, which could be a result of French influence (*-ant* does not tend to be used as a suffix in Breton, but is a productive participial suffix in French). This therefore appears to be the one example in the corpus of a new lexeme having been formed by means of clipping, possibly with the influence of French morphology.

The example of *egor* shows that clipping can act as a process for word-formation in Breton, independent of its use in other languages. However, the fact that this occurs only once in the corpus suggests that it is not a particularly common way of forming new lexemes. Where clipping is used in other contexts, it tends to be linked to French either via the clipped word itself, often a borrowing from French, or through the choice to clip those words that have clipped equivalents in French, indicating that French influence is at work here. It should also be noted that the use of clippings in French, and indeed in English, typically occurs in less formal registers, and the corpus appears

109. *Rouedad* already exists as a separate lexeme, but here is used in the specific sense 'internet' in the same way as equivalent terms in other languages, and so can be considered a clipping.

to follow the same pattern, with the largest number of clippings occurring in the radio subcorpus and the smallest number occurring in the print subcorpus. Clippings in Breton thus fulfil the same function as in French, connoting a more familiar style of language in the same way.

### 4.2.3 Semantic expansion

Section 3.3.3 considered semantic expansion in terms of general trends; here, it will be examined with respect to some specific examples. Semantic expansion typically occurs through the use of metaphor. For example, *distagañ* (from *di-*, a negative prefix, and *stagañ*, 'attach') was originally used to mean 'detach'. This came to mean 'declaim' through the metaphorical association with pronouncing words clearly and distinctly, i.e. detached from one another. In modern Breton it has evolved further and is typically used to mean 'pronounce', hence the derived noun *distagadur* ('pronunciation'). This sort of metaphorical semantic expansion is cross-linguistically common (McMahon 1994:182). "Néo-breton" has been criticised for its use of this sort of lexical expansion process (Rottet 2014:214; Hewitt 2016:16; see also *yezh* as a supposed case of semantic expansion above), but it is hard to see why this is the case, given that it is a common method for lexical expansion in all languages.

This scholarly opposition is based on the assertion that semantic expansion is a stereotypically "néo-breton" process in that it relies on the existing lexicon of Breton rather than accepting borrowings. This process can again be examined with the aim of identifying any influence from French on a structural level.

In the same way as the calques examined above, numerous examples of semantic expansion occur "internationally", i.e. with parallels in both French and English, as well as other languages. *Kreizenn*, for example, is defined in Le Gonidec and Villemarqué's dictionary as "centre, le point du milieu d'un cercle, d'un globe" (Le Gonidec 1850:218), corresponding to the French word *centre*, as defined in *Le Robert* (1989:438): "point tel que tous les points d'une figure soient symétriques deux à deux par rapport à lui". However, in the corpus, *kreizenn* occurs multiple times with a different meaning, more equivalent to a later *Robert* definition for *centre*, "lieu

où se concentrent certaines activités" (*Le Robert* 1989:439). *Kreizenn* thus appears to have developed additional meanings in a way that parallels French and English *centre*; users of "néo-breton" do not seem to avoid this similarity to French. A similar process has occurred in the case of *ezel*, defined by Le Gonidec and Villemarqué as "membre, partie extérieure du corps de l'animal, comme le pied, la main" (Le Gonidec 1850:310), but which appears frequently in the corpus with the meaning 'member' as in 'part of an organisation'.

As with previous cases, the fact that these semantic expansions have international equivalents may mean that they are seen as more acceptable. However, examples where the parallel is restricted to French can also be identified in the corpus. *Tonenn* ('tonne') is one such example. Devri gives the definition "tonne (recipient)" and a single attestation in Grégoire de Rostrenen's 1732 dictionary; historically, therefore, *tonenn* can be seen to have been equivalent to 'tonne' in the sense of 'large barrel'. In the corpus, however, it occurs in the phrase "dek tonenn eoul-maen ba'r mor" ('ten tonnes of petroleum in the sea'). The use of *tonenn* as a measurement is also advised in all the modern dictionaries and terminology databases that were examined for the purpose of this research, i.e. Hemon's dictionary, the Mouladurioù Hor Yezh dictionary, TermOfis, and brezhoneg21.[110] In this case, therefore, the semantic expansion appears to have occurred in parallel with French only, and yet, despite this influence from French that cannot be attributed to international usages, is not stigmatised by purist sources such as Hemon's dictionary.

Semantic expansion that mirrors French, while perhaps seeming an easy target for criticism, thus seems to escape this in the literature. While some examples, such as *distagañ*, appear to be specific to Breton, it has indeed been seen that French is a model in most cases, particularly when an international equivalent is available. This process thus resembles a way of avoiding French borrowings on the surface, but is more influenced by French at a structural level. Again, this fits with the "néo-breton" stereotype whereby French influence is outwardly avoided but is present in the linguistic structure.

110. It does not appear in *Brezhoneg ... buan hag aes*, but this is unsurprising given that this publication lists only basic vocabulary.

### 4.2.4 Affixation

As with semantic expansion, affixation was considered on a broad level as part of chapter 3, while this section will examine it in terms of specific examples. As noted in section 3.5.2.2, most of the affixes found in the corpus have Celtic origins; those derived from French tend to be restricted to use with elements that are also derived from French. An exception to this is the suffix *-ach* (< French *-age*), which occurs once in the corpus in the word *brezelach*, found among the Facebook data. This word is used in a comment on the then minister of defence in the French government, who is referred to as minister of *brezelachoù*. The derogatory tone of the overall post, as well as the ironic use of *brezel* ('war') rather than standard *difenn* ('defence'),[111] indicates that *-ach* functions as a pejorative suffix here: *brezelachoù* could be translated as something like 'pathetic little wars'. While *-age* does not have this pejorative meaning in French (a comparable suffix in French might be *-aille* or *-âtre*), this use in Breton clearly does bear such a connotation. The use of a French-derived suffix undoubtedly strengthens this dismissive tone, playing on the fact that French is typically avoided in "néo-breton", and conveying a nuance that would otherwise be less elegantly expressed.

The rest of this section will consider the more numerous Breton-derived affixes, concentrating on suffixes as these are more relevant to the questions at hand. These can be divided into two categories: firstly, we can consider those that change the part of speech of the word they are applied to. These typically come from a small stock: for example, *-añ* and *-iñ* for verbs (e.g. *votiñ*, 'to vote', < French *vote*), and *-el* and *-us* for adjectives (e.g. *prezidantel*, 'presidential'; *dedennus*, 'interesting', < *dedenn*, 'to attract'). Some affixes, as well as changing word class, have specific semantic properties: *-aat* works in the same way as French *-iser*, so *resisaat* (< *resis*, 'precise') means 'to make precise'. These suffixes can be applied broadly in semantically relevant contexts, enabling their use with borrowings to result in words such as *vakañsiñ* ('to holiday', < *vacances*) and *adlibiañ* ('to ad-lib'); these are both found in the corpus. This is thus a productive way of expanding the Breton lexicon that also serves to adapt borrowings to Breton morphology, making them able to inflect in a way that fits with the language's general morphological structure. In cases where Breton suffixes of this type are ap-

111. The Breton translation given by TermOfis for *ministre de la défense* is *ministr an difenn*.

plied to Celtic elements, rather than using an existing French borrowing, however, this is subject to criticism, being viewed as unnecessarily purist (see section 1.4.2); again, though, it is a cross-linguistically common process.

The second category of suffix also serves to adapt borrowed elements to Breton morphology, but without changing word class. This type of affixation creates lexemes that are more morphologically and semantically complex, taking Breton beyond the lexical fields to which traditional speakers are supposedly restricted, and hence are particularly subject to criticism from scholars. Examples in the corpus are numerous, and include *brezhoneger* ('Breton speaker', < *brezhoneg* 'Breton' and *-er*, an agentive suffix), and *prederouriezh* ('philosophy', < *preder* 'thought', *-our*, an agentive suffix, and *-iezh*, a suffix used to denote abstract nouns). Examples involving borrowed elements include *ekonomiezh*—cases of this type are less stigmatised.[112]

In some cases, using multiple affixes makes it possible to express a fairly complicated concept within a single word. TermOfis gives *kengevredigezh* as the term for *fédération d'associations*, which can be decomposed as follows:

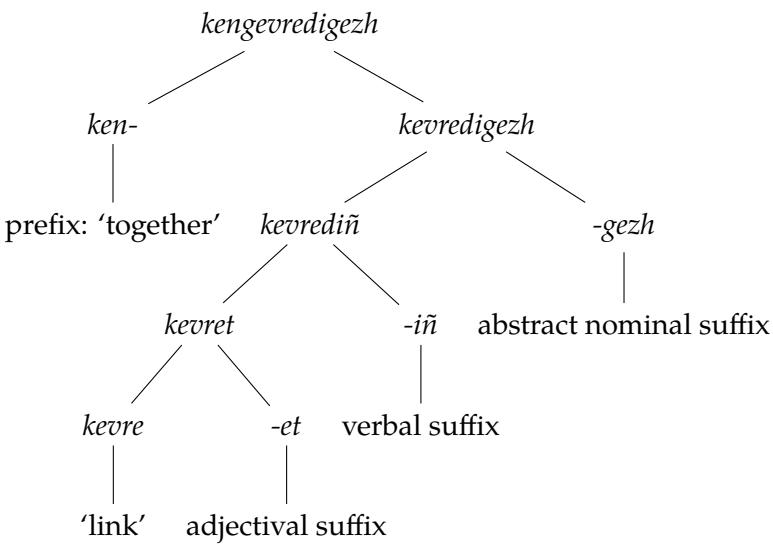


Figure 4.1: Morphological decomposition of *kengevredigezh*

It is rare to find a translation of a French term that condenses a phrase into a single

112. For example, Timm (2001:456) contrasts *pourmenadenn*, from French *promenade* plus the Breton suffix *-enn*, with *baleadenn*, from Breton *bale* ('to walk'), assigning *baleadenn* to "néo-breton" and *pourmenadenn* to traditional Breton. However, see section 4.3.5 for further discussion of this example.

word in this way: most of the translations for similar terms in TermOfis translate each word individually. Indeed, a way of translating French terms into Breton that avoids affixation and the creation of new lexemes is the use of phrasal equivalents for a single word, such as *saver tiez* ('erector of houses') for 'architect' (Rottet 2014:236), which is in this respect effectively the inverse of *kengevredigezh*. Words that use multiple affixes in this way have of course been criticised by scholars: one example is "néo-breton" *skriverezerez*, a word supposedly recommended for 'secretary' (Hornsby 2009b:76), which adds four suffixes to the verbal root *skriv-* ('to write'). However, *skriverezerez* is not found in modern dictionaries and terminology databases: both TermOfis and the Mouladurioù Hor Yezh dictionary recommend the French/international borrowing *sekretour*.[113] This is perhaps a sign that "néo-breton" (in its most official form) has more recently gained increased tolerance of French borrowings.

*Skriverezerez* has been noted as an example of excessive suffixation, but remains a case where each suffix contributes a particular semantic addition. Some examples in the corpus, however, appear to apply unnecessary suffixes. *Brallerezh* occurs twice in the Facebook subcorpus (produced by different speakers, but part of the same conversational exchange). No other examples of the word can be found online, nor in any of the published sources examined for the current research. In its two occurrences in the corpus, it appears to mean 'movement', and is composed of the root *brall* plus *-erezh*, a suffix used for abstract nouns. However, *brall* equally means 'movement': it is unclear why the addition of *-erezh* has occurred in this instance, unless it is a ludic suffix to give an air of pomposity (this phenomenon will be discussed further below). Similarly, *jahinerezh* occurs in the print subcorpus:

> *... lod Euskariz a zalc'h soñj eus reuzidi ar strolladoù damsoudardel, ar bac'hadurioù tidek, ar jahinerezh a zo bet da goulz ar renad Franco ha war-lerc'h.*
>
> 'some Basques remember the bad times of paramilitary groups, arbitrary imprisonments, the torture [= *jahinerezh*] that occurred at the time of Franco's reign and afterwards.'

113. This is not quite a just comparison, as the final *-ez* on *skriverezerez* is a feminine suffix and so the word specifically means a *female* secretary, for which *sekretourez* would be used. This still does not go as far as *skriverezerez*, however, whose multiplicity of suffixes provoked Hornsby's correspondent to decry it as a "monstrosity" (Hornsby 2009b:86).

Again, the abstract nominal suffix *-erezh* is used here, along with the root word *jahin*. However, similarly to *brallerezh*, *jahin* already has the meaning 'torture', so it is unclear why the additional suffix is required. Unlike *brallerezh*, however, *jahinerezh* is attested elsewhere, including published sources such as *Vocabulaire du breton d'aujourd'hui* (Gouélian 2002:102). TermOfis condones both words equally, giving *jahin*(*erezh*).

More examples can be identified. In another in the print subcorpus, the words *inuktituteg* and *swahilieg* are used, making use of the suffix *-eg*, which is often employed for names of languages.[114] In this case, TermOfis makes different recommendations, giving *inuktitut* and *swahili*. While the addition of *-eg* means that the borrowed words conform better with Breton morphological patterns, this is irrelevant here as Breton nouns inflect only for number and gender, and as abstract nouns this is typically invariable.[115]

Affixation thus appears to be used in various ways in modern Breton, sometimes contributing nuances that would otherwise be difficult to express. As this section has discussed, it is similar to semantic expansion in that both are identified as supposedly "néo-breton" practices, despite being common across languages. This demonstrates how minoritised languages can be subjected to a level of scrutiny that is not typically applied to major world languages, further jeopardising their already fragile position.

### 4.2.5 Ludic uses of language

*Brezelach* was noted in the previous section as a specific type of borrowing from French that conveys a dismissive attitude. This particular use of non-standard language functions as a "newspeak"-like type of wordplay, ironically substituting *difenn* for its antonym *brezel* as well as adding the pejorative suffix. Similarly playful uses of language can be observed elsewhere in the Facebook subcorpus, where users constantly demonstrate their high level of linguistic self-awareness in the use of wordplay and the discussion of others' lexical choices. This can result in word formation, and this section will highlight some instances of these practices.

114. *-eg* is the nominal form, while *-ek* is the adjectival form.
115. Except, perhaps, in very specific contexts: for example, it is not impossible that a comment on regional varieties of Swahili would read "There are many Swahilis ...".

Wordplay based on names is a recurring feature within the Facebook subcorpus; two such examples play on the names of politicians. Firstly, Jean-Yves Le Drian, minister for Brittany at the time the post was written, is rendered in Breton as "Ifig an Drian". *Ifig* is not a straightforward Breton translation of Yves (*Erwan* might be more usual), but uses the diminutive suffix *-ig*, indicating an affective reaction to Le Drian: the remainder of the post implies that in this case the diminutive serves to express contempt, similarly to *brezelach*. As a second example, we find *hamoneg*, used to mean something like 'in the linguistic style of Benoît Hamon' (a politician from Brittany and at the time the *Parti socialiste* candidate for presidency of France). This word resembles *hemoneg*, a term used as an informal descriptor of the *peurunvan* orthography, now the most commonly used way of spelling Breton, and indeed an explicit comparison is drawn between the two words by the writer of the post. It can be noted that mention of orthography is somewhat taboo in the other two subcorpora, given that it has the potential to reopen a debate of extremely long standing that also entails accusations that some parts of the Breton revitalisation movement in the early twentieth century supported collaboration during the second world war, as previously noted (see Wmffre 2008). Again, therefore, this play on words, difficult to translate, has a mocking function.

A further example of ludic word-formation based on names is the word *marsialig*. This occurs in response to a user who posts a link to Devri, which was initially compiled by the Breton linguist Martial (*Marsial* in Breton) Ménard. *Marsialig* is used as part of a response that parodies a well-known Breton song: in translation, it reads 'formerly there were plastic magic tricks and now there is the *marsialig* dictionary'. It can be noted that *marsialek* would be a more expected spelling, as this would make use of the suffix *-ek*, the adjectival equivalent of the *-eg* seen in *hamoneg*, and in the previously noted *swahilieg* and *inuktituteg*, which normally refers to language: *marsialek* could hence be translated as '(in) Marsial's language'. Using *-ig* instead, however, has two functions: to rhyme with *plastik*[116] earlier in the same sentence (following the rhyme scheme of the parodied song) and to invoke the diminutive suffix *-ig*, seen above in the example of *Ifig an Drian*.

Other examples of ludic language use are less linked to processes of word formation,

116. Final voiced plosives are devoiced in Breton.

but are also worth mentioning at this point. Exchanges in other languages sometimes occur in the Facebook subcorpus: as section 4.4.2 will suggest, these often occur for ludic effect. Orthography is also played with: in one case, the word *hiperbrezhoneK* is spelt with a final capital K, thus highlighting the frequent use of this grapheme in Breton compared with French, and so rendering this word, meaning 'hyper-Breton', "hyper-Breton" itself. Similarly, we find the term *Facebouc'h*—a version of *Facebook* respelt to include the Breton word *bouc'h* ('goat', < French *bouc*), for no discernible reason other than wordplay.

Frequently in the Facebook subcorpus words are borrowed from French and respelt to match Breton spelling conventions: for example, *seksist* and *racist*.[117] While these examples may not be ludic themselves, there are certain cases where this is taken to the extreme, particularly that of *drapo* (< French *drapeau*). The borrowing of *drapo* is certainly not done in order to avoid using a Celtic-derived neologism as a result of an anti-"néo-breton" position, as the standard Breton word for 'flag' is *banniel*, itself an early borrowing from French and an uncontested part of the Breton lexicon. It is also clearly not done to avoid the use of technical vocabulary, as the usual word for 'flag' cannot be classed as such. The unusual occurrence of this word in Breton is most likely, therefore, to be intended for ludic effect, perhaps mocking other participants' use of French borrowings and hence alluding to a more stereotypically "néo-breton"-aligned identity on the part of the speaker.

These ludic uses of language illustrate the highly self-aware and metalinguistic content of the Facebook subcorpus. By manipulating language in these ways, users are able to show their high level of competence in Breton, using it with a function that goes beyond the merely communicative, but becomes performative, emphasising their identities as highly competent Breton speakers in a way that counteracts insecurities about the authenticity of new speaker language. This reinforces the supposition that minoritised language use often occurs for reasons that have just as much to do with identity as they do with communication. In this context, where users are able to use Breton in a manner that expresses their personal identity, rather than as part of an organisation, it

117. Technically, the grapheme <c> does not exist in Breton, so this should strictly be spelt <rasist>. The removal of the final <e> from French *raciste*, however, indicates that some degree of respelling has occurred.

makes sense that we find such examples of wordplay.

### 4.2.6 Summary

This section has looked in particular at the topic of structural influence on "néo-breton" from French, showing that this is indeed evident in the case of clippings and semantic expansion, which tend to follow French and international patterns. Calques, however, appear not to be subject to such influence. This shows that the dominance of French clearly has effects on the structure of "néo-breton", but that this is perhaps not as widespread as some scholars have suggested. Investigating these phenomena has also highlighted how criticism of "néo-breton" can hold the language to perhaps unrealistically high standards: both affixation and semantic expansion are stigmatised in the literature despite being cross-linguistically common ways of increasing a language's lexicon. Denying Breton these mechanisms in a way restricts the language's vitality by removing a way for it to adapt to new lexical fields; as noted above, it seems odd that Breton should not be allowed to make use of prevalent word-formation processes that are generally permitted without comment in major world languages—particular attention is typically paid to the lexicon of revitalised languages,[118] making this especially contentious despite potentially restricting the way these languages can develop.

The processes examined in this section in some cases offer ways of making subtle semantic distinctions, such as in some of the cases of affixation mentioned, and are also able to act as expressions of the speaker's identity, particularly in the cases of ludic language use cited in section 4.2.5. This contributes to a rich linguistic variety that enables the expression of disparate concepts and identities. The variety of ways of using the processes observed indicates again that "néo-breton" cannot be seen as a fully standardised monolith, and that different speakers engage with the language in ways that reinforce their own identities.

118. A recent example of this, although relating to borrowings rather than internal processes, was the media reaction to Catalan singer Rosalía's use of Spanish loanwords in a song released in the summer of 2019: see https://www.theguardian.com/world/2019/jul/12/rosalias-spanishisms-upset-catalonias-language-purists, accessed 13 Aug 19.

## 4.3 Specific categories and individual words of interest

While tagging the corpus, some words emerged that were of interest in ways that cannot be investigated through large-scale analysis. This section presents a selection of those words, showing how they fit with the patterns already identified in the corpus data. For most of the words examined in this section, sample sizes are very small: while generalisations cannot be made on this basis, they nonetheless raise salient points.

### 4.3.1 Discourse markers

Discourse markers, a category comprising both pragmatic particles with semantic content and semantically empty fillers,[119] presented some difficulties in terms of statistical analysis due to their occasional non-language-specific characteristics. For example, the same sound can be rendered in English as *huh* or in French as *hein*, these in effect being language-specific spellings of a non-language-specific "word" or sound. In a context where the origin of a word in a specific language is a principal way of categorising and analysing tokens, it would be impossible to tag these utterances in this way, particularly when any language-specific attribute that may be assigned to them depends on the way they are written down or transcribed. For this reason, instances of *hein* in the radio subcorpus, as well as cases of similar sounds such as *hm* and *ah*, were not tagged.

The case of *setu* is discussed in section 3.3.3, where it is noted that its use as a discourse marker can be interpreted as a case of semantic expansion following a French model. Of the other discourse markers in the corpus, many are shared with French: sometimes these are non-language-specific markers that can be spelt according to French conventions, as in the case of *hein* above. Other such examples in the corpus include *ba*(*h*), *bof*, *brr*, and *opala* (< French *hop là*). The pragmatic nature of these words means that they do not tend to appear in dictionaries: out of this list, *opala* is the only one that is likely to be recognised as a well-established part of Breton vocabulary, to the extent that it is

119. These should not be considered binary categories but instead two ends of a continuum. For example, Beeching (2007:148) points out that *quoi* used as a discourse marker is "virtually desemanticised": over time, it has progressively lost semantic content as its discursive use has increased. However, for some speakers, and in some utterances, a degree of semantic content will remain present in this use of *quoi*.

now the name of a Breton literary magazine (in the variant spelling *hopala*) and is also used as part of Facebook's Breton interface.

Other discourse markers in the corpus borrowed from French include the pragmatic particles *donc* and *voilà*, and the fillers *fin*, *bon*,[120] *beñ* (< French *ben* or *bien*), and *kwa* (< *quoi*). Again, these tend not to appear in more formal linguistic registers and are consequently not found in historical corpora or dictionaries. They are joined by a number of discourse markers that originate in Breton: the pragmatic particles *eta* ('so'), *arsa* ('well'), and *kea* ('go') and the fillers *hat*, *hug*, and *ata*.

Given the complications caused by tagging discourse markers discussed at the beginning of this section, a statistical analysis of their distribution would be less accurate than similar analyses of other tokens. Certain fillers were tagged in written contexts if they could be interpreted as conveying something about the writer's attitude towards borrowed and coined vocabulary, but were not tagged in the radio subcorpus because the lack of orthographic representation meant that they did not convey such information (as in the case of *hein* above). The statistics around the use of discourse markers are therefore presented with the caveat that in comparison with the other two subcorpora, the number of discourse markers in the radio subcorpus is likely to be under-reported. This means that while the tagging indicates a large number of discourse markers in the radio subcorpus, there would in fact be even more by certain measures.

152 discourse markers are tagged in the corpus: 72.4% are derived from French and the other 27.6% derived from Breton (going by the definitions given in section 3.2.3). This is the inverse of the usual trend whereby the tagged tokens come more often from Breton than from French. In this case, variation among the subcorpora is particularly noticeable: even though the presence of discourse markers in the radio subcorpus is likely to have been under-reported, it contains the vast majority of them: 83.8% of tokens, based on normalised sizes for the subcorpora. 15.6% of them occur in the Facebook subcorpus, and just 0.6%, i.e. one single token, in the print subcorpus. This shows that their occurrence is linked to medium and register, with the spoken medium and lower registers more likely to contain them, these being the contexts in which utterances are

120. Like *quoi*, *bon* is an example of an originally semantically meaningful word that has lost this semantic content in its transition to discourse marker: see Beeching (2009).

most spontaneous.

Within the radio subcorpus, there is little variation between public and associative programmes. However, there are nearly twice as many discourse markers in feature programmes than in news programmes: 66.4% of such tokens in the radio subcorpus occur in feature programmes, with the other 33.6% in news programmes, based on a normalised sample size. This, again, can be accounted for with reference to scripted versus spontaneous speech: as the news programmes mostly consist of presenters reading a bulletin, they are less likely to be speaking spontaneously, while the higher prevalence of interviews in the feature programmes leads to more spontaneous speech and hence a higher rate of discourse markers.

Based on the fact that discourse markers occur overwhelmingly in the radio subcorpus, it may seem reasonable to assume that the large number of discourse markers that are derived from French could be due to the fact that most of them occur in the radio subcorpus, where, as section 3.2.3 showed, the tagged tokens are more likely to derive from French than in the other subcorpora. This would imply that in the Facebook subcorpus more discourse markers should derive from Breton; however, only 56.0% of them do. While this is over half, it is not skewed as far as the overall data in chapter 3, and shows that discourse markers do tend to be more likely to come from French regardless of medium, even in the generally more stereotypically "néo-breton" context of Facebook.

This acceptance of French-derived discourse markers may reflect the fact that in the majority of cases they act as semantically empty fillers, even when derived from existing language-specific words (e.g. *fin* and *beñ*). The example of *kwa* (< *quoi*) is particularly interesting here: in the radio sample used for piloting transcription for this research, one speaker used *per* as a Breton-derived replacement for *quoi*, either coining this word as an equivalent based on the fact that Breton interrogative words tend to begin with *p*, or using a dialectal version of *petra* ('what') as a calque of *quoi*. This case of what appears to be extreme "néo-breton" avoidance of French *quoi* did not occur in the eventual corpus, however. It appears that *quoi/kwa* is typically considered sufficiently semantically empty to be used in Breton, despite the fact that it is phonetically identical

to French *quoi*.

### 4.3.2 Idiom and swearing

The use of Breton-specific idiom can in effect be considered the inverse of calquing, as it demonstrates a knowledge of which linguistic structures are unique to particular languages, as well as, perhaps, a desire to maintain this distinction. Use of Breton-specific idiom can therefore be considered a less stereotypically "néo-breton" practice. Indeed, such idiom appears restricted in the corpus: two examples occur frequently. The first is the phrase *mont ar maout gant*, which literally translates as 'the ram go with' and is used to denote winners of competitions. A canonical example occurs in the sample from *Ya*:

> *Aet eo ar maout 13 gwech gante*
> 'They won 13 times' (literally 'the ram went with them 13 times')

This idiom originates in a traditional form of Breton wrestling, called *gouren*, where winners traditionally received a ram as a prize. In the corpus, it occurs seven times, restricted to the radio and print subcorpora, although this is an effect of subject matter, as these are the only subcorpora where sports are discussed.

Five of the seven examples include the word *mont* ('to go'), while the other two deviate from this pattern:

> Radio (news programme, associative station): *emañ ivez ar maout gant Brest*
> 'Brest has also won' (literally 'the ram is also with Brest')

> Print (*Ya*): *Kavet o deus c'hoarierien ar Mor-Bihan an nerzh evit kavout en-dro ar maout*
> 'The players of Morbihan have found the strength to win again' (literally '... to have the ram again')

Given that the canonical use of this idiom includes the verb *mont*, it might initially seem that these two examples attest a use of the idiom that deviates from traditional

speaker practices. However, examples of traditional speakers using *maout* with this meaning without the verb *mont* have been recorded: a recording from the *Banque sonore des dialectes bretons*[121] presents a speaker who says "neuze e oa ar maout deoc'h": literally 'then the ram was to you', i.e. 'then you had the ram'. While the expression came to be best known as *mont ar maout gant*, therefore, this was not its only form. Expressions produced by new speakers that use this idiom without *mont* are hence not at odds with traditional speaker practices, but instead replicate their diversity. In fact, traditional Breton uses *maout* even more liberally than occurs in the corpus.[122]

The second idiom from traditional Breton found in the corpus is *mont d'an Anaon*, which is used as a euphemism meaning 'die' and literally means 'go to the spirits'. This figure of speech occurs five times in the corpus: once among the Facebook data and four times in the print sources. The verb *mervel* ('to die'), on the other hand, does not appear at all in the corpus, suggesting that *mont d'an Anaon* has in effect taken the place of *mervel* in the way that euphemistic expressions are often used in preference to their more direct equivalents.

*Mont d'an Anaon* sometimes occurs without the definite article (i.e. in the form *mont da Anaon*), and this is the case for two of the five instances in the corpus, one on Facebook and the other in the *Brud Nevez* sample. The examples with the definite article occur in *Bremañ* and *Ya*. Having seen that these two publications tend to use a more standardised form of Breton, we may expect *mont da Anaon* to be a non-standard or traditional-Breton-aligned version. However, it also appears in the Mouladurioù Hor Yezh dictionary, and conversely, all the pre-1900 examples in Devri are of *mont d'an Anaon*. This variation appears therefore not to signify alignment with more or less standard Breton.

Occurrences of swearing in the corpus can also be examined briefly. Unsurprisingly, these appear exclusively within the Facebook subcorpus. The most well-known obscenity in Breton is *gast*—the equivalent of *putain*—which has become so celebrated as to become a signifier of Breton identity, featuring on souvenirs from the region.[123]

121. http://banque.sonore.breton.free.fr, accessed 14 Jun 18, segment PYKAD-29274-LM-0068.
122. See for example http://banque.sonore.breton.free.fr, accessed 14 Jun 18, segment JP-29262-EQ-0116.
123. *Gast* is probably one of the most well-known of all Breton words, reflected in its use in the headline of a national news article about Breton vocabulary: (https://www.nouvelobs.com/rue89/rue89-rue89-culture/20100427.RUE6264/gast-le-telegramme-propose-un-lexique-breton-du-net.html, accessed 15

However, it does not appear in the corpus, perhaps because it has become something of a stereotype, associated with the image of Breton-speaking peasants or *ploucs*,[124] i.e. a pejorative view of traditional speakers as backward. The high frequency of *gast* also no doubt means that it has lost some of its semantic force and no longer has the emotional impact or shock value with which swearing is associated.

There are three examples of swearing in the corpus, all of which use borrowed forms of French *foutre*. In one case, this occurs as *foeltr*, an early borrowing from French, which evidently took place before the complete velarisation of syllable-final /l/. It is likely that this word actually derives from *foudre* rather than *foutre*, as when it stands alone, it typically means 'lightning'. However, the phrase *foeltr-forzh* involves the use of *foeltr* with the meaning of *foutre* and thus could have a different origin, or could be a result of speakers reanalysing *foeltr* as related to *foutre* due to the proximity of *foutre* and *foudre*. *Ober foeltr-forzh* is defined in Devri as "n'en avoir rien à foutre", and it is in the phrase *foeltr-forzh* that *foeltr* appears in the corpus.

The other two examples of swearing in the corpus use more recently borrowed forms of *foutre*: once in its uninflected form (as *foutr*) and once as the past participle *foutu*, borrowed into Breton with no orthographic modification. The fact that both words keep their French morphology—the addition of Breton inflection would perhaps have given us **foutrañ* and **foutet*—shows that they are not considered particularly integrated into Breton by their users. To swear in Breton, then, speakers appear to rely on borrowings, and perhaps especially on recent or perhaps even one-off borrowings that are unlikely to be considered fully assimilated parts of Breton vocabulary. Swearing is perhaps not considered part of the linguistic repertoire available in Breton, which may make borrowing from French more acceptable.

Returning to idiom, this appears similarly restricted in the corpus. While *mont ar maout gant* and *mont d'an Anaon* both occur fairly frequently, these are the only two examples of idiom attested. Both originate in traditional Breton, but the fact that they are the only instances of this in the corpus suggests that the range of traditional Breton idioms that have made their way into the language of new speakers is limited; other traditional

Jun 18).

124. The word *plouc* is a derogatory term for 'peasant' that derives from the Breton word for 'parish', *plou*.

expressions may not be known to new speakers. However, it has been noted that both examples are used in various ways, rather than with restricted syntactic patterns, indicating that they have escaped complete standardisation as part of the creation of "néo-breton". Moreover, speakers do not appear to calque French idioms, showing an awareness of what is language-specific, and, as observed in the case of calques, not imposing French syntactic structures on Breton, contrary to some writers' claims (e.g. Madeg 2011:64). It is undeniable, however, that in terms of idioms and swearing, the diversity of expressions present in traditional Breton does not appear to be replicated in the language of new speakers. As with discourse markers, this may be a product of their typical absence from dictionaries and pedagogical materials, as well as the fact that they will not occur naturally to speakers for whom Breton is a second language.

### 4.3.3 *biken james*

The phrase *biken james* appears once in the corpus, and is an interesting illustration of the nuances available in Breton as a result of borrowing. The phrase is attested in Breton prior to the revitalisation of the language, appearing in a religious work published in 1727.[125] Both *biken* and *james* have the same meaning, 'never'; *james* derives from French *jamais*, while *biken* is a Celtic word. The use of two words for 'never' thus allows an intensifying effect, similar to English 'never ever'. *Holl-razh* is a similar case, also occurring once in the corpus. In this case, both components mean 'all', but *holl* is the standard Breton or KLT word, while *razh* is used in Vannetais.

These examples show how the use of borrowed words can help express a particular nuance, similarly to the cases of ludic word-formation discussed in section 4.2.5. Examples of this type show that borrowed words and Celtic words need not replace each other, but instead that the two can be used in conjunction with an intensifying effect. The combination of multiple linguistic repertoires thus allows a greater level of expressivity. Section 4.4 will return to this idea.

125. This publication, *Pedennou hac instructionou christen evit servichout de heuryou Brezonec ha Latin*, is among those used as sources for Devri.

### 4.3.4 *publik* and *foran*

The word *publik* serves as a conspicuous example of the lexical variation possible in Breton. The choice to use this obvious borrowing in the name of the *OPLB* (*Ofis publik ar brezhoneg* in Breton), as well, indeed, as *ofis*, identifies this official institution with a more relaxed stance towards the use of borrowings than is commonly assumed among those who have been involved in Breton language planning. In naming itself in this way the *OPLB* perhaps made a conscious decision to align itself with this more tolerant ideology, seeking to position itself as an organisation concerned with speakers outside the "néo-bretonnant" category as well as those within it.

French terms involving the word *public* in French are translated in a number of ways in TermOfis. Over 250 of the terms in the database are concerned here, most with Breton equivalents that also include a specific word functioning as a translation of *public*. Nine different words or phrases are used in this capacity, the most frequent being *publik* (137 uses) and *foran* (85 uses). These are both borrowings from French, *publik* perhaps more transparently so: *forain* is no longer common in French, tending to be used in a more restricted sense. *Foran*'s less obvious link with French may therefore render it more acceptable in "néo-breton".

In earlier traditional Breton there appears to have been a small semantic difference between *publik* and *foran*. Devri glosses *publik* simply as "public", while *foran* is glossed as "public, ouvert à tous"; this definition comes from a dictionary published in 1872. In TermOfis, a distinction of meaning between *publik* and *foran* is retained: *foran* retains the meaning it is given in Devri, while *publik* is used to mean 'public' in the sense of 'related to the state'. This seems to be a change of meaning from older Breton, as in older texts, *publik* appears to have the same broader meaning as French *public*. The potential semantic overlap between *foran* and some uses of *publik* in earlier Breton thus seems to have been removed in TermOfis, which restricts the meaning of *publik* so that it fulfils only those functions not fulfilled by *foran*. This preserves the lack of one-to-one mapping between French and Breton lexemes, but alters the meaning of the existing Breton words to bring about what could be considered a more efficient relationship between word and meaning in the case of *publik*, removing its polysemy.

