

---

# Investigating language spoken in the UK - what data can I use?

Dr Alita Nandi, Essex (Understanding Society)

Dr Oliver Duke-Williams, UCL (UK Data Service)

Dr Jemima Stockton, UCL (CeLSIUS)

25<sup>th</sup> March 2020

# ESRC Data and Methods Services: supporting research and providing training



The ESRC funds a range of data and methods services to support research and study. It provides resources needed to access high quality socio-economic data and to develop research skills. These services are:



**Administrative  
Data Research  
Network**  
[www.adrn.ac.uk](http://www.adrn.ac.uk)



**Census & Administrative  
Data Longitudinal  
Studies Hub**  
[www.calls.ac.uk](http://www.calls.ac.uk)



**Centre for  
Longitudinal  
Studies**  
[www.cls.ucl.ac.uk](http://www.cls.ucl.ac.uk)



**Consumer Data  
Research  
Centre**  
[www.cdrc.ac.uk](http://www.cdrc.ac.uk)



**CLOSER**  
[www.closer.ac.uk](http://www.closer.ac.uk)



**ESRC Business and  
Local Government  
Data Research  
Centre**  
[www.blgdataresearch.org](http://www.blgdataresearch.org)



**National Centre  
of Research  
Methods**  
[www.ncrm.ac.uk](http://www.ncrm.ac.uk)



**UK Data Service**  
[www.ukdataservice.ac.uk](http://www.ukdataservice.ac.uk)



**Understanding  
Society**  
[www.understandingsociety.ac.uk](http://www.understandingsociety.ac.uk)



**Urban Big Data  
Centre**  
[www.ubdc.ac.uk](http://www.ubdc.ac.uk)

# Forthcoming Events

- You may also be interested in future webinars in this series
  - Mental health: (9 December 2019)
  - Religion: (18 March 2020)
  - Language spoken: (25 March 2020)
  - Political behaviour (1 April 2020)
- See: <https://www.ukdataservice.ac.uk/news-and-events/events> for details

# The webinar today

Alita

- Understanding Society

Oliver

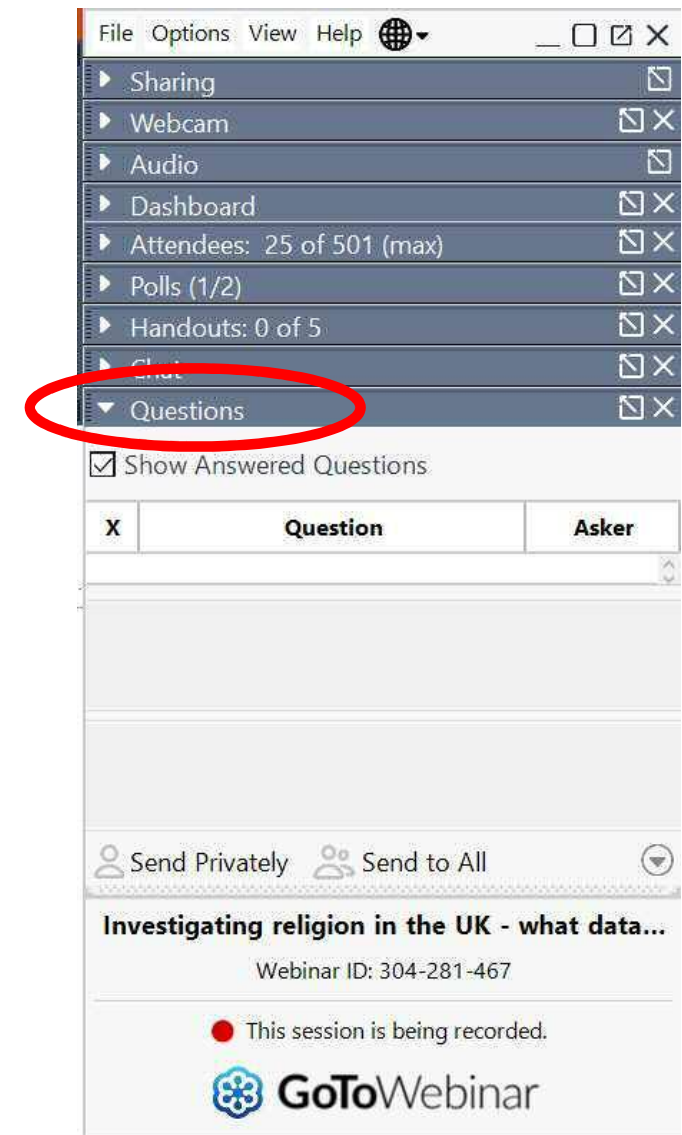
- Cross-sectional census data

Jemima

- Longitudinal census data

Questions

- Type your questions into the questions box - we will answer them at the end



# Languages in the census

Dr Oliver Duke-Williams, UCL

[o.duke-williams@ucl.ac.uk](mailto:o.duke-williams@ucl.ac.uk)

@oliver\_dw

Webinar: March 25<sup>th</sup> 2020

## Background

- GCHQ: "There are 42 languages spoken across our workforce"
  - Is this higher or lower than we might expect?



