

# Investigating the “new speaker” stereotype: Lexical variation in the Breton media

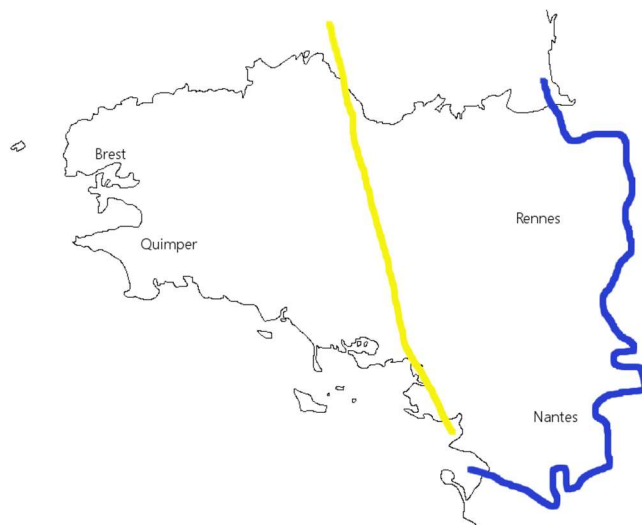
Merryn Davies-Deacon

Part of MEITS Strand 3 panel at the Modern Languages Core Disciplinary Research Group Showcase, Queen’s University, Belfast, 11 May 2018.

## 1

The previous paper presented research on French in urban context. For this paper, we’re staying in France, but turning our attention to a regional minority language. My PhD project focuses on the lexicon of Breton and what it can tell us about “new speakers” and the sort of language they use.

## 2 Breton



Those of you who attended this event last year may remember this map from my part of the MEITS presentation. The historical territory of Brittany is west of the blue line on the map. West of the yellow line is the area where Breton is traditionally spoken. Major cities are marked. There are now around 200000 speakers of Breton (Broudic, 2013:10) following a rapid decline throughout the twentieth century. Academic

literature tends to divide them into *new speakers* and *traditional speakers*. The basic definition of this distinction can be summed up as follows: traditional speakers have acquired Breton as a result of an uninterrupted chain of intergenerational transmission, while new speakers have acquired it by other means.

### 3 *New versus traditional speakers*

In addition to the basic definition, various other criteria have been cited in the literature to distinguish new from traditional speakers; they are listed in the following table. Those criteria above the double line relate to the speakers themselves, while those below it relate to their language.

New speakers	Traditional speakers
Younger, middle-class, mobile, and well-educated	Older, working-class, usually farmers
Located in larger towns and cities	Located in rural hamlets
Live all over Brittany	Restricted to western Brittany
Acquire Breton as a second language	Acquire Breton through intergenerational transmission
Acquire Breton in formal education	
Militant promotion of Breton	Speak Breton only because of linguistic ability
Literate in Breton	Unable to read and write Breton
Proud of Breton identity	No strong sense of Breton identity
Use “standardised” Breton	Use dialectal Breton
French-influenced linguistic structure	Linguistic structure not influenced by French
Celtic-derived lexicon	French-influenced lexicon

(See e.g. Jones, 1995; Hornsby, 2005; Timm, 2001; German, 2007)

For the purposes of my research, certain criteria are particularly salient: new speakers are assumed to be younger, more mobile, and literate in Breton, and to use Breton in as many contexts as possible. These factors combine to indicate that it is the language of new speakers we would expect to find in the media, which is the context under examination.

### 4 *The lexicon of Breton*

My research focuses on the lexicon, picking up from the claim that new speakers use more Celtic-derived vocabulary in order to “purify” the language and avoid French influence. For example, new speakers may use *baleadenn* (‘walk’) instead of French-derived *pourmenadenn* (< *promenade*). They may use *abeg* (‘reason’) instead of French-derived *rezen* (< *raison*). They may use *baraerezh* (‘bakery’) rather than French-derived *boulangerezh* (< *boulangerie*) (Timm, 2001:456).

## 5 Research aims

Given that we expect to find the media to be the preserve of new speakers, we therefore expect the lexicon found in the media to display these characteristics. To investigate this, I'm looking at the lexicon in three different media contexts: magazines, radio broadcasts, and Facebook group posts. This range will help indicate whether there is any variation as a result of medium and register within this domain.