Not all affected terms in TermOfis obey this distinction, however: there appear to be some inconsistencies. The terms *fonction publique* and *administration publique* are both translated as *melestradurezh publik*, indicating, as expected, that this means 'public' in the sense of 'managed by the state'. However, the category in which these terms are listed is named "melestradurezh foran", which seems to use *foran* in a way that contradicts its advised use on TermOfis. This term is also used for the translation of *plan de normalisation linguistique des administrations publiques*, which is given as *steuñv normalaat ar yezh er melestradurezhioù foran*. The presence of this sort of variation is testament to the fact that as TermOfis is incremented gradually, by a number of terminology researchers, different opinions and decisions over time may lead to divergent lexical choices. This is more reflective of actual language use than an entirely invariant word list would be: the current study has shown that variation is indeed prevalent within modern Breton. However, speakers looking for prescriptive advice may be confused to see these terms used inconsistently. If TermOfis wishes to encourage variation, it would perhaps be more useful to explain the difference between *publik* and *foran* explicitly and provide a range of options for translating the relevant terms, rather than, apparently arbitrarily, assigning one of two options to each case.

When French *public* is used as a noun, other words are given as Breton equivalents in TermOfis. Four terms translate French *grand public* or *tout public* as *an holl*, meaning 'all [people]', e.g. *espace tout public* = *lec'h an holl*. Terms relating to sports and the arts often use *arvesterien*, which literally means 'spectators': in fact, the "arts" and "sports" categories of TermOfis give *arvesterien* as a translation for *public* as a standalone term. Another term used multiple times in the database, particularly in the category of terms relating to infrastructure, is *tud diavaez*, literally 'people from outside', used in terms such as *moned berzhet ouzh an dud diavaez* (French: *accès interdit au public*). Similarly, in other contexts, *tud* ('people') on its own is used for some terms whose French equivalents include *public*: for example, *accueil du public* is listed as *degemer an dud*.

This variety of translations for *public* shows that the compilers of TermOfis are sensitive to the existence of differences in lexicosemantic correspondence in French and Breton in this case. *Public* has not undergone the same type of semantic assimilation that was observed in cases of semantic expansion (see section 4.2.3), even despite the existence

of a Breton borrowing (*publik*) with clear relation to the French word. In the case of *publik* and *foran*, the different word–meaning correspondence has even been heightened as a result of its systematisation.

In the corpus, there are three examples of *foran* and three of *publik*. Their distribution follows the pattern prescribed in TermOfis with one exception, found in the radio subcorpus:

> *met dre ma vez cheñchet ar rolloù cheñchet roll ar publik ivez*
> 'but through [this] the roles are changed, the role of the public is also changed'

Here *publik* is used as a noun, as *public* would be in French; a TermOfis-conformant rendering would be *an dud*. The occurrence of this non-standard language use among the radio data is unsurprising, when this source has been seen to be the least conformant with standard/"néo"-Breton, as well as the context with the most French borrowings. In this instance, the use of *publik* as a noun shows influence from French linguistic structure and does not appear to occur in traditional Breton. It is therefore not a marker of divergence from "néo-breton" in a way that aligns the speaker with traditional Breton, and is instead likely to have occurred due to influence from French, particularly in this context where the speaker produces spontaneous oral speech. This demonstrates how the prescriptions of TermOfis may not be reflected in actual language use. The more "néo-breton" stance ignores the reality of speaker bilingualism, advocating fine points of semantic detail that may be unrealistic in this context.

### 4.3.5 *pourmen(adenn)* and *bale(adenn)*

Timm (2001:456) cites the words *pourmenadenn* and *baleadenn* as Breton equivalents of French *promenade*, assigning the former term to traditional Breton and the latter to "néo-breton". The fact that *pourmenadenn* resembles its French equivalent makes its status as a borrowed word obvious. However, Jean Le Dû (pers. comm.) points out that both words are used by traditional speakers in different parts of Brittany, and that *baleadenn* also derives from a French word, *balader*. Given that *balader* is less commonly used than

(*se*) *promener*, and the borrowing is less transparent, Timm's rationale for assigning the words to these different categories can be understood. However, her inclusion of *baleadenn* in a list of "proposed lexical innovations" (Timm 2001:456) seems erroneous given the word's appearance in traditional Breton. This case illustrates how it can be difficult or indeed impossible for linguists, especially those from outside Brittany, to gain a full understanding of the specificities of each Breton dialect, and hence hard to comment on what is "néo-breton" in a way that does not ignore the diversity of speaker practices. This was also shown in the example of *yezh*, discussed in section 3.3.3. By making such generalisations, scholars may be committing the same errors they criticise in language planners, and failing to recognise their own ideological biases.

As with *public*, we can explore this further with reference to occurrence in dictionaries. *Baleadenn* and *pourmenadenn* can be considered in relation to the verbs on which they are based, i.e. *bale* and *pourmen*, both meaning 'to walk'. The *Atlas linguistique de la Basse-Bretagne*'s entry for *marcher* (map 379) shows that among traditional speakers, *bale* is used in the western half of the Breton-speaking area, i.e. Léon and western Cornouaille, while another verb, *kerzhet*, is used in the eastern half, i.e. the rest of Cornouaille, Trégor, and the Vannetais area. Map 449 of the *Nouvel atlas linguistique de la Basse-Bretagne*, published nearly 75 years later, shows that this is still the case. Neither map contains *pourmen*, which suggests that for traditional speakers, it may not be semantically equivalent to French *marcher*.

This semantic difference is indeed implied in the Mouladurioù Hor Yezh dictionary, which glosses *bale* as both *se promener* and *marcher*, *kerzhet*[126] as *marcher*, and *pourmen* as *se promener*. This explains the absence of *pourmen* from the *ALBB* and *NALBB*. The distribution of terms suggests that *pourmen* is perhaps used to mean *se promener* in the same area that *kerzhet* is used for *marcher*, and that elsewhere, *bale* is used for both. The dictionary does not give information on the use of dialects, as this is beyond its scope. However, this arguably obscures the nuances of this vocabulary.

As *bale* is used in Léonais, the dialect most similar to standard Breton, and as this word covers the meanings of both *se promener* and *marcher*, it seems reasonable to expect it to be found more widely in the corpus than *pourmen* and *kerzhet*, and this is indeed

126. This is spelt as *kerzhout* in the dictionary, but this variation can be disregarded here.

the case. *Bale* occurs 21 times in the corpus, three of these occurrences being on the radio (uttered by a single speaker in the feature programme taken from public radio), two on Facebook, three in *Ya* and the remaining 13 in *Brud Nevez*, many of these occurring in a single article, its high frequency here related to subject matter. *Pourmen* occurs four times in the corpus, once on Facebook and three times in *Brud Nevez*, in the same article that contains the multiple instances of *bale*, while *kerzhet* occurs only once, again in the same *Brud Nevez* article. It is *Brud Nevez*, therefore, that tends to use the less standard terms *kerzhet* and *pourmen*; the other sources tend to avoid these words in favour of *bale*. This fits with *Brud Nevez*'s general alignment with more traditional forms of Breton. However, the way in which the three words are used does not quite fit with traditional Breton: we have seen that among traditional speakers *bale* tends to be used in certain areas, while *kerzhet* and probably *pourmen* are used in others. The writer of the article, on the other hand, uses all three, thus not aligning herself with any dialect in particular. This may be done as a way of avoiding repetition, given that the article discusses walking and therefore requires words meaning 'to walk' frequently; perhaps a longer discussion of walking in one of the other parts of the corpus would also have used the words *kerzhet* and *pourmen*, although as observed, the use of these less standard terms does fit with the typical practices of *Brud Nevez*. While preserving the use of these less common words, it nonetheless does so in a way that may not be considered appropriate by traditional Breton speakers or "native authenticists".

Returning to *baleadenn* and *pourmenadenn* and turning to their occurrence in the corpus, they do not fit the pattern established by *bale* and *pourmen*. Again, both terms are found in *Brud Nevez*, which contains two instances of *pourmenadenn* and one of *baleadenn*. In other parts of the corpus, however, there are two occurrences of *pourmenadenn*, spoken by different radio station employees as part of the same news programme, and no further occurrences of *baleadenn*. It is difficult to generalise based on only two instances of *pourmenadenn*, and negative evidence in the case of *baleadenn*, but perhaps the use of the former rather than the latter relates to the erroneous perception of *baleadenn* as a "néo-breton" coinage, as seen in Timm (2001:456). While there is no such stigma around *bale*, perhaps the more morphologically complex *baleadenn* is understood as a neologism by new speakers. This could have encouraged a more recent generation to avoid it in an

attempt to make their language more comprehensible to traditional speakers: this is of particular importance on the radio, as section 5.6.4 will discuss. This example shows how perceptions of words as borrowings or coinages can be more important than their actual linguistic provenance, and why it is important to acknowledge such perceptions: they are instrumental in the construction of "néo-breton".

### 4.3.6 *arvar* and *dañjer*

This is another example of words that are of different provenance but have a very similar semantic quality. The words *arvar* and *dañjer* both appear in Devri. *Dañjer* is clearly a borrowing from French and is equated in Devri to French *danger*, while *arvar* is given a wider range of meanings, including *doute* as well as *danger*. Additionally, among the pre-1900 texts indexed in Devri, *arvar* is found only in Vannetais sources. This is an unusual case, given that Vannetais tends to be more lexically influenced by French than the KLT dialects are, given its geographical location: we would normally expect to find the borrowed term in Vannetais and the Celtic term in KLT. Indeed, this is part of the rationale for Léonais having been chosen as a basis for standard Breton: as its territory is the furthest west, its vocabulary and structure have undergone the least amount of influence from French. It can therefore be considered the "purest" of the dialects by this measure, and hence the most conformant to stereotypical "néo-breton" requirements.

*Arvar* thus presents a problem for users who conform to such stereotypes. Being from Vannetais, it should be considered non-standard and avoided. However, its KLT equivalent, *dañjer*, is a borrowing, which such speakers would also supposedly wish to avoid. It is perhaps for this reason that both *dañjer* and *arvar* appear in modern standard Breton. Again, both can be found in the Mouladurioù Hor Yezh dictionary, which again glosses the two words slightly differently: *dañjer* is glossed as *danger*, while *arvar* receives the wider gloss of "doute, incertitude, danger", corresponding to its attested meanings as listed in Devri. The dialectal distinction is again not made clear.

TermOfis also includes both words, using them for different purposes as with *publik* and *foran*. *Arvar* tends to be used with the sense of "endangerment", with the result

that the phrase *en arvar* is used as an equivalent for *menacé* or *en danger*. We therefore see terms such as *yezh en arvar* for *langue en danger* and *ever diavar* for *usager à non-risque*. Elsewhere, we see *dañjer*, as in *dañjer: dour fall da evañ* (*danger : eau non potable*) and *panell dañjer* (*panneau danger*). This semantic difference becomes more obvious in the case of the derived adjectival forms *dañjerus* and *arvarus*. *Dañjerus* is used in TermOfis to mean 'dangerous', i.e. 'causing danger', while *arvarus* is used to mean 'affected by danger', i.e. 'at risk' or 'threatened'. Hence we see *lastez dañjerus* (*déchets dangereux*) but *metoù arvarus* (*milieu à risque*).

In the corpus, there is one instance of *dañjer*, in the Facebook subcorpus, and nine of *arvar*: three spoken by the same guest on public radio, three on Facebook, one in *Bremañ* and two in *Ya*. The higher rate of occurrence of *arvar* shows that this formerly dialectal word has become widespread in modern Breton, as its use in TermOfis would suggest. All but one of its occurrences fit with the pattern established by TermOfis: the other, one of the radio occurrences, appears to go against this. Likewise, the one example of *dañjer* in the corpus does not conform to the prescriptions of TermOfis, according to which we should expect *arvar*.

*Dañjerus* appears four times in the corpus: once on Facebook, twice in *Bremañ* and once in *Brud Nevez*. These all fit the definition given in TermOfis. *Arvarus* appears three times, twice in *Brud Nevez* and once in *Ya*: however, in all of these cases the word is used to mean 'dangerous', thus contradicting TermOfis' prescription.

The occurrences of *dañjer*, *arvar* and their derived adjectives in the corpus show that these words are not always used in the same way as in TermOfis and dictionaries. The dialectal quality of *arvar* appears to have been lost: none of its occurrences is produced by a Vannetais speaker, and the word seems no less likely to occur in more stereotypically "néo-breton" contexts. Its use in the corpus also shows that the neat lexicosemantic distinction between *arvar* and *dañjer* attested in TermOfis is not matched by real usage: instead: the semantic line between *arvar* and *dañjer* is more blurred than terminology databases and dictionaries can adequately convey.

### 4.3.7 *skolaer* and *skoliad*

This is a further example of two similar words whose roles have changed in modern standard Breton. *Skoliad* is in fact not attested in Meurgorf or Devri prior to 1900; it appears to have been coined in the early days of "néo-breton". The *-ad* suffix suggests it could originate in Vannetais, where this suffix is common. This would result in a situation similar to *dañjer/arvar*, where one of the pair of words, initially dialectal, has entered the standard. However, in this case, both Meurgorf and Devri contain a number of Vannetais sources and *skoliad* is not attested in any of them, so it appears not to come from Vannetais.

It is hence more probable that *skoliad* was coined in order to take on one of the two distinct meanings of *skolaer*, which appears in Meurgorf and Devri meaning both 'schoolmaster' and 'scholar' depending on the source. In Le Gonidec and Villemarqué's dictionary, it receives the latter definition, but Le Gonidec notes that Grégoire, an earlier dictionary compiler, "prétend que *skôlaer* doit signifier maître d'école, et que, pour écolier, il faut écrire *skólier*"[127] (Le Gonidec 1850:526). Le Gonidec however goes on to state that he has in fact heard *skolier* used only in the sense of 'schoolmaster'. The word does appear in the much earlier *Le Catholicon* with the meaning 'scholar'; *skolaer* does not appear in this dictionary. Evidently, the closeness in form of the two words seems likely to have led to their being mixed or conflated. *Skolier* seems not to appear other than in the three sources mentioned (i.e. *Le Catholicon*, Grégoire's dictionary, and Le Gonidec and Villemarqué's dictionary), according to Meurgorf and Devri, and is also absent from modern dictionaries, presumably having fallen out of use because of its resemblance to, and semantic confusion with, *skolaer*. In the nineteenth century, then, despite Le Gonidec's attempt to keep *skolaer* restricted to one of its two meanings, it appears that it would have been used for both.

Twentieth- and twenty-first-century dictionaries attest a second attempt at restricting the meaning of *skolaer*, confining it to 'schoolmaster' in line with Grégoire. The word is glossed this way in Hemon's dictionary, the Mouladurioù Hor Yezh dictionary, and TermOfis. While it is cognate with French *écolier*, indicating that the 'scholar' meaning

127. Diacritics have been retained from the original, although the acute and circumflex are usually no longer used in Breton.

would be etymologically more appropriate, the presence of the agentive suffix *-aer*, suggesting the meaning 'someone who conducts schooling', was perhaps instrumental in determining which of the two meanings of *skolaer* should prevail in modern standard Breton. Additionally, the purism of early "néo-breton" could have meant that the etymological connection with French was in fact deliberately broken. In order to account for 'scholar', therefore, a new word had to be introduced, hence the appearance of *skoliad*, which can be found in Hemon's dictionary, the Mouladurioù Hor Yezh dictionary, and TermOfis.

In the corpus, these words occur infrequently. Other than in the Facebook subcorpus, there is only one attestation of both *skolaer* and *skoliad*, both in *Ya*. Their use conforms with what is prescribed by modern dictionaries. Among the Facebook data, however, an extended comment thread on the subject of teaching Breton in schools results in seven instances of *skolaer*, produced by four different users, and six of *skoliad*, from three different users. In every case, the word in question matches its definition in modern standard Breton. In this case, unlike that of *dañjer* and *arvar*, new speakers of Breton seem to have accepted the distinction created in the modern standard language. As this is an example where the two Breton words can be mapped onto distinct French words, unlike in the earlier cases, this adherence to the standard may be easier for speakers for whom French is a dominant language. In other cases, where the lexicosemantic distinction does not match French, this will be less intuitive.

### 4.3.8 Summary

This section has focused on various case studies from the corpus, ranging from categories of words to individual lexemes. While the sample sizes examined in this section were generally too small to be analysed quantitatively as in chapter 3, they contributed to the overall analysis in ways that added further nuance to the patterns observed in that chapter, particularly relating to the structure of "néo-breton" and its degree of proximity to French.

Discourse markers were found to occur mostly in the radio subcorpus, unsurprisingly given the oral medium. It was additionally found that while words containing bor-

rowed elements tend to be less common among the tagged tokens (see section 3.2.3), in this category they are far more common. This appeared to relate to the semantic emptiness of such tokens, and their use in spontaneous speech, rendering the use of French borrowings both natural and acceptable.

The investigation of idioms revealed that only two examples from traditional Breton, *mont ar maout gant* and *mont d'an Anaon*, persist in the corpus. In terms of idiom, "néo-breton" therefore appears to lack much of the repertoire available to traditional speakers, reflective of a pattern whereby minoritised languages under pressure from a dominant language often lose specialist terms and expressions without close equivalents in other languages: Ó Dochartaigh (1992:31) notes this phenomenon in relation to Irish. However, there seems to be no converse attempt to calque idioms into Breton from French, a process that would easily attract criticism from some scholars due to its imposition of a French linguistic structure on Breton. While speakers therefore take care to avoid this, the result is that they lack a broad range of available idiom.

Similarly to discourse markers, borrowings tend to be more prevalent in the case of swearing. Again, this may be connected to the semantic emptiness of such words. It may be the case that no words native to Breton are considered to have the same force as swearing in French, hence necessitating such borrowings.

The example of *biken james* showed how borrowings and native terms can be combined to convey a particular nuance. The final four parts of this section, focusing on pairs of semantically similar words, also showed how words from different origins can express different nuances in "néo-breton". The cases of *pourmen* and *bale*, and particularly *dañjer* and *arvar*, dealt with the role of dialects in "néo-breton", showing how dialect words can be preserved in the modern standard. The *pourmen* and *bale* example also showed that scholars and speakers can hold beliefs about word origins that may contradict historical fact, but ultimately become more salient, showing the importance of perceptions in the construction of "néo-breton".

These sections also showed that despite attempts to regulate Breton, the definitions found in modern dictionaries and the prescriptions of TermOfis do not always match actual usage. Again, this also occurs in other languages, with language academies and

similar prescriptive bodies often having little effect on actual use: in the case of French in France, "while the AF [*Académie française*] is well known, people are far less certain what it does" (Estival and Pennycook 2011:329). Here, for three out of the four pairs of words examined, at least one instance was identified in the corpus where usage differed from what is advised in these sources; *skolaer/skoliad* was the only pair where this did not happen. The variation observed here does not reveal any systematic pattern of language change, but does highlight how variation has been present in Breton throughout the modern period.

## 4.4 The effects of multiple linguistic repertoires

The previous section noted some ways in which the structural influence of French could be at work in "néo-breton". This section will explore the role of French more explicitly. Indeed, the analysis presented up to this point focuses on the use of Breton; however, this disregards the fact that at some points, other languages are also present to a greater degree than in the occasional lexical borrowing. In contexts where a large amount of such borrowing occurs, such as the use of discourse markers discussed in section 4.3.1 above, claiming that there are clear boundaries between languages and that the text under analysis is exclusively Breton could even obscure the linguistic subtleties of the discourse. Additionally, in a context where many speakers have Breton as a second language, and all are or are presumed to be fluent in French, the specific characteristics of this multilingual situation and the way in which they inform the resulting discourse should be acknowledged. This section highlights some multilingual features of the corpus and how they fit into contemporary "néo-breton". The Facebook subcorpus will be of particular relevance here; while its heteroglossic nature meant it was difficult to examine using quantitative analysis, some of its specificities will instead come to light in this section.

### 4.4.1 Translanguaging

In this section it is relevant to discuss translanguaging, "the flexible use of linguistic resources by bilinguals in order to make sense of their worlds" (García and Leiva 2014:200). Like the concept of the new speaker, it has been applied in various contexts, and often in different ways. Here, the term translanguaging will be defined according to the work of Otheguy, García, and Reid (2015:297), who argue that "translanguaging refers to using ... one's linguistic repertoire without regard for socially and politically defined language labels or boundaries". This distances it from code-switching, references to which "still constitut[e] a theoretical endorsement of the idea that what the bilingual manipulates, however masterfully, are two separate linguistic systems" (ibid.:282).

As this part of the thesis focuses more on the level of discourse than on individual lexemes, it is appropriate to consider translanguaging at this point, given that it is a discursive phenomenon. Additionally, up to this point in the thesis, this idea of Breton and French as separate linguistic systems has been taken as the norm. Blommaert and Rampton (2011:4) argue that "named languages ... are ideological constructions historically tied to the emergence of the nation-state"; however, this ideological status does not mean that the concept of the named, bounded language should be rejected entirely. If it is considered salient in the minds of speakers, it will exert influence over their ways of using and perceiving language and thus be crucial in any analysis of those uses and perceptions. Indeed, the concept of the nation-state is of undeniable importance where the relationship between Breton and French is concerned, given that the French state is commonly seen, particularly by militant speakers, as a motor of oppression of the Bretons and their language. Likewise, the revitalisation of Breton is frequently bound up with the struggle to gain a greater level of official recognition of the idea of Breton nationhood (McDonald 1989:157). Translanguaging "attempts to wipe out the hierarchy of languaging practices that deem some more valuable than others" (García and Leiva 2014:200); however, in the Breton movement the typical strategy has not sought to "wipe out" this hierarchy but rather to reorder it, placing Breton in a position of higher importance than French.

It is therefore perhaps unsurprising that work on translanguaging tends to focus on heritage languages spoken in contexts where the various individual named languages concerned are recognised as national languages with large speaker populations; indeed, "studies of spontaneous translanguaging have mainly focused on cases of bilingual speakers who speak an additional language in English-speaking countries" (Cenoz and Gorter 2017:904).[128] Where there is a more salient power imbalance, as in the case of regional minoritised languages such as Breton and other revitalisation contexts, translanguaging's ability to "challeng[e] the view of languages as autonomous and pure" (García and Leiva 2014:204) may be undesirable: we have seen that this notion of "purity" has, on the contrary, often been viewed as a target for Breton, particularly in the earlier part of the revitalisation period, and, as the corpus shows, this appears to continue to be desirable in some of the more stereotypically "néo-breton" contexts. Blackledge and Creese (2017:250) note that translanguaging "is contingent on local attitudes to, and beliefs about, communicative practice", and given that in the case of Breton the community's attitudes have historically favoured a clear conceptual line between Breton and French for the reasons noted above, it is not surprising that translanguaging is not a common concept in studies of contemporary Breton use.

However, Creese and Blackledge (2015:20) point out that translanguaging is a characteristic feature of "late modern societies", and also emphasise the potential of "advances in digital technology" to increase the presence of translanguaging practices in communication. As a result, therefore, we may expect to identify translanguaging in the corpus, particularly among the Facebook data. As this section progresses, it will be considered whether this is the case, and whether we in fact do find any examples of "the speakers' construction and use of original and complex interrelated discursive practices that cannot be easily assigned to one or another traditional definition of a language" (Blackledge and Creese 2017:251).

128. Cenoz and Gorter (2017) focus on the minoritised language context, but specifically in an education setting. Osterkorn and Vetter (2015) offer a study of Breton in this setting with a slightly different theoretical approach, noting that "during informal moments or unobserved situations" (ibid.:125), pupils in the Diwan immersion education system mix linguistic repertoires. As this section will argue, this practice appears restricted in the corpus.

### 4.4.2 Multilingualism within the corpus

The use of languages other than Breton does indeed occur most often in the Facebook subcorpus. Despite a bilingual warning, displayed prominently in *Facebook e brezhoneg*'s header image, that advises users "kasit ho kemmenadennoù e brezhoneg mar plij ha chomit dereat ha doujus—merci de rédiger vos commentaires en breton et de rester courtois et respectueux", users frequently post in French, suggesting that while those running the group may be guided by an ideology of monolingualism that seeks to instate Breton as the sole preferred language, this belief is not universally shared by the members of the group; these ordinary members are perhaps less militant (see section 5.4) than the group's administrators. Subsequent replies to posts tend to be in the same language as that in which the original post was written, so those posts that are written in French usually attract further French replies. In some cases, users post in Breton and then include a French translation; in such instances, users tend to reply in Breton, perhaps because the Breton content of the original post precedes its French equivalent. Users can therefore be observed to treat each post and its ensuing replies as either French- or Breton-specific in most cases. This suggests that in general, users view Breton and French as separate repertoires that need not normally be mixed; this cannot be considered a case of translanguaging.

As well as French, other languages occasionally appear on Facebook, the most frequent being English. This is sometimes used as part of otherwise Breton-language conversations in instances of *expressions figées*: "no comment" and "game over" occur in the sample. Other than this, the most notable use of English in the subcorpus is where two users facetiously begin a conversation using only English. This continues over the course of several comments, during which both users, clearly not fluent English speakers, do little more than insult each other. It is unlikely that they would have done this in Breton or French, languages that for them would have a more ordinarily communicative purpose and would be understood easily by most other users of the group. The use of English is highly performative, providing the users with the opportunity to engage in a kind of linguistic posturing, appearing to attack each other personally without actually doing so in a meaningful way: it can be considered an example of the ludic language discussed in section 4.2.5.

Some of the more unusual examples of multilingualism in the Facebook subcorpus involve writing Breton using different writing systems rather than using other languages. At one point a user writes "Σalud" instead of the usual *salud* ('hello'). At another, a user writes a Breton phrase in Cyrillic, and another replies with a phrase written in Japanese katakana; in both of these cases the users provide glosses in the standard orthography immediately afterwards. As another example of ludic language, this appears to function as a way for the users to play with language and demonstrate their own linguistic competence by showing their skills in other writing systems. The playful nature of this type of language use is confirmed by the use of smilies in the second example. These examples conform more to the concept of translanguaging as discussed above: they mix repertoires in a way that break down usual interlanguage barriers. However, as they are restricted to writing systems rather than entire languages, these still cannot decisively be called examples of translanguaging.

There is some use of French in the radio subcorpus. Sometimes, it is used for links between the Breton-language programme and a preceding or following French-language segment, where the French-speaking presenter is presumably not a Breton speaker. At other times, French is used in cases where speakers appear to be unsure of words in Breton; this will be discussed in section 4.4.3.

In the print subcorpus, there is very little use of other languages. Two instances of English in this subcorpus act as glosses for place names in minoritised languages, which, as we saw in section 3.2.3, tend to be used in this subcorpus in preference to those from dominant languages. In *Ya*, we read "Balley Keeill Eoin", followed immediately by the English gloss "Saint-Johns" (sic). In *Bremañ*, more interestingly, we see the Irish "Doire" followed by the English "Londonderry", implying that all inhabitants of Northern Ireland are either Irish-speaking nationalists or English-speaking unionists. This fits with *Bremañ*'s more overtly political ideology, shown in the way in which it tends to link minoritised language use to political separatist movements in its coverage of minorities around the world: here it is implied that true Irish nationalists must be Irish speakers. The only use of French in *Bremañ*, other than in its opening pages (mentioned in section 3.2.1 above), is in a gloss, which will be explored further below. *Ya* contains no instances of French at all, despite including quotations in English, Manx

and Dothraki (an invented language used in the television series Game of Thrones); this fits with its more lenient attitude towards internationally-used vocabulary, as well as, on the other hand, its general conformance to the attitudes and practices of "néo-breton".

Despite being the least typically "néo-breton" publication of the three, *Brud Nevez* uses French only once, in a discussion of the Spanish word *espumador*, misunderstood by the husband of the article's author as being semantically identical to the related French word *écumeur*; in this case, French is used to explain the resulting confusion, in effect taking on the function of metalanguage. *Brud Nevez* also uses the word *coach* twice in a single article, both times in italics, which indicates that the word is not to be understood as part of the Breton lexicon. In this instance the author appears not to have identified a word in Breton that she found capable of expressing the semantic qualities of French/English *coach*, and has resorted to borrowing, but has indicated that the borrowed word should not be considered part of the Breton language through the use of italics. The fact that this occurs in *Brud Nevez* seems significant, as this is the least stereotypically "néo-breton" publication and hence the least concerned with purism: obvious one-off borrowings of this type will therefore be more tolerated here.

Related to the use of different linguistic repertoires is the simultaneous use of pairs of words associated with different attitudes towards "néo-breton". In *Brud Nevez*, one article uses the word *surf* throughout, but at first mention the author notes that this should be *faouta*[129] *ar mor* "e brezoneg brao"—'in good Breton'. In the same article, she uses the word *kordenn* ('rope'), and then follows this with "Digarezit ! Arabad komz euz kerdin war vourz, nemed euz 'pennou-fard', 'giz 'm-eus kavet em geriadur": 'Sorry! One should not speak of *kerdin* [plural of *kordenn*] on board ship, but of *pennou-fard*, as I found in my dictionary". Subsequently, she uses *pennfard*/*pennou-fard* throughout the article, in this respect conforming with what is prescribed in the standard language, unlike in the case of *surf*, which she continues to use after acknowledging that *faoutañ ar mor* is more standard. These explicit references to lexical choice foreground the fact that this is something that the speaker must navigate in modern Breton: in many cases they must make conscious decisions that may result in their identification as new or tra-

129. *faoutañ* in the orthography used in most of the corpus

ditional speakers. Similarly, in *Ya*, an article makes reference to "reoù 'dispartiañ'—pe baraj": '*dispartiañ* or *baraj* [matches]'. *Dispartiañ* and *baraj* are two terms for 'play-off'; the first is an adaptation of an existing Breton word, while the second is a borrowing from French. By placing *dispartiañ* within quotation marks, the author implies that *baraj* is his preferred term; *dispartiañ* is perhaps a more official word, but one that he would rather not use. However, by juxtaposing the two terms, he both conforms to the standard and subverts it, producing a text that can please users of both borrowings and coinages, as well as making readers aware of both forms and potentially facilitating their own navigation of a wider range of Breton texts.

On the radio, similarly, an announcer names "handball", but immediately follows this with the equivalent neologism "mell-dorn". Another instance in this subcorpus involves terms for the *caisse d'allocations familiales*, commonly referred to in French as *le CAF*. The speaker uses the officially prescribed Breton term, *kef goproù tiegezhel*, but then follows this with "ar c'haf". These cases of the use of doublets function as a way of ensuring that audiences understand what is being referred to: in the particular example of *le CAF*, those who do not deal with public administration in Breton are unlikely to have come across the official term. This technique can thus be used as a communicative strategy, as well as pleasing speakers who have different views on what counts as "correct" Breton. This will be discussed again in section 5.6.4.

### 4.4.3 Speakers with reduced competence

Given that it is typically assumed that most speakers who use Breton in the contexts examined are new speakers, and that the typical "néo-bretonnant" has been characterised as lacking competence (Madeg 2011:64; Hewitt 2016:16), we should expect to find examples in the corpus of speakers who are not fluent in Breton, i.e. to be unable to carry out a conversation in Breton without hesitation or to have difficulty understanding what is said to them. While this is unlikely to be the case among those specifically employed to use Breton, i.e. radio and magazine employees, it is possible that we may find such non-fluent speakers among the radio guests and Facebook users. As a result, some of the language in the corpus will be produced by non-fluent speakers, while

some will be produced by fluent speakers trying to make themselves comprehensible to those who lack full competence. It was decided not to account for this in the overall analysis of the corpus, given that there is no clear way to divide learners from fluent speakers according to linguistic ability. Additionally, if a large number of new speakers of Breton can be classed as learners or non-fluent speakers, their language variety should be considered as valid as that of fluent speakers when new speaker language is the object of investigation. Consequently, language produced by or directed at non-fluent speakers is included in the analysis as normal. This section nonetheless presents some instances of such language to highlight some of its specificities.

The largest amount of discourse that appears to come from non-fluent Breton speakers is unsurprisingly in the Facebook subcorpus, as there is no restriction on the competence of users. Many such speakers communicate in French, which was excluded from the sample in most cases. Sometimes, learners ask for translations from or into Breton; due to the overtly linguistic nature of such posts, they also tended to be excluded. However, learners' comments in French on other posts originally written in Breton can be found in the corpus, although not in great number.

In addition to this, posts and comments that are aimed at learners can be identified. Posts written in Breton with a translation into French immediately following can often be classed as part of this category, as the aim of the French translation is presumably to make the message clear to those who may not fully comprehend the Breton version. There is one instance in the corpus where a user glosses three of the words in their own comment, presumably in order to make what they have written more easily understood by the person to whom they are replying, whom they apparently understand to be a learner. Only one of the words is glossed in French, however; the others are glossed with Breton synonyms, upholding the group's policy of favouring Breton over French. A small number of French glosses are used on the radio: in a discussion of cooking using seaweed, the speaker refers to "bezhin-saout, hag a zo dulse e anv e galleg"—"*bezhin-saout* ['dulse'], which is called *dulse* in French". Here, this is clearly done because the Breton word is an item of specialist vocabulary, as in the case of *kef goproù tiegezhel* above. French also occurs in the radio subcorpus at a point where a guest uses a technical Breton word, and then checks its validity with the interviewer:

**Guest:** *fin n'eo ket skiantourien met euh tudou– tudourien* ('well, they aren't scientists but er, *tudou– tudourien*')

**Interviewer:** *mm*

**Guest:** *evit ethnologues quoi* ('for *ethnologues*, right')

**Interviewer:** *ya* ('yes')

In this situation the guest appears to perceive the interviewer as a source of authority on "correct" Breton vocabulary, probably because of her status as an employee of the radio station: she is understood to use Breton daily as part of her job. Because radio employees are frequent and institutionally sanctioned users of Breton, guests may see the language varieties used by their interviewers as target varieties, particularly if they themselves are not fluent speakers. The employees are therefore in a particularly influential position as regards type of language used.

Instances of glossing can also be found in *Bremañ*. Four words are glossed in the sample, two occurring in each of two articles. Like in the Facebook example, most of these cases use Breton synonyms: three out of the four in this instance. However, again, one of the words (*rebourserezh*) is glossed with its French equivalent (*représailles*). This is the only point in the entire print subcorpus where a French word is used to explain the meaning of a Breton word. It is interesting that this occurs in *Bremañ*, given that this publication has been seen to be that which tends to follow the more conservative "néo-breton" trend most closely in many respects, and so ought to be the most hostile to the use of French in an otherwise Breton context. *Bremañ* is also clearly not aimed at learners, as its sister publication *#brezhoneg* has this function. The four words that carry glosses were therefore evidently deemed to be particularly esoteric: other than *rebourserezh*, they are *neridigezh* ('(over)population'), *hoali* ('lifespan'), and *Ac'huberezh* ('(the) Occupation'). Each of these is glossed with a Breton phrase.