## A related query

- Where is the most linguistically diverse place in the UK?
  - What do we mean by this?
  - What data can be used to answer this?
  - How easy is it to find an answer?

## Questions on language used

- The questions asked in different parts of the UK in 2011 were not the same
- Questions
  - Longstanding questions on traditional languages
    - Use of Welsh / Scottish Gaelic / Irish / Ulster-Scots
  - From 2011, questions on use of other languages
    - Main language used / other language used
    - For non-English speakers: proficiency in English





# Getting data on languages used

- InFuse



# Getting data: England and Wales, Northern Ireland

File Edit View History Bookmarks Tools Help

InFuse 2011 England, Wales, Scot

infuse2011.ukdataservice

Showing 6 topic combinations for topic(s) (Language, main language used (detailed) [E][N][W])

Page 1 of 1

Filters

- Language, main language used [E][N][W]
- Language, main language used (detailed) [E][N][W]
- Language, proficiency in English
- Language, English as a household language [E][N][W]
- Language, skills in Irish [N]
- Language, knowledge of Irish [N]
- Language, skills in Ulster-Scots [N]
- Language, knowledge of Ulster-Scots [N]
- Language, main language spoken in Northern Ireland [N]
- Language, skills in English and/or Scots [S]

Selected category combinations

- Age
- Language, main language used (detailed) [E][N][W]
- Age
- Qualification, highest level of
- Language, main language used (detailed) [E][N][W]
- Age
- Language, main language used (detailed) [E][N][W]
- Language, proficiency in English
- Age
- Population (daytime/workday) [E][N][W]
- Economic activity
- Language, main language used (detailed) [E][N][W]
- Age
- Non-UK born short-term residents [E][W]
- Language, main language used (detailed) [E][N][W]
- Age
- Language, main language used (detailed) [E][N][W]
- Population (out of term-time) [E][W]

File Edit View History Bookmarks Tools Help

InFuse 2011 England, Wales, Scot

infuse2011.ukdataservice

Selected category combinations

- Age : Age 3 and over - Main language (detailed) : Total: Main language (detailed) - Unit : Persons
- Age : Age 3 and over - Main language (detailed) : English (English or Welsh if in Wales) - Unit : Persons
- Age : Age 3 and over - Main language (detailed) : Welsh/Cymraeg (in England only) - Unit : Persons
- Age : Age 3 and over - Main language (detailed) : Other UK language - Unit : Persons
- Age : Age 3 and over - Main language (detailed) : Irish Gaelic - Unit : Persons
- Age : Age 3 and over - Main language (detailed) : Scottish Gaelic - Unit : Persons
- Age : Age 3 and over - Main language (detailed) : Manx Gaelic - Unit : Persons
- Age : Age 3 and over - Main language (detailed) : Gaelic (not otherwise specified) - Unit : Persons
- Age : Age 3 and over - Main language (detailed) : Cornish - Unit : Persons

You have selected 137 category combinations

File Edit View History Bookmarks Tools Help

InFuse 2011 England, Wales, Scot

infuse2011.ukdataservice

Selected Geographic Areas

- All Wards and Electoral Divisions (8546 areas) in Countries and Groupings (7 areas) - England and Wales
- All Merging Wards and Electoral Divisions (43 areas) in Countries and Groupings (7 areas) - England and Wales
- All Wards and Electoral Divisions (582 areas) in Countries and Groupings (7 areas) - Northern Ireland

# Getting data: Scotland

The screenshot shows a web browser window with the URL `infuse2011.ukdataservice`. The page title is "InFuse 2011 England, Wales, Scotland". The main heading is "Topics". Below the heading, there is a "Show Guidance" button and a note: "Some topics are only available for certain countries; E - England, N - Northern Ireland, S - Scotland, W - Wales".