## 6 Divisions of the corpus

I collected roughly equal amounts of Breton from the three sources, two of which were split into different parts. The radio data can be split in two ways: either by station—public (state-financed) or associative (independent)—or by programme type. Half the programmes I collected were news bulletins and the other half were longer “feature” programmes. The magazine sources came from three different publications, *Bremañ*, *Brud Nevez*, and *Ya!*. *Bremañ* and *Brud Nevez* publish once every two months, while *Ya!* is published every week. *Bremañ* and *Ya!* use the most commonly used Breton orthography, and *Brud Nevez* uses a different one. All three have slightly different focuses in terms of their subject matter. These differences may be reflected in the type of lexicon they use.

The Facebook data are all taken from the same public group, *Facebook e brezhoneg*. The sample contains varying amounts of text written by 99 different users.

## 7

I compiled these data (having transcribed the radio sample orthographically) into a corpus, which was tagged according to the Text Encoding Initiative. Words salient to the investigation were tagged in a way that encoded various attributes. The choice to tag a word was based on a combination of factors that I assembled into a workflow.

If you're familiar with HTML, the structure of the resulting corpus will make sense to you. Here is a sample of what it looks like, this being one paragraph from the magazine subcorpus:

```
<p>Petra e oa <w lemma="batailhonoù ar yaouankiz" ana="#brand #brt #noun">Batailhonoù ar  
Yaouankiz</w> ? Aozet e vezent dreist-holl gant <w lemma="rezistant" ana="#ga #orthchange #noun #new  
#termo #inflbr #frfrac">rezistanted</w> yaouank-flamm. Koshañ <w lemma="ezel" ana="#stb #brt #semexp  
#old #noun #slegon #mhy #b21 #termo #hemon64">ezel</w> ar strollad a oa 26 vloaz hepken. <w  
lemma="tro-dro" ana="#brt #brbrc #adv #old #mhy #termo #antitermo #sb21">Tro-dro</w> da 20 vloaz e  
oa ar re all. A-raok ar brezel o doa kalz anezhe kemeret perzh ur wech bennak en <w lemma="aozadur"  
ana="#stb #brt #brbraf #noun #early #semexp #termo #b21 #inflbr #mhy #shemon">aozadurioù</w> <w  
lemma="politikel" ana="#ga #frbraf #adj #early #mhy #termo #orthchange #hemon64">politikel</w> tost
```