The most obvious use of non-fluent Breton in the corpus involves another radio interviewee, whose speech is more hesitant than that of other speakers. As well as using uncommon French-derived borrowings, presumably in cases where she does not know the usual Breton word, she also uses French to a greater extent as part of her discourse. The following segment illustrates both of these:

> **Interviewer:** *peseurt anv eo ar bier-se* ('what is this beer called?')
>
> **Guest:** *La Corbelle. euh ni a vev en Neizh Vran. ar vran a zo ar corbeau hag euh la corbelle a zo feminin* ('*La Corbelle*. Er, we live in Neïz Vran ['crow's nest']. The *bran* is the *corbeau* and, er, *la corbelle* is feminine')

In this case, *feminin* is a borrowing of French *féminin* that does not usually occur in Breton, and appears in this case to be used because the speaker cannot spontaneously recall the usual word for 'feminine' (*benel*). The use of *corbeau* can be related to the fact that the beer's name is in French, and the speaker is explaining the link between *corbeau* and *corbelle*. However, it is unnecessary for her to draw the explicit link between *bran* and *corbeau* in the first place, as it is unlikely that listeners would not know the meaning of *bran*; they would surely see the link between *bran* and *corbelle* without this needing to be explained further. Her recourse to French here is therefore further indication of a lack of fluency in Breton.

The speaker's use of the word *bran* as well as *corbeau* shows that she is not using *corbeau* as a replacement for *bran* in the same way that she uses *feminin* instead of *benel*: *feminin* is being used as a Breton word, while *corbeau* functions as a French gloss. As a result, *feminin* was tagged as a French borrowing and analysed along with the other tagged tokens, while *corbeau* was interpreted as French and hence excluded from the analysis.

This use of French segments and uncommon borrowings means that this speaker can be identified as the user of the most mixed variety in the entire corpus, a result of her lack of fluency. Mixing languages in this way can be interpreted as a sign of lack of competence, and thus is avoided where possible in the corpus. Exceptions to this in the Facebook subcorpus, as identified earlier, often occur with ludic intent or to show the user's linguistic abilities; here, however, communication is more important, and in a situation where listeners will understand both Breton and French, the speaker is able to employ both languages without the risk of not being understood. The inclusion of her speech in the radio programme accordingly bestows legitimacy upon her language variety, making it clear that this particular media platform is accepting of learners and users of non-standard varieties. While the use of mixed varieties ought to be taboo according to the ideologies associated with "néo-breton"—and indeed appears to be

so in much of the corpus—it appears in this instance that inclusivity is more important: an ideological approach that promotes the expansion of the speaker community wins out over those ideologies more usually considered to prevail in such contexts.

### 4.4.4 Summary

This section showed that languages other than Breton typically occur infrequently in the corpus, other than among the Facebook data; however, even in this context Breton and other languages tend to be kept apart rather than mixed as part of the same conversation. Facebook might be expected to be a context where languages are mixed more freely, due to the lack of editorial control on the platform and the fact that the community is not a Breton-only space in practice. However, section 4.4.2 argues that Facebook cannot be considered a locus of translanguaging; Breton and French are typically considered distinct repertoires and kept apart. The closest we come to translanguaging in the corpus is in fact the example of the non-fluent speaker on the radio, whose speech shows evidence of French ways of shaping and perceiving language: however, the variety she uses would be considered undesirable by fluent speakers. Translanguaging therefore does not seem to occur as a rule in the case of Breton, despite the bilingual nature of the speech community. As suggested in section 4.4.1, this is likely to be because of the imbalance in power relations between Breton and French and the persistence, to some extent, of a purist ideology.

The use of different languages and writing systems on Facebook allows users to show off their linguistic abilities and gain status as members of a group where language is a principal topic. However, when users doing this include glosses, this allows them to retain comprehensibility for those who are not familiar with the language or writing system they are using, thus striking a balance between communication and the expression of identity, as with the use of dialects seen in chapter 3. This focus on ease of communication can also be seen in the (less frequent) use of glosses or extended translations in posts that may be aimed at learners.

In the print subcorpus, other languages are used very little, and French is one of the least used despite being the most likely to be understood, mirroring the sources' gen-

eral preference for internationally-used vocabulary, as befits "néo-breton". *Brud Nevez* is the only magazine to use a French word other than in a gloss, fitting with its stance that is typically more tolerant of borrowings, although even here this happens only once.

Occasionally in the corpus we see the simultaneous use of two synonymous terms associated with different varieties of Breton, e.g. *handball/mell-dorn*. As with the use of glosses, this enables what is said to be understood by a wider audience, creating a more inclusive atmosphere. In cases where the speaker self-deprecatingly disparages their own use of Breton in comparison with the standard, as in the cases of *surf/faoutañ ar mor* and *kordenn/pennfard* in *Brud Nevez*, this can help reassure speakers of non-standard varieties that their variety is legitimate and create a sense of familiarity and comradeship with the speaker (in this case, writer) in question. As noted above, a similar atmosphere of inclusivity and encouragement can be interpreted in the use of glosses elsewhere, helping those whose Breton is not entirely fluent to understand more difficult items of vocabulary. When these glosses take the form of Breton synonyms rather than French words, they not only conform to the "néo-breton" ideology of avoiding the use of French, but also encourage learners to continue making the effort to prioritise Breton. Learners are also made to feel welcome by the occasional inclusion of non-fluent speakers on the radio, who are still able to make themselves understood without difficulties because of the bilingualism of the community: such speakers can meet their communicative needs by using strategies that involve the use of French. In a community where speaker numbers are at risk, these efforts to encourage participation from learners are crucial, and form an exception to the general avoidance of French as a strategy for facilitating the inclusion of non-fluent speakers.

## 4.5 Conclusion

This chapter examined aspects of the data less easily quantifiable than those investigated in chapter 3, showing some ways in which they reinforce the patterns already established in that chapter while also elucidating some additional features of the corpus. As in chapter 3, differences were observed among and within the subcorpora, and some

of the distinctive characteristics of the Facebook subcorpus in particular were highlighted. This was the only context in which ludic language and the use of non-Latin writing systems occurred, showing that users of this medium are highly performative, employing these strategies as a way of demonstrating their linguistic skill in this less formal, largely unmoderated environment.

Many of the features examined in this chapter emphasised the effect of community bilingualism in the form of the use of French and influence from French structural features. The general avoidance of French in the corpus can be associated with the beliefs and practices typically associated with "néo-breton". Semantic expansion and clippings were seen to occur mostly in parallel with French, again adhering to the "néo-breton" stereotype in terms of the influence of French on a structural level. The frequent occurrence of discourse markers borrowed from French can also be attributed to this, although it can equally be used to assert that the type of Breton found in the corpus does *not* conform to "néo-breton" stereotypes, as this involves lexical borrowing from French. This example highlights how the category of "néo-breton" itself, founded on stereotypes and contradictions, is somewhat unstable: the example of *baleadenn* and *pourmenadenn* also attests to this.

Other features were clearer indications of difference from stereotypical "néo-breton": the avoidance of French idiom and non-international calques showed that at this particular level, influence from French structure is minimal in the corpus. A departure from prescribed standards was also noted in the cases of pairs of semantically similar lexemes, indicating that contemporary Breton displays more variation and less rigid adherence to the standard than some work has suggested, and indicating that prescriptive sources may be paid little heed by speakers.

Another point raised in this chapter was the inclusivity of the contexts examined, vital for maintaining the diverse speaker community. While Facebook users were seen to show off their linguistic skills with the use of ludic language and multiple linguistic repertoires, this did not mean that this was an environment where the participation of learners was discouraged: despite the warning that posts should be in Breton, those in French were seen to be accepted by users, and instances were observed of users taking

care to gloss words for those they perceived as learners.[130] The acceptance of those with weaker linguistic skills was also seen in the use of glosses in *Bremañ* and the occasional inclusion of non-fluent speakers on the radio. In some respects, this defies the stereotype of militant "néo-bretonnant" refusal to use French, showing that communication and comprehension are crucial. However, it also lends weight to the idea that many speakers of "néo-breton" are learners who speak Breton imperfectly (Hewitt 2014). In lexical terms, however, it is difficult to draw a clear line between learner varieties and non-standard varieties in general if, as we have observed, learners employ strategies that involve large amounts of borrowing from French, which is claimed in the literature to be a feature of the language of traditional speakers. If large amounts of French borrowing are characteristic of the language of learners, and many speakers of "néo-breton" are learners, the idea that "néo-breton" speakers generally avoid French borrowing is untenable: again, the contradictions built into the "néo-bretonnant" category are exposed.

130. Indeed, the inclusiveness of this online context is notable, given that social media are known for heated arguments and abuse, and that the history of the Breton revitalisation movement is certainly not immune from this itself (see Wmffre 2008). Some of the uses of ludic language could be seen as examples of "trolling", but they are relatively few and tend not to involve insults or abuse; even when they do, the primary purpose of such content is evidently wordplay and linguistic posturing, such as in the exchange in English mentioned in section 4.4.2.

# Chapter 5

# Attitudes and practices among employees of Breton-language media

## 5.1 Introduction

Chapters 3 and 4 concentrate on findings from the corpus, highlighting attributes of the lexical features of Breton in the media. This chapter focuses instead on the interviews carried out with employees of Breton-language media, broadening the picture revealed so far with the addition of sociolinguistic information not so easily accessible from the corpus alone, and thus responding in particular to research question 2, also contributing material relevant to questions 3 and 4. The following final chapter will bring together data from both the corpus and the interviews in the form of conclusions.

This chapter is structured around the major themes discussed in the interviews, and is composed mostly of a full descriptive account. Firstly, section 5.2 explores the speaker profiles of the participants. Their acquisition and use of Breton is discussed in section 5.3. Section 5.4 discusses the term "militant" and how it links to the concept of the "néo-bretonnant". The operation of the media organisations is described in section 5.5, followed by a discussion of how lexical variation and dialects are seen by the in-

terviewees and their organisations in section 5.6. The conclusion of the chapter will identify patterns observed in the interviews.

## 5.2 Speaker profiles

Participants were asked various demographic questions in conformance with the methodology developed by the MEITS project, of which this thesis is a part: section 2.3.4 discusses the consequences of using this cross-project methodology for measuring participants' identity.

As stated in section 2.3, nine interviews were carried out, eight of which were with employees of the media organisations whose output is examined in the earlier parts of this thesis: two employees of public and four of associative radio, and one each from *Brud Nevez* and *Bremañ*. The ninth interview was with an employee of the *OPLB*.

Quotations from participants will be marked with codes denoting the specific interviewee. These codes are presented in Table 5.1, along with a summary of the speaker profiles, which will be elucidated more thoroughly below.

| Code | Organisation | Gender | Age group | Region | French identity | Breton identity | Acquisition |
|---|---|---|---|---|---|---|---|
| P1 | Public radio | M | > 55 | West[131] | 0 | 1 | Family and community |
| P2 | Public radio | M | < 40 | West | 0.65 | 0.99 | School |
| P3 | Associative radio | F | < 40 | West | 0.03 | 0.97 | Immersion course |
| P4 | Associative radio | F | < 40 | West | 0.48 | 1 | Immersion course |
| P5 | Associative radio | F | < 40 | West | 0.52 | 0.99 | Family and school |
| P6 | Associative radio | F | < 40 | West | 0.68 | 0.99 | Immersion course |
| P7 | *Bremañ* | M | < 40 | East | 0.37 | 0.97 | Family and school |
| P8 | *Brud Nevez* | M | > 55 | West | 0.48 | 0.48 | Family and community |
| P9 | *OPLB* | M | > 55 | East | 0 | 1 | Books and evening classes |

Table 5.1: Summary of speaker profiles

Participants were evenly distributed in terms of gender: five were male and four female. They spanned a wide range of ages, with the youngest being 25 and the oldest 72. However, these were not distributed consistently: in effect, they can be put into two groups, with six participants under forty (25, 28, 31, 33, and two aged 37), and the other three over 55 (59, 69, and 72), leaving a gap of over twenty years in between. By chance, this split into two groups mirrors the new/traditional speaker distinction, and is reminiscent of the claim that these two types of speaker come from disparate age groups, with a "missing" generation in the middle (Jones 1998b:321; Adkins 2013:58). Indeed, the interviewees in the younger category were all born after the establishment of the first

131. In this table, "west" refers to Basse-Bretagne and "east" to Haute-Bretagne.

Diwan school in 1977, which suggests that there may be a relationship between the greater number of speakers born after this point working for the media and the first steps towards making Breton properly available in education and consequently more visible in society. However, only one of these six interviewees was educated at Diwan,[132] and it would be dangerous to make generalisations about the age of the Breton-speaking population based on this very small sample. Additionally, it is not the case that the three older speakers can easily be called "traditional speakers" based on their acquisition of Breton: for many of the participants, including some of the younger interviewees, the situation is much more complicated, as later sections will discuss. This case highlights the difficulty of defining the new speaker (see section 1.4.1): different research contexts have tended to use a range of definitions, as pointed out by O'Rourke and Pujolar (2019:18). All the participants, however, are affected by the new speaker paradigm in that they use Breton in a context where struggles regarding legitimate and authentic language play out as a result of tensions between the supposed authority of the "native" traditional speaker and any speaker denied access to this category. This use of the new speaker concept "as a social category ... subject to social negotiation and variation" (ibid.), rather than as an analytical category tightly defined by linguistic criteria, matches the recommendations made by O'Rourke and Pujolar (ibid.), as to do otherwise "could potentially lead to the misuse of academic research in the delegitimation or exclusion of speakers" (ibid.:18), as can be seen in the stigma and stereotypes attributed to the "néo-bretonnants".

Seven of the interviews took place in Basse-Bretagne, the territory in which Breton is traditionally spoken, while the remaining two occurred in Rennes, in the eastern part of the region, known as Haute-Bretagne. However, both the interviewees spoken to in Rennes had familial connections with western Brittany. One worked in Rennes but lived further west, and had spent much of his childhood in Basse-Bretagne; the other had lived near Rennes all his life, but his parents both came from Finistère and he had made frequent visits there as a child. While it has been claimed that younger speakers of Breton and "néo-bretonnants" need not have these connections with Basse-Bretagne, their choice to learn Breton instead being motivated by a sense of pan-Breton identity (Jones 1998b:321), none of the speakers interviewed can thus really be cited as

132. Two others attended bilingual schools, however.

an example of this phenomenon, as eight out of nine had these links to the western part of the region. The ninth interviewee, however, had acquired Breton in childhood from her father, who himself had come from Haute-Bretagne; *he* therefore was an example of a person who had chosen to learn Breton despite a lack of connections with Basse-Bretagne or Breton-speaking relatives, although of course he did not feature among the sample of participants.

Among the interviewees who consciously chose to learn Breton, this tended to proceed from family or community connections, as the next section will relate. Of course, it is impossible to generalise based on this extremely small sample—and exceptions such as the father cited above can be identified—but it is interesting to note that the role of Basse-Bretagne and traditional speakers from that part of the region is still an important factor for younger speakers.

Most interviewees had lived in Brittany all their lives, although all had moved around the region. One had spent part of his early childhood in Dijon, and three of the other younger participants had spent time outside Brittany as part of their education: one had gained a master's degree in Tours, while the other two had lived abroad, one in Northern Ireland, and the other in Canada. All undertook the majority of their university studies in Brittany, however. It can additionally be noted here that all interviewees had attended university, matching the stereotype of the well-educated "néo-bretonnant" (Jones 1995:428; Timm 2001:454).

As described in section 2.3.4, interviewees were asked to mark their feelings of Breton and French identity on a line, corresponding with the cross-strand methodology adopted by the MEITS project. Use of terms such as "identité nationale/régionale" was avoided in favour of the questions "à quel point vous sentez-vous breton.ne/français.e ?", in order not to imply that this question was trying to elicit interviewees' opinions on the constitutional status of Brittany or on Breton independence. Each participant was given a piece of paper as reproduced in Appendix E. The position of participants' marks on each line was then measured and divided by the length of the line to obtain a numeric measure of identity (with 0 equating to "pas du tout breton.ne/pas du tout français.e" and 1 equating to "entièrement

breton.ne/entièrement français.e"). Each interviewee's response to the two questions can therefore be plotted on a graph with French and Breton identity each mapped onto one of the two axes. For a participant self-reporting as 0.75 for French identity and 0.25 for Breton identity, we would see the following:

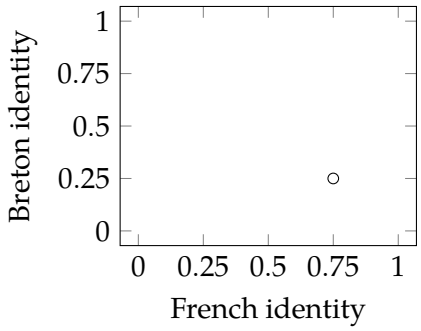


Figure 5.1: Example graph of Breton and French identity

In the event, participants reported their identities as shown on this graph:

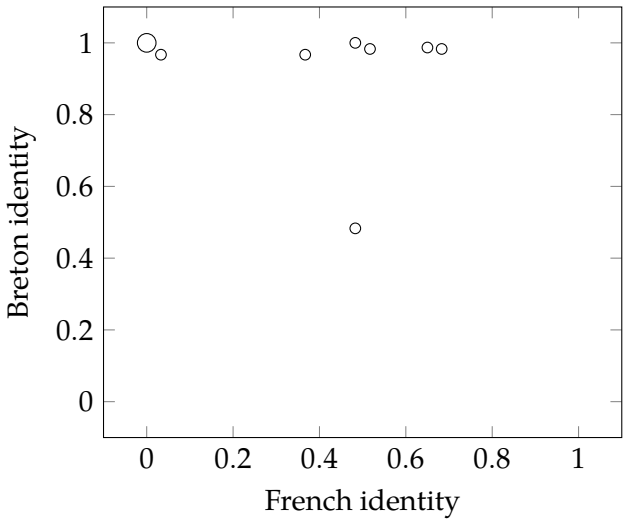


Figure 5.2: Participants' Breton and French identity

As the graph indicates, most participants rated themselves highly for Breton identity, while ratings for French identity varied, although none was higher than 0.7. The larger mark in the top left corner represents two participants who both rated themselves at 1 for Breton identity and 0 for French identity. They were both among the older group of participants, with the third member of this group marking himself at around 0.5 on

both axes. This may imply that among the older group of speakers, identity of this type is considered in more black and white terms: one can be entirely Breton and not at all French, or (presumably) vice versa; or instead, one's identity can be made up of both in equal measure, this being interpreted by the participant as a rating of around 0.5 for both, in order to sum to 1 in total. For most of the younger speakers, identity seems to be more nuanced, with Breton identity remaining close to 1 (although not usually as high as 1), while French identity clusters around the middle of the scale. This perhaps reflects a less rigid separation of French and Breton identity on the part of the younger participants, acknowledging that both can be experienced simultaneously and interact with each other.

Some interviewees noted that the strength of their Breton identity was heightened by their frequent use of Breton, demonstrating that this feeling of identity is fluid and can evolve over time. One noted that before gaining employment in Breton-language media and consequently making the effort to improve his linguistic skills, he had not felt such a strong sense of Breton identity. Another mentioned that having learnt Breton in early adulthood, she no longer had to prove or justify her Breton identity:

> **P4** depuis que je parle breton mon identité bretonne a beaucoup changé, a beaucoup évolué. J'ai beaucoup moins besoin de la revendiquer ... de ne pas parler breton, c'était comme si j'étais obligée d'en faire beaucoup [trails off]
>
> **MDD** De faire la preuve que tu[133] étais bretonne ?
>
> **P4** Oui c'est ça

Claiming that it is the strength of participants' Breton identity that inspired them to learn the language would therefore be an oversimplification. Instead, the salience of that identity will most probably have increased as the participants found themselves using more and more Breton in their working lives. Rather than identity being a determining factor for language choice in this instance, then, it is more appropriate to claim that both these factors exert mutual influence on each other, echoing the approach

133. *Vous* was used during the interviews by default, but some participants suggested using *tu*.

taken by third-wave sociolinguistic research, which "places speakers not as passive and stable carriers of dialect, but as stylistic agents, tailoring linguistic styles in ongoing and lifelong projects of self-construction and differentiation" (Eckert 2012:97–98).

Having enumerated some of the basic elements of these speaker profiles, we can consider how these data intersect with the supposed demographics of the "néo-bretonnant" population. All participants have been seen to be well-educated and mobile, and all are literate in Breton, characteristics that would typically place them in this category. The distribution of ages also mirrors the supposed new/traditional speaker divide, although this would place the three older speakers into the traditional speaker category, which contradicts their status as educated, literate Breton speakers who use the language in contexts wider than their own local communities. Likewise, determining whether participants are "néo-bretonnants" based on their use of standard or dialectal language is also difficult, as many of them are hard to classify one way or the other: this will be discussed in more depth below. Based on these attributes, the speakers cannot firmly be assigned to one or the other of the two categories, as many of them occupy a somewhat contradictory position. The principal factor that differentiates new from traditional speakers is of course the method by which they acquired Breton, however, which will be discussed in the following section.

## 5.3 Interviewees' acquisition and use of Breton

### 5.3.1 Acquisition of Breton

As stated in the previous section, all but one of the interviewees had connections with the traditionally Breton-speaking part of Brittany. Out of the nine participants, six had at least one Breton-speaking parent. In five cases, these were traditional speakers who had themselves grown up with Breton in the home; the sixth, mentioned in the previous section, had learnt the language as an adult. The other three participants all had at least one Breton-speaking grandparent. With the exception of the interviewee whose father had learnt Breton in adulthood, whose family came from Haute-Bretagne, no participant was thus more than two generations removed from traditional Breton.

Despite six interviewees having a Breton-speaking parent, only two were raised with Breton as a main language from early childhood. One of these cases was the participant whose father who had himself acquired Breton in adulthood: this fits the stereotypical characterisation of "neo-bretonnants" as more motivated and hence more determined to pass Breton on to their own children. In the other case, both parents were traditional speakers of Breton, and used only Breton in the home during the interviewee's childhood. In terms of acquisition, then, this participant must be called a traditional speaker. However, he was also educated in the bilingual school system, an attribute more commonly assigned to "neo-bretonnants". Again, if we go by the categories of new and traditional speaker posited in the literature (see section 1.4.2), different elements of the participants' upbringing present conflicting evidence about which category they fit into, and show that the stereotypical clusters of attributes often do not in fact necessarily occur in conjunction.

In the other four cases where participants had at least one traditional speaker parent, these parents did not set out to use the home as a locus for the transmission of Breton. In three out of four cases, however, once the participant made the conscious decision to acquire and use Breton as a teenager, the parents and other family members began using more Breton with them. One participant suggested a number of reasons that his mother did not speak Breton to him during his childhood:

> **P2** il y a le fait que mon père ne le parlait pas. ... Il y a eu aussi le fait que nous on apprenait le breton à l'école, le breton du Léon, donc du nord-Finistère, et elle, elle parle breton Pourlet, de centre-Bretagne, et ... on considérait que ce n'était pas le même breton, donc elle ne pouvait pas communiquer. Elle ne voulait pas, je pense, un peu nous pénaliser à l'école en nous parlant un autre breton. ... Et en plus il y a eu une histoire familiale assez compliquée ; il y avait beaucoup d'autres choses à s'en occuper au niveau de la santé de mes frères, et donc le breton finalement et le combat pour la langue sont passés un petit peu à côté

The remaining participants in this category formed the older group of interviewees. Two grew up in Basse-Bretagne and were able to acquire Breton through hearing it in

the community, both from their parents and from others. One noted that as the oldest child, he gained more exposure to Breton than did any of his siblings, who acquired only "une connaissance passive" (**P8**). The second explained that initially his parents spoke to him and his brother in French, but he was nonetheless able to acquire Breton easily:

> **P1** nos parents exprimaient entre eux en breton. La vie de la communauté, c'était essentiellement du breton ... mais ils nous parlaient en français. En fait, on est devenus bilingues très vite parce que beignés dans un [milieu où] les gens, on les entendait parler breton ... on a appris en même temps les deux [langues], et un jour moi j'ai décidé ... j'étais en quatrième ou en troisième ... j'avais dit « on est capable de parler breton », ... et après on n'a utilisé que le breton

The third of the older speakers, however, grew up in Haute-Bretagne where this Breton-speaking community was not present; his parents were not in the habit of using Breton in the presence of him and his siblings. Like the participant quoted above, he turned towards Breton in his early teens, but lacking the community environment had to rely initially on self-teaching from books, then on a minimal level of provision at school, and finally on night classes, where he was able to perfect his knowledge of the language: "vers 17, 18 ans j'arrivais à m'exprimer à peu près correctement en breton" (**P9**).

All interviewees also encountered Breton in formal education. For the three older speakers, provision of Breton in schools was poor at the time, but all had a minimal amount of exposure to the language in secondary school. The same was true for some of the younger participants, one of whom emphasised the lack of time dedicated to Breton in the French state school system:

> **MDD** tu a appris le breton à l'école comme langue étrangère ?
>
> **P3** Oui, c'était même moins que ça. C'était en initiation, une heure par semaine

For three of the younger interviewees, therefore, the acquisition of Breton chiefly oc-

curred in adulthood. The other three attended schools where Breton was used as a medium of instruction. One underwent secondary education at a Diwan school until brevet level, while two more were educated at bilingual schools in the state education system. Interestingly, two of these three participants were those who grew up with Breton consciously transmitted in the home. Again, they therefore do not fit the stereotypical perception of the pupil in immersion education as acquiring Breton only in the context of formal education. In fact, all the participants who attended schools that used a degree of Breton-medium education had at least one parent who spoke Breton.

It has already been seen that among the interviewees who were exposed to Breton as children, their continued use of the language as adults stemmed in many cases from a conscious decision made as teenagers; indeed, on a larger scale, "it is widely accepted that the teenage years are a critical period in the development of young people's attitudes towards a minority language" (Cunliffe, Morris, and Prys 2013:341). Similarly, the three participants for whom the acquisition of Breton mostly occurred in adulthood appear to have undergone an equivalent experience at a similarly formative moment in their lives, this being the beginning of adulthood, around the time of attending university. One interviewee in this category described how she had attended a course in Breton the summer before commencing her university studies. While she had originally planned to study another subject at university, she enjoyed the Breton course so much that she made the late decision to change the subject of her degree and continue with Breton in the university setting. The second participant in this group explained how she made the decision to learn Breton after finishing her degree, characterising this as a crucial moment for acquiring the language:

> **P4** à ce moment-là c'était vraiment un besoin, et je me suis dit qu'en finissant mes études ce serait le seul moment où j'aurais l'occasion et assez de temps devant moi pour le faire

She therefore decided to attend a six-month intensive course, which took her from no knowledge of Breton to a level of competence sufficient to gain work as a radio presenter. For the third member of this category, learning Breton began at night classes while undertaking a university degree, and after this, she similarly attended an exten-

ded immersion course, in this case of nine months' duration. For all three of these participants, therefore, the acquisition of Breton happened in early adulthood, at a point when these speakers were likely to have been consolidating and renegotiating their Breton identity as they began their adult lives. Similarly, one of the two participants educated bilingually recounted how he had consciously decided to use Breton more in his everyday life in early adulthood.

> **P2** [j'ai] des amis d'enfance avec qui j'étais à l'école et avec qui je jouais justement quand j'étais enfant plutôt en français, et bien maintenant notre langue c'est le breton. ... Donc avec eux aussi j'ai basculé ... et j'ai des amis qui apprenent le breton aussi, et donc en fait, c'est assez rigolo mais mon environnement s'est bretonnisé.
>
> **MDD** Donc avec tes amis vous avez décidé de changer consciemment ?
>
> **P2** Oui, c'était conscient

For many participants, then, speaking Breton was the result of these conscious choices, made at a formative period in their lives. This can be related to the concept of the linguistic *muda*, typically used in the context of new speakers of Catalan to denote "how ... native speakers of Spanish become users of Catalan at specific moments of their lives" (Pujolar and Puigdevall 2015:171). While there are differences from the Catalan context in that the latter is a more visibly bilingual wider society, the concept of a particular moment at which the *muda* occurs seems to be applicable across languages, emphasising how certain events during a speaker's lifetime, typically involving broader changes such as starting a new school, job, or a family (ibid.:172), can also be catalysts for a change relating to language choice. Additionally, the potentially huge consequences of making this decision for the individual, such as bringing one's children up to speak Breton (see section 5.3.2), changing degree course, and indeed, choosing to work in a Breton-speaking environment, emphasise the common discourse among minoritised language speakers and activists that language revitalisation "is a moral duty" (Costa 2017:36), something that requires significant personal changes for the good of a "common human heritage" (ibid.:34).

Alongside these overarching moral concerns, some participants mentioned a more pro-

saic reason for developing an interest in Breton: that their parents or grandparents spoke it secretly and they wanted to know what was being said. Perhaps ironically, then, the fact that older traditional speakers were ashamed of their use of Breton and made the effort not to use it around younger members of the family was a factor that led to these younger relatives wanting to speak the language themselves.

For this younger group, despite Breton having largely having ceased to be spoken as a community language, a greater number of formal ways of acquiring Breton were available. Diwan and bilingual schools had not been established when the older participants were growing up, meaning that for those who could not benefit from being surrounded by a Breton-speaking community, acquisition depended on a high level of self-motivation. When the younger interviewees attended school, these options were more widespread, and for those still unable to benefit from Diwan and similar systems, Breton could be acquired in adulthood through the immersion courses that three of the participants attended. These courses often have a practical focus on preparing attendees for jobs that require the use of Breton (Adkins 2014:110), making them ideal methods of acquisition for the participants in this study at the point of beginning their professional careers.

As the present section discusses participants' acquisition of Breton, and this is typically considered the most fundamental criterion for defining speakers as "new" or "traditional" (see section 1.4.1), it is worth exploring at this point whether the participants can easily be assigned to these categories based on the information already discussed, in order to probe the validity of this binary. Based solely on transmission, one of the interviewees can certainly be classed as a traditional speaker, having grown up with parents who used only Breton with him in the home. Four of them can be called definite new speakers: three of these had parents who did not speak Breton, and learnt the language as adults, while the fourth did have parents who spoke the language, but nonetheless had to acquire it from books and later from night classes.

For the remaining four participants, the lines between new and traditional speaker are more blurred. Three of them had at least one traditional speaker parent, but these parents did not begin to prioritise the use of Breton with the participant until a point later

than early childhood, once the participant had already made the decision to use Breton actively. One of these three could perhaps be referred to as a new speaker due to the fact that his initial acquisition of Breton occurred through school, but the other two initially acquired the language in the community around them. The fourth interviewee in this more nebulous category did learn Breton in the home, but it was from a parent who was himself a new speaker. Speakers of this type are not typically prominent in work that seeks to examine the new speaker specifically, and receive very little discussion in research on Breton such as that discussed in section 1.4. They have nonetheless been cited in the literature in the wider field of minoritised language research, often characterised as "neo-native speakers" and appearing in multiple language contexts including Cornish (Payton 2000; Sayers 2012:101), Hawaiian (Wilson 1998:325), Irish (Ó Giollagáin 2004:74), and Manx (Thomson 2000).

While these speakers' method of acquisition is similar to that of traditional speakers in that they acquire the language in the home during childhood, the fact that they acquire it from new speakers means that the language they speak will be linguistically closer to new speaker varieties. Such speakers will thus have native-speaker intuitions about the language, but these may be different from those of traditional speakers as a result of a different type of language being acquired. As the revitalisation of Breton continues, more and more speakers of this type will emerge in this particular language context, meaning they must be accounted for in descriptions of the sociolinguistic state of the language. This is an example of a situation that the typical binary opposition fails to describe adequately. Going by the working definition of the new speaker stated in section 1.4.1, this type of speaker meets the criteria for the new speaker category because the intergenerational transmission is not the result of a continuous chain, but will lack some of the characteristics of the "néo-bretonnant", particularly in terms of linguistic influence from French.

The difficulty of assigning even this small sample of speakers to the new or traditional speaker categories for the reasons discussed in this section thus indicates a difficulty inherent in this categorisation: as also pointed out by O'Rourke and Pujolar (2019), relying on a rigid framework for defining such categories ignores the fact that they are in fact fluid and contested, and speak more to notions of perceived competence and

social capital than to facts about language acquisition or linguistic practices.

### 5.3.2 Participants' use of Breton today

All nine interviewees evidently use Breton to a large extent today, as it is the language in which they work. However, most admitted that they also use a large amount of French in the workplace, whether for administration, liaising with external suppliers, or communicating with colleagues: this last example was given by an employee of a radio station that broadcasts some programmes in French, and where some employees are consequently not Breton speakers. The interviewee from the *OPLB* claimed however to use only Breton in the workplace. He additionally stated that he has numerous Breton-speaking contacts among his local community, despite living in the Rennes area: "je pars faire un jogging au bord du canal, je rencontre une autre personne qui parle breton, on se parle en breton, voilà, il y a plein d'occasions comme ça" (**P9**). This perhaps surprising assertion no doubt reflects this participant's somewhat closer alignment to the stereotypical characteristics of the "néo-bretonnant". His emphasis on the ease of finding Breton speakers even in Haute-Bretagne reflects his ideological position that Breton should be used in all possible domains, and his desire for "une vraie société bretonne avec des gens qu'on rencontre vraiment dans la rue pour parler la langue" (**P9**); therefore, emphasising that such meetings already occur in his daily life acts as an indication that this aim is not as unrealistic as some would argue. As both this participant and another mentioned, Breton speakers tend to make themselves known to each other and can bond more easily over their shared language: "s'ils sont là, ils nous diront très rapidement qu'ils sont là, parce que ça leur offre plaisir de voir d'autres gens qui parlent breton" (**P5**).

All participants also use Breton with friends, although some expressed regret at not having many friends who spoke Breton. Others also pointed out that the internet was often crucial in maintaining these Breton-speaking friendships, as Breton-speaking friends often did not live in their local area and so had to be communicated with over longer distances. Speaking Breton with friends was often the result of meeting that friend in a Breton-speaking environment: "une fois que la relation est établie en breton

avec telle ou telle personne, elle perdure dans cette langue-là" (**P8**). All participants therefore had friends with whom they spoke only Breton: it was not the case that French was used in certain contexts and Breton in others. However, most interviewees also stated that this depended on who else was present in the conversation: if people who did not speak Breton were present, French would be used in order to ease comprehension. Only one participant gave an example of a situation that did not fit with this pattern: as a member of a Breton cultural and political organisation, she recalled that meetings among members were usually conducted in Breton, despite the fact that two out of the eleven members were not speakers; summary translations would then be produced for these members. This appears more stereotypically "néo-bretonnant": in this context where Breton identity and ways of promoting it are at the forefront of discussion, speakers may tend towards this attitude even in contexts where using Breton may impede some interlocutors' understanding. In general, though, the interviewees claimed not to do this, instead using French if it was better suited to their communicative purposes. This focus on the importance of communication is reminiscent of some of the strategies identified in section 4.4.3.