On the left side, there is a "Filters" section with a "Clear filters" button. The filters include:

- Language, skills in Scots [S]
- Language in Scotland [S]
- Language other than English used at home [S]
- Language, skills in Welsh [W]
- Language, Welsh profile [W]
- Language, ability to speak Welsh [W]
- Language, number of people who can speak Welsh [W]
- Language, ability of parents to

The main content area shows "3 topic combination(s) found for topic(s) (Language other than English used at home [S])". It indicates "Page 1 of 1" and "First < 1 > Last". There are three result cards, each with a "Select" button:

- Card 1:
  - Age
  - Language other than English used at home [S]
- Card 2:
  - Age
  - Language other than English used at home [S]
  - Long-term condition, type of [N][S]
- Card 3:
  - Age
  - Language other than English used at home [S]
  - Sex

## Numbers of different languages tabulated

Country	Languages (inc. English and 'Other')
England and Wales	92
Scotland	4
Northern Ireland	13

- Additional tables in Scotland
  - [http://www.scotlandscensus.gov.uk/documents/additional\\_tables/AT\\_002\\_2011.xls](http://www.scotlandscensus.gov.uk/documents/additional_tables/AT_002_2011.xls) - Scotland, all languages
  - [http://www.scotlandscensus.gov.uk/documents/additional\\_tables/CT\\_0267\\_2011.xlsx](http://www.scotlandscensus.gov.uk/documents/additional_tables/CT_0267_2011.xlsx) - Scotland, by datazone, 17 languages

## Top ten languages used at home, Scotland 2011

Language category (by descending number of people)	Scotland
All people aged 3 and over	5,118,223
English only	4,740,547
Scots	55,817
Polish	54,186
Urdu	23,394
Punjabi (Not otherwise specified)	23,150
Chinese (Not otherwise specified)	16,830
French	14,623
Gaelic (Not otherwise specified)	14,531
British Sign Language	12,533
German	11,317

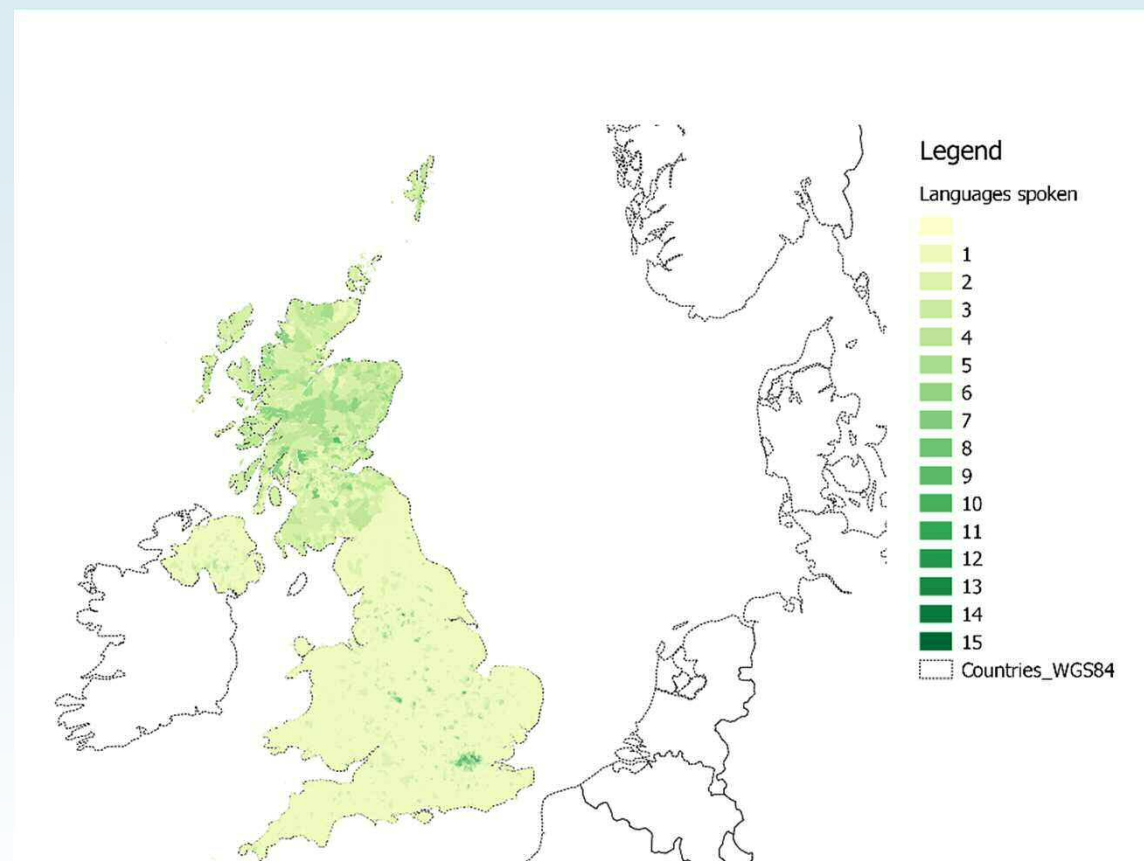
Source: 2011 Census, NRS, Table AT\_002\_2011