d'an <w lemma="ideologiezh" ana="#ga #inter #new #noun #frbraf #orthsame #termo">ideologiezh</w> <w lemma="komunour" ana="#adj #ga #frbraf #early #orthchange #mhy #termo">komunour</w> met ne oant ket <w lemma="ezel" ana="#stb #brt #semexp #old #noun #slegon #mhy #b21 #termo #hemon64">ezel</w> eus ar <w lemma="strollad komunour gall" ana="#noun #brand #brt #expl">Strollad komunour gall</w> (<w lemma="pcf" ana="#brand #ga #init #noun">PCF</w>) evit kement-se war a lâre Gilbert Brustlein, unan anezhe. Pa darzhas ar brezel e voe kalz anezhe o nac'h ar <w lemma="faezhidigezh" ana="#noun #old #brt #brbraf #termo #sb21 #hemon64">faezhidigezh</w> hag a gemeras perzh er <w lemma="rezistañs" ana="#tb #ga #noun #new #orthchange #termo #sb21 #frfraf">Rezistañs</w> o <w lemma="kregiñ" ana="#tb #brt #verb #old #mhy #scath #termo #b21 #hemon64">kregiñ</w>. D'ar choulz-se e oa <w lemma="kammed" ana="#tb #noun #brt #old #brbraf #inflbr #mhy #cath #legon #termo #sb21 #hemon64">kammedoù</w> kentañ ar <w lemma="rezistañs" ana="#tb #ga #noun #new #orthchange #termo #sb21 #frfraf">Rezistañs</w> hag a oa dreist-holl un afer <w lemma="propaganda" ana="#engl #noun #early #orthsame #termo">propaganda</w>, <w lemma="titouriñ" ana="#brt #early #verb #mhy #brbraf">titouriñ</w>, tec'hout pe <w lemma="tarverezh" ana="#noun #brt #brbraf #old #termo">tarverezh</w> <w lemma="hiniennel" ana="#adj #old #brt #brbraf #termo #b21 #hemon64">hiniennel</w>. Ne oa na <w lemma="peuraozañ" ana="#verb #old #brt #brbraf #inflbr">peuraozet</w> nag unanet c'hoazh. D'an 22 a viz Mezheven 1941, gant <w lemma="aloubadeg" ana="#noun #brt #brbraf #old #mhy #termo #b21 #shemon">aloubadeg</w> an <w lemma="urss" ana="#kerk #ga #noun #init">URSS</w> a voe divizet gant pennoù bras ar <w lemma="pcf" ana="#brand #ga #init #noun">PCF</w> mont e-barzh ar <w lemma="jeu" ana="#tb #ga #noun #old #orthsame #mhy #antitermo #sb21 #hemon64">jeu</w> en un doare ofisiel. Diwar neuze e c'houlennas an <w lemma="urss" ana="#kerk #ga #noun #init">URSS</w> digant ar <w lemma="komunour" ana="#noun #ga #frbraf #early #orthchange #mhy #termo #mut #inflbr">gomunourien</w> a <w lemma="pep lec'h" ana="#brt #mwebt #adv #old #mut">bep lec'h</w> sailhañ war an <w lemma="alaman" ana="#noun #inflbr #early #ga #orthchange #mhy #termo #hemon64">Alamaned</w>. Ar <w lemma="strollad komunour" ana="#brt #brand #noun">Strollad komunour</w> kuzh a zivizas tuta e-touez e stourmerien yaouank, kentoc'h gant ar re o doa prenet skiant er <w lemma="propaganda" ana="#engl #noun #early #orthsame #termo">propaganda</w> kuzh hag ar re a veve dindan guzh. Evel ar Roazhonad Louis Coquillet <w lemma="da skouer" ana="#brt #old #mwebt #adv #mhy #hemon64">da skouer</w>, a oa klask warnañ hag a oa tec'het kuit diouzh polised <w lemma="roazhon" ana="#noun #kerk #kero #brt">Roazhon</w> o stankañ anezhe en e ranndi ! Ouzhpenn-se, un nebeud anezhe o doa stourmet a-enep arme Franco e <w lemma="spagn" ana="#noun #kerk #kero #brt">Spagn</w> er <w lemma="brigadennoù etrebroadel" ana="#noun #brand #brt">Brigadennoù etrebroadel</w>. Ul lodenn all anezhe a oa <w lemma="estrañjour" ana="#tb #noun #ga #orthchange #old #inflbr #frfraf #mhy #antitermo #hemon64">estrañjourien</w> deuet eus broioù aloubet gant an <w lemma="nazi" ana="#noun #new #ga #inter #orthsame #termo #inflbr #hemon64">nazied</w>.</p>

## 8 Potential characteristics of new/trad lexicon

Once the words were tagged their distribution was examined according to various criteria. So far I've been focusing on large-scale quantitative analysis. There's no easy way of classing a word as "new" or "traditional", so I've been considering various criteria to see whether patterns can be identified among these through the different parts of the corpus: fortunately, they can!

As the language of new speakers is characterised as purist, and is said to tend to keep borrowed elements separate and unintegrated, the following potential criteria for "new" versus "traditional"

lexicon were based on this principle. For a few, there are intermediate positions that may signal a more nuanced attitude.

New?	In the middle?	Traditional?
Words with Celtic roots (usually Breton, sometimes Welsh)	Words used in multiple world languages	Words borrowed from French
New coinages	Semantic expansion of existing vocabulary	Borrowings
Borrowed elements kept apart in compounding and affixation	Borrowed elements morphologically combine with native elements	
Borrowed elements keep original inflection	Borrowed elements inflect and mutate as Breton words	
Borrowed elements keep original spelling and pronunciation	Borrowed elements are phonetically and/or orthographically adapted	
Found more commonly in new dictionaries	Found more commonly in pre-1900 dictionaries/not in new dictionaries	
Found in official terminology databases	Not considered part of official vocabulary	
Non-dialectal	Dialectal	

## 9 Findings: Radio

Having analysed the data at a broad level, according to some of the criteria listed above, some preliminary findings can be presented.