Some participants recounted cases of couples comprising one person who spoke Breton and one who did not, but where the non-speaker had not felt an impetus to learn the language. Two of the three older interviewees stated that their own partners were not Breton speakers, and despite being married to someone who had used Breton throughout their career and had many Breton-speaking friends and colleagues, had evidently felt no need to learn the language themselves. This had also had an effect on the linguistic upbringing of the couples' children, who in both cases were raised speaking only French. The third older participant, more stereotypically "néo-bretonnant", had followed a different trajectory, befitting the ideological position of a stronger preference for Breton over French:

> **P9** j'ai rencontré celle qui allait devenir mon épouse en breton, donc on ne s'est jamais parlé que le breton. Et quand nos deux enfants sont arrivés, on ne s'est pas posé la question, on les a élevés en breton aussi. Donc ça se produit trop rarement, mais dans notre famille la chaîne s'est renouée de justesse

Two of the five younger participants also had children, and both were raising these children with Breton. In one case, the interviewee himself had been brought up speaking Breton, and so for him it felt natural to use Breton with his own child:

> **P7** on a une fille depuis un an, et elle ne parle pas encore, mais elle entend le breton avec nous. Nous l'élevons en breton.
>
> **MDD** Donc vous faîtes l'effort de ne pas utiliser le français dans la maison ?
>
> **P7** Oui. C'est surtout ma femme qui fait l'effort, parce que ce n'est pas sa langue maternelle. Vous savez que pour moi c'est quelque chose qui est en quelque sorte naturelle. ... Il aurait été plus un effort d'utiliser le français, étant donné que la vie famillale pour moi, de mon côté de la famille, se passe en breton

The other participant with children, who had herself chiefly acquired the language as an adult, used a mixture of Breton and French with them. The father of one child also spoke Breton, while the other child's father did not. This did not prevent the interviewee from using Breton with the child herself, but neither did she insist on using only Breton with the children, instead using French in some situations. This signifies a more fluid and relaxed approach, allowing for simultaneous acquisition of the two languages without the worry that one will come to take precedence over the other. While the older participants appeared to require both parents to be Breton speakers as a prerequisite for using Breton with children, this was not the case for this interviewee. This points to a more fluid state of affairs, acknowledging the fact that as globalisation and individual mobility increase, the two parents may not have the exact same combination of linguistic repertoires, and multiple languages can be used in the familial setting, potentially for different purposes. While more stereotypically "néo-bretonnant" speakers may make the effort to use only Breton within the family, speakers who do not identify with this category have perhaps shifted from using French by default to mixing the two languages in a way that could potentially break down rigid interlinguistic boundaries. For the younger participants, this fits with their typically more complex perception of the intersection between their Breton and French identity seen in the previous section.

## 5.4 Militantism and politics

Section 1.4.1 gave an account of how the concept of the "néo-bretonnant" developed from the "militant" speaker identified by McDonald (1989). Indeed, while the term "néo-bretonnant" was not used in the interviews, the word "militant" was used by participants numerous times: five of the nine interviewees used the terms "militant(s)" or "militantisme" unprompted. This section will illustrate how different participants had different views on this concept, relating to both its meaning and their own level of identification with it. Some of their uses of the term are reproduced below:

> **MDD** avec tes enfants, tu fais des efforts pour utiliser un breton plus correct ?
> **P3** Non, c'est notre langue de famille et je veux qu'elle reste spontanée. Donc mon militantisme, il ne va pas jusqu'à dans ma communication avec mes enfants ... non, je parle complètement naturellement

> **P9** même si on est militant, même si on se dit « c'est bien, mettre du breton partout »

> **P5** je ne me trouve pas militante, et pourtant je pense que le sujet [Breton politics] est quand même un peu ... parce que c'est ce sujet-là et parce que j'utilise le breton ... de faire ça, c'est déjà militant

These three participants clearly have different perceptions of what it means to engage with militantism. The first recognises herself as a militant, but appears to link the type of Breton a person speaks with how militant they are, suggesting that the most militant of Breton speakers would use a more correct variety of Breton: in the context of this interview, this refers to a variety that adheres more closely to a single dialect and tends to avoid French borrowings—i.e. closer to stereotypical "néo-breton". When she speaks Breton, on the other hand, "je rajoute des mots de français ou des mots d'anglais ... il est mélangé entre plusieurs dialectes et le breton standard" (**P3**); a more militant speaker, she implies, would not engage in these practices. For this speaker, then, a

degree of militantism is expressed in the fact that she speaks Breton and uses it with her children, but this is not perceived as the strongest form of militantism because of the linguistic characteristics of her variety. For her, the "militant" category equates to the concept of the stereotypical "néo-bretonnant" as used in this thesis, which matches the correspondence in the literature between McDonald's (1989) "militant" and Jones' (1995) "néo-bretonnant".

In this case, the interviewee contrasts the way that militants would speak with how she uses Breton—"naturellement"—implying that for her, speaking the more "correct", "néo-breton" variety of the militants would be unnatural. She hence distances herself from the category of militant speaker or "néo-bretonnant" by emphasising her use of language that is "spontanée".

The second participant does not make an overt judgement on his own militantism, but equates militantism with the sentiment "c'est bien, mettre du breton partout". Elsewhere in the interview, he suggests that he feels this way himself, affirming his "objectif d'avoir une vraie société bretonne" (**P9**). He is also the most conformant of the participants to the characteristics of the stereotypical "néo-bretonnant" overall, living in eastern Brittany and having initially acquired Breton from books, as well as bringing up his children with Breton and making it the language of his daily life: he also makes a point of using Breton on Facebook, even though "parfois, ça agace certaines personnes" (**P9**). Again, therefore, militantism can be linked to the qualities of the stereotypical "néo-bretonnant".

The third participant explicitly declares herself not to be a militant, but nonetheless acknowledges that speaking about Breton politics on the radio, in Breton, is already a militant act. Later, she demonstrates her belief that even the act of listening to the radio in Breton is somewhat militant:

> **P5** [ceux] qui font l'effort de nous écouter, déjà ils font un petit peu un effort, et du coup c'est des gens qui sont un peu militants. Ils ont une volonté d'être liés à cette communauté-là ... je pense que forcément c'est des auditeurs qui ont choisi d'écouter

Likewise, while declaring herself not to be a militant she nonetheless acknowledges that by creating Breton-language radio programmes she does exhibit a certain degree of militantism. Indeed, this participant appeared to be one of the more politically engaged in non-linguistic terms: she reads the magazine of the *Union démocratique bretonne*, the largest Breton regionalist party, and also described her membership of a recently established association that aims to go into schools and give lessons on Breton history:

> **P5** pour expliquer à tous ces gens qui ne font pas partie du milieu breton quelle richesse on a culturellement, parce que ... les gens qui ne sont pas dans le monde breton par choix, ils n'y sont pas du tout, parce qu'ils ne connaissent rien, ils ne savent pas pourquoi ils sont bretons. ... C'est triste, parce que c'est quelque chose qui se perd, et du coup nous, on aimerait bien faire ça ... la France est un pays centralisé qui ne respecte pas ces droits [des populations autochthones], qui pour moi sont fondamentaux

Despite not personally identifying as militant, this participant appears to fit the term in the sense of motivation and political activism. Her reluctance to openly describe herself as such may be linked to the association of the militant speaker with the stereotypical "néo-bretonnant": as a speaker of a variety that includes certain dialectal features, this interviewee may wish to distance herself from a category of speakers commonly thought to use an artificially standardised variety.

Political issues also sometimes came up during the other interviews. One that frequently recurred was the status of the *département* of Loire-Atlantique, part of the historical region of Brittany. The assignment of the *départements* of France to regions in the mid-twentieth century resulted in Loire-Atlantique being made part of the Pays de la Loire region: it is not part of Brittany in administrative terms. The *département* nonetheless retains a sense of Breton identity, particularly strong in its largest city, Nantes, which was the seat of the medieval dukes of Brittany. Today, its departmental administration is involved in various cultural and linguistic initiatives that proclaim this Breton identity, such as sending a representative to the administrative board of the *OPLB*.[134]

134. http://www.brezhoneg.bzh/160-kuzul-meran.htm, accessed 28 Nov 18.

Campaigns to reorganise the current structure of the French regions and group Loire-Atlantique with the other Breton *départements* are ongoing, and interviewees frequently referred to Nantes and Loire-Atlantique in ways that showed they considered them to be part of Brittany, usually without pointing out that administratively speaking, this is untrue. Indeed, acceptance of Loire-Atlantique as part of Brittany could be considered a belief typical among those who consider themselves to have a strong sense of Breton identity, a category into which most of the interviewees have been seen to place themselves. Those who implied that Loire-Atlantique was part of Brittany were therefore unlikely to intend this as a political statement within the context of the interview, but rather to be referring to something that they held as uncontroversial fact, despite this being at odds with the structures imposed by the French state.

This section has already hinted at links between Brittany and other minority populations, in the case of the interviewee who pointed out France's lack of respect for its own minorities. As the analysis of the corpus data pointed out, *Bremañ* in particular often carries articles discussing other minority peoples. This can also be interpreted as a form of militantism, given that it is explicitly political. The representative of *Bremañ* gave some more information about this practice during the interview:

> **P7** il n'y a pas de politique rédactionnelle précise ou stricte là-dessus. Ce sera plus par rapport à la sensibilité des gens. Dans l'histoire du magazine et des gens qui ont participé au magazine depuis le début ... les peuples sans état en quelque sorte auront une place importante dans les articles, mais on n'a pas décidé que forcément dans cette rubrique-là il faudra parler dans tous les numéros des peuples minorisés, mais c'est quelque chose que presque tous les rédacteurs ont en tête.
> **MDD** Donc c'est devenu une tradition peut-être ?
> **P7** Oui, quelque chose comme ça, oui. Peut-être qu'au départ, surtout avant Internet ... dans les années 80 et 90, il était plus difficile d'obtenir des informations sur d'autres peuples ailleurs, et donc les gens avaient des contacts ... [avec] par exemple les gens du pays de Galles, donc il y avait plus d'articles peut-être de manière plus consciente. ... Donc tout ça a un petit peu formé une sorte d'esprit où, depuis que les contacts sont plus faciles

> pour voir d'autres personnes, ça s'est un peu élargi petit à petit au monde entier en quelque sorte

As *Bremañ* developed, it may thus have gained a representation for taking an interest in these issues and a concomitant association with a militant or typically "néo-bretonnant" stance, and so this tradition would have been perpetuated in the recruitment of other writers with similar priorities. This de facto editorial tradition therefore associates *Bremañ* with a more militant outlook in this respect, even if, as the participant implies, its current staff makes no effort to be overtly political.

*Brud Nevez*, a less typically "néo-bretonnant" publication, certainly in its linguistic aspects, also seeks to include news and opinions from outside Brittany:

> **P8** à l'occasion des dernières élections européennes on a publié un dossier sur les jeunes et l'Europe, et comme un certain dans l'équipe connaissait des Italiens, et on connaissait un Espagnol, moi je connaissais un Irlandais, donc il est arrivé que les étrangers ont pu écrire directement en breton. ... Ça donne une bouffée d'air différente pour le lecteur, qui ne s'attend pas à voir une Lithuanienne ou un Croate écrire et publier en breton, donc ça donne une couleur différente à la revue en même temps

There are some differences between *Brud Nevez* and *Bremañ* in their approach to content relating to other populations, however. Firstly, the articles in *Brud Nevez* were written for a special issue, rather having some presence in every issue of the magazine. Additionally, the *Brud Nevez* articles were written by members of the groups in question; for *Bremañ*, members of the usual pool of writers report on external matters. Moreover, the *Brud Nevez* issue did not focus on minorities specifically, but instead concentrated more generally on "les jeunes et l'Europe". While *Brud Nevez* can cover similar topics to *Bremañ*, then, and may feature political discourse, it does not approach these subjects within the same, more militant, tradition.

This section has shown that militantism is indeed present among those who work for Breton-language media, in ways that demonstrate some of the non-linguistic stereotypes about "néo-bretonnants", showing that many of those stereotypical features can

be identified among present-day Breton speakers. However, it has also illustrated how speakers accept or reject different aspects of this militantism with relation to their own identity, some being reluctant to identify with "militant" as a label. Likewise, different organisations deal with political issues in different ways, with *Bremañ* taking a more militant position, although one that appears to be the result of tradition rather than active decision, and *Brud Nevez* following a less militant line. As with many of the other aspects of the use of Breton examined in this thesis, both linguistic and non-linguistic, the experience of militantism and interest in political topics cannot be described in a particular way that all speakers can identify with, even within this small sample of nine participants. Again, this shows the breadth of opinions and practices among speakers and organisations, showing that not all can be assumed to be guided by the same ideologies in the choices they make concerning Breton.

## 5.5 The organisations

Interviewees had a variety of roles in their organisations: they included the editor-in-chief of a publication, co-directors of a radio station, journalists, and presenters. This allowed insights into a range of aspects of what working for these organisations entailed. Multiple interviewees emphasised however that such roles were often not formally imposed:

> **P8** l'organigramme n'est pas aussi prestigieux que ça. ... Mes camarades m'ont demandé d'[être] président ... puis j'ai joué de fait le rôle de rédacteur en chef : c'est à dire que comme j'ai une connaissance professionnelle de la manière dont les choses doivent se passer pour la maîtrise, le dossier, comment classer les articles en fonction de leur intérêt, le titrage et tout ça, il y a la légende de photo, il y a solliciter les auteurs, il y a collaborer avec eux pour quand c'est nécessaire pour un peu de rewriting

> **P7** il y a une sorte de comité de rédaction où il y a cinq participants avec un directeur de rédaction mais qui est bénévole, donc ce n'est pas un directeur

> de rédaction d'une structure plus grande. Il va décider certaines choses, mais pas de manière très stricte, ou ce sera un peu plus son avis

Indeed, a recurring aspect of the organisations was that they were often run informally, with the defined roles of staff not precluding them from undertaking other work. Institutional structures were therefore fairly flexible, and particularly in the case of the magazines, there was also a large voluntary component. This is in fact likely to have had practical implications for this project in the form of the lack of interviewee from *Ya*, the third publication: participants pointed out that this could have been because *Ya* employs only one or two permanent paid members of staff, meaning the pool of possible interviewees was limited.

The participants representing *Bremañ* and *Brud Nevez* described the editorial processes of the two magazines, which were similar across both publications. In each case, the editorial staff would be in contact with an established pool of writers: this was "préférable si on veut assurer la continuité" (**P8**), rather than attempting to find new writers for each issue. For *Brud Nevez*, issues tend to be more strictly thematic, although some articles diverge from the overall theme; for *Bremañ*, a wider range of themes is included in each issue, although conversely, some groups of articles may share a thematic link. In order to gather the articles for each issue, the editorial staff contact the pool of writers by email, asking for volunteers to contribute articles on specific subjects; they also encourage the writers to contact them spontaneously to offer any other articles they may be interested in writing. The participant from *Bremañ* stressed that in many cases, there is little interest among the writers in contributing articles on the desired themes: when this happens, the editorial staff reach out to their personal contacts, or write the necessary articles themselves. In some cases, if no Breton-speaking experts on a subject can be found, articles are written in French or English and then translated by the editorial staff. This does however create extra work for a very small team: *Brud Nevez* employs only three paid members of staff.

Interviewees from the associative radio stations described how the four stations that make up the Radio Breizh network work together. While each station is in charge of producing its own programmes, these are typically syndicated across the network and

broadcast by two or three of the stations at different times during the week. No station is therefore obliged to produce a full timetable's worth of programming. Furthermore, the stations attempt to work together to ensure that the times at which Breton-language programmes are broadcast are coordinated across the network, meaning that Breton can be heard on at least one station as often as possible. Listeners who tune in via the internet, including on the Radio Breizh mobile app, are therefore able to choose a station based on this, increasing the probability of being able to listen to a programme in Breton at a given time. The radio stations also aim to transmit a wide range of programmes in order to appeal to speakers from diverse backgrounds and different levels of competence in Breton: interviews with traditional speakers that focus on traditional local practices, more general news and talk programmes, and music-focused programmes. The larger amount of broadcasting time, compared with public radio, allows this flexibility, meaning that diverse types of listeners with different requirements can be catered for:

> **P5** il y a un animateur qui fait de la musique ... ça, je pense que c'est plutôt cool pour les apprenants. ... [Il y a un] qui interviewe les anciens, et là alors tous les anciens écoutent celle-là, ils adorent cette émission-là ... et puis il y a moi qui fais l'émission du matin, ... là, il y a une volonté compliquée d'être abordable pour les gens qui sont en train d'apprendre le breton, et en même temps d'être suffisamment intéressant d'avoir du contenu pour les gens qui veulent juste qu'on entende du breton le matin, et qui ont envie d'être un peu au courant, d'être liés avec la communauté

This strategy acknowledges speakers' diverse backgrounds and repertoires, creating a maximally inclusive atmosphere.

Interviewees from the radio stations were also asked specifically about news programmes. Each of the four Radio Breizh stations employs one journalist, and every day a news bulletin is produced that is broadcast on all four stations, with one journalist in turn reading the bulletin each day. This team of journalists collaborates using conference-call technology, allowing them to work together despite working in separate locations. With one journalist based in each station, it is also easier to arrange

interviews for the news programmes, as the journalist based in a particular area will have less distance to travel to visit interviewees. This means that the journalist based at the station that serves a particular area will typically be the one who does any external reporting on stories from that area, including interviewing spokespeople for organisations or members of the public.

For both news bulletins and more general broadcasts, many participants emphasised that they scripted programmes as much as possible prior to presenting them. For every programme, this writing would be done by the person who was to present it, in order to ensure that the language used would be a stylistic match for the presenter's dialect and speech style. This was the case at both the public and the associative stations. One participant recalled that he had been required to do this from when he first started in the job, despite his lack of confidence in his level of Breton of the appropriate register:

> **P2** au tout début j'étais beaucoup aidé par les autres de l'équipe. Il y a tout un banc du vocabulaire en breton que je ne connaissais pas ... donc ça, c'était un gros travail, même le vocabulaire de la météo, savoir les expressions, les tournures de phrase. Et tout ça serait vraiment des expressions et des tournures de phrases bretonnes et pas traduites du français

This concept of avoiding direct translation from French has already been seen in chapter 4, which showed that calquing and borrowing idiom from French were rare in the corpus. It will also recur in section 5.6.4 below.

While most radio programmes are scripted, there are some exceptions to this, including those taking a more casual discussion format. One interviewee gave an example:

> **P3** il y a certaines émissions qui sont plutôt des émissions où on rigole autour de la table. On fait par exemple un quiz audio, donc il y a une personne de l'équipe qui prépare le quiz ... et rien n'est écrit à part la personne qui prépare. ... Donc la personne qui prépare l'émission, elle la rédige elle, mais après c'est de l'interaction, c'est de l'echange, et donc ça, ce n'est pas rédigé, et c'est bien. Ça fait du bien, et le public aime beaucoup d'avoir les jeux qui ne sont pas qu'informatifs

The participant clearly saw this as an exception to the norm, though, remarking that "sinon, même quand on a l'impression que ce n'est pas écrit, c'est écrit à la radio" (**P3**).

Interviews are evidently another context where scripted content cannot be relied on, producing more spontaneous speech: some effects of this seemed apparent in the corpus data. Participants stated that in these cases, they would prepare a rough outline of questions to ask, but would be guided by the interviewee in the event. Finding interviewees was cited as a difficult task, as with finding writers to provide articles in the case of the magazines, and this could have a limiting effect on potential subject matters. One participant cited psychology as a subject she would like to discuss more on her own morning talk show, but was unable to find professionals in the field who spoke Breton.

In another participant's recollection, the smaller pool of available interviewees had an effect on gender balance, meaning that interviewees tended to be male. She suggested that this may be because "peut-être [les femmes] sont moins nombreuses dans les associations" (**P6**). Indeed, news stories in Brittany often involve the major primary industries of farming and fishing, and organisations in these fields may be more traditionally structured, meaning spokespeople will more often be male.

The same participant also noted that interviews may occasionally take place in languages other than Breton, with a Breton voiceover added during the editing phase. She gave the example of reports on stories from Catalonia, which can involve interviews being conducted in Catalan. The mention of Catalan rather than French implies that French would not be used in this context, despite surely being an obvious lingua franca for international stories. This apparent preference for other minoritised languages over French mirrors findings from the corpus, such as the preference for place names in minoritised languages and the greater acceptability of international loans: the general avoidance of French seems paramount.

Interviewees described various ways in which their organisations make use of internet-based technologies. The use of conference-call software by the Radio Breizh journalists has already been highlighted, as has the fact that all the radio stations concerned broadcast their programmes simultaneously online, making them accessible from outside

their specific broadcast ranges within Brittany. Programmes remain accessible via the relevant websites for long periods after their initial transmission, creating an archive of broadcasts that can serve as a pedagogical tool. However, the internet may not often occur to potential listeners as a source for listening to radio programmes: interviewees mentioned that they would more typically do this "dans la voiture" (**P7**) or "dans la salle de bain" (**P9**).

The four Radio Breizh stations broadcast their programmes online via a collective website, http://radiobreizh.bzh, but each also runs its own website containing more information on its programmes and staff. One interviewee described how the content of these websites could enhance the experience of listeners. She described a programme for children that she believed was also listened to by adults in the early stages of learning Breton, and went on to explain how the station's website was used to help those who listened to this programme: "on met le vocabulaire de l'émission sur le site Internet, avec pas de traduction mais une explication des mots compliqués" (**P4**). This, then, is an example of how more traditional forms of media can harness the potential of the internet in order to create a more inclusive and enriching environment for audiences who may be unsure of a word's meaning.

*Bremañ* and *Brud Nevez* differ markedly from each other in their engagement with the internet. The interviewee from *Brud Nevez* described his own enjoyment of maintaining websites, including that of the magazine itself (http://www.brudnevez.org) as well as his own personal blog. He noted that *Brud Nevez* had been the first Breton-language magazine to offer its subscribers the option of downloading each issue as a PDF file for reading on digital devices. In this case, his personal motivation to use such technologies was clearly high, which had beneficial effects for the magazine, increasing its online presence in the form of a frequently updated website and offering subscribers a range of ways to read content. For *Bremañ*, however, these digital options seem not to have been explored as thoroughly. The interviewee from this publication believed that the editorial board were mostly unaware of how the internet could be harnessed to its fullest potential. He explained that social media were used to promote *Bremañ*, but that its Facebook page was mostly limited to sharing content that accompanied articles in the published editions, such as extra photographs. The magazine had experimented

with an online version around ten years ago, but its staff had been unable to dedicate enough time to maintaining it. As he pointed out, the rapid evolution of technology means that a similar attempt today would probably take a very different form, but also that it was difficult for the staff to keep abreast of new technological developments and they would be unsure how best to implement a similar system again. He acknowledged that it was important to consider the possibilities brought about by the internet, however, as *Bremañ* was struggling to attract readers:

> **P7** il y a beaucoup d'abonnés âgés ... je pense qu'ils se sont abonnés avant et qu'ils sont toujours abonnés, la génération qui militait dans les années 80, 90, et donc qui est maintenant dans la soixantaine, 50, 60. Je pense que ça doit être une grosse partie des abonnés aujourd'hui ... je ne crois pas qu'il y ait énormément de jeunes abonnés ... je pense qu'ils lisent d'autres choses, et surtout peut-être qu'ils lisent d'une autre manière, je pense, beaucoup plus tout ce qui est numérique

On this topic, more can be said about the audiences of Breton-language media. As stated above, one participant pointed out that the very act of listening to the radio in Breton was somewhat militant, requiring the motivation to tune in to the right station at the right time.

> **P5** c'est un auditeur qui a de la volonté. Pour celui qui écoute ... toute la société est en français ... et même quand on parle breton et qu'on est habitué au breton, il est plus facile d'écouter du français que d'écouter du breton

Another participant considered listeners to fall into two age groups:

> **P4** je me dis qu'on a deux sortes d'auditeur, qui reflètent l'état de la langue bretonne aujourd'hui. Il y a des personnes qui sont à la retraite, qui nous écoutent parce qu'ils aiment la langue, et qui sont probablement aussi militants pour la langue ... et aussi une tranche d'âge qui va jusqu'à 35, 40 ans, des personnes qui ont été à Diwan et qui ont appris.

This mirrors the usual split between new and traditional speakers, with a gap between the two groups that matches the period between the general cessation of intergenerational transmission and the establishment of the Diwan schools. Indeed, the participant appears to characterise the younger group as stereotypical "néo-bretonnants" in her belief that they are "des personnes qui ont été à Diwan et qui ont appris", However, she also characterises the older generation as "militants pour la langue", which does not match the usual descriptions of traditional speakers. Again, this relates them to the way listeners were described by the previous participant: they must be making this conscious effort to tune in. If they are traditional speakers, as their age would imply, it follows that some such speakers also fit these militant characteristics, and that sizeable numbers of them enjoy hearing Breton on the radio, being used in new contexts to discuss new topics. Indeed, previous research has indentified older traditional speakers of Breton who consider neologisms and related features of modern standard Breton "clear and logical" (Hornsby and Quentel 2013:78).

Other differences between the two groups did however resemble the supposed differences between new and traditional speakers: the participant suggested that older listeners tended to be more working-class than the younger ones and come from a more restricted range of occupations: "il y a beaucoup qui étaient agriculteurs ou pêcheurs" (**P4**). She perceived a similar split among the interviewees that she typically spoke to, and, reflecting another typical new/traditional speaker difference, also noted that younger interviewees used less dialectal forms of language. As a result, programmes featuring this type of interviewee can be broadcast more easily by all four Radio Breizh stations instead of having to be limited to a particular station and associated dialect area. In the case of programmes that concentrate on the more dialectal speech of older people, this restriction becomes more necessary:

> **P5** un collègue [qui] va faire des interviews avec les anciens essentiellement, donc il rencontre des personnes qui ont 80 et 95 ans, et il y a un accent qui est très très marqué ... [il] a une émission, Kreiz Mintin. Elle n'est diffusée que sur RBG [Radio Bro Gwened], parce que si c'était diffusé ailleurs, ça serait très très compliqué. Ça pourrait être intéressant, mais je ne sais pas trop si les autres radios la prendraient, parce que l'accent est tellement différent

> qu'ils ne la voudront peut-être pas

Participants also spoke about listeners' and interviewees' attitudes to language: one noted that younger people tended to be more willing to adopt new terminology than older people. Another participant stated that listeners were encouraged to phone the station to query any unusual vocabulary, but admitted that this happened only very rarely. This implies that listeners may not wish to engage actively with the stations, preferring to remain uninvolved as passive consumers.

A number of attributes came to light during the interviews that could be considered characteristic of Breton media, and presumably of media in minoritised languages in general. Firstly, there are certain problems that tend to hinder these media organisations. The difficulty of finding writers or interviewees from diverse fields has already been pointed out; likewise, audience numbers are low. One participant described the situation for print publications in Breton:

> **P8** les médias en breton, c'est une dizaine de périodiques dont il n'y a aucun qui tire en plus de 2000 exemplaires. Il y a un hebdomadaire [*Ya*] qui a 1500 ou 1600, peut-être 1700 maintenant. Les revues autrement, c'est quelques centaines. ... Donc la presse écrite en breton, ça n'est pas immense, on ne va pas trouver une périodique en breton à des kiosques, à des bureaux de presse, ça n'existe pas. Ça n'est diffusé que sur abonnement

Such low sales figures evidently ensure that the organisations tend to be underfunded, making them able to employ only a small number of staff, which further restricts their ability to produce content in larger quantities or on more varied subjects.

On the side of the radio stations, another participant pointed out the large amount of paperwork required to run an associative station, much of which has to be done in French. She made a point of completing this bilingually in cases where this was permitted, despite this taking even more time. This made her unable to be involved with producing programmes, and she found that as a result, her Breton had deteriorated since taking the role of station director:

> **P3** je pense que peut-être [mon breton] s'est un peu appauvri, parce que ... quand j'étudiais, je faisais vraiment beaucoup d'effort pour parler avec beaucoup d'accent, et avant, à la radio, j'avais interviewé des personnes âgées aussi beaucoup.[135] Donc j'avais un breton qui était peut-être plus coloré, et maintenant je fais plutôt de l'administratif pour la radio comme je suis directrice, donc je fais beaucoup de papiers en français, donc je pratique moins avec des anciens

Having to use French was not always perceived as a negative aspect of work, however: one interviewee, despite generally coming across as more stereotypically "néo-bretonnant", meaning he might be expected to want to use Breton as often as possible, instead stressed the importance of using French to liaise with external suppliers, and did not seem to regret having to do so:

> **P7** j'utilise énormément le français aussi. J'utilise le breton avec mes collègues, les gens qui travaillent ici ... mais pour réaliser le travail, je suis en contact avec des gens qui ne parlent pas breton. Donc le français a une place importante aussi dans le travail tous les jours forcément

French bureaucracy could also be detrimental to Breton-language media in other ways. One participant explained the system within the public radio network that meant a reporter could be sent to anywhere in France. In his case, he had been fortunate that a reporter at France Bleu Breizh Izel had retired, which had allowed him to obtain a permanent post. Had this not occurred at the right time, he would have remained within "ce système de concours, qui nous maintient dans la précarité pendant entre 4, 5, 6 ans parfois" (**P2**). This kind of nationwide employment system, that neglects to consider the specificities of France's regions, can therefore be detrimental to those regions' languages, as in the case of the Breton-speaking teachers sent out of Brittany in 2016, leaving some schools within the public bilingual education system unable to provide teaching in Breton for certain subjects.[136]

135. It can be seen from this that the participant considers a more dialectal form of Breton to be desirable: this will be discussed in the next section.

136. See https://www.letelegramme.fr/finistere/brest/brest-faute-de-professeur-bretonnant-les-eleves-du-college-de-l-iroise-manifestent-11-10-2016-11250681.php, accessed 7 Aug 19.

Some more beneficial aspects of Breton-language media were also mentioned over the course of the interviews. Multiple participants noted that as long as one had the necessary language skills, it was easier to find work in Breton-language media organisations than French-language ones. Similarly, it was easy to make connections when starting out, due to the smaller size of the Breton-speaking community and links among different organisations: many interviewees had contributed to other Breton-language media by, for example, writing an article for another magazine, or being interviewed on a different radio station. One participant recounted how the use of Breton could sometimes add further nuance to a news report or interview with a particularly striking example from his early career:

> **P1** quand l'Amoco [Cadiz] s'est échoué ... la Marine faisait des exercices, ils mobilisaient un pétrolier. ... On avait donc un officier supérieur de la Marine ... il était aussi le porte-parole officiel de la Marine, et c'était lui que l'admiral avait désigné pour parler avec nous. Il raconte en français, il raconte très formel, « exercice parfait ». Quand il est interrogé en breton, il dit tout le contraire : « ce genre d'exercice, c'est nul ». Et on l'a vu, il l'a dit en breton, et je me suis dit, les auditeurs qui ce jour-là ont écouté les infos qui comprenaient à la fois le français et le breton, ils ont eu la vrai info, mais ils l'ont eu en breton. Voilà, c'était rigolo

Many interviewees also pointed out the important role of the media in making Breton more visible:

> **P8** les médias en breton ... ça donne une visibilité à la langue qu'elle n'a pas beaucoup autrement. Alors maintenant si, on voit des panneaux de signalisation bilingues ... mais c'est plus symbolique qu'utilitaire. Par contre, la radio est accessible à ceux qui n'ont que la connaissance passive de la langue

The existence of media in Breton can therefore act to make the language known to those who do not speak it; especially in the case of certain radio stations, which also broadcast programmes in French, meaning that a listener with no knowledge of Breton

who typically tunes in to hear these may come across some Breton on the stations, which could spark a degree of curiosity about the language.

The media was also recognised as a vector for new terminology. Interviewees reported cases where a new term had been created or adopted to be used in a radio programme, and its use had subsequently spread to interviewees and to other areas of the media. This example of community-led language planning, and of hearing ordinary speakers use such new terms, is an alternative approach to more traditional top-down and prescriptive methods, and may be more successful in reaching the community. The following section continues this discussion of more specifically linguistic points.

## 5.6 Dialects and lexical variation

Linguistic topics in the interviews covered dialect, style and register, competence, neologisms and borrowings, and the use of dictionaries. This section presents points of interest from the ensuing discussion.

### 5.6.1 Individual variation and change

Nearly all participants confirmed that they had noticed the style of their own Breton changing over time. Many emphasised that this was not a result of deliberate decision, but rather of coming to use Breton in different situations and with different interlocutors from a variety of backgrounds, which had caused them to adapt their own way of speaking unconsciously.

There were two patterns of diachronic change, depending on the context in which the participant had acquired Breton. Those who had grown up with regular exposure to Breton from traditional speakers within their family or community reported that their language had become less dialectal over time, as a result of greater exposure to people who used other varieties of Breton in a more diverse range of environments than during their childhood. Two of these speakers, both of whom came from the Trégor area, both reported that their speech had become closer to the standard as a result of

working in radio and customarily having to communicate in much wider contexts than their own immediate communities, but that it maintained Trégorrois features to this day. The third speaker in this group, who had spent his childhood in areas where Breton was not spoken as a community language, had had to contend with this sort of situation since childhood, when he had attended a bilingual school in Rennes. He had initially moderated his accent in a school setting, believing that his classmates would not understand him if he used the type of Breton he was accustomed to speaking at home. However, this had lasted only for the first year or two of his education in the bilingual state school system:

> **P7** j'ai dû voir que ça se passait bien pour être compris, et donc du coup je n'avais plus trop peur de prononcer les choses à la manière dont on les prononce dans la famille plus généralement. ... Je pense que mes professeurs en CM2 et ensuite au collège avaient dû me dire que ce n'était pas grave, qu'il allait mieux parler comme à la maison pour que les gens s'adapteraient aussi à ma manière de parler

The concept of the *muda* as discussed in section 5.3.1 is thus again relevant here, as changes in the participant's life and circumstances had effects on the type of Breton he used. He related that as he had grown older, and spent more time away from his family, he had nonetheless found himself beginning to speak a less dialectally marked variety: like the other two participants in this category, he estimated that this was because more of the time that he spent speaking Breton was spent with people from different areas. However, he stressed that while he had moved towards a less dialectal variety of Breton that was more similar to that of other speakers around him, this was not necessarily the same thing as the "official" standard: he was guided more by community norms than prescription.

The second group of interviewees comprised those who had acquired Breton in contexts outside the traditional speaker environment: through education, from a new speaker parent, or in adulthood. On the whole, these participants had witnessed the opposite effect from that experienced by the previous group: their Breton had become more dialectal over time. Participants also often associated this with an increase in

competence. One speaker noted that his Breton had developed immensely since he started working in radio five years earlier:

> **P2** les tous premiers journaux ... ce n'est pas du tout la même chose. Je ne parle pas de la même manière, je n'ai pas le même accent, les mêmes accents toniques, le même vocabulaire non plus, je ne construis pas pareil mes journaux

Other interviewees in this category noted that they had moved from one dialect to another since starting employment. One explained that having initially acquired Breton from her father, who himself had learnt the language as an adult, she had grown up using KLT Breton, a term that in this context indicates the general supradialectal variety used in most teaching materials. This lack of specification of a particular dialect within the KLT area suggests that that the variety she used in childhood would be considered closer to the standard than to any dialectal variety. Despite having grown up nearer the *pays vannetais*, and attending a Diwan school in Vannes, she had not started using Vannetais until she began working for the radio, at a station broadcasting in the Vannetais area. Similarly, another participant, who had initially learnt a Léonais variety when she acquired Breton as an adult, had shifted towards Cornouaillais since starting work at the station broadcasting in Cornouaille. As mentioned earlier, this indicates that in general, dialects are perceived favourably by new speakers, contrary to their stereotypical description: those working in radio seem to deliberately seek to use a dialect that will be familiar to their listeners, showing the continued value of traditional speakers and their varieties in these contexts, despite claims about new speakers' use of stereotypical "néo-breton".

Two participants fell into neither of these categories. While they had acquired Breton in ways that would have assigned them to the latter group, i.e. not surrounded by multiple family members or a community where Breton was the usual language, and so can more easily be classed as new speakers, they did not follow the usual narrative of members of this group in terms of change to their variety of Breton. One of these participants stated that she had observed no change to her Breton in the years since she had acquired the language, through a long-term immersion course for adults. She had

learnt Vannetais at this course and continued to use the same Vannetais variety now that she worked in radio. While the station she worked for broadcast in the area where Vannetais was spoken, she regularly communicated with employees of the other stations, who would use different dialects, due to her position as a journalist, and would also occasionally interview speakers of other varieties. However, she estimated that this had had no effect on her own speech: while she had become more at ease with understanding other dialects, she did not feel she would have the ability to use them herself, and believed they had had no influence on the way she spoke. Conversely, neither did she believe that working for the radio had increased the strength of her Vannetais dialect. She was the only interviewee who believed her language variety not to have changed at all since she first learnt Breton.

The second participant who did not fit either of the two general trajectories, from the *OPLB*, presented a different sort of exception. Having begun learning Breton from books as a teenager, going by other participants' experiences he would have been expected to gain a more dialectal variety as he gained competence and began engaging more often in oral conversation. However, while he had initially wanted to acquire his parents' dialectal variety, his attitude towards dialects changed as he became an adult. He characterised this as "une maturation, une réflexion" (**P9**), explaining further:

> **P9** je ne vis pas dans la région d'origine de mes parents, je vis ici à Rennes, et les personnes avec qui je suis en contact sont originaires soit des régions différentes, des régions qui parlaient breton, soit même des Rennais pur souche qui ont appris la langue, qui n'ont dans leurs ancêtres aucune personne qui ne parle le breton, ce qui fait qu'on arrive à une sorte de langue standard, moyenne. Et en fait je ne vois pas de justification à rester sur une forme hyper-dialectale ... et je pense aussi qu'il ne faut pas faire une sorte de fétichisme autour de telle ou telle forme linguistique. Pour moi la langue, c'est d'abord une façon de communiquer avec d'autres personnes. La langue en elle-même, ça n'est rien en fait, tant qu'il n'y a pas de personnes en train de se parler. Donc ce qui m'apporte, c'est de parler breton, un breton le plus correct que possible. Quand je dis breton standard, ça ne veut pas dire ni breton sans couleur ni breton sans accent, je pense qu'un

> breton standard peut être un vrai breton correct, mais c'est un breton qui va un petit peu s'uniformiser, parce que justement les occasions de parler la langue se décloisonnent. Alors que mes parents ne parlaient breton que dans un lieu très précis et passaient au français dès qu'ils allaient à 30 kilomètres, moi je parle breton avec des gens de partout. Donc de ce fait-là-même il y a une espèce d'autorégulation ... et une sorte de breton moyen qui émerge naturellement. Et je pense que c'est ni bien ni mal, c'est tout simplement absolument naturel, et voilà, la seule chose est d'essayer de faire en sorte que cette langue effectivement se conserve au maximum. ... Je pense il faut accueillir ça comme une évolution naturelle et pas chercher à lutter contre, parce que de toute façon on a perdu à l'avance

This case was therefore quite different from the others. The participant's attitude towards the legitimacy of standard Breton is one of various attributes, including his method of acquiring Breton and the fact that he lives in eastern Brittany, that confirm him as the most stereotypical "néo-bretonnant" out of the interviewees. Other participants did not tend to share this view of the standard: it has already been mentioned that they typically place more value in the dialects, and further evidence of this will be discussed below.

With the exception of these two participants who did not fit the general pattern, the trajectories of the other interviewees form a type of convergence; not towards a standard variety, however, meaning that it cannot properly be called an example of levelling. Instead, those speakers who originally used more dialectal varieties had moved towards a variety where dialectal features were retained, but were less strong; conversely, those who had used more standard varieties had found their language becoming more dialectal. This results in a situation where most speakers have identifiably dialectal features, but not to an extent that impedes communication with speakers from different areas.

This phenomenon is worth discussing in more detail, as it does not appear to have been identified in existing literature on Breton. In this respect, the concept of the dialect continuum as represented by a cone model (Berruto 2018:508) can be borrowed

to illustrate how speakers move through this space. As noted, this is not an example of dialect levelling as commonly theorised in situations of language contact and obsolescence, involving "the loss of localised forms in favour of others of wider currency" (Hornsby 2009a:165): dialectal speakers are indeed losing some specific local features as they move to a more neutral variety, but conversely, speakers who begin by acquiring a more standard form of Breton also find themselves moving in the other direction.

Applying the cone model to Breton, static conceptions of speakers and their dialects, following the typical binary between traditional speakers and "néo-bretonnants", can be represented according to the following image.

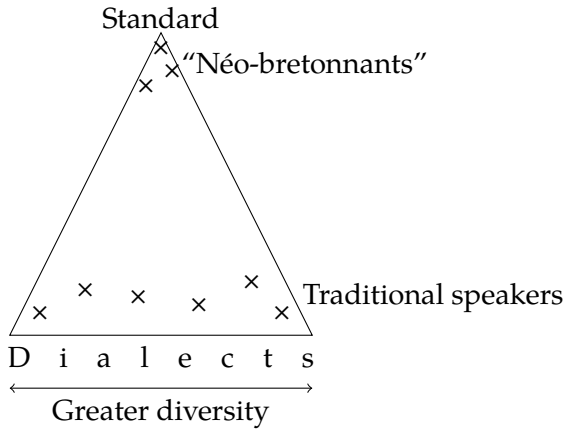


Figure 5.3: Cone model showing typically reported divide among Breton speakers

Representing speakers' reported mobility according to the same model, we can modify the image to produce the following diagram, in which certain speakers move in both directions towards a mutually intelligible zone depending on their personal histories of Breton acquisition.

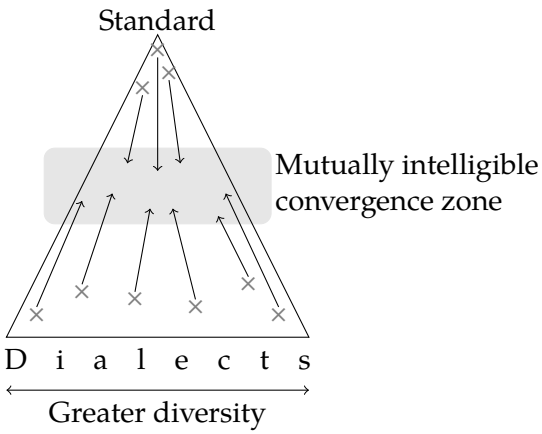


Figure 5.4: Cone model showing dialect convergence as reported by participants

As noted, this can be conceptualised using the term "dialect convergence", a broader concept than that of levelling, still conveying the sense of disparate dialects moving closer together but not specifying that the converged dialect indicates greater proximity to the prestige standard for all speakers. Dialect convergence has been theorised in previous work, notably Hinskens, Auer, and Kerswill (2005) and, in the case of regional varieties in France, Hornsby (2009a). In these contexts, though, it is still discussed mostly with reference to levelling, in the sense of speakers of dialects moving towards a more standard variety. Hinskens, Auer, and Kerswill (2005:6) do note that the reverse can occur, referring to this as "downwards" convergence, "when members of the local elite speak dialect".

The situation of Breton as a minoritised language perhaps means convergence of this type is more frequent than in many languages. It can be connected to the fact that the standard remains contested and unused by a large group of speakers: new speakers can thus seek consciously to move away from this point of reference. Similar processes can be noted in other languages in a situation of revitalisation: O'Rourke and Walsh (2015) note examples of new speakers of Irish who take an "essentialist" position, favouring the dialectal varieties of traditional speakers and making an effort to replicate them. The overall idea of dialect convergence is also supported by pragmatic factors: citing Siegel (1985), Hinskens, Auer, and Kerswill (2005:6) note that dialect convergence "can be described as the linguistic manifestation of speakers adapting 'to the speech of others to reduce differences"'. This highlights the need for intercomprehen-

sion across the diverse speaker community, especially in these cases of speakers who work in media and are charged with communicating in a way that appeals to the broadest possible cross-section of that group, including both new and traditional speakers. This practical motivation, along with the lack of adoption of the standard by traditional speakers, results in a convergence situation rather than a more straightforward case of levelling. Again, the importance of communication in a diverse community is prioritised.

As noted, two out of the seven participants, both of whom acquired Breton as adults, are nonetheless exceptions to this model, and reasons for this must be explored. The first of these two participants initially acquired Vannetais, the most strongly marked of the four major regional dialects of Breton. As a speaker who is present in the corpus, she displays notable Vannetais characteristics in her speech, often more strongly than others. Going by her claim that she still uses the variety of Breton that she first acquired as a result of an immersion course, it appears that this variety was already strongly marked as Vannetais, although not a traditional speaker variety given her method of acquisition: it was hence already located in the "convergence zone" marked on the cone model, and thus already provided the opportunity to balance communication and identity. Indeed, she characterised her variety as "vannetais *soft*", noting "je fais exprès de ne pas dire des mots qui sont inconnus ou compliqués pour ennuyer des gens" (**P6**).

The second exception to this trajectory, the participant from the *OPLB*, showed different ideological characteristics from the other interviewees. While others typically saw dialectal varieties as more valid than the standard, he insisted on the value of the latter, clearly attributing to it more prestige. As a result, he had no reason to move into the "convergence zone", as the standard remained more desirable for him. For this particular speaker, Breton follows the model of major languages in that the prestige of the standard variety persists. The fact that he grew up and lives in eastern Brittany is also relevant, and has parallels with new speakers of Scottish Gaelic: McLeod and O'Rourke (2015:161) note that in their study, "new speakers with no close family connections to Gaelic or any particular dialect area ... perceived that endeavours to acquire a particular local dialect would be somewhat arbitrary or artificial".

This participant's greater emphasis on the value of standard Breton can be related to the concepts of anonymity and authenticity, as posited by Woolard (2016). While anonymity, speaking a language in a way that does not index one's local origins, is typically considered a target for major world languages—hence the generally accepted prestige of standard English and French—authenticity is more commonly valued among speakers of minoritised languages. This ideology, on the other hand, "locates the value of a language in its relationship to a particular community. Within this logic, a speech variety must be perceived as deeply rooted in social and geographic territory to have value" (ibid.:22). This authenticity is clearly valued by most of the participants, as they seek to use a recognisable dialectal variety of Breton, often motivated by the particular geographical area in which they work. However, the participant from the *OPLB* instead appears to value anonymity, seeking to mould Breton in the image of major languages. This anonymity brings with it certain advantages to which the *OPLB* ascribes particular importance. These include ease of acquisition: "anonymous languages supposedly can be learned by anyone, but authentic languages can be learned by no one" (ibid.:24); as well as a hegemonic position in modern society: ""in the modern and modernist public, the citizen-speaker ... is supposed to sound like an Everyman, using a common, unmarked public language" (ibid.:25). As the *OPLB* thus seeks to mould Breton in the image of major world languages by pursuing its regulation and standardisation, it is not surprising that the participant representing this organisation values anonymity over authenticity, despite the fact that other interviewees appear to take the opposite approach.

Participants were more united in their thoughts on synchronic variation. All stressed that their Breton varies in register depending on the situation, subject matter, and interlocutor, showing that variation according to register and style persists among new speakers. Some, although not all, also noted that their dialect varied according to their interlocutor: they would, often without intending, adapt their speech slightly towards the other person's dialect. Reports of this phenomenon were largely restricted to the participants who had acquired Breton in the community or as children: those who had learnt as adults, who had spent less time immersed in the language, were perhaps not as comfortable using a wider range of dialectal features. However, one participant

who had learnt Breton in adulthood noted that while she maintained the same dialect, the *strength* of her accent varied according to her interlocutor: with older people who often used a more dialectal variety, her own use of dialect would become stronger.

### 5.6.2 Dialects in the media

Participants were asked about how dialectal Breton was viewed by the media organisations they worked for; the general opinion was that the use of dialects was seen favourably. As can be expected, the nuances of these views differed, particularly depending on the type of media: print and radio will therefore be considered separately in this section.

The representatives from the two print publications came to similar conclusions about how their magazines treat dialectal submissions: both stated that the use of dialects was permitted in the publication as long as the dialect was represented in the orthography the magazine used, and the use of dialect was consistent throughout the article. Dialect of course tends to be less visible in writing than in speech, as phonetic variation may not be represented orthographically. Written articles will therefore appear less markedly dialectal than their spoken equivalents: this was discussed in section 3.4.3. Additionally, one of the participants from the magazines noted that among contributors, "il y a une forme de standardisation qui s'est installée" (**P8**), demonstrating that authors may feel the need to write using a less dialectal form of Breton in the context of an article that will be read by speakers of disparate dialects, again showing the importance of communication.

While no representative from *Ya* was interviewed, a participant confirmed that she had occasionally contributed articles to the publication, using her own Vannetais variety, and that these had been published without undergoing any form of standardisation. This therefore suggests that *Ya* has a similar attitude towards dialects to the other publications.

Interviewees from the radio stations stressed that dialectal speech was seen as an advantage for employees. One explained that speaking with an "accent français" was

seen as less desirable than speaking with a mixture of dialectal features from different areas:

> **P4** [il y a des accents] qui se sont plus ou moins mélangés et qui sont quand même colorés mais qui n'ont pas de spécificités particulières.
> **MDD** Et c'est toujours considéré comme authentique si on est comme ça ?
> **P4** Je ne sais pas ... pour moi c'est authentique ... c'est plus authentique qu'un breton—on critique beaucoup ce breton-là—qui est avec un accent français. Ce breton-là n'est vraiment pas bien vu, mais il existe. Il y a beaucoup de gens qui parlent breton avec l'accent français, donc on les diffuse, nous, on les interviewe, on ne va pas du tout ne pas leur donner la parole, parce qu'ils ont des choses à dire et c'est très intéressant. Mais nous, c'est vrai qu'on essaie de ne pas parler comme ça

Another interviewee gave similar opinions:

> **P5** c'est un peu mal vu de parler un breton qui est, comme on dirait, *di-vlaz*,[137] qui n'est pas marqué ... ce n'est pas très bien vu aujourd'hui. On a un accent plus classe quand on arrive à choper un vrai accent ; là, c'est vraiment la classe, parce que ce n'est pas facile.
> **MDD** Mais si c'est quelqu'un qui vient d'une partie de la Bretagne, qui parle le breton d'une autre partie, c'est toujours considéré comme valable si ce n'est pas l'accent de son propre territoire ?
> **P5** Tout à fait. Au moment où il a fait un choix et qu'il s'est tenu à son choix, qu'il a bossé sur son choix, c'est trop cool

From this, it is evident that the use of dialects appears to be favoured on the radio above an unmarked standard, even when it amounts to a mixture of features from multiple dialects, or when it is not the dialect that speakers would be expected to use based on their origins. It is interesting to note that the second interviewee claims that unmarked standard varieties are "pas très bien vu *aujourd'hui*", implying perhaps that a shift in

137. From Breton *di-*, 'without', and *blaz*, 'flavour'.

attitudes has taken place since the earlier days of the Breton revitalisation movement. This would seem to tally with the findings of this chapter when they are compared with older literature on new speakers of Breton. It does, however, strongly contradict the "néo-bretonnant" stereotype that suggests that speakers in this category speak a standard Breton with strong phonological influence from French (Timm 2003:42; Madeg 2011:64; Hewitt 2016:16).[138] In this setting, speakers appear to reject this stereotypical variety, demonstrating their desire not to be associated with features of "néo-breton" that have been stigmatised.

### 5.6.3 Inclusion of different dialects and abilities

Interviewees were asked how they accounted for the potential breadth of abilities and dialects among their audiences. Previous sections of this chapter have already noted how the wide range of types of programmes broadcast on the associative radio stations can be beneficial in this respect, and pointed out the function of the internet for explaining more complicated vocabulary in programmes targeted at children or learners. On the same subject, the interviewee from *Bremañ* described how the magazine deals with instances of more advanced or unusual vocabulary, explaining that articles are often accompanied by a panel containing a "lexique" (**P7**) for these words. Based on the sample gathered from *Bremañ* for this research and its use of this technique as discussed in section 4.4.3, it can be noted that this strategy is slightly different from that used by the radio station, in that *Bremañ* sometimes resorts to using French; the participant from the radio station, on the other hand, specified that on its website, words were listed "avec pas de traduction, mais une explication des mots compliqués" (**P4**).

The "lexiques" in *Bremañ* can also contain dialectal words, but without specifically marking that they are from a certain dialect:

> **P7** généralement, on ne précise pas que c'est dialectal … c'est juste considéré que ça peut être une autre manière de dire la même chose, sans devoir pré-

138. On the other hand, the fact that mixing dialects is apparently considered acceptable demonstrates that these speakers are not attempting to approximate the language of traditional speakers in this respect. This therefore does characterise them as new speakers, if not as stereotypical "néo-bretonnants", who would be expected to be ideologically opposed to the use of dialectal features.

> ciser nécessairement de quel dialecte c'est attaché : plus pour laisser libre. Les gens vont l'utiliser ou pas, donc ils se disent « ah non, je ne vais pas utiliser ce mot-là, parce qu'il est attaché à un dialecte en particulier »—c'est juste considéré que tous ces mots-là sont bretons

This holistic approach allows the language's existing resources to be drawn on to maximum effect, minimising the need for either borrowings or coinages. At the same time, it could encourage speakers to use vocabulary drawn from multiple dialects in their own speech, which other users of Breton may consider invalid. The participant explained that this practice stemmed from the fact that *Bremañ* tends not to include linguistic information; this is typically restricted to its sister publication for Breton learners, *#brezhoneg*. *Bremañ* is therefore able to avoid dwelling on linguistic topics, showing that the use of Breton need not be self-reflexive. The radio stations operate under similar principles:

> **P4** le premier [principe], c'est de parler de tout, et de ne pas parler uniquement en breton de la langue bretonne, et de ce qui est fait pour soutenir la langue bretonne. On veut parler du monde autour de nous, qui ne tourne pas forcément autour du breton, et par contre, on le fait uniquement en breton

This could be a factor in the general lack of provision of radio programmes aimed specifically at learners: without them, it is easier to provide the rare resource of a Breton-only environment for competent speakers.

Many of the participants from the radio stations mentioned that they made a particular effort to speak clearly and concisely when presenting programmes, and that they used "des mots assez courants" (**P6**) where possible. However, several of them pointed out that this was not deliberately done in order to accommodate speakers accustomed to different dialects, and still less as an aid to learners; rather, it conformed with a principle of clarity that was promoted in radio broadcasting within France more generally.

Participants nonetheless spoke about some aspects of radio broadcasting affected by the presence of substantial dialectal variation. As noted, the radio stations tend to

encourage a moderate level of dialectal use, as this is typically preferred to a neutral variety. One interviewee pointed out that it was important to maintain this even in programmes aimed at learners, who may be listening to the radio for the very purpose of gaining exposure to dialectal Breton and acquiring it themselves:

> **P5** les gens qui apprennent, il y en a beaucoup qui font un travail très intéressant d'écoute pour attraper l'accent. C'est très important pour eux, pour beaucoup d'entre eux, et ils aiment bien quand on passe des anciens, parce que ... ils écoutent pour ça, ils écoutent pour réussir à choper l'accent

Another participant, who characterised his own dialect as "plutôt un mélange" (**P2**), explained how he might adapt his use of dialectal vocabulary according to the situation:

> **P2** il n'y a que dans le Léon où ils utilisent le mot *kouer* pour un 'paysan'. Et du coup, je l'utilise rarement, mais une fois de temps en temps je l'utilise, notamment quand le paysan vient du pays de Léon. Je me dis, « ça va un peu rendre hommage à lui, qui utilise ce mot-là »

As a concession to learners and speakers of other dialects, he would employ a strategy involving the use of doublets:

> **P2** si j'utilise le mot *kouer* que je sais que beaucoup ne vont pas comprendre, je vais un peu compenser dans ma petite brève en mettant le terme *peizant*.[139] ... Donc, par exemple, si je présente quelqu'un : Yvon Bouttier, la semaine dernière ... je ne sais plus si j'ai dit *kouer*, mais en tout cas j'aurais pu dire *ar c'houer Yvon Bouttier a zo da gaozeal diwar-benn al labour ar peizant* ['the *kouer* Yvon Bouttier will be talking about the work of the *peizant*']. Du coup les gens comprennent que le *kouer*, il va nous parler de son métier de paysan, c'est qu'il doit être paysan

139. Used in dialects other than Léonais, including standard Breton.

The use of borrowings and neologisms is also pertinent here, as it intersects with the use of dialects: according to the stereotypes about new and traditional speakers, borrowings ought to be associated with dialects and neologisms with the standard. The following section will take up the discussion of this lexical choice.

### 5.6.4 Vocabulary

#### 5.6.4.1 Borrowings and neologisms

Interviewees were not informed that borrowings and neologisms were the main focus of the study, as it was felt that this might influence their answers, given that much research on the topic has tended to repeat the usual list of stereotypes about new and traditional speakers. Nonetheless, where it was appropriate during the interviews, participants were asked about their own attitudes towards borrowings and neologisms, as well as the attitudes exhibited towards these by their organisations.

Interviewees generally had nuanced views of borrowings and neologisms; many pointed out that both were necessary depending on the context of their use. Two participants noted that if a neologism was taken up by those interviewed on the radio, and subsequently used in the wider Breton-speaking community, this was an indicator that the term in question was well-crafted and appropriate:

> **P8** j'ai remarqué que les personnes que nous interviewons fréquemment tendent à acquérir eux-mêmes certains termes qu'ils n'avaient pas originellement dans leur propre vocabulaire. Ils les captent à la radio, ils les entendent dans nos émissions, et donc—je ne sais pas si le nombre de termes serait conséquent ou pas : plusieurs dizaines de termes au minimum sont aujourd'hui maîtrisés par des gens qui sont sollicités pour intervenir à la radio. ... Au fur et à mesure, il [the interviewee] fera attention à comment il parviendra à utiliser, à placer ce mot-là, dans le sens qu'il faut, au moment où il le faut. Donc les médias audiovisuelles ont une capacité d'influence sur ce point de vue-là

Only one interviewee raised the point that borrowing occurs in every language, and so there is no reason to reject it in Breton. It is interesting to note that this participant also noted her own propensity towards playing with multiple linguistic repertoires: she could be considered to have a less bounded view of languages than the others.

The most commonly expressed view was that while both borrowings and neologisms had their place in Breton, it was often better to avoid both. In the case of neologisms, this could depend on how the word was formed and became established in Breton: one participant stated that he generally avoided technical words created through neologism, but cited the example of a word he had used recently, *rannembregerezh*.[140] As well as the semantic transparency of this word, rendering it easier to understand, his use of it was motivated by careful research that relied more on attestation and personal experience than on prescription:

> **P2** 'filiale', en français, les gens le connaissent, mais en breton il y a un mot qui a été créé pour désigner ça. Et moi je ne le connais pas, et souvent j'ai le principe, si moi je ne connais pas, c'est qu'il y a énormément de gens qui ne vont pas le connaître non plus, et qu'il vaut mieux chercher un autre mot. Et donc moi j'ai trouvé un autre mot, c'est *rannembregerezh*. ... Je cherchais dans les archives en français 'filiale', j'ai regardé les jours, voir quand eux ils en avaient parlé dans les infos en breton, pour savoir comment dire

While the participant from the *OPLB* suggested that neologism was preferable to periphrasis, noting a desire to be "plus précis" (**P9**), the latter was otherwise a popular strategy for avoiding having to choose between neologisms and borrowings. One participant stressed the importance of not merely calquing from French, matching the findings in section 4.2.1. Another illustrated this by giving the example of a news report he had presented on the subject of a rise in the cost of doctor's appointments.

> **P1** je me souviens d'un collègue qui avait une brève sur les consultations qui augmentaient chez les médécins : quand on va au médécin, on payait telle somme, ça s'augmentait de 1 ou 2 ou 5 euros. Et je me souviens, il

140. 'Branch' (of a company), from *rann* ('part') and *embregerezh* ('business organisation').

> m'avait fait une brève en breton, j'avais écouté, j'avais trouvé ça complètement déplacé. Il avait traduit mot-à-mot quasiment ce que l'on l'avait dit en français. Moi ... je l'ai dit en breton : *koustañ a raio kerroc'h mont ti ar medisin*, il coûtera plus cher d'aller chez le médécin. Avec *pemp euro a gresk*, par exemple, avec cinq euros d'augmentation

The same participant also stressed that it was possible to discuss most technical subjects in Breton without needing to create or borrow words, citing a report he had presented on the subject of embryonic implantation in calves. He noted that the topic had been discussed without recourse to borrowings, and neither had complex neologisms been used: instead, the report had been in "un breton que tout le monde pouvait comprendre" (**P1**). This is therefore likely to have involved periphrasis and the avoidance of calques in the same way as the above example.

Another strategy mentioned by a number of interviewees was the use of doublets, similarly to the above example of *kouer*: this was deemed particularly appropriate for discussions of public administration, in cases where the standard term used in official contexts, often as a result of recommendation in TermOfis, was judged to be a technical neologism. Using this technique, the presenter would favour this officially prescribed term throughout the discussion, but accompany it at its first mention by its equivalent in "français bretonnisé" (**P3**). This would ensure that traditional speakers, or any speakers less accustomed to the use of certain neologisms, would be able to follow the discussion. This method conforms with the use of standard Breton in a lexical field where it is particularly appropriate, and yet does not exclude those who may not be acquainted with certain standard terms. As such, it is a pragmatic attitude towards neologisms that prioritises communication over an alignment with either new or traditional speakers, with the objective of becoming what Pennycook (2012:99) terms a "resourceful speaker", one who is "good at shifting between styles, discourses and genres" in order to communicate with speakers with diverse backgrounds and repertoires; the primary goal is not approximation to the language of the traditional native speaker, nor necessarily to convey a particular ideological stance but rather "the need to learn how to negotiate and accommodate" (ibid.). Similar pragmatic concerns are also evident in the case of dialectal words, as discussed in section 3.4 as well as in the

present chapter.

#### 5.6.4.2 Dictionaries and terminology sources

The previous section indicated that the use of neologisms and borrowings tends to occur on a case-by-case basis, often depending on context and the relevant lexical field. Similar opinions were frequently expressed regarding the roles of dictionaries: many participants named multiple dictionaries, noting that they would choose a particular one depending on the context in which the word was to be used.

The most frequently cited dictionaries were those of Francis Favereau and Martial Ménard,[141] although participants often mentioned the different strengths of the two: Favereau's dictionary was seen as "le plus courant" (**P8**) and indeed "plutôt dialectale" (**P3**), tending more towards "français bretonnisé" (**P4**), "plein de formules différentes dialectales" (**P7**), while Ménard's was "extrêmement rigide" (**P8**), "avec des néologismes et un niveau de breton plus académique, littéraire" (**P3**). Choosing between them would thus depend on context; additionally, "ça dépend de la sensibilité de la personne" (**P4**). Radio stations did not typically have a specified policy on the use of dictionaries, but instead used a range of resources as employees saw fit, including Favereau and Ménard, but also other publications such as dialectal dictionaries. This was particularly the case for speakers of Vannetais, as a number of published Vannetais dictionaries are in existence. Online resources, including Devri, were also cited, as accessing them was often more convenient. TermOfis was mentioned, but only in the context of terms relating to public administration, and interviewees stressed that it was not a major source.

The print publications have more developed policies on the use of dictionaries, as can be expected from the nature of the medium, where consistency of orthography is important. *Brud Nevez* makes use of Favereau and Ménard's dictionaries, but also of the dictionary published by Emgleo Breiz, this being one of the few that use the particular orthography employed in the magazine. It is therefore likely that lexical choices in *Brud*

141. Neither of these dictionaries was examined in relation to the appearance of tokens in the corpus investigated in this study: both are extremely comprehensive, and the statistical indications of trends across the subcorpora were unlikely to have been as obvious.

*Nevez* will be influenced by the decisions made in the compilation of this dictionary.

*Bremañ*, on the other hand, uses the monolingual dictionary published by An Here as a principal source of reference. This is perhaps a contentious choice, given the controversy surrounding the dictionary's political position: on publication in 2001, it gained notoriety for its anti-French political views, acting as an example of how language revitalisation can be ideologically linked to political separatism.[142] The interviewee from *Bremañ* gave two reasons why this dictionary is favoured: firstly, it was written by a team, rather than a single person, which suggests greater scientific rigidity; secondly, its monolingual nature "apportait quelque chose de différent, une sorte de base un peu plus solide" (**P7**). Once the decision had been made to use this dictionary as a main source, this had been subsequently adhered to for the sake of consistency. Unsurprisingly, the participant made no mention of the dictionary's controversial status. As this chapter has shown, political militantism is often associated with the stereotypical characteristics of the "néo-bretonnant": the use of this dictionary therefore fits the observed preference for a more conservative type of "néo-breton" in *Bremañ*. Additionally, the fact that the publication favours a monolingual dictionary indicates alignment with the stereotypical "néo-bretonnant" who wishes to avoid French as much as possible, although, as observed, the participant from *Bremañ* noted that using French is a necessity in the workplace.

Strategies other than using dictionaries were also cited as ways of finding terminology. The collaboration among the journalists at the Radio Breizh stations using online distance call technology has already been mentioned: interviewees noted that if one of the journalists was unsure about how to express a concept in Breton, they would consult their colleagues at the other stations. Similarly, the participants from France Bleu Breizh Izel stressed that there was no de facto principal reference dictionary at the station, but that instead, staff would pass terms on to each other informally. One noted that in the initial stages of his employment, he had "fait un petit carnet" (**P2**) to serve as a personal reference guide. However, he stressed that knowledge of actual usage was a preferable source:

142. "Cet outil linguistique dissimule, sous couvert de citations pédagogiques, des slogans militants. Ainsi, le verbe « être » est illustré par « La Bretagne n'existera vraiment que lorsque le français sera détruit en Bretagne » ; la préposition « entre » par « Il faut choisir entre la Bretagne et la France », etc." http://www.lexpress.fr/informations/le-coup-de-balai_642102.html, accessed 26 Oct 17.

> **P2** [je n'aime pas] rester trop attaché aux dictionnaires. ... regarder dans le dictionnaire, c'est l'un des dernières choses à faire. Si tu ne connais pas le mot, vas-y, mais il faut faire un peu le ménage avant ... ce sont des choses que l'on aurait gardées dans l'oreille, qui étaient dites par des gens, et c'est souvent plus pertinent que si on peut les trouver dans les dictionnaires

Finally, at *Bremañ* there was an additional alternative way of finding vocabulary: as the magazine is produced by Skol an Emsav, which also provides Breton courses, the editorial staff work in physical proximity to language teachers. The interviewee from *Bremañ* described how he would tend to use these colleagues, who often have a more advanced linguistic knowledge of Breton, as a source for vocabulary in the first instance, rather than resorting to the use of dictionaries.

These alternative strategies show the importance of the speaker community, including new speakers, in transmitting appropriate vocabulary. It also attests to the primacy of bottom-up usage rather than official recommendations, as also indicated by the importance of the media as a vector for new vocabulary and the ability to test the success of a neologism by seeing the extent to which it is taken up in the broader community. The focus on community norms rather than top-down standards shows how speakers and their practices are typically accorded more prestige than are authorities.

## 5.7 Conclusion

The way in which the information presented in this chapter interacts with stereotypes about new and traditional speakers and their language, and with the linguistic features of the corpus highlighted in the previous two chapters, will be discussed in the final chapter, which presents overall conclusions. Before doing so, however, it is worth drawing together some of the general trends identified during the interviews, beginning with observations relating to sociolinguistic variables.

### 5.7.1 Patterns across sociolinguistic profiles

As the sample of participants was small, it could be misleading to generalise about the entire Breton-speaking population based on these data. However, some patterns can nonetheless be identified; these are worth summing up here before investigating their role in the stereotypical new/traditional speaker divide (in the final chapter).

#### 5.7.1.1 Age

Age emerged as a relevant factor correlating with participants' expression of their Breton and French identity: older participants were more likely to view this in more binary terms, considering themselves either wholly Breton and not at all French, or an equal amount of the two. It was also two of the older participants who had not brought up their children using Breton, as their spouses in both cases did not speak it: in the case of the younger participant with a child whose father did not speak Breton, this did not stop her using Breton (as well as French) with the child anyway (see section 5.3.2). It is difficult to generalise based on these few examples, but this could reflect the generational divide, being a possible result of shifting attitudes towards multilingualism and the utility of speaking multiple languages in the home, in the light of recent trends that have encouraged child bilingualism as a way of encouraging connections to a minoritised language and even of developing certain cognitive abilities: see Madeg (2011) for a Breton-focused example of literature aimed at parents that emphasises these benefits.

#### 5.7.1.2 Gender

No strong patterning was observed based on gender, although in the sample, this correlated with age: the three participants in the older group were all men. This itself could be a result of the fact that in the past, men would have been more likely to take up authoritative roles in Breton-speaking society, which, being rural, would have taken a more traditional approach to customs involving gender roles. This may also explain why all the female participants worked for associative radio, a context less linked to the establishment than public radio.

A potential correlation between gender and language attitudes is that all those who expressed opinions on "incorrect" Breton were men: one participant stressed his desire to improve his own knowledge of the language (see section 5.6.1), while another characterised other radio employees as insufficiently competent, noting that "le niveau n'est pas bon" (**P1**). On the other hand, those who stressed the validity of all dialects were women, as was the interviewee who spoke about her own mixing of Breton and French in bringing up her children. Not all participants expressed opinions on this subject, however, so it is difficult to state with certainty that this is a trend among women.

Indeed, if this is the case, it must be asked whether it fits with general sociolinguistic observations about gender and language: specifically, the established tenet that women "use linguistic forms associated with the prestige standard more frequently than men" (Trudgill 1972:179). These observations appear to contradict this tendency, as the prestige standard in this case can be interpreted either as standard "néo-breton", or as a specific dialect, and certainly not a mixed variety (see discussion of dialect mixing in section 3.4.2 and of the avoidance of translanguaging in section 4.4.2). Again, the small sample size may be a factor here; alternatively, this could be an indicator that women are less ready to impose onto language a masculine discourse of established hierarchies (Connell 1987:109), and may consequently see less relevance in a standard variety explicitly imposed by authoritative bodies. Indeed, Cheshire (2002:431) notes that "women's behaviour in face-to-face communication has been shown to be more cooperative and listener-oriented than men's, and that women construct interaction around the objectives of providing support and solidarity"; in this situation, where the very notion of a prestige standard is contested due to the coexistence of conflicting ideologies, and speakers often perceive a high level of difficulty in mutual comprehension across dialects, the need to create this solidary atmosphere is of particular importance, as other findings within this thesis have suggested. Similarly, Tagliamonte (2012:34) stresses that the commonly cited trends in language use by gender are "oversimplifications" and that it is necessary to consider "what makes one form prestigious and another form stigmatised"; again, in this context, this is not a question with a definitive answer.

#### 5.7.1.3 Location

Three of the four interviewees based further east (either in the eastern part of Basse-Bretagne or in Rennes) could be considered more typical of the "néo-bretonnant" in their identification with militantism: one was more active in Breton cultural groups and expressed the strongest political opinions out of all the participants; another used the term "militants" in a general sense to refer to active speakers, suggesting that he identified with the category to a certain extent himself; the third was the most stereotypically "néo-bretonnant", in his preference for standard Breton, the way he framed the act of having raised his children to speak Breton, and his preference for using new terms over periphrasis. The speaker who referred to "mon militantisme" (**P3**), on the other hand, who was based in a more western location, did not appear to fit so well with these "néo-bretonnant" characteristics, despite manifesting some strong political opinions; she had a more relaxed attitude towards language use and the boundaries between Breton and French, and could certainly be classed as a language activist, if not as a "militant" in the sense found in McDonald (1989). As stated above, this demonstrates how the term "militantism" can be interpreted and identified with in different ways.