The radio subcorpus is the least conformant to the new speaker stereotype. This fits with its status as a spoken medium containing some low-register conversational data: as traditional speakers are said to use only these lower oral registers, this is the type of media that they are most likely to have access to.

While the speakers in the sample are new speakers, they may be aware of this potentially large traditional speaker audience and adapt their language accordingly.

While the radio subcorpus was split according to station and to programme type, few systematic differences among these categories can be found. There is greater presence of dialectal words on the associative stations than on the public station, but otherwise, variation is mostly minimal: that observed can probably be attributed to differences of subject matter, as this can have a large effect on the lexicon.

It can however be observed that employees of the radio stations tend to use a more stereotypically “new speaker” variety than guests: this may be the result of editorial influence, or, more likely, because more of their speech is scripted, meaning that it is further from the lowest conversational registers.

### ***10 Findings: Facebook***

This subcorpus is in an intermediate position in terms of register: it is neither speech nor formal writing. This is reflected by an overall moderate conformance to the stereotypical new speaker lexicon. However, both extremes can be identified within this: some very purist vocabulary and spelling (more so than elsewhere in the corpus), and some very dialectal posts. I’m currently examining the data on a smaller scale in order to highlight some of these features.

The Facebook sample also contains a lot of metalinguistic discourse, as well as the ludic use of language and translanguaging, which provides an interesting insight into how multilingual individuals perform their linguistic repertoires on the internet.

### ***11 Findings: Print***

Overall, this subcorpus is the most conformant to stereotypical new speaker language, befitting its higher register. However, *Brud Nevez* is noticeably different, and can be seen to be more aligned with traditional Breton in certain ways. While *Bremañ* and *Ya!* are both unequivocal loci of new speaker language, they take slightly different stances: *Bremañ* is in a more “classical” position, using coinages and keeping borrowings morphologically distinct to mark them as non-Breton, while *Ya!* uses more international borrowings (i.e. from languages such as English) and is more tolerant of dialectal vocabulary. This difference goes against the stereotype of a single, monolithic new speaker variety that is unable to express linguistic nuances.

### ***12 Next steps***

Having completed this broad-scale analysis, the next step for me is to continue the analysis in greater depth, examining a selection of individual words and noting how they contribute to the trends shown in the different media under examination. After that, I plan to interview a number of Breton speakers who produce and consume these media, finding out the rationale behind their linguistic decisions and whether their backgrounds and practices fit the model of the new speaker. Finally, I’ll be attempting to tie all these things together into a nice PhD thesis!

### 13 References

- Broudic, F. (2013). Langue bretonne : Un siècle de mutations. *International Journal of the Sociology of Language* 223, 7–21.
- German, G. (2007). Language shift, diglossia and dialectal variation in western Brittany: The case of southern Cornouaille. In H. L. C. Tristram (Ed.), *The Celtic languages in contact: Papers from the workshop within the framework of the XIII International Congress of Celtic Studies*. Bonn: Universitätsverlag Potsdam, pp. 146–192.
- Hornsby, M. (2005). *Néo-breton* and questions of authenticity. *Estudios de sociolingüística* 6, 191–218.
- Jones, M. C. (1995). At what price language maintenance? Standardisation in modern Breton. *French Studies* 49(4), 424–438.
- Timm, L. A. (2001). Transforming Breton: A case study in multiply conflicting language ideologies. *Texas Linguistic Forum* 44(2), 447–461.

**MULTILINGUALISM:**  
Empowering Individuals  
Transforming Societies

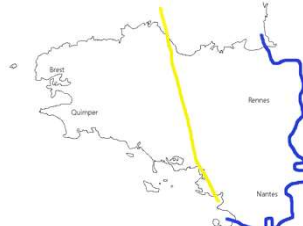
MULTILINGUALISM: EMPOWERING INDIVIDUALS,  
TRANSFORMING SOCIETIES

Investigating the “new speaker” stereotype: Lexical  
variation in the Breton media  
Merryn Davies-Deacon