With the caveat of the extremely small sample size, it therefore seems from the interviews that the more typical characteristics of militantism and of the "néo-bretonnant" may be more apparent in areas further to the east, potentially because these areas have a weaker connection to traditional speaker communities, allowing them to develop their own practices relating to the use of Breton and the performance of Breton identity, which they may feel to be more vulnerable in this eastern region. A parallel, although more geographically extreme, is the case of Irish-American learners of Irish in the United States: Walsh and Ní Dhúda (2015:191) point out that in this context, "strong intrinsic motivation and considerable dedication are required to become and remain a new speaker of Irish, particularly abroad where the opportunities for socialisation are even more limited than they are in Ireland". This distance from the traditional speaker community thus encourages, indeed requires, a greater sense of motivation from the speaker, which can result in a more stereotypically "néo-bretonnant" positioning.

It must nonetheless be noted that even the participants based outside the traditionally Breton-speaking territory still had familial connections to this western area, which calls into question the idea that many new speakers are motivated to learn Breton despite a lack of connections to the language in their family: all participants had older relatives who spoke Breton, and all but the participant whose father was a new speaker had either grown up in Basse-Bretagne, or had parents who had done so. This suggests that family traditions still play an important role in the decision to learn Breton, this being more pertinent than a sense of pan-Breton identity that ignores the family and local community. While militantism may link to certain practices associated with the "néo-bretonnant", this need not be related to how the speaker actually acquired Breton, as sections of this conclusion will show.

#### 5.7.1.4 Education

No correlation can be observed involving participants' level of education, but this is because all participants were university-educated: most had studied languages, with Breton being the most frequently studied, and English also being common. Two participants had undertaken master's degrees; no particular effect of this could be observed.

At school level, more potentially relevant differences can be identified. Three participants had attended Diwan or bilingual schools. The only one to have attended a Diwan school was the participant with a new speaker parent; his consequently greater level of militantism than that of other parents may be a factor affecting this, as other parents may have been more concerned that their children acquire French in an educational context. In terms of correlation between education and attitudes or practices, it is difficult to find any common thread setting apart the three interviewees educated in Breton from the others. One commonality among them is that they all indicated a certain level of ease with adapting their dialect, mentioning diachronic change in some cases. However, this was also the case for some of the older speakers, implying that this is more likely to relate to age of acquisition regardless of whether this occurred in an informal or formal context, seeming to be easier for those who began acquiring Breton in childhood.

#### 5.7.1.5 Breton and French identity

As the section on age above indicated, the levels of Breton and French identity reported appear to have been influenced by participants' ages. Beyond this, Breton and French identity does not seem to be connected to any other findings from the interviews. Apart from the one participant who marked his Breton and French identity both at around 0.5, the other eight all expressed similar degrees of Breton identity, all above 0.9, with French identity ranging between 0 and 0.7; it could be expected that level of French identity may inversely correlate with militantism, but this was not the case. Because of this consistency beyond the one outlier, it is difficult to claim that any other aspects of the data correlate with participants' Breton and French identity, other than that their status as active Breton speakers is typically matched by a high level of Breton identity.

#### 5.7.1.6 Method of acquisition

Acquisition is of course the principal benchmark for distinguishing between new and traditional speakers; as previously noted, this does not always match with whether a speaker displays the characteristics of the stereotypical "néo-bretonnant" in terms of practices relating to militantism. Moreover, even purely in terms of acquisition, it is sometimes difficult to classify speakers. The participants can be grouped into sets according to their principal method(s) of acquisition as follows:

- Family and community: 2
- Family and school: 2
- Primarily school: 1
- Primarily adult education: 4

The fact that the group of nine interviewees can be split into as many as four sets emphasises the diversity of ways of acquiring Breton, depending on location, family background and time: acquisition within the community would have been impossible for the younger participants, while acquisition through school and adult education would

have become much easier to access in recent years. This is indeed reflected in the intersection of the distribution of participants' ages and their acquisition methods: those who acquired Breton in the community both fall into the older group, and those who learnt it through adult education into the younger group. However, again, as the small group of participants is made even smaller by this fragmentation, it is difficult to suggest any potential connections between perceptions and/or practices and method of acquisition, beyond the earlier observation that those speakers who acquired Breton before adulthood may find it easier to switch between dialectal varieties. When the stereotypical attributes of the two types of speaker are brought into the mix, it becomes still more complicated: all participants fit in some ways with the stereotypical "néo-bretonnant", regardless of their method of acquisition of Breton, but all defy the stereotypes in other ways, as the final chapter will summarise.

### 5.7.2 Other points of interest

The media were recognised as an important vector for new terminology, highlighting the importance of vocabulary use in media contexts as being of great influence on the general speaker community. The range of programmes on different subjects and in different styles on the Radio Breizh stations came across as an asset: in cases where this range is more limited, such as on public radio and television, potential audiences can be discouraged. If a listener or viewer happens to dislike a programme, and it is the only programme broadcast, the listener's options will be limited; if it is one of a range of programmes, there will still be plenty of other opportunities to hear Breton in other forms.

The media organisations themselves are similar in that they all employ a very small staff; employees must therefore possess multiple skills. In the case of the radio programmes, staff plan, script, present, and edit their own material, meaning they must be highly motivated. Employees' different interests and skills can be extremely influential over the way in which an organisation produces and distributes material: this can be seen in the contrasting attitudes taken towards the internet at *Brud Nevez* and *Bremañ*. Other differences between the organisations, particularly the different style

and content of the magazines, are often not a result of any strict editorial line, but rather of a set of traditions developed over the years. This is no doubt strengthened by the fact that articles tend to be written by a particular group of authors in each case, who will have become accustomed to writing about certain subjects, using a certain style.

As the chapter shows, there was a general sense throughout the interviews that dialects are to be valued over standard Breton. This is markedly at odds with the stereotypical perception of "néo-bretonnants". However, it can be noted that participants often characterised their own dialects as not particularly strong, in opposition to the dialects used by older traditional speakers. As noted, their accounts suggested a form of convergence: those who initially used stronger dialects had developed a more standard variety, while those who had initially learnt standard Breton had become more dialectal. This attests to the fact that the Breton-speaking community, aided by new technology, is increasingly interconnected, made up of speakers from all over Brittany with a range of dialects. Developing a distinct but comprehensible dialect therefore allows a speaker to be understood throughout the pan-Breton community, while nonetheless displaying their own local identity.

Participants indicated nuanced attitudes towards the lexicon of Breton, their consultation of specific dictionaries often depending on the lexical field in question. The general preference for periphrasis and strategies involving the use of doublets indicated a sensitivity to the broad range of repertoires and competences among audiences by avoiding the choice between neologisms and borrowings as much as possible. Overall, the stereotype, suggesting that new speakers, and by extension speakers in the contexts under examination, will take a "néo-bretonnant" stance and prefer Celtic-derived vocabulary and neologisms, appears to be a significant oversimplification.

# Chapter 6

# Conclusions: Contemporary Breton and its speakers

## 6.1 Introduction

This final chapter returns to the research questions posed in section 1.6, considering each one in turn in light of the data presented in chapters 3–5. Section 6.2 deals with stereotypes around "néo-bretonnants" and their language, determining whether these are attested in the corpus and interview data, and thus covering research questions 1 and 2. Section 6.3 looks at the corpus data in more detail, examining how differences in medium and register among the three subcorpora are reflected in the type of language they contain, focusing on research question 3. Section 6.4 discusses whether the results of the investigation point to any particular changes in the linguistic content of modern Breton or in the way speakers use the language, dealing therefore with research question 4. Finally, section 6.5 suggests directions for future research prompted by this thesis.

## 6.2 Stereotypes around "néo-breton" and "néo-bretonnants"

The principal subject of enquiry in this thesis was new speakers of Breton and their language, and in particular whether these can be defined with reference to the stereotypical characteristics of the "néo-bretonnant" and "néo-breton". It is important to recall that traditional speakers were not a direct subject of investigation in this research: the contexts examined were specifically those where we would expect to find the language of new speakers. It may be the case that traditional speakers of Breton overwhelmingly conform to the stereotypes that are assigned to them: based on the current research, it is impossible to offer any evidence that either supports or weakens this assertion. On the other hand, having focused on new speakers, it *is* possible to draw conclusions about this category. Consequently, this section will examine only stereotypes around new speakers and their language, rather than attempting to draw any conclusions about traditional speakers. Research questions 1 and 2 will both be covered here, in sections 6.2.1 and 6.2.2 respectively, in order to draw general conclusions relating to both in section 6.2.3.

### 6.2.1 Linguistic stereotypes

A number of linguistic stereotypes about "néo-breton" were listed in section 1.4.2. As this research focused on lexicon, some of these, such as the phonetic qualities of this type of language, were beyond its scope. However, lexical aspects were examined in detail across the corpus, and the results of this investigation can be supplemented by findings from the interviews. Some other supposed features of "néo-breton", particularly concerning morphology and syntax, can also be discussed here, as relevant findings were made during the investigation of lexical aspects of the corpus data.

#### 6.2.1.1 Standard and dialectal Breton

The stereotypical perception of "néo-breton" is as a standardised, non-dialectal variety (e.g. Jones 1995:428; German 2007:146; Rottet 2014:212). Before determining whether this is the case, it is important to distinguish these two characteristics from each other:

as interviewees pointed out, speaking a non-dialectal form of Breton does not entail using an explicitly prescribed standard. The two participants who believed their Breton to be broadly non-dialectal both stated that it was not particularly influenced by overt prescription, but instead by the fact that they communicated with diverse contacts from different backgrounds.

The supposedly non-dialectal quality of "néo-breton" will hence be considered first. Findings from the interviews in particular in fact suggested the opposite: those working for the radio stations stated that the use of a moderate amount of dialect was usually preferred over a non-dialectal variety. In print, too, dialects are tolerated, even in the more stereotypically "néo-breton" publications *Bremañ* and *Ya*: we see evidence of this in the corpus, where Vannetais vocabulary appears in *Ya*, and in the interviews, where the representative of *Bremañ* affirmed that the magazine accepts articles written in dialectal Breton as long as they use the *peurunvan* orthography.

While the dialects are tolerated and indeed favoured by many of the interviewees, and dialectal features appear in the corpus, some dialects are more visible than others. In the corpus, Vannetais was by far the most commonly occurring dialect, followed by Léonais, whose prevalence was explained by its being the basis for much of standard Breton, and thus in fact often not marking dialectal features at all. Vannetais therefore seems to be the dialect that has persisted most successfully; this is no doubt due to its greater linguistic distance from the others, resulting in a history of separate codification and more visible indexical features.

While interviewees expressed generally positive attitudes towards dialects, the occurrence of dialectal tokens in the corpus is in fact fairly low. Many contributors to the corpus use dialectal features, but as section 3.4 shows, this tends to involve only a restricted set of lexical items. As the section discusses, these are typically closed-class words such as prepositions, and it seems likely that this could be another way of asserting a local, dialectal identity while not limiting comprehension. Similarly, an interviewee's description of her own variety as "vannetais *soft*" also contributes to this impression. Overall, then, the use of dialects is generally favoured, but normally restricted to a specific small set of features. Heavier examples of dialect use were seen in

the Facebook subcorpus, but these posts attracted less engagement, with replies also being dialectal. In these instances, the concern for expressing local identity has won out over the need to attract engagement from the entire virtual community, but they are the exceptions to a rule that generally balances both in the form of only moderate dialect use, discussed as "dialect convergence" in section 5.6.

As stated above, even those interviewees who emphasised the non-dialectal quality of their Breton nonetheless denied that what they spoke was dictated by a prescribed standard. Indeed, the corpus also attests to this: attempts in TermOfis to impose strict semantic boundaries on pairs such as *publik/foran* and *dañjer/arvar* do not appear to have been successful. Indeed, many interviewees confirmed that they would rely on the *OPLB*'s recommendations only in specific lexical fields, such as public administration. It is therefore evident that the concept of standard Breton, particularly in the sense of a variety imposed from above, is not typically accorded the prestigious position that stereotypical accounts of "neo-breton" suggest: even among those who do not use dialectal varieties, community norms are considered more valuable than imposed standards.

#### 6.2.1.2 Neologisms and borrowings

The use of purist neologisms with Celtic roots, rather than borrowings from French, is one of the most visible and most frequently cited supposed attributes of "néo-breton". This dates back to before the significant decline of the traditional speaker population, to the nineteenth century and earlier lexicographical traditions (Constantine 1996:13), but is characterised as having come to a head in "néo-breton".

The characterisation of neologisms as based on Celtic roots suggests that new Breton vocabulary denoting recently invented concepts will typically be Celtic-derived. In some cases, we would expect this vocabulary to be used exclusively by new speakers: given the stereotypical depiction of traditional speakers, it would be unusual to find them discussing technology, for example. New vocabulary in these domains is thus particularly likely to derive from Celtic sources, if the stereotypes around both types of speaker are accurate. However, as section 3.3.2 shows, 78.3% of the tagged tokens in

the corpus containing French-derived elements are not found in Breton prior to 1900, indicating that borrowing has remained a frequent process in the twentieth and twenty-first centuries, and that words created in this way are used frequently in contexts where "néo-breton" would be expected.

Celtic-derived neologisms nonetheless also appear frequently, tending to occur in greater quantity in contexts that can be characterised as more typically "néo-breton" (see section 6.3). Interviewees pointed out, however, that the acceptance of neologisms tends to be less firm among older speakers, implying that traditional speakers conform to their stereotypical attribute of preferring borrowings. One strategy mentioned in response to this was the use of doublets in radio programmes in order to ensure that audiences with diverse repertoires can be included. Participants explained that both the neologism and the borrowing would be used at first mention, followed subsequently by the neologism alone. This suggests that neologisms ultimately tend to be favoured over borrowings on the radio: however, this was in the specific context of terms relating to public administration, and as interviewees affirmed with reference to dictionary consultation, different lexical fields can demand a different approach to the use of neologisms or borrowings.

While this research does not group words by their semantic field, similar observations can be made based on the corpus: for example, discourse markers, a more informal and spontaneous feature, often come from French. This could be influenced by the fact that French will be a dominant language for many of the speakers in the corpus; however, it can also be interpreted as a marker of a more casual register of Breton that may be perceived as more similar to the language of traditional speakers. Indeed, the role of perception is more generally important in constructing "néo-breton", as discussed in section 4.3.5. Many interviewees spoke about avoiding the controversial matter of borrowings versus neologisms altogether by using periphrasis. In general, at least among interviewees, there appears to be a high level of awareness of the ideologically charged nature of the subject, resulting in alternative strategies that avoid having to explicitly take a position while remaining comprehensible to audiences from diverse backgrounds.

Section 3.5 also investigated the integration of neologisms in the corpus, considering morphological, orthographic and phonological aspects. It was noted that orthography tends to be altered more readily than phonology: it seems more important to alter borrowings so that they conform to Breton spelling rather than to its sounds. This reflects the fact that all Breton speakers within Brittany will also speak French, meaning French phonology is easy to deal with. It also perhaps reflects the historic importance of orthography in the Breton revitalisation movement.

#### 6.2.1.3 Calques, affixation, and compounding

Some methods of word formation were investigated in section 4.2. Lexical calques from French were found to be infrequent in the corpus, and a number of interviewees expressed negative attitudes towards them, stressing the importance of not translating directly from French. As this attitude would suggest, French idioms tend not to be calqued in the corpus. However, the use of traditional Breton idioms is also low, with only a few more well-known examples occurring. Less well-known traditional Breton idioms are presumably unknown to many new speakers, resulting in a situation where the range of idioms available is extremely restricted.

Clippings and semantic expansion are used in the same way as in French, and often operate on the same words or their equivalents, such as *jedo* from *jedoniezh* (French *maths*, *mathématiques*—see section 4.2.2) and *kreizenn* (French *centre*—see section 4.2.3), cases of clipping and semantic expansion respectively. This can be connected to the longstanding contact situation between Breton and French, showing that while French influence may be consciously avoided, this is not always possible at a less visible structural level.

Affixation is another common word-formation process. While excessive application of affixes has been criticised in the literature (Hornsby 2009b:86), section 4.2.4 shows that in some cases it can convey subtle semantic differences that would not be expressed as elegantly otherwise. Indeed, while literature on "néo-breton" can condemn affixation and semantic expansion through metaphor (German 2007:186), it must be noted that these are common cross-linguistic processes for expanding a language's lexicon: the

fact that such processes are criticised when used in minoritised languages highlights that the cause in this case is ideological rather than motivated by linguistic function, displaying a preference for traditional linguistic forms and practices even when this may hinder adaptation to new domains of use.

In general, French structural influence can be observed in word-formation processes, particularly in cases of semantic expansion: this reflects the fact that French is a first language for many speakers and a dominant language within the Breton-speaking environment. However, in more surface structures such as idioms and calques, efforts are clearly made to avoid French influence. In a way, this is similar to the stereotypical "néo-breton" characteristic of consciously avoiding French influence on a surface level while unavoidably resembling it on a structural one. However, it is commonly claimed that the surface level in this situation equates only to the lexicon, while syntax is generally influenced by French. This case demonstrates, on the other hand, that speakers of "néo-breton" also take care to avoid French influence in terms of certain syntactic structures, suggesting that their level of linguistic understanding may be more advanced than stereotypical claims would suggest.

#### 6.2.1.4 Errors and competence

Some research has characterised "néo-bretonnants" as lacking competence in the language (e.g. Hewitt 2016). Indeed, while phonology and morphology were not considered in this thesis beyond the specific aspects mentioned in the previous section, some examples of apparent errors in pronunciation and mutations were observable in the radio and Facebook subcorpora. However, speech errors would be expected even among traditional speakers, and moreover, such apparent mistakes may in fact be cases of dialectal divergence from the standard, which is not criticised in the speech of traditional speakers. Further investigation of this would require direct comparison with the language use of traditional speakers.

An easier area to investigate in this regard is the extent to which learners of Breton are catered for in the media. Two interviewees spoke about the provision of glossaries: one associative radio station makes them available for certain programmes via its website,

while *Bremañ* can include glossaries in its articles. The former case is an example of how technology can be used to enhance the experience of certain audience members, while not encroaching on that of more competent speakers: only those who require the extra information need engage with it. However, it also requires an internet connection and relies on listeners being aware of the service's existence.

In *Bremañ*, as the interviewee explained, glossaries tend to gloss words in Breton rather than resorting to French (although one example does use French in the corpus), and to include a minimum of grammatical information, meaning they are not particularly suited to learners. Indeed, *Bremañ*'s sister publication *#brezhoneg* is targeted specifically towards learners, meaning that *Bremañ* is able to focus completely on competent speakers. Similarly, while learners often post in *Facebook e brezhoneg*, the group's instruction to members to post in Breton rather than French discourages this. In general, then, there is a sense that the particular sources investigated are intended as spaces where competent speakers can use Breton without having to make allowances for those less competent with the language, creating a protected space where French can be avoided as much as possible. For learners, completely separate media spaces are available, such as *#brezhoneg*, other Facebook groups, and occasional radio programmes in French that aim to teach Breton phrases.[143] While this echoes how linguistic spaces in media are constructed for dominant world languages, it nonetheless perpetuates the idea that French should be excluded from Breton spaces, perhaps ignoring the reality that nearly all Breton speakers are competent in French. However, this state of affairs does contradict the characterisation of new speakers as lacking full competence in Breton: if this were the case, we would expect more allowances to be made for reduced competence.

It was nonetheless evident from the interviews that competence in Breton can entail more than a familiarity with the standard. Interviewees emphasised that dialectal forms of Breton were often more prestigious, particularly on the radio. The ability to adapt one's variety and communicate with speakers of other dialects is thus an additional skill required by fully competent speakers. Among participants who had acquired Breton through education or in adulthood, the sentiment was expressed that working in media had allowed them to make significant progress in this regard, par-

143. *Le p'tit cours de breton* is a short daily programme on France Bleu Breizh Izel with this purpose: see https://www.francebleu.fr/emissions/le-p-tit-cours-de-breton, accessed 13 Mar 19.

ticularly in their ability to communicate with speakers from different dialectal backgrounds. Speakers who do not use Breton in these particular contexts may therefore have fewer opportunities to communicate with a range of interlocutors, and particularly with traditional speakers who use more diverse dialects. Going by this criterion, the assertion that some new speakers may lack competence is therefore perhaps true, going by some of the interviewees' own standards. However, it is important to note that not all speakers view this type of competence as a necessary goal: different speakers characterise competence in different ways, which can lead to a variety of goals for learners. It is also evident from the interviews that speakers are often aware when their own competence fails to meet their expectations, and work actively to improve it; this again shows that new speakers can be more linguistically aware than they are stereotypically depicted.

#### 6.2.1.5 Multiple repertoires

Stereotypical depictions of "néo-bretonnants" as favouring purism imply that most new speakers would be keen to preserve the boundaries between Breton and French, despite the fact that all Breton speakers within Brittany will be bilingual: the political separatism typically associated with the "néo-bretonnant" influences the language ideologies of the same group, inspiring a need to emphasise the distinctiveness of Breton by keeping it apart from French. Indeed, the previous section attests to this in its discussion of the discouragement of French on Facebook and in the glossaries found in *Bremañ*. This is also evident from *Bremañ*'s explanation of its inclusion of French bibliographic information (see section 3.2.1). Likewise, from both the corpus and the interviews, the general importance of keeping Breton separate from French at the level of discourse became apparent. In the corpus, French is used on Facebook, but is kept apart from Breton, rarely mixed within the same sentence or even the same comment. In the interviews, nearly all the participants confirmed that they would use Breton with friends and contacts who spoke it and French when communicating with those who did not, rather than mixing languages, even in cases of shared bilingualism. For the most part, despite bilingualism within the community, translanguaging therefore cannot be observed, as section 4.4.1 concluded. As with other minoritised languages,

as also pointed out in that section, the dominance of French and the historical purism within the Breton revitalisation movement are likely to contribute to this situation, promoting the conception of Breton as a clearly bounded language in order to emphasise its typological distinction from French, as well as to be taken more seriously by outsiders. Moreover, questions of identity are relevant here: for those who wish to affirm Breton identity, using Breton as much as possible would be particularly desirable.

Within Breton, however, multiple repertoires can be combined, most often in the use of synonyms from different dialects. An institutionally prescribed example is the reshaping of dialectal words to take on particular semantic qualities, such as the use of *arvar* in TermOfis with a slightly different meaning from that of *dañjer* (see section 4.3.6). Similarly, compound words such as *holl-razh*, with synonymous elements from KLT and Vannetais respectively, allow greater expressivity, as discussed in section 4.3.3. Another slightly different case of the combination of multiple repertoires within *Bremañ* is the acceptance of words from multiple dialects without specifying to the reader which dialects they belong to, thus tacitly encouraging readers to accept them all as valid Breton, enabling the expansion of the language's vocabulary without resorting to either borrowings or coinages.[144] Indeed, section 3.4.2 notes some examples, from the Facebook subcorpus, of users apparently drawing vocabulary from multiple dialects. The combination of multiple repertoires within Breton thus seems more acceptable than that of multiple languages, implying that boundaries between languages are considered more important than those between dialects; again, this is likely to be related to perceptions of Breton identity as non-French, implying a general position that fits the stereotypical "néo-bretonnant" more closely. Accepting dialectal words may mark a departure from the historical purism of "néo-breton", but on the other hand, accepting words from *all* dialects without attempting to ringfence them to particular contexts is a practice that traditional speakers and "native authenticists" may interpret as inauthentic: this is perhaps more accurately represented as a change in "néo-breton" practice rather than a departure from it altogether. Nonetheless, this move away from both traditionalist and purist ideologies shows the value of the new speaker paradigm for interpreting

144. A similar example, though based on diachrony rather than synchrony, is the concept of *tota Cornicitas* in revived Cornish, operating according to the principle that "since the entire remains of the Cornish language are meagre, it is legitimate to use any source of any period for the revived language" (Williams 1997:5).

the type of Breton found in these contexts, as it moves away from reliance on notions of the ideal native speaker and of native varieties as necessary targets for language revitalisation (O'Rourke and Pujolar 2019:17).

### 6.2.2 Non-linguistic stereotypes

Section 1.4.2 also listed a number of non-linguistic stereotypes around "néo-bretonnants". The corpus itself was analysed on a purely linguistic level, as obtaining information about speaker profiles here would have been difficult for reasons of both ethics and practicality. This section therefore mainly summarises findings from the interviews, although some features of the corpus data will be mentioned when relevant.

It must also be reasserted at this point that the interviews deliberately formed a more minor part of the research than the analysis of the corpus, and the sample size was small, with nine interviewees. Consequently, it is impossible to make generalisations about the entire Breton-speaking population based on the results. However, as chapter 5 showed, some trends can be identified, which may be salient in the understanding of how practices and perceptions interact.

#### 6.2.2.1 Acquisition as a second language

This cannot properly be referred to as a "néo-bretonnant" stereotype, as it is typically one of the core attributes of the new speaker in general. Indeed, nearly all the interviewees can be classed as having acquired Breton as a second language: five acquired it after French, and another two acquired it concurrently with French, but initially in a situation where it was not used actively within the family. However, the remaining two are exceptions to this, for different reasons. One acquired Breton as a first language, as a result of continuous uninterrupted intergenerational transmission. In terms of acquisition, he must be classed among traditional speakers, but other attributes (his age, militantism, and mobility, having spent part of his upbringing in eastern Brittany) align him with the "néo-bretonnant" overall. The other participant grew up

using Breton within the family, having acquired this from her father, who learnt it as a second language himself. She too therefore does not meet the criterion of having acquired Breton as a second language; however, she would typically be counted among new speakers for not having acquired Breton as a result of intergenerational transmission dating back indefinitely. This case demonstrates how not all new speakers can strictly be considered L2 speakers of Breton: a more nuanced definition is necessary, hence the working definition of the new speaker posited in section 1.4, in which case six participants can be classed as new speakers. Even after making allowances for this, however, this does not always correlate with "néo-bretonnant" characteristics.

#### 6.2.2.2 Acquisition in an educational setting

Stereotypical "néo-bretonnants" are supposed to have acquired Breton through education, and specifically through the Diwan immersion schools (Jones 1995:428; Adkins 2013:59; Le Dû and Le Berre 2013:53; Rottet 2014:213). However, this creates an interesting conflict with another typical "néo-bretonnant" attribute, that of having been motivated to learn Breton because of a strong sense of Breton identity; children acquiring Breton through education would have done so as a result of their parents' decisions, at a time before their own agency became a relevant factor. While the parents' identity is clearly relevant in such cases,[145] the child must also independently decide whether to continue using Breton later in life, regardless of whether they have undergone education in the language. This is reflected in the fact that many participants reported that they had begun to speak Breton in earnest at a specific point in their lives, which was related in section 5.3.1 to the concept of the linguistic *muda*.

Only one of the nine participants attended Diwan, and this was nonetheless not her primary means of acquisition. This indicates that the stereotype of "néo-bretonnants" learning Breton exclusively through the Diwan school system is questionable. As section 5.3.1 showed, the other participants acquired Breton in various ways, many of which can be classed as formal education in some form—but in total, only three parti-

145. However, education at a Breton-medium school does not necessarily have the primary goal of developing the child's fluency in Breton: Hornsby (2017:100) notes the example of a Breton speaker whose parents sent her to a Diwan school in order to take advantage of smaller class sizes. Through this postvernacular practice, the language is commodified and used as a vehicle for accessing other non-linguistic benefits rather than as a goal in itself.

cipants attended schools where lessons were taught in Breton, one at a Diwan school and the other two at bilingual state schools. In only one of these cases was formal education the primary method of acquisition.

#### 6.2.2.3 Age

"Néo-bretonnants" are claimed to be young, in opposition to older traditional speakers (Jones 1998b:321; Hornsby 2005; German 2007:153; Adkins 2013:58; Rottet 2014:213). However, again, the interviews revealed exceptions to this pattern. The distribution of participants' ages did reveal clustering into two groups, as this stereotype would suggest, but it was not the case that the younger group consisted exclusively of canonical "neo-bretonnants" and the older group of traditional speakers. While two of the older speakers acquired Breton from their families and communities, they nonetheless exhibit some "néo-bretonnant" attributes, and the third participant is undeniably a new speaker, having learnt Breton from textbooks and evening classes. Conversely, the younger speakers do not all fit the stereotypical "néo-bretonnant" profile, notably in the case of the participant who acquired Breton through continuous intergenerational transmission, being the only participant to have both acquired Breton in this way and used it as a main language since early childhood. While age can provide a guideline, it cannot be used as a way of delineating new and traditional speakers.

Some of the remarks participants made about their audiences are also relevant here. Interviewees tended to reinforce the fact of there being two groups of speakers, divided by age, but the claims they made, particularly about the older group, did not suggest that this group was populated solely by stereotypical traditional speakers: its members were characterised as militant by two participants. This echoes another participant's insistence that audiences must be strongly motivated to seek out Breton-language media. While these remarks do not necessarily contradict the fact that this older group may be composed mostly of traditional speakers, they do show that militantism and motivation are not characteristics restricted to the younger group, nor to new speakers.

#### 6.2.2.4 Urbanity

One of the stereotypes around "néo-bretonnants" is that they are based in more urban areas, while traditional speakers are said to live in the countryside (Jones 1995:428; Abalain 2000:76; Timm 2001:454; Timm 2003:34; German 2007:153; Hornsby and Quentel 2013:75). In this case, the stereotype is largely upheld by the distribution of participants. In terms of their provenance, those more closely linked to Breton-speaking families and communities often come from more rural backgrounds, while those whose parents did not speak Breton often came from towns. However, the two interviewees based further east do not fit this pattern; both had Breton-speaking parents from Basse-Bretagne but grew up in urban areas, mostly outside traditionally Breton-speaking territory, as the parents had moved away from their original localities.

At the time of the interviews, all the participants lived in comparatively more urban areas, ranging from small towns to small cities; none lived in the stereotypical traditional speaker environment of the small farming village. This is therefore one attribute of the typical "néo-bretonnant" that appears to be generally present among Breton speakers in the media, based on the small sample.

#### 6.2.2.5 Education and class

Education is also an attribute where the interviewees appeared to match the "néo-bretonnant" stereotype of being well-educated (Jones 1995:428; Timm 2001:454; Adkins 2013:58; Rottet 2014:213); indeed, all the participants had university degrees. However, it must be noted that all worked in similar professions; those in other careers may not be educated to the same level. In the same way, the social class of all the interviewees, as measured by their education and careers, would have been similar. While these attributes do fit the "néo-bretonnant" stereotype, this is because participants working in the media were specifically sought out; a broader sample of new speakers may have identified exceptions. It is therefore impossible to draw conclusions on this matter.

#### 6.2.2.6 Provenance within Brittany

According to the stereotypes around "néo-bretonnants", they may come from the eastern part of Brittany, and have no connection with the western part of the region (Jones 1998b:321). This was not typically the case among the interviewees. Three had grown up in the eastern part of Brittany, but two of these had parents who were originally from Basse-Bretagne. The other participant had a familial connection to Breton in that her father spoke the language, although he did not come from Basse-Bretagne himself. This is the one case where the extension of Breton to a pan-regional context can be identified, but it must be noted that it was not the participant herself who instigated this. While speakers may live outside Basse-Bretagne, it therefore appears possible that they will still have connections to the western region or to Breton-speaking family members. Notwithstanding the caveat, again, of the small sample size, the role of eastern Brittany as a locus of new speakers may therefore be less prominent than it appears.

#### 6.2.2.7 Breton identity

"Néo-bretonnants" are said to feel a particularly strong sense of Breton identity, in contrast with traditional speakers (Jones 1995:428; Le Nevez 2013:92; Rottet 2014:213). This has ties both with militantism, which will be discussed in the next section, and with regional provenance, suggesting that a pan-Breton identity is more significant to new speakers than a sense of belonging to a local area or to western Brittany. The previous section has shown that this may not be the case.

Nonetheless, interviewees did express a strong Breton identity, with all but one rating it at above 0.9 on the scale, compared with lower ratings for French identity, ranging from 0 to 0.7. As section 5.3.1 pointed out, many interviewees had experienced a decisive moment at a formative time in their lives, such as as teenagers or after finishing university, at which they had decided to learn Breton or committed to using it in preference to French where possible. This too attests to the importance of identity in guiding their decisions to speak Breton. It therefore does seem appropriate to class the participants in general as having a strong sense of Breton identity, linked to their use of the language, and in some cases, heightened by it, as some participants acknowledged:

this is another "néo-bretonnant" stereotype that seems a reality among users of Breton in a media context.

#### 6.2.2.8 Militantism and politics

"Néo-bretonnants" tend to be characterised as militants, keen to promote Breton widely and also involved in local political activities (Jones 1995:428; German 2007:187); indeed, the link between language revitalisation and regionalist politics in France was set out in section 1.2. Section 5.4 also discussed the meaning of "militantism" in the Breton context, showing that interviewees identify with the term in different ways: one of the most politically active rejected the label, while others were happy to use it despite not conforming with some other prominent characteristics of the "néo-bretonnant". One participant conformed closely to the typical militant model in his discourse and his desire to use Breton as much as possible (see below): overall, he was the most typically "néo-bretonnant" of the interviewees. However, the others expressed more nuanced views on militantism, showing different levels of conformity with the category; two participants in particular appeared to identify with the term, but are less clear-cut cases.

Nonetheless, many of the participants engaged with local politics to some extent. The most frequent political topic was the status of the Loire-Atlantique department as part of Brittany: it was clear that participants took the view that it should be considered part of the region, with some making sure to state this explicitly. Similar attitudes could be seen in the corpus: the sample from *Bremañ* made it explicitly clear that the magazine not only viewed Loire-Atlantique as part of Brittany, but assumed that its readers shared this position. This particular facet of political militantism therefore appears to be widespread among speakers of Breton in these contexts. However, beyond this, militant attitudes seem less global than some literature has implied, particularly where this involves uncritically self-identifying as militant. This may signal a loosening of the connection between language revitalisation and political activism in this context, allowing speakers with a greater diversity of political views to feel part of the Breton language community.[146] Like the identification of moderate use of dialects in the cor-

146. O'Rourke and Pujolar (2019:15) similarly point out that in many cases, despite stereotypical beliefs

pus and by interviewees (see section 6.2.1.1), this emerging depoliticisation of Breton can be linked to the growth of the new speaker community, which motivates, and to a certain extent requires, such a diversification of ideologies and practices, acknowledging that as the speaker population grows, more room needs to be made for those who do not hold the stereotypical political or linguistic beliefs of the "néo-bretonnant".

#### 6.2.2.9 Desire to use Breton as much as possible

The final non-linguistic stereotype surrounding "néo-bretonnants" is their supposed desire to use Breton as much as possible (German 2007:187; Hornsby and Quentel 2013:75–76). In the corpus, this was mostly borne out, with very little use of French on the radio and in the printed sources, particularly the latter: even in cases of defining difficult vocabulary, there was very little recourse to French. Overall, these contexts conveyed a sense of being Breton-only spaces.

On Facebook, the situation was more nuanced; despite the group's header image requesting participants not to write in French, frequent use of French could be observed. However, this was generally not mixed with Breton, suggesting that users valued linguistic boundaries rather than using the group as a translanguaging space (see discussion in section 4.4.1), despite general shared bilingualism. Overall, then, the use of Breton within the corpus implies that in these specifically delineated Breton-speaking contexts, it is indeed used as much as possible. Interviewees also confirmed this, by explaining that media in Breton typically create a context where competent speakers can expect to be fully immersed in the language.