Arts & Humanities Research Council UNIVERSITY OF CAMBRIDGE THE UNIVERSITY OF EDINBURGH The University of Nottingham QUEEN'S UNIVERSITY BELFAST

**MULTILINGUALISM:**  
Empowering Individuals  
Transforming Societies

Breton



- Around 200000 speakers (Broudic, 2013:10)
- Academic literature divides them into *new speakers* and *traditional speakers*

Arts & Humanities Research Council UNIVERSITY OF CAMBRIDGE THE UNIVERSITY OF EDINBURGH The University of Nottingham QUEEN'S UNIVERSITY BELFAST

**MULTILINGUALISM:**  
Empowering Individuals  
Transforming Societies

New versus traditional speakers

New speakers	Traditional speakers
Younger, middle-class, mobile, and well-educated	Older, working-class, usually farmers
Located in larger towns and cities	Located in rural hamlets
Live all over Brittany	Restricted to western Brittany
Acquire Breton as a second language	Acquire Breton through intergenerational transmission
Acquire Breton in formal education	Restrict use of Breton to intimate settings
Use Breton as much as possible	Speak Breton only because of linguistic ability
Militant promotion of Breton	Inability to read and write Breton
Literate in Breton	No strong sense of Breton identity
Proud of Breton identity	Use dialectal Breton
Use “standardised” Breton	Linguistic structure not influenced by French
French-influenced linguistic structure	Celtic-derived lexicon
Celtic-derived lexicon	French-influenced lexicon

(See e.g. Jones, 1995; Hornsby, 2005; Timm, 2001; German, 2007)

Arts & Humanities Research Council UNIVERSITY OF CAMBRIDGE THE UNIVERSITY OF EDINBURGH The University of Nottingham QUEEN'S UNIVERSITY BELFAST

**MULTILINGUALISM:**  
Empowering Individuals  
Transforming Societies

The lexicon of Breton

- Identified as prominent feature of new versus traditional speaker distinction

French	TB	NB
<i>promenade</i> ‘walk’	pourmenadenn	baleadenn
<i>raison</i> ‘reason’	rezen	abeg
<i>boulangerie</i> ‘bakery’	boulangerezh	baraerezh

adapted from Timm, 2001:456

Arts & Humanities Research Council UNIVERSITY OF CAMBRIDGE THE UNIVERSITY OF EDINBURGH The University of Nottingham QUEEN'S UNIVERSITY BELFAST

**MULTILINGUALISM:**  
Empowering Individuals  
Transforming Societies

Research aims

- Language associated with new speakers
  - In the media
- Three domains of media language
  - Magazines
  - Radio broadcasts
  - Facebook group posts
- Focusing on lexicon

Arts & Humanities Research Council UNIVERSITY OF CAMBRIDGE THE UNIVERSITY OF EDINBURGH The University of Nottingham QUEEN'S UNIVERSITY BELFAST


**MULTILINGUALISM:**  
Empowering Individuals  
Transforming Societies

Divisions of the corpus

Radio		Magazines	Facebook
Public (state) station	Associative (independent) stations	Bremañ	Facebook e brezhoneg (99 “speakers” in sample)
		Brud Nevez	
News bulletins	Longer programmes	Ya!	

Arts & Humanities Research Council UNIVERSITY OF CAMBRIDGE THE UNIVERSITY OF EDINBURGH The University of Nottingham QUEEN'S UNIVERSITY BELFAST



Potential characteristics of new/trad lexicon			 <b>MULTILINGUALISM</b> Empowering professionals Transforming Societies
New?	In the middle?	Traditional?	
Words with Celtic roots (usually Breton, sometimes Welsh)	Words used in multiple world languages	Words borrowed from French	
New coinages	Semantic expansion of existing vocabulary	Borrowings	
Borrowed elements kept apart in compounding and affixation	Borrowed elements morphologically combine with native elements		
Borrowed elements keep original inflection	Borrowed elements inflect and mutate as Breton words		
Borrowed elements keep original spelling and pronunciation	Borrowed elements are phonetically and/or orthographically adapted		
Found more commonly in new dictionaries	Found more commonly in pre-1900 dictionaries/not in new dictionaries		
Found in official terminology databases	Not considered part of official vocabulary		
Non-dialectal		Dialectal	