It must additionally be asked whether this supposed desire to use Breton as much as possible may lead to situations where the language is used to an extent that would impede communication, such as in the presence of non-speakers. Among interviewees, this was generally reported not to be the case: participants' accounts of their use of Breton showed that the practicalities of communication would typically take priority over the performative assertion of their identities as Breton speakers. Participants con-

about new speakers' motivations, "learners are not engaging with the language for nationalistic reasons" but instead often for reasons connected to new circumstances and opportunities; they suggest that the *muda* is a more suitable way of conceptualising this (see section 5.3.1).

firmed that in most cases, they used Breton only when in the exclusive company of Breton speakers, changing to French as soon as communicative reasons made this necessary. Exceptions to this rule were mentioned by two participants: one related to the context of a campaign group composed of people sympathetic to the cause of Breton speakers, even if they did not speak the language themselves; the other was from the most stereotypically "néo-bretonnant" participant, who used Facebook as a space for Breton, despite some of his contacts not understanding it. However, it can be noted that beyond these more extreme cases, ease of communication typically takes precedence.

### 6.2.3 Summary

From this list of stereotypes around "néo-bretonnants" and their language, it is apparent that some are largely adhered to, while others appear less secure. While the participants were indeed well-educated, located in more urban areas, and typically strong in their Breton identity, there are ways in which they defied other stereotypes: acquisition of Breton had occurred in a wider range of contexts than we are told to expect, communication was shown to be typically prioritised over the widespread use of Breton, speakers tend to have links to western Brittany, and interviewees showed a range of ways of identifying with, or in some cases rejecting, the concept of militantism. Similarly, while an age divide was observed among interviewees, it did not map well onto other aspects of the supposed new/traditional speaker dichotomy. Taken together, the various ways in which the participants defied both "néo-bretonnant" stereotypes and many of the usual definitional criteria for the new speaker demonstrate the dangers of relying on such definitions as a starting point for analysis, as emphasised by O'Rourke and Pujolar (2019): the usefulness of the concept lies instead in its ability to shed light on social factors such as issues surrounding legitimacy and perceptions of authentic varieties, and how speakers position themselves in order to access the resources and capital that are associated with being a competent speaker of the language in question.

In terms of language, stereotypical features are again present to different extents. Attributes of stereotypical "néo-breton" can be observed in the persistence of French influence at a level less easily consciously modified, such as the use of particular discourse

markers and of neologisms created through semantic expansion; it can also be seen in the continued need to keep Breton separate from French at the level of discourse. However, other findings show that the extreme purism associated with "néo-breton" is an unfair representation of new speaker language, with speakers generally aware of this cliché and often making conscious efforts to escape it. This can involve strategies to communicate with diverse audiences and not show an overt alignment with either side of the stereotypical new/traditional speaker binary, such as the use of doublets or periphrasis, as well as the consideration of a word's lexical field on a case-by-case basis when deciding on the use of particular dictionaries. As interviewees attested, the development of the Breton-speaking community into a more interconnected and diffuse body allows new speakers to move away from dictionaries and prescribed standards and towards community norms and practices. Perhaps most strikingly, local dialects are valued and in many cases seen as desirable: this revalorisation of local varieties has been no doubt motivated by the growth of the Breton new speaker community, leading speakers to seek authenticity in local varieties. Certain parts of the corpus also attest less typically "néo-breton" language than others, as the following section will discuss.

Overall, the picture that emerges is more complicated than previous research suggests: it would seem an oversimplification to class all users of Breton in these contexts as "néo-bretonnants", conforming with all the stereotypes previously prevented. Characteristics that both reinforce and break down these stereotypes are present, as are conscious efforts to engage critically with Breton and move away from the clichés of purism or militantism, despite the difficulties caused in this regard by the dominance of French. Likewise, it appears that there is communication across the new/traditional speaker boundary, such as in the form of radio broadcasts. The prescriptions of the *OPLB* are required only in specific circumstances: in many cases, the community can sustain itself by operating on the basis of inter-speaker norms rather than prescribed standards. This overall situation must be taken into account when considering new speakers of any language: the separation of this category from traditional speakers will in some respects be artificial, ignoring shared experiences and community ties. Theorising a set of new speaker characteristics can lead to overgeneralisation—as O'Rourke and Pujolar (2019:18) also point out—as well as failure to account for the range of motivations

that may lead to the use of a minoritised language.

## 6.3 Variation across media and registers

This section sums up findings that respond to the third research question, which asked whether patterns can be distinguished in lexical variation among and within the three subcorpora. It was initially observed that as expected, the three map onto different linguistic registers according to the medium they employ: various aspects of the subcorpora indicate this, such as the higher rate of discourse markers in the radio data, and the lack of interjections in the print data. As expected, the print subcorpus is the most formal in register, while the radio subcorpus, consisting of spoken and sometimes spontaneous language, is the most informal. The Facebook subcorpus, as an example of computer-mediated communication, is of a written medium but displays characteristics of speech; it occupies an intermediate position in terms of register.

Among all the parts of the corpus, *Brud Nevez* contained the lowest rate of tagged tokens: this reflects its thematic focus on subjects that are more pertinent to Brittany specifically and do not require the use of newer vocabulary. This distances the *Brud Nevez* sample, thematically and linguistically, from the other two print publications, showing how different publications can cater to more or less typically "néo-bretonnant" speakers. It is also likely that *Brud Nevez* avoids the controversial choice between neologisms and borrowings by using simpler vocabulary or periphrasis.

Some differences across the corpus were observed in terms of word origins. The radio subcorpus showed the highest proportion of tagged tokens derived from French. This was particularly the case among speakers who were guests on the programmes, who differed from employees: employees often read scripted material, while guests are usually contributing spontaneous speech. It therefore seems that spontaneous speech is the medium in which French-derived words occur most frequently. There are several potential reasons for this: spontaneous speech will supply the lowest linguistic register found in the corpus, so simpler vocabulary is more likely to be used, and it may be the case that much of this comes from French in the form of long-established borrowings.

However, if this were the case we would expect a lower annotation rate, as in the case of *Brud Nevez*; long-established borrowings were not typically tagged unless they were controversial or had a Celtic-derived synonym in the corpus.

Alternatively, the higher occurrence of French-derived words could reflect the fact that speakers have less time to consider what vocabulary to use in this situation of spontaneous speech production. They are thus influenced by French as a dominant language, so French-derived words are more likely to occur to them. This too raises questions, though, as the use of French-derived vocabulary is a practice that is stereotypically attributed to traditional speakers for whom French is a second language, rather than new speakers as we expect to find in a radio context. If this is the case, therefore, the use of French-derived vocabulary in spontaneous speech would seem instead to be a trait we should associate with both new *and* traditional speakers. Indeed, the fact that traditional speakers are typically restricted to spoken Breton, if we accept the typical assertion that they are illiterate in the language, means that the greater use of borrowed vocabulary may be related to register rather than to their status as traditional speakers: if they were able to write in Breton, we might find that they would be more accepting of neologisms. On the other hand, the greater use of neologisms in written "néo-breton" may be specific to the medium of writing, as it is in such contexts that users of the language have more time to reflect on vocabulary choices and the opportunity to avoid borrowings. Without an examination of contemporary written Breton from traditional speakers, a variety that we are told is no longer available, it is difficult to evaluate this. It is possible, however, to conclude that informal spontaneous speech, even in this supposedly "néo-breton" context, contains more borrowings and fewer neologisms than stereotypical reports would lead us to expect, suggesting that if we are to accept this stereotype as valid, it should nonetheless be considered to be particularly characteristic of written registers.

Other parts of the corpus also displayed notable characteristics relating to the use of words of different origins. The Facebook subcorpus had the highest rate of Breton-derived words among the tagged tokens, implying that according to this criterion, users of Facebook produce the most stereotypical "néo-breton". However, in other ways, the Facebook subcorpus diverged from this descriptor, as subsequent

paragraphs will show. For now, it is worth noting that despite the informal register of Facebook posts when compared with print publications, it is nonetheless a context where Breton-derived words are common, suggesting that given the opportunity to reflect on their own vocabulary choices in using language in a non-spontaneous way, many speakers still prefer to use neologisms even in a less formal context.

The print subcorpus displays a similar adherence to "néo-breton" in that this is the subcorpus with the highest occurrence of Welsh-derived words: this befits the higher register found here. It also uses more "international" words: typically either learned borrowings that have entered Breton from French, or words originally from English, initially borrowed into French and then into Breton. This again reflects the higher register of this subcorpus, responding to its need for more specialised and technical vocabulary, and for words denoting more recently invented concepts. *Ya* in particular has a large proportion of "international" tokens. While the fact that these words are used in French may appear to distance them from stereotypical "néo-breton", section 2.2.4 argues that instead, these may characterise a newer form of "néo-breton" that is more inclusive of international vocabulary, while still tending to avoid words that are specifically native to French. Indeed, *Ya* also has the highest number of Breton-derived words among its tagged tokens, suggesting that it still conforms to the typical patterns of "néo-breton". The use of more "international" words thus implies that these are tolerated by virtue of their use in languages other than French.

Associative radio also attested significantly more international words among the tagged tokens than public radio, although in this case this was not accompanied by a lower proportion of French-derived words, so it may have been more influenced by subject matter; indeed, guests on associative radio also use more French-derived words than their counterparts among the public radio data. It may thus be the case that guests on public radio programmes may feel more pressure to use a more standard form of Breton in the more professionalised, better funded setting of public radio, compared with the more relaxed and localised associative radio stations; the prestige of the standard according to official bodies such as the *OPLB* may mean that in contexts less connected to grassroots activities and conversely more integrated into state structures, that variety is deemed more appropriate.

The morphological processes examined typically did not vary significantly among subcorpora. However, semantic expansion did, with the highest rate of occurrence again appearing in *Ya*. This contributes to the impression *Ya* gives as favouring a more innovative form of "néo-breton": in this case, it is drawing on existing resources to expand the language's vocabulary in a way done less by other parts of the corpus. Again, *Ya* seems more ready to use a wider range of sources in its lexicon, while remaining stereotypically "néo-breton" in general.

Some trends were observed regarding the use of dictionaries and the attestation of tagged tokens from the corpus in a selection of sources, adding further nuances to the overall picture. The print sources had a higher rate of the use of words found in the two sources that represented mid-twentieth-century "néo-breton" (Hemon's dictionary and *Brezhoneg ... buan hag aes*; see section 3.3.1 for discussion of why these sources were selected), showing their overall alignment with this variety. However, as noted above, only *Bremañ* can be considered to truly represent this more conservative type of "néo-breton", as *Ya* makes a more innovative use of the language, with more international words and semantic expansion, while *Brud Nevez* tends to avoid complicated or controversial vocabulary; indeed, it uses more vocabulary in ways not advised by TermOfis, showing its lesser adherence to the standard.

Out of the three subcorpora, Facebook had the highest proportion of tagged tokens used in ways not recommended in TermOfis. Again, this gives us a slightly different perspective on Facebook's position: it was stated above that Facebook tended to conform to the "néo-breton" stereotype in terms of word origins, favouring those with Celtic roots. With the addition of this information, it seems that this subcorpus is both stereotypically "néo-breton" and non-standard: that it contains a high proportion of users who conform more strongly to typical "néo-breton" linguistic trends, but who are not guided by official sources in their doing so. Again, this emphasises the important role of the community in creating norms, while also highlighting users' creativity. The radio data, on the other hand, included more tagged tokens that were attested in Le Gonidec's dictionary and in TermOfis, suggesting an alignment with standard Breton in general, but not with the overly purist tradition of early "néo-breton" sources. The part of the corpus with the lowest rate of dictionary attestation overall was made up of

news programmes from the associative radio stations, suggesting that this is the locus of the most non-standard language. This is likely to be connected to the greater prevalence of dialectal Breton here (see below), particularly in the case of one particular announcer who used a far larger number of dialectal tokens than any other.

The findings from the interviews are also relevant here, as participants were asked whether their organisations made use of particular dictionaries. In general, participants from radio stations stressed that they availed themselves of a large number of dictionaries, choosing specific ones based on the programme's context and lexical field, and that they often prioritised discussion among colleagues over dictionaries in deciding on vocabulary to use. The magazines, on the other hand, both had a particular dictionary that they relied on, which ensured orthographic consistency for each publication. Thus the comparative non-standardness of *Brud Nevez* is likely to be reflected in the particular dictionary that it uses, while the fact that *Bremañ* uses more standard language shows that its preferred dictionary also conforms with a more standard variety of Breton, following the tradition of Hemon's dictionary and *Brezhoneg ... buan hag aes*. An organisation's choice of a particular dictionary can help perpetuate a certain linguistic and political tradition. In the examples of *Bremañ* and *Brud Nevez*, the dictionaries chosen are appropriate: the An Here dictionary used by *Bremañ* is monolingual and known for its controversial political views, suiting the militant "néo-bretonnant"; the one used by *Brud Nevez*, like the magazine itself, uses an orthography other than the de facto standard, occupying a more marginal position in terms of circulation and proximity to the "néo-breton" community.

The use of dialectal words also varied among the subcorpora. Vannetais and central Breton were more frequent in the radio subcorpus, while Léonais and Cornouaillais appeared more prevalent in the print subcorpus. However, as section 3.4 showed, when the tokens in question are examined, those tagged as Léonais and Cornouaillais in this sample have in fact also become part of standard Breton. Again, this shows that the print subcorpus, in general, uses more standard language. Out of the three print sources, *Ya* was the most dialectal, using significantly larger amounts of Vannetais vocabulary in particular. This is another facet of the more modern form of "néo-breton" that it uses, accepting dialectal vocabulary as valid language: while the interviewee

from *Bremañ* noted that this publication does the same, no particular instances of this were observed in the sample. This suggests that while both publications may include dialectal language, *Ya* perhaps does so more often, and so is maybe seen by would-be contributors as a more welcoming space for their more dialectal varieties, while *Bremañ* is perceived as more conservative, and reflects this perception in its output. The context with the greatest amount of Vannetais is associative radio: again, this was particularly skewed by one announcer, who works for Radio Bro Gwened, but whose programmes are broadcast across all four Radio Breizh stations.

Aspects of tokens' morphological integration also show variation among the different parts of the corpus. Less affixation occurs among tagged tokens in the radio subcorpus, indicating its more informal register, entailing less complicated vocabulary; conversely, among the print sources, *Bremañ* employs the most affixation, as its more formal register would suggest. In parts of the corpus where French-derived elements occur less frequently, those that do occur are less likely to be morphologically integrated, showing that they are less likely to be considered valid parts of the Breton lexicon that obey its morphological rules. Such integration occurs more often in *Ya* than in *Bremañ*, again indicating its more inclusive approach.

Mutation and affixation were examined specifically: it was found that less mutation occurred in feature programmes on the radio, this being the context with the greatest amount of spontaneous speech. This appears to confirm the theory that new speakers find it less easy to avoid French influence in spontaneous speech (Madeg 2011:64; Hewitt 2016:16); but mutation also varies among traditional speakers depending on local dialect (Kennard and Lahiri 2015:5),[147] so it would be misleading to claim this is necessarily symptomatic of a "mistake" specific to new speaker language without further investigation to determine whether such apparent errors occur systematically. In general, words from non-Celtic sources mutated less often, suggesting again that they are not always considered fully integrated parts of the Breton lexicon. This was particularly the case on Facebook, which may suggest that borrowings found in the Facebook subcorpus tend to be less long-established or even one-off borrowings. In contrast, the print subcorpus includes more words of non-Celtic origin that *do* mutate,

147. This is also the case in other Celtic languages: in Irish, variation affects the realisation of nasalisation "both within an dialect and across different dialects" (Frenda 2006:9).

indicating that borrowed words used here may have a longer tradition of use in Breton. It also appears that Breton inflections are more likely among borrowed words in the print subcorpus, while on Facebook, the original inflections are borrowed: again, this suggests that print publications tend to be a locus of more established borrowings, as well as greater conformity with the standard linguistic structure of Breton, and hence to a more standard form of Breton overall.

Variation of multiple kinds can thus be observed in the corpus. Often this relates to register, with the print subcorpus showing characteristics that relate to its use of the highest register of language in the corpus, while the radio subcorpus conversely contains characteristics that attest to its use of the lowest register. However, they also reveal different levels of conformity with stereotypical "néo-breton" and with standard Breton. Within the print corpus in particular, three distinct attitudes are visible among the three publications, each potentially catering to a different audience, and most interestingly, the emergence of a newer, arguably more inclusive form of "néo-breton" that takes words from more sources can be identified in *Ya*. On Facebook too, while some aspects of the data show a general conformity with stereotypical "néo-breton" features, more international and morphologically unintegrated borrowings show the inclusion of vocabulary from multiple languages and linguistic repertoires. Overall, the identifiable differences according to registers and ideologies indicate that in this context that we expect to be largely populated by new speakers, Breton does not come in a single uniformly standardised and non-dialectal form, as some research has suggested (e.g. Jones 1995:428; German 2007:146; Rottet 2014:212): it instead varies according to register, subject matter, medium, audience, and perhaps additional factors. Contexts of this type can provide a rich source of data for the linguistic analysis of contemporary minoritised languages.

## 6.4 Directions of language change

It must nonetheless not be forgotten that this research focused on language used in the media, meaning that any conclusions drawn about the linguistic nature of contemporary Breton are specific to this particular context. In the print and radio subcorpora,

the language used is subject to editorial control and targeted towards wide audiences: this means it will be different in nature from the Breton used within families or among friends. The Facebook subcorpus is less governed by these considerations, but its contents are still a particular type of language that is constructed with the aim of communicating with a large, geographically dispersed community, in a specific Breton-only space. The examples of ludic use of language highlighted on Facebook shows that this is a context where users may be heavily performative, highlighting the fact that again, this is not representative of face-to-face communication. While the type of language used on Facebook could be an indicator of how Breton is developing, it would be naïve to assume that all the characteristics of the Facebook subcorpus are making their way into other domains of Breton use. However, the performative and ludic nature of much of the language found in the Facebook subcorpus does emphasise speakers' high level of self-awareness and use of Breton in a way that constructs their personal identity, a theme that is also apparent in the use of dialectal features, as this section will discuss.

While the language found in these contexts will thus not be representative of Breton as a whole, the media nonetheless have the important role of acting as a vector for new terminology, as interviewees pointed out: the widespread use of a new term in the media will both indicate that it has been accepted as a valid part of the Breton lexicon, and introduce it to other speakers in the community, who may not directly access sources such as TermOfis. However, even media sources are unlikely to reach the whole Breton-speaking population. Even among more motivated audiences, those living in the eastern part of Brittany will find it difficult to access Breton radio programmes, as they must make the specific effort to listen to the radio via the internet in many cases. The print media are similarly affected by low circulation figures, particularly among younger people, as an interviewee pointed out. Future years will no doubt see these figures dwindle still more as speakers come to rely increasingly on content that is accessible online, and the need for new technologies to be harnessed effectively by speakers of minoritised languages is paramount.

Notwithstanding the fact that the language investigated in this research is unique to the media, some trends can be observed that may indicate how Breton is developing in general. Some characteristics that match stereotypical "néo-breton" can indeed be

identified, particularly in terms of more structural, less semantically heavy elements such as morphological integration. Similarly, standard vocabulary is often favoured: interviewees' reports of using doublets on the radio at first mention but then defaulting to the standard term attests to this. However, this strategy is clearly part of an awareness of the fact that audiences have multiple repertoires caused by diverse methods of acquiring and using Breton: when communication occurs in this wide context that transcends local boundaries, speakers continuously make allowances for this with the use of doublets and periphrasis, ensuring the greatest possible reach of their language and highlighting the fact that rather than two discrete communities of new and traditional speakers, there is room for significant overlap. In the same way, interviewees explained how they consider specific words individually, making use of multiple dictionaries, with their final choices guided by the context in which the word will be used and its semantic field. Users of Breton in the media are certainly aware of the stereotypes around the "néo-bretonnant" and keen to avoid them in their lexical choices, even if the dominance of French sometimes makes this difficult at a structural level.

The emergence of a more innovative type of "néo-breton", noted particularly in *Ya*, can also be seen to relate to the presence of diverse speakers with multiple individual repertoires. In this variety of Breton, semantic expansion seems more productive, reducing the role of outright neologism; dialectal features seem more likely to occur, acknowledging the validity of non-standard varieties; and, while borrowings from French remain stigmatised, those of "international" provenance, often learned words or technical terms with cognates in other major world languages, appear more frequently. While this variety still matches some of the characteristics of stereotypical "néo-breton", it equally recognises the diversity of the Breton-speaking community and its place within an increasingly globalised world, where more and more Breton speakers will have some knowledge of other languages, particularly English.

The continued value of dialects has emerged as an especially important theme: most interviewees affirmed that they considered themselves to use a variety with at least some dialectal characteristics, even if they had acquired Breton through education or in adulthood. This again suggests a greater desire for a connection with the traditional speaker community than previous research has suggested, and functions as another

way of avoiding the "néo-bretonnant" stereotype. As interviewees pointed out, the radio in particular is an essential tool for acquiring a more dialectal form of Breton. The role of dialects as an identity marker is evident from how they are used: as section 5.6 discussed at length, the collective reports of interviewees suggested a form of "dialect convergence" whereby among speakers who communicate in this context that transcends geographical limitations, those with a more localised dialect end up speaking a more pan-dialectal variety, while those who initially learn a more standard variety will become more dialectal, allowing all speakers to use a degree of dialect but remain comprehensible to users of other varieties. Section 3.4.3 also showed that prepositions feature overwhelmingly among dialectal tokens, implying that specific indexical words with lower levels of semantic content have become the typical markers of dialectal Breton: again, this functions as a way of marking allegiance to a particular dialect while remaining comprehensible to other speakers. Both the corpus data and the interview data thus point out the importance of dialects—contrary to the typical claims that "néo-breton" is non-dialectal—and how they index identity in a context where communication among speakers from diverse backgrounds must be prioritised. The use of specific words to mark particular dialects mirrors how, in the wider community, Breton people who do not speak Breton will know and make use of specific Breton words, such as *gast* (*putain*) or *kenavo* ('goodbye'). Within the Breton-speaking community, though, the valorisation of dialects hints that the new speaker community is growing and becoming more disparate such that value must be sought in more localised linguistic varieties. It is no longer the case that all new speakers of Breton are acquainted with each other, and so the mere fact of being a Breton speaker is not enough to mark oneself out: one must display a more local identity. Dialects thus re-emerge as an important factor, not only for this reason, but also as a way of avoiding negative stereotypes about older forms of "néo-breton". With this moderate use of dialectal features, speakers balance concerns of identity and communication, as well as demonstrating their linguistic competence in their ability to acquire a variety of Breton that differs from the kind found in elementary learning materials.

## 6.5 Theoretical implications and directions for further research

### 6.5.1 Theory around new and traditional speakers

One of the main aims of this thesis, particularly relevant to research question 2, has been to examine the extent of the applicability of the "néo-bretonnant" category, which has appeared numerous times in existing literature, sometimes with little criticism or interrogation of its validity (Jones 1995:428–429; Timm 2001:454; Timm 2003:34; German 2007:152–153; Adkins 2013:58; Le Nevez 2013:92–93; Hewitt 2016). At this point, it is clear that while some of the stereotypically invoked features of "néo-breton" and the "néo-bretonnant" are present in the corpus, this does not contribute to a situation where all instances of Breton use in media contexts conform with these characteristics. Rather, a multiplicity of practices can be observed, affected by context, medium, register, personal attitudes and ideologies, and individual speakers' trajectories regarding the acquisition and use of Breton. Strikingly, several interviewees did not fit easily into either the new or traditional speaker category, showing that there is more communication and mobility between the two groups than some research has implied, perhaps particularly in the media context, and highlighting the dangers of using this category as a basis for linguistic analysis, as pointed out by O'Rourke and Pujolar (2019).

This thesis thus adds to a growing body of literature that demonstrates that the definition of the new speaker of Breton can be more complicated than would be suggested by its portrayal in much research, which has often relied unproblematically on a binary divide between stereotypical "néo-bretonnants" and traditional speakers. Even when the third category of "native authenticists" (Hornsby and Quentel 2013:78) is acknowledged, this tends to be invoked in opposition to the "néo-bretonnants", suggesting again that the different groups are clearly delineated and act in discrete spheres.

This research does not attempt to claim that the new speaker category is invalid: on the contrary, certain aspects of the language examined in the corpus appear to be a product of the fact that very many speakers of Breton in the contexts examined have acquired it through means other than intergenerational transmission. The category

of new speaker is particularly useful when examining community practices and motivations: the emphasis on identity and self-reflexivity brought about by the dominance of new speakers is an important factor in considering the direction minoritised languages will take in the future. However, this research shows that scholarly work should be careful in defining the new speaker and reflect critically on any characteristics this type of speaker is believed to have, as well as noting that some speakers may straddle the implied binary division.[148] In the specific case of Breton, assumptions about the new speaker that map the category onto the stereotypical "néo-bretonnant" should be thoroughly interrogated, as should similar cases in other languages where conceptual divisions of the same kind tend to be posited: the examples of Manx and Irish were noted in section 1.4.1.

### 6.5.2 Future directions

Resources such as corpora of Breton are modest in number, but do exist. However, the current study is believed to be the first corpus-based research that focuses specifically on a context in which the language used is expected to come mostly from new speakers: indeed, most existing corpora have focused specifically on dialects of traditional Breton, with a view to preserving examples of these endangered varieties, as noted in section 2.2.2. This is also one of the first studies to focus in part on the use of Breton on the internet,[149] a context that has become increasingly important in recent years.

For these reasons, more research in the area using corpus techniques would no doubt be beneficial, perhaps focusing on larger samples of the Breton lexicon. Some apparent trends within the current corpus were restricted to a small number of instances as the corpus was fairly small: the development of tools that would enable the automatic annotation of Breton would be beneficial in this respect, enabling the creation, annotation, and analysis of larger corpora. Part of speech taggers for Breton do exist,

148. Again, O'Rourke and Pujolar (2019:18) make similar points relating to the utility of the new speaker concept as a "lens" that allows research to focus on "the contemporary dynamics of multilingual communities", rather than as a method of typological classification. The "sociolinguistics of the speaker" (ibid.:27) that they advocate shifts the focus to issues surrounding authority, legitimacy and access to capital in a way that demonstrates how anyone perceived as a member of the new speaker category can be excluded from these, regardless of whether they meet the sometimes restrictive criteria for being a new speaker in any typological sense.

149. Baxter (2009) is another example, also focusing on lexicon, but does not take a corpus approach.

but tools geared more towards lexicographical investigation, such as digitised versions of well-known dictionaries, would also be useful. These would help track changes in Breton, and particularly in standard Breton, over time. Such tools and studies would also provide useful perspectives on other languages with significant new speaker populations, much work on which has hitherto concentrated on ideologies and attitudes rather than linguistic data: for this reason, the papers in Smith-Christmas et al. (2018), many of which investigate the linguistic characteristics of new speaker varieties, are particularly welcome.

The investigation of the Facebook subcorpus revealed a number of interesting features, meriting further exploration that was not possible within the scope of this study. It has emerged from this work that further examination of the use of Breton and other minoritised languages on social media would undoubtedly shed light not only on contemporary linguistic practices but also on the way that minority identity is performed and negotiated in an online context, furthering the work of groups such as the Digital Language Diversity Project,[150] which has shown that digital tools are immensely useful to minoritised language communities: further research could show what type of benefits they provide, to which types of speaker, and how they could be more effectively harnessed to support the entire community.

The continued presence of attitudes and practices that do not fit with the way standard Breton is institutionally depicted, or with the stereotypes around the "néo-bretonnant", has been a particular focus in this thesis. More research concentrating specifically on speakers who do not fit easily into one side of the new/traditional binary divide would undoubtedly be an asset to the understanding of the Breton language community, and indeed those of other minoritised languages. Such speakers could be "neo-native" speakers (i.e. L1 speakers of Breton with new speaker parents); educated traditional speakers who rose to positions of authority within the community, such as the participants recruited by Ó hIfearnáin (2013:120); members of the "native authenticist" group; or merely new speakers who do not identify with the "militant" or "néo-bretonnant" labels. This would provide further understanding of the interplay between new and traditional speakers of Breton. It would be similarly beneficial to

150. http://www.dldp.eu, accessed 20 Mar 19.

conduct research focusing on non-standard varieties, in order to investigate how and why they are used today, whether language policy adequately deals with their use, and whether a principle such as *polynomie*, as implemented in Corsica (Jaffe 1999), could be beneficial within the Breton language community. In general, more work on non-standard Breton would help highlight speakers' multiple repertoires and motivations, showing that not all speakers need fit the "néo-bretonnant" model, and that as the Breton new speaker community expands, there is room for speakers of multiple backgrounds and attitudes to engage with Breton in the way most individually suited to them. This can only be an asset to the language, encouraging diversity, adaptability, and vitality.

# Appendix A

# Abbreviations and specialist terms

Abbreviations are typically spelt out and terms glossed at first mention, but this appendix provides a summary to aid reading.

| | |
|---|---|
| ***ALBB*** | *Atlas linguistique de la Basse-Bretagne*, compiled during the 1920s |
| **Basse-Bretagne (Breizh-Izel)** | The western part of Brittany, where Breton has traditionally been spoken into the modern period |
| **brezhoneg21** | Online database of Breton terminology specialising in scientific terms |
| **Devri** | Online historical corpus of Breton containing texts dating back to the ninth century, developed by researchers at Université Rennes 2 |
| **Diwan** | Associative network of schools teaching through the medium of Breton |
| **ECRML** | European Charter for Regional or Minority Languages, drawn up by the Council of Europe in 1992 and unratified by France |
| **Haute-Bretagne (Breizh-Uhel)** | The eastern part of Brittany, where Breton ceased to be spoken during the medieval period |

| | |
|---|---|
| **KLT** | Collective term for the Cornouaillais, Léonais and Trégorrois dialects of Breton, linguistically closer to each other than to Vannetais |
| **Meurgorf** | Online historical corpus of Breton maintained by the **OPLB** |
| ***NALBB*** | *Nouvel atlas linguistique de la Basse-Bretagne*, compiled in the later twentieth century |
| ***OPLB*** | *Office public de la langue bretonne*: body officially designated to oversee Breton language planning, funded by regional and local government |
| ***peurunvan*** | The most widely used supradialectal orthography for Breton |
| **Redadeg** | Sponsored run taking place over a number of days every two years, raising money for **Diwan** and other Breton language organisations |
| ***skolveurieg*** | Supradialectal orthography for Breton with two variants: one for **KLT** and the other for Vannetais |
| **TermBret** | Division of the **OPLB** responsible for corpus planning and maintaining **TermOfis** |
| **TermOfis** | Online database of recommended Breton terminology, focusing particularly on public life and administration |
| **UNESCO** | United Nations Educational, Scientific and Cultural Organisation |

# Appendix B

# Maps

As this thesis is submitted for the degree of PhD in French, French is used throughout the thesis to refer to departments, dialects, and other concepts where French and Breton words exist, but there is no English equivalent. In this appendix, both the French and Breton names are provided.


Figure B.1: Map of the departments of Brittany, including Loire-Atlantique as a fifth department, part of the historical region but not officially considered part of present-day administrative Brittany (see section 5.4)

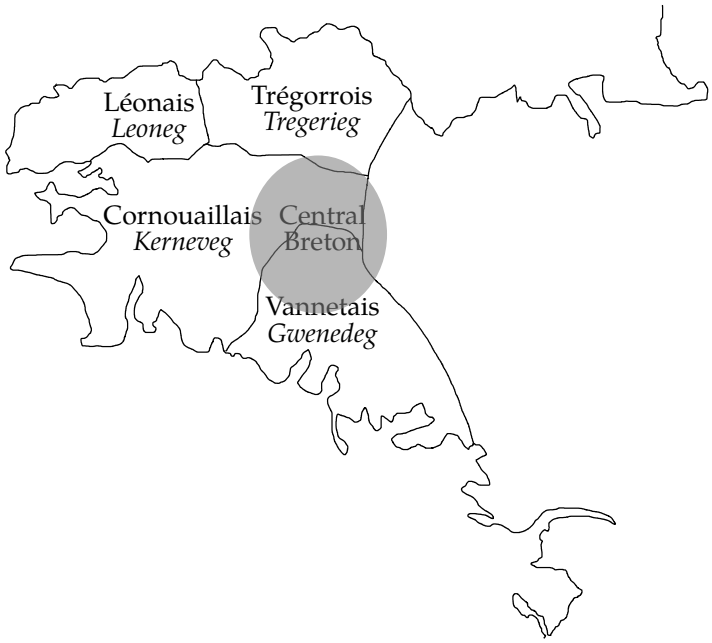


Figure B.2: Map of the four traditionally cited Breton dialects, plus the "central Breton" region (see section 1.3), roughly marked by the grey ellipse covering the eastern Cornouaillais and northern Vannetais areas

# Appendix C

# Accessing the corpus and interview transcriptions

A full version of the corpus was provided to examiners at the time of thesis submission, along with interview transcriptions. The corpus and interview transcriptions are also accessible at https://github.com/merryndd/thesis-public-files. The publicly accessible version of the corpus available via this link is a redacted version that does not contain the data from social media: ethical issues relating to these data are discussed in section 2.2.2.1.

# Appendix D

# Consent form distributed to participants

## Informations sur le projet

Ce projet a pour but d'examiner l'utilisation du breton dans les médias (radio, magazines et journaux, Internet) et comment ceux qui travaillent avec les médias en langue bretonne utilisent la langue dans leur vie professionnelle et personnelle.

Étant donné que vous travaillez pour une organisation qui utilise le breton dans ce contexte professionnel, vous êtes invité(e) à participer à un entretien d'une durée d'environ 30 minutes qui aura lieu en langue française.

Pendant l'entretien, ce que vous dites sera enregistré en format audio. Ensuite, cet enregistrement sera transcrit et les données seront conservées sans identification du participant. Ces données ne seront pas utilisées pour des fins commerciales, mais des extraits anonymes apparaîtront dans une thèse doctorale et dans des éventuelles publications qui seront accessibles dans le domaine public.

L'entretien sera conduit par Merryn Davies-Deacon, chercheuse doctorante à Queen's University, Belfast, Irlande du Nord. Vous êtes invité(e) à lui poser des questions avant ou après l'entretien si vous en avez. Si vous avez des questions ou des commentaires

après l'entretien, vous pouvez la joindre par mail, [redacted]@qub.ac.uk, ou par téléphone, 0044 7[redacted].

## Attestation de consentement

Je confirme :

1. que j'ai reçu et compris les informations sur le but de la recherche, que j'ai eu l'occasion de poser des questions à son sujet, et que j'ai eu des réponses à mes éventuelles questions ;
2. que ma participation est volontaire et que j'ai le droit de me retracter de l'étude à tout moment sans avoir à donner d'explication ;
3. que les données que je fournis seront conservées d'une façon sécure sans identification personnelle ;
4. que tout ce que je dis pendant l'entretien sera enregistré et transcrit, y inclus les informations personnelles (âge, sexe, etc.), et que tout ceci pourra être cité dans une thèse doctorale, et dans des publications éventuelles et communications orales, de façon totalement anonyme et sans identification de la personne qui a fourni ces données ;
5. qu'aucune partie des données fournies ne sera utilisée pour des fins commerciales.

J'accepte de participer à cette étude.

Nom et prénom du participant : _____

Signature du participant : _____ Date : _____

Signature de la chercheuse : _____ Date : _____

# Appendix E

# Identity scales

| | | |
|---|---|---|
| Pas du tout breton.ne | _____ | Entièrement breton.ne |
| Pas du tout français.e | _____ | Entièrement français.e |

# Appendix F

# Interview question sheet

## Questions démographiques

(Ces questions sont requises par le projet de recherche dans le cadre duquel je travaille.)

1. Âge
2. Genre (homme, femme, autre)
3. Où est-ce que vous habitez et depuis quand ?
4. Est-ce vous avez déjà habité ailleurs ?
   (a) Où, quand et pendant combien de temps ?
5. A quel point vous sentez-vous breton· ne ? A quel point vous sentez-vous français· e ?
6. Est-ce que vous utilisez des langues à part le breton et le français ?

## Histoire personnelle

1. Est-ce qu'il y avait des membres de votre famille qui parlaient breton ?

2. Comment avez-vous acquis le breton ? (dans la maison, à l'école, apprentissage adulte)

3. Quand utilisez-vous le breton aujourd'hui ? Où, avec qui ?

4. Est-ce que vous participez à des activités culturelles de la région ? (musique, danse, associations culturelles ou politiques ...)

   (a) Et est-ce que vous y parlez breton ?

5. Est-ce vous modifiez le style de votre breton (quand vous parlez avec ...) ?

## Utilisation des médias et d'Internet

1. Est-ce que vous lisez les journaux et les magazines/regardez la télé/écoutez la radio ?

2. Est-ce que vous utilisez Internet ?

   (a) De quelle façon ?

3. Et est-ce que vous faîtes tout ça en breton ?

4. Est-ce que vous avez contribué à ou travaillé pour d'autres médias en langue bretonne ?

## L'organisation, l'individu

1. Qu'est-ce vous faîtes pour l'organisation, et depuis quand ?

2. Comment les thèmes pour les émissions/les articles sont-ils décidés ?

3. Qui produit les bulletins/les articles ?

4. Ce sont qui les auditeurs/les lecteurs ? Est-que vous pouvez décrire un· e auditeur· trice/un· e lecteur· trice « typique » ?

5. Est-ce que vous croyez qu'il y ait des problèmes de compréhension pour certains auditeurs/certains lecteurs (ceux qui parlent des dialectes différents, les débutants, ceux qui ne connaissent pas le vocabulaire spécialiste, etc.) ? Comment vous faîtes pour éviter à ces problèmes et assurer que tout le monde puisse comprendre ?

6. Est-ce qu'il y a des fois quand vous ne connaissez pas un mot ? Est-ce que vous utilisez certains dictionnaires ou d'autres ressources ?

# Appendix G

# Chi-square tests

Chi-square tests measure correlation between categorical (i.e. non-numeric) variables by measuring the difference between observed and expected values. As the tags assigned to the tokens in the corpus represent categorial variables, it is appropriate to use this test to show correlation.

Chi-square tests generate a "test statistic" which can be compared to a critical value specified by a table. If the test statistic is higher than the critical value, there is a significant difference between observed and expected values, indicating that there is a significant difference across categories.

The chi-square formula is as follows:

$$\chi_c^2 = \sum \frac{(O_i - E_i)^2}{E_i}$$

where $\chi_c^2$ is the chi-square test statistic, $O_i$ is the observed value at position $i$ in the table of data and $E_i$ is the expected value at the same position. The test statistic is therefore found by finding the difference between the observed and expected values at each position in the table, squaring this, dividing by the expected value, and adding together the values obtained at all positions.

As all the tests dealt with either two or three categories (e.g. the two radio programme

types or the three subcorpora), a worked example of both of these types is presented below.

## G.1 Chi-square tests with two categories

This example tests the distribution of tokens in the radio data by station type, focusing on language of origin. In terms of raw data, there were 278 Breton-derived tokens in the public radio data and 281 in the associative radio data. We can therefore begin to fill in a table of the required values for the chi-square test.

| | $O_i$ | $E_i$ | $\lvert O_i - E_i \rvert$ | $(O_i - E_i)^2$ | $\frac{(O_i - E_i)^2}{E_i}$ |
|---|---|---|---|---|---|
| **Public** | 278 | | | | |
| **Associative** | 281 | | | | |

Table G.1: Table for two-category chi-square test, partially filled

Calculating the expected value depends on the size of the part of the corpus in total. If both had exactly the same word count, the expected value for both would be the same, i.e. half the total number of tokens, i.e. $\frac{278+281}{2} = 279.5$. However, the sizes of the two parts of the corpus are slightly different. There are 5183 words in the public radio sample and 5660 words in the associative radio sample. The expected values must therefore be proportionally equivalent to the sample sizes. These are calculated using the following formula:

$$\frac{S_i}{\sum S} \times \sum O$$

where $O$ is again the observed value and $S$ is the total sample size, $S_i$ being the sample size for the part of the corpus in question.

For the public radio data, this gives:

$$\frac{5183}{5183 + 5660} \times 559 = 267.20$$

and for the associative radio data:

$$\frac{5660}{5183 + 5660} \times 559 = 291.80$$

With these figures, the rest of the table can be filled in.

| | $O_i$ | $E_i$ | $\lvert O_i - E_i \rvert$ | $(O_i - E_i)^2$ | $\frac{(O_i - E_i)^2}{E_i}$ |
|---|---|---|---|---|---|
| **Public** | 278 | 267.20 | 10.80 | 116.64 | 0.44 |
| **Associative** | 281 | 291.80 | 10.80[151] | 116.64 | 0.40 |

Table G.2: Table for two-category chi-square test, wholly filled

The figures in the final column are added together to give a test statistic of 0.84. This must then be compared with the critical value. At one degree of freedom,[152] the critical value is 3.84 at a significance level of 0.05. As 0.84 is smaller than 3.84, the difference between the two is insignificant.

## G.2 Chi-square tests with three categories

This example tests the distribution of French-derived tokens across the three subcorpora. In this case, there were 559 observed in the radio subcorpus, 442 in the Facebook subcorpus, and 469 in the print subcorpus. We can thus begin to fill in the table.

151. As there are only two variables in this test, the two values in this column will always be the same.
152. The number of degrees of freedom is the number of variables minus 1.

| | $O_i$ | $E_i$ | $\lvert O_i - E_i \rvert$ | $(O_i - E_i)^2$ | $\frac{(O_i - E_i)^2}{E_i}$ |
|---|---|---|---|---|---|
| **Public** | 559 | | | | |
| **Associative** | 442 | | | | |
| **Print** | 469 | | | | |

Table G.3: Table for three-category chi-square test, partially filled

Given that the total sample sizes were 10843, 11566, and 11936 respectively, we can calculate the expected values as above.

| | $O_i$ | $E_i$ | $\lvert O_i - E_i \rvert$ | $(O_i - E_i)^2$ | $\frac{(O_i - E_i)^2}{E_i}$ |
|---|---|---|---|---|---|
| **Public** | 559 | 464.09 | 94.91 | 9007.91 | 19.41 |
| **Associative** | 442 | 495.04 | 53.04 | 2813.24 | 5.68 |
| **Print** | 469 | 510.88 | 41.88 | 1753.93 | 3.43 |

Table G.4: Table for three-category chi-square test, wholly filled

Adding together the figures in the final column, we find the test statistic, 28.52. At two degrees of freedom and a significance level of 0.05, the critical value is 5.991. This time, the test statistic is greater than the critical value, and so in this case there is a significant difference.

# References


Abalain, H. (2000). *Histoire de la langue bretonne*. Paris: Gisserot.

Adkins, M. (2013). "Will the real Breton please stand up? Language revitalisation and the problem of authentic language." *International Journal of the Sociology of Language* 223, 55–70.

Adkins, M. (2014). "An American abroad: My adult language immersion studies in Brittany." *Journal of Celtic Language Learning* 19, 109–128.

Ager, D. E. (1999). *Identity, insecurity and image: France and language*. Clevedon: Multilingual Matters.

Arksey, H. and Knight, P. (1999). *Interviewing for social scientists*. London: SAGE.

Baker, P. (2010). *Sociolinguistics and corpus linguistics*. Edinburgh: Edinburgh University Press.

Baxter, R. N. (2009). "New technologies and terminological pressure in lesser-used languages: The Breton Wikipedia, from terminology consumer to potential terminology provider." *Language Problems and Language Planning* 33 (1), 60–80.

Beacco, J.-C. and Cherkaoui Messin, K. (2010). "Les politiques linguistiques européennes et la gestion de la diversité des langues de France." *Langue française* 167, 95–111.

Beeching, K. (2007). "Social identity, salience, and language change: The case of post-rhematic *quoi*." In *The French language and questions of identity*. Ed. by W. Ayres-Bennett and M. C. Jones. London: Legenda, pp. 140–149.

Beeching, K. (2009). "Sociolinguistic factors and the pragmaticalisation of *bon* in contemporary spoken French." In *Sociolinguistic variation in contemporary French*. Ed. by K. Beeching, N. Armstrong, and F. Gadet. Amsterdam/Philadelphia: John Benjamins, pp. 215–239.

Berruto, G. (2018). "The languages and dialects of Italy." In *Manual of Romance sociolinguistics*. Ed. by W. Ayres-Bennett and J. Carruthers. Berlin: de Gruyter, pp. 494–525.

Bert, M. (2011). "Situation sociolinguistique du francoprovençal : L'étude FORA." *Langues et cité* 18, 5.

Blackledge, A. and Creese, A. (2017). "Translanguaging and the body." *International Journal of Multilingualism* 14 (3), 250–268.

Blackwood, R. J. (2004). "Compulsory Corsican language classes in school as a method for reversing the language shift to French?" *Transactions of the Philological Society* 102, 307–333.

Blackwood, R. J. (2011). "The linguistic landscape of Brittany and Corsica: A comparative study of the presence of France's regional languages in the public space." *Journal of French Language Studies* 21, 111–130.

Blackwood, R. J. (2014). "The top-down revitalisation of Corsican: Considering the reversal of a language shift in the linguistic and semiotic landscapes of Ajaccio." *French Studies* 68 (1), 61–77.

Blanchet, P. (2012). "L'identification des langues : une question clé pour une politique scientifique et linguistique efficiente." *Modèles linguistiques* 66, 17–25.

Blommaert, J. and Rampton, B. (2011). "Language and superdiversity." *Diversities* 13.

Broudic, F. (2007). "La puissante ténacité de l'obstacle de la langue bretonne." In *L'éecole française et les langues régionales XIXe–XXe siècles*. Ed. by H. Lieutard and M.-J. Verny. Montpellier: Presses universitaires de la Méditerranée, pp. 177–195.

Broudic, F. (2009). *Parler breton au XXIe siècle : Le nouveau sondage de TMO-Régions*. Brest: Emgleo Breizh.

Broudic, F. (2011). "Quels médias pour la langue bretonne ?" In *Les médias de la diversité culturelle dans les pays latins d'Europe*. Ed. by A. Lenoble-Bart and M. Mathien. Brussels: Bruylant, pp. 85–97.

Calvez, R. (2012). "Ce que parler du breton veut dire." *Ethnologie française* 42 (4), 647–655.

Carruthers, J. and Fisher, L. (in preparation). "Conducting interdisciplinary research in modern languages: Theory and practice in establishing 'common ground'."

Cenoz, J. and Gorter, D. (2017). "Minority languages and sustainable translanguaging: Threat or opportunity?" *Journal of Multilingual and Multicultural Development* 38 (10), 901–912.

Chalm, E. (2009). *La grammaire bretonne pour tous*. 2nd ed. Lannion: An Alarc'h Embannadurioù.

Cheshire, J. (2002). "Sex and gender in variationist research." In *The handbook of language variation and change*. Ed. by J. K. Chambers, P. Trudgill, and N. Schilling-Estes. Oxford: Blackwell, pp. 423–443.

Colonna, R. (2015). "La standardisation à l'épreuve de la minorisation : Le cas du corse." *Sociolinguistica* 29, 39–59.

Connell, R. W. (1987). *Gender and power: Society, the person and sexual politics*. Cambridge: Polity Press.

Conrad, S. (2015). "Register variation." In *The Cambridge handbook of English corpus linguistics*. Ed. by D. Biber and R. Reppen. Cambridge: Cambridge University Press, pp. 292–308.

Constantine, M.-A. (1996). *Breton ballads*. Aberystwyth: CMCS Publications.

Cornwall Council (21 Apr. 2016). *Cornwall Council condemns government decision to cut funding for Cornish language*. Press release.

Costa, J. (2014). "Must we save the language? Children's discourse on language and community in Provençal and Scottish language revitalisation movements." In *Endangered languages: Beliefs and ideologies in language documentation and revitalisation*. Ed. by J. Sallabank and P. K. Austin. Oxford: Oxford University Press, pp. 195–214.

Costa, J. (2017). *Revitalising language in Provence: A critical approach*. Transactions of the Philological Society 114 (51). Oxford: Blackwell.

[CoE] Council of Europe (1992). "European charter for regional or minority languages." *European Treaty Series* (148). URL: https://rm.coe.int/CoERMPublicCommonSearchServices/DisplayDCTMContent?documentId=0900001680695175 (Accessed 19 July 2018).

Creese, A. and Blackledge, A. (2015). "Translanguaging and identity in educational settings." *Annual Review of Applied Linguistics* 35, 20–35.

Cucarull, J. (2002). *Histoire économique et sociale de la Bretagne*. Paris: Gisserot.

Cunliffe, D., Morris, D., and Prys, C. (2013). "Young bilinguals' language behaviour in social networking sites: The use of Welsh on Facebook." *Journal of Computer-Mediated Communication* 18, 339–361.

Devlin, K. (27 Nov. 2015). "Gaeldom warns of threat to television channel as funding of £1m withdrawn." *The Herald* (*Glasgow*), 4.

Dowling, E. A. M. (2014). "Anglo-American influence on the contemporary French lexis: The role of domain, text type and register." PhD thesis. Queen's University, Belfast.

Eckert, P. (2012). "Three waves of variation study: The emergence of meaning in the study of sociolinguistic variation." *Annual Review of Anthropology* 41, 87–100.

Estival, D. and Pennycook, A. (2011). "L'Académie française and Anglophone language ideologies." *Language Policy* 10, 325–341.

Ferrara, K., Brunner, H., and Whittemore, G. (1991). "Interactive written discourse as an emergent register." *Written Communication* 8 (1), 8–34.

Fife, J. (2010). "Typological aspects of the Celtic languages." In *The Celtic languages*. Ed. by M. J. Ball and N. Müller. 2nd ed. London: Routledge, pp. 3–21.

Fishman, J. A. (1991). *Reversing language shift: Theoretical and empirical foundations of assistance to threatened languages*. Clevedon: Multilingual Matters.

Frenda, A. S. (2006). "Gender and mutation in Irish: A preliminary account for further investigation." *The ITB Journal* 7 (2), 4–42.

Fridrichová, R. (2013). "Quelques observations sur les mots tronqués dans le français contemporain." *Romanica Olomucensia* 25 (1), 1–13.

García, O. and Leiva, C. (2014). "Theorising and enacting translanguaging for social justice." In *Heteroglossia as practice and pedagogy*. Ed. by A. Blackledge and A. Creese. Dordrecht: Springer, pp. 199–216.

Gardin, B. (1975). "Loi Deixonne et langues régionales : représentation de la nature et de la fonction de leur enseignement." *Langue française* 25, 29–36.

German, G. (2007). "Language shift, diglossia and dialectal variation in western Brittany: The case of southern Cornouaille." In *The Celtic languages in contact: Papers from the workshop within the framework of the XIII International Congress of Celtic Studies*. Ed. by H. L. C. Tristram. Bonn: Universitätsverlag Potsdam, pp. 146–192.

Gillham, B. (2005). *Research interviewing: The range of techniques*. Maidenhead: Open University Press.

Gouélian, H. J. (2002). *Vocabulaire du breton d'aujourd'hui*. Paris: Gisserot.

Grinevald, C. and Bert, M. (2011). "Speakers and communities." In *The Cambridge handbook of endangered languages*. Ed. by P. K. Austin and J. Sallabank. Cambridge: Cambridge University Press, pp. 45–65.

Gros, J. (1970). *Le trésor du breton parlé : Éléments de stylistique trégorroise*. Vol. 1: *Le langage figuré*. 2nd ed. 3 vols. Saint-Brieuc: Les presses bretonnes.

Gros, J. (1966–1989). *Le trésor du breton parlé : Éléments de stylistique trégorroise*. 3 vols. Brest, Lannion, and Morlaix: Emgleo Breiz, Barr-Heol, and Skol Vreizh.

Guiomar, J.-Y. (1992). "La Révolution française et les origines celtiques de la France." *Annales historiques de la Révolution française* 287, 63–85.

Guivarch-Tonnard, A. (21 Nov. 2014). "Facebook se met au breton : « plijus ! »" *Ouest-France*. URL: http://www.ouest-france.fr/bretagne/facebook-se-met-au-breton-plijus-2997311 (Accessed 24 Jan. 2017).

Hemon, R. (1984). *Grammaire bretonne*. Brest: Al Liamm.

Hewitt, S. (1977). "The degree of acceptability of modern literary Breton to native Breton speakers." Diploma of Linguistics thesis. Cambridge University.

Hewitt, S. (2014). "Breton orthographies: An increasingly awkward fit." Paper presented at *Fourth Cambridge Conference on Language Endangerment*, University of Cambridge, 4–5 July. URL: http://www.academia.edu/7537879 (Accessed 19 July 2018).

Hewitt, S. (2016). "The problem of neo-speakers in language revitalization: The example of Breton." Paper presented at *International Conference on Endangered Languages*, Tbilisi State University, 20–24 October. URL: https://www.academia.edu/29610759 (Accessed 9 Nov. 2016).

Hewson, C. (2003). "Conducting research on the internet." *The Psychologist* 16 (6), 290–293.

Hinskens, F., Auer, P., and Kerswill, P. (2005). "The study of dialect convergence and divergence: Conceptual and methodological considerations." In *Dialect change: Convergence and divergence in European languages*. Ed. by P. Auer, F. Hinskens, and P. Kerswill. Cambridge: Cambridge University Press, pp. 1–48.

Hornsby, D. (2009a). "Dedialectalisation in France: Convergence and divergence." *International Journal of the Sociology of Language* 196–197, 157–180.

Hornsby, M. (2005). "*Néo-breton* and questions of authenticity." *Estudios de sociolingüística* 6 (2), 191–218.

Hornsby, M. (2009b). "Globalisation processes and minority languages: Linguistic hybridity in Brittany." PhD thesis. University of Southampton.

Hornsby, M. (2010). "The ideology of monolingualism as 'standard' in Brittany." In *From vestiges to the very day: New voices in Celtic Studies*. Ed. by M. Watson and L. Milligan. Aberdeen: AHRC Centre for Irish and Scottish Studies, pp. 67–75.

Hornsby, M. (2015). "The 'new' and 'traditional' speaker dichotomy: Bridging the gap." *International Journal of the Sociology of Language* 231, 107–125.

Hornsby, M. (2017). "Finding an ideological niche for new speakers in a minoritised language community." *Language, Culture and Curriculum* 30 (1), 91–104.

Hornsby, M. and Quentel, G. (2013). "Contested varieties and competing authenticities: Neologisms in revitalised Breton." *International Journal of the Sociology of Language* 223, 71–86.

[IEO] Institut d'études occitanes (2015). *Occitan ... qu'es aquò ?* URL: http://www.ieo-oc.org/IMG/pdf/quesaoiFrVpdf-2.pdf (Accessed 28 Aug. 2019).

Jackson, K. H. (1967). *A historical phonology of Breton*. Dublin: Institute for Advanced Studies.

Jaffe, A. (1999). *Ideologies in action: Language politics on Corsica*. Berlin: Mouton de Gruyter.

Jamieson, S. (2004). "Likert scales: How to (ab)use them." *Medical Education* 38, 1217–1218.

Jones, M. C. (1995). "At what price language maintenance? Standardisation in modern Breton." *French Studies* 49 (4), 424–438.

Jones, M. C. (1998a). "Death of a language, birth of an identity: Brittany and the Bretons." *Language Problems and Language Planning* 22 (2), 129–142.

Jones, M. C. (1998b). *Language obsolescence and revitalisation: Linguistic change in two sociolinguistically contrasting Welsh communities*. Oxford: Clarendon Press.

Jung, A. and Urvoas, J.-J. (2012). *Langues et cultures régionales : en finir avec l'exception française*. Paris: Fondation Jean Jaurès.

Kallweit, D. (2018). "New media: New Romance varieties?" In *Manual of Romance sociolinguistics*. Ed. by W. Ayres-Bennett and J. Carruthers. Berlin: de Gruyter, pp. 386–404.

Kennard, H. J. (2014). "The persistence of verb second in negative utterances in Breton." *Journal of Historical Linguistics* 4 (1), 1–39.

Kennard, H. J. and Lahiri, A. (2015). "Mutation in Breton verbs: Pertinacity across generations." *Journal of Linguistics*, 1–33.

Kennedy, H. (2006). "Beyond anonymity, or future directions for internet identity research." *New Media and Society* 8 (6), 859–876.

Kergoat, L. (2012). "Terminologie bretonne : Rétrospective, débats et enjeux." *Recerc* 5.

King, N. and Horrocks, C. (2010). *Interviews in qualitative research*. London: SAGE Publications.

Lagadeuc, J. [1499] (1867). *Le Catholicon*. Ed. by R. F. Le Men. Lorient: Éditions Corfmat.

Laroussi, F. and Marcellesi, J.-B. (1993). "The other languages of France: Towards a multilingual policy." In *French today: Language in its social context*. Ed. by C. Sanders. London: Routledge, pp. 85–104.

Le Coadic, R. (2013). "A propos des relations entre langue et identité en Bretagne." *International Journal of the Sociology of Language* 223, 23–41.

Le Dû, J. (2001). *Nouvel atlas linguistique de la Basse-Bretagne*. 2 vols. Brest: Centre de recherche bretonne et celtique.

Le Dû, J. (2009). "Can a language be 'saved'? The case of Breton." In *Kodukeela ja keele kodu/Home language and the home of a language*. Ed. by M. Kallsmaa and V. Ijan. Tallinn: Eesti Keele Sihtasutus, pp. 77–89.

Le Dû, J. (2014). *Enquêtes dialectologiques en vue de constituer le Nouvel atlas linguistique de la Basse-Bretagne*. URL: http://purl.org/poi/crdo.vjf.cnrs.fr/crdo-COLLECTION_NALBB.

Le Dû, J. and Le Berre, Y. (2013). "La langue bretonne dans la société régionale contemporaine." *International Journal of the Sociology of Language* 223, 43–54.

Le Gonidec, J.-F. (1850). *Dictionnaire breton-français*. Ed. by T. Hersart de la Villemarqué. Saint-Brieuc: Éditions L. Prudhomme.

Le Guyader, A. (2002). "Les sources de l'histoire du Réseau breton." *Revue d'histoire des chemins de fer* 24–25, 328–341.

Le Nevez, A. (2013). "The social practice of Breton: An epistemological challenge." *International Journal of the Sociology of Language* 223, 87–102.

Le Nigen, V., Gouerou, C., Chartier-Le Floch, E., and Nono (2014). *Un modèle politique breton ?* Spézet: Coop Breizh.

Le Pipec, E. (2013). "Les trois ruptures sociolinguistiques du breton." *International Journal of the Sociology of Language* 223, 103–116.

*Le Robert* (1989). *Le Robert*. Vol. 2: *Bip–Cout*. 2nd ed. 9 vols. Paris.

Le Roux, P. [1924] (1977). *Atlas linguistique de la Basse-Bretagne*. 2nd ed. 6 vols. Brest: Éditions armoricaines.

Leach, D. (2012). "*Biographie-Memoires*, by Célestin Lainé (Neven Henaff)." *e-Keltoi* 4, 39–180.

Lerat, P. (1988). "Les internationalismes dans les langues romanes." *Annexes des Cahiers de linguistique hispanique médiévale* 7, 483–491.

Lodge, R. A. (1993). *French: From dialect to standard*. London: Routledge.

Madeg, M. (2011). *J'élève mon enfant en breton*. Brest: Emgleo Breiz.

Manchec-German, G. (2018). "Which linguistic model for Breton?" *Studia Celto-Slavica* 9, 1–30.

Markham, A. and Buchanan, E. (2012). *Ethical decision-making and internet research: Recommendations from the AoIR Ethics Working Committee*. Report. Association of Internet Researchers. URL: http://aoir.org/reports/ethics2.pdf (Accessed 27 Mar. 2017).

Martin, J.-B. (2011). "Le franco-provençal." *Langues et cité* 18, 2–3.

McDonald, M. (1989). *"We are not French!": Language, culture and identity in Brittany*. London: Routledge.

McLeod, W. and O'Rourke, B. (2015). "'New speakers' of Gaelic: Perceptions of linguistic authenticity and appropriateness." *Applied Linguistics Review* 6 (2), 151–172.

McMahon, A. (1994). *Understanding language change*. Cambridge: Cambridge University Press.

Milin, R. (2018). "From a wooden clog to a monkey skull." Paper presented at *Third Poznań Conference of Celtic Studies*, Adam Mickiewicz University, Poznań, 9–10 July.

Mills, J. (2015). "Etymology and the relexification of Cornish in the twentieth century." Paper presented at *Etymological thinking in the nineteenth and twentieth centur-*

*ies*, Taylor Institution, Oxford, 7 November. URL: https://youtube.com/watch?v=pOiwUqQHeWE (Accessed 2 Feb. 2016).

Milroy, J. (2007). "The ideology of the standard language." In *The Routledge companion to sociolinguistics*. Ed. by C. Llamas, L. Mullany, and P. Stockwell. Abingdon: Routledge, pp. 133–139.

Milroy, J. and Milroy, L. (1997). "Varieties and variation." In *The handbook of sociolinguistics*. Ed. by F. Coulmas. Oxford: Blackwell Publishing, pp. 47–64.

Milroy, L. and Gordon, M. (2003). *Sociolinguistics: Method and interpretation*. Oxford: Blackwell.

Moal, S. (2004). "TV Breizh, two years on." *Journal of Celtic Language Learning* 8, 24–34.

Morvannou, F. (2004). "Regards sur 700 ans de breton écrit." In *Corona Monastica. Moines bretons de Landévennec : histoire et mémoire celtiques. Mélanges offerts au père Marc Simon*. Ed. by L. Lemoine and B. Merdrignac. Rennes: Presses universitaires de Rennes, pp. 365–379.

Nance, C., McLeod, W., O'Rourke, B., and Dunmore, S. (2016). "Identity, accent aim, and motivation in second language users: New Scottish Gaelic speakers' use of phonetic variation." *Journal of Sociolinguistics* 20 (2), 164–191.

Nicolas, M. (2007). *Histoire de la revendication bretonne*. Spézet: Coop Breizh.

Ó Dochartaigh, C. (1992). "The Irish language." In *The Celtic languages*. Ed. by D. MacAulay. Cambridge: Cambridge University Press, pp. 11–99.

Ó Giollagáin, C. (2004). "A contrastive view of Irish language dynamics." *Collegium antropologicum* 28 (Supplement 1), 73–81.

Ó hIfearnáin, T. (2013). "Institutional Breton language policy after language shift." *International Journal of the Sociology of Language* 223, 117–135.

Ó hIfearnáin, T. (2015). "Back to the future: Standard and language standards in contemporary Manx Gaelic." *Sociolinguistica* 29, 99–120.

O'Rourke, B. and Pujolar, J., eds. (2019). *From new speaker to speaker: Outcomes, reflections and policy recommendations from COST Action IS1306 on New Speakers in a Multilingual Europe: Opportunities and Challenges*. Newcastle Emlyn: IAITH: Welsh Centre for Language Planning.

O'Rourke, B. and Walsh, J. (2015). "New speakers of Irish: Shifting boundaries across time and space." *International Journal of the Sociology of Language* 231, 63–83.

[OPLB] Office public de la langue bretonne (2014). *Enquête d'initiative locale sur la pratique du breton à Carhaix : Résultats*. Report. Office public de la langue bretonne. URL: http://www.fr.brezhoneg.bzh/include/viewFile.php?idtf=1868&path=2d%5C%2F1868_937_Enklask-Karaez-2012.pdf (Accessed 5 Feb. 2016).

[OPLB] Office public de la langue bretonne (2016). *Geriaoueg kod an hentoù / Vocabulaire du code de la route*. URL: http://www.fr.brezhoneg.bzh/uploads/Document/38/2619_214_Geriaoueg_Kod_an_hentou.pdf (Accessed 19 July 2018).

Osterkorn, P. K. and Vetter, E. (2015). "« Le multilinguisme en question ? » – The case of minority language education in Brittany (France)." In *The multilingual challenge: Cross-disciplinary perspectives*. Ed. by U. Jessner-Schmid and C. J. Kramsch. Trends in Applied Linguistics 16. Berlin: De Gruyter.

Otheguy, R., García, O., and Reid, W. (2015). "Clarifying translanguaging and deconstructing named languages: A perspective from linguistics." *Applied Linguistics Review* 6 (3), 281–307.

Payton, P. (2000). "Cornish." In *Languages in Britain and Ireland*. Ed. by G. Price. Oxford: Blackwell, pp. 109–119.

Pennycook, A. (2012). *Language and mobility: Unexpected places*. Clevedon: Multilingual Matters.

Perrot, M.-C. (1997). "La politique linguistique pendant la révolution française." *Mots* 52, 158–167.

Pihlaja, S. (2017). "More than fifty shades of grey: Copyright on social network sites." *Applied Linguistics Review* 8 (2–3), 213–228.

Pujolar, J. and Puigdevall, M. (2015). "Linguistic *mudes*: How to become a new speaker in Catalonia." *International Journal of the Sociology of Language* 231, 167–187.

Région Bretagne (2012). *Une politique linguistique pour la Bretagne : Rapport d'actualisation mars 2012*. Report. Région Bretagne. URL: http://www.bretagne.bzh/upload/docs/application/pdf/2012-04/rapport_dactualisation_de_la_politique_linguisique_2012.pdf (Accessed 26 Nov. 2015).

Région Rhône-Alpes (28 May 2015). *Rhône-Alpes et la Vallée d'Aoste signent la charte de coopération pour le francoprovençal*. Press release.

Rottet, K. J. (2014). "Neology, competing authenticities, and the lexicography of regional languages: The case of Breton." *Journal of the Dictionary Society of North America* 35, 208–247.

Roulston, K. (2013). "Analysing interviews." In *The SAGE handbook of qualitative data analysis*. Ed. by U. Flick. London: SAGE Publications, pp. 297–312.

Saugera, V. (2012). "How English-origin nouns (do not) pluralise in French." *Linguisticæ investigationes* 35 (1), 120–142.

Sayers, D. (2012). "Standardising Cornish: The politics of a new minority language." *Language Problems and Language Planning* 36 (2), 99–119.

Siegel, J. (1985). "Koines and koineisation." *Language in Society* 14 (3), 357–378.

Smith-Christmas, C., Ó Murchadha, N. P., Hornsby, M., and Moriarty, M., eds. (2018). *New speakers of minority languages: Linguistic ideologies and practices*. London: Palgrave Macmillan.

Tagliamonte, S. A. (2012). *Variationist sociolinguistics: Change, observation, interpretation*. Oxford: Wiley-Blackwell.

Ternes, E. (1992). "The Breton language." In *The Celtic languages*. Ed. by D. MacAulay. Cambridge: Cambridge University Press, pp. 371–452.

Thomson, R. L. (2000). "Manx." In *Languages in Britain and Ireland*. Ed. by G. Price. Oxford: Blackwell, pp. 58–69.

Timm, L. A. (1989). "Word order in 20th-century Breton." *Natural Language and Linguistic Theory* 7 (3), 361–378.

Timm, L. A. (2001). "Transforming Breton: A case study in multiply conflicting language ideologies." *Texas Linguistic Forum* 44 (2): *Proceedings from the Eighth Annual Symposium about Language and Society, Austin, April 20–22, 2001*, 447–461.

Timm, L. A. (2003). "Breton at a crossroads: Looking back, moving forward." *e-Keltoi* 2, 25–61.

Timm, L. A. (2005). "Language ideologies in Brittany, with implications for Breton language maintenance and pedagogy." *Journal of Celtic Language Learning* 10, 34–42.

Timm, L. A. (2010). "Language, culture and identity in Brittany." In *The Celtic languages*. Ed. by M. J. Ball and N. Müller. 2nd ed. London: Routledge, pp. 712–752.

Tobin, G. A. and Begley, C. M. (2004). "Methodological rigour within a qualitative framework." *Journal of Advanced Nursing* 48 (4), 388–396.

Tracy, S. J. (2012). *Qualitative research methods: Collecting evidence, crafting analysis, communicating impact*. Oxford: Wiley-Blackwell.

Trépos, P. (1960). "Francis Gourvil. *Théodore-Claude-Henri Hersart de la Villemarqué (1815-1895) et le Barzaz-Breiz (1839-1845-1867). Origines. Editions. Sources. Critiques. Influence.* I–VI, 609 p. Rennes, Imprimeries Oberthur, 1960." *Annales de Bretagne* 67 (4), 389–397.

Trudgill, P. (1972). "Sex, covert prestige and linguistic change in the urban British English of Norwich." *Language in Society* 1 (2), 179–195.

Valera-Hernández, S. (2014). "Dictionary- and corpus-based research in applied and descriptive linguistics." In *Linguistic insights: Diachrony and synchrony in English corpus linguistics*. Ed. by A. Alcarez-Sintes and S. Valera-Hernández. Oxford: Peter Lang, pp. 161–170.

Vetter, E. (2013). "Teaching languages for a multilingual Europe—minority schools as examples of best practice? The Breton experience of *Diwan*." *International Journal of the Sociology of Language* 223, 153–170.

Vigers, D. (2013). "Signs of absence: Language and memory in the linguistic landscape of Brittany." *International Journal of the Sociology of Language* 223, 171–187.

Vurpas, A.-M. (1993). "Peut-on observer l'émergence de koinès dialectales en francoprovençal de France depuis le XVIe siècle à nos jours ?" In *Écriture, langues communes et normes : Formation spontanée de koinès et standardisation dans la Galloromania et son voisinage*. Ed. by P. Knecht and Z. Marzys. Geneva: Droz, pp. 171–184.

Walsh, J. and Ní Dhúda, L. (2015). "'New speakers' of Irish in the United States: Practices and motivations." *Applied Linguistics Review* 6 (2), 173–193.

Walther, J. B. (2002). "Research ethics in internet-enabled research: Human subjects issues and methodological myopia." *Ethics and Information Technology* 4, 205–216.

Williams, N. J. A. (1997). *Clappya Kernowek: An introduction to Unified Cornish Revised*. Portreath: Agan Tavas.

Wilson, W. H. (1998). "The sociopolitical context of establishing Hawaiian-medium education." *Language, Culture and Curriculum* 11 (3), 325–338.

Wise, H. (1997). *The vocabulary of modern French: Origins, structure, and function*. London: Routledge.

Wmffre, I. (1998). *Central Breton*. Languages of the World: Materials 152. Munich: LINCOM Europa.

Wmffre, I. (2008). *Breton orthographies and dialects: The twentieth-century orthography war in Brittany*. 2 vols. Oxford: Peter Lang.

Woehrling, J.-M. (2013). "Histoire du droit des langues en France." In *Histoire sociale des langues de France*. Ed. by G. Kremnitz. Rennes: Presses universitaires de Rennes, pp. 71–88.

Woolard, K. A. (1992). "Language ideology: Issues and approaches." *Pragmatics* 2 (3), 235–249.

Woolard, K. A. (2016). *Singular and plural: Ideologies of linguistic authority in 21st century Catalonia*. Oxford: Oxford University Press.

Wright, S. (2007). "Language policy and language planning." In *The Routledge companion to sociolinguistics*. Ed. by C. Llamas, L. Mulling, and P. Stockwell. London: Routledge, pp. 164–171.

Zimmer, M. (2010). "'But the data is already public': On the ethics of research on Facebook." *Ethics and Information Technology*.