## Joining these together

- Data need to be assembled from several sources to attempt to answer the question linguistic diversity
  - Separate tables in InFuse for different parts of the UK
  - Data for England and Wales and Northern Ireland at ward level
  - Data for Scotland replaced with more detailed NRS data at datazone level
  - Data mapped using boundaries from <https://borders.ukdataservice.ac.uk/>
- **Mapped: the number of languages for which speakers are at least 1% of the ward / datazone population**

## Number of languages spoken by at least 1% of ward population

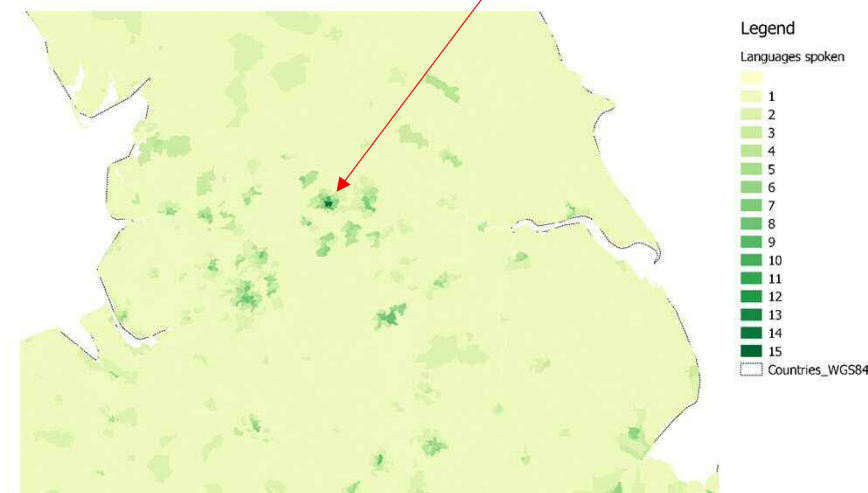
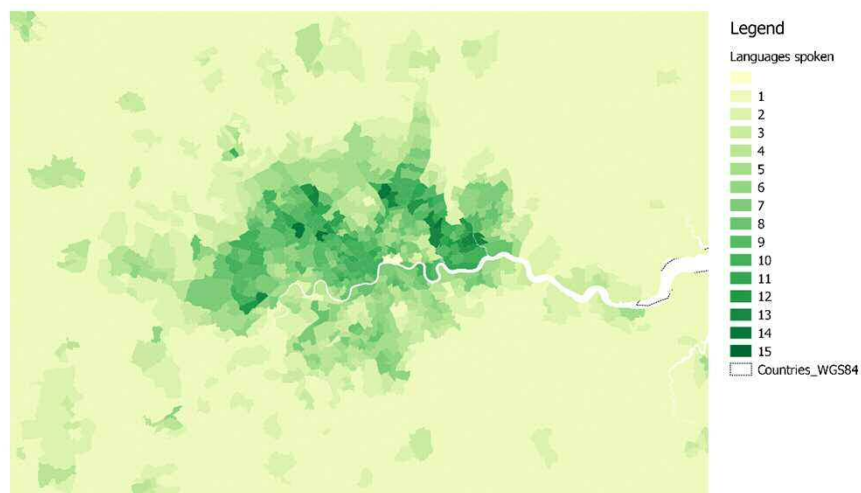
- Key thing to note
  - Different question asked in Scotland – results not necessarily comparable
  - Northern Ireland data: same question as E&W, but range of result categories smaller



Source: 2011 Census, safeguarded microdata



City ward, Bradford



Source: 2011 Census, safeguarded microdata

## Census microdata

- Microdata files contain almost all fields from the census forms except personal identifiers such as name and address
  - Some values are broad coded etc
- Secure files have stronger access restrictions but offer greater detail
- Safeguarded files are easy for academic researchers to use, have restricted detail in some cases

# Census microdata files

		Open	Safeguarded	Secure
1961 / 1971 / 1981	Individuals	1% (GB)	5% (GB)	9% (GB)
	Households		0.95% (GB)	
1991	Individuals		2% (GB,NI)	
	Households		1% (GB,NI)	
2001	Individuals		3% (UK)	
	Households			1% (EW)
	'Small Area Microdata'		5% (UK)	
2011	Individual		5% (EW,S,NI)	10% (EW,S,NI)
	Households			10% (EW,S,NI)
	Individual ('Teaching')	1% (EW,S,NI)		

## 2011 microdata files and languages

	Open	Safeguarded	Secure
England and Wales	-	9	600+
Scotland	-	11	600+
Northern Ireland	-	13	600+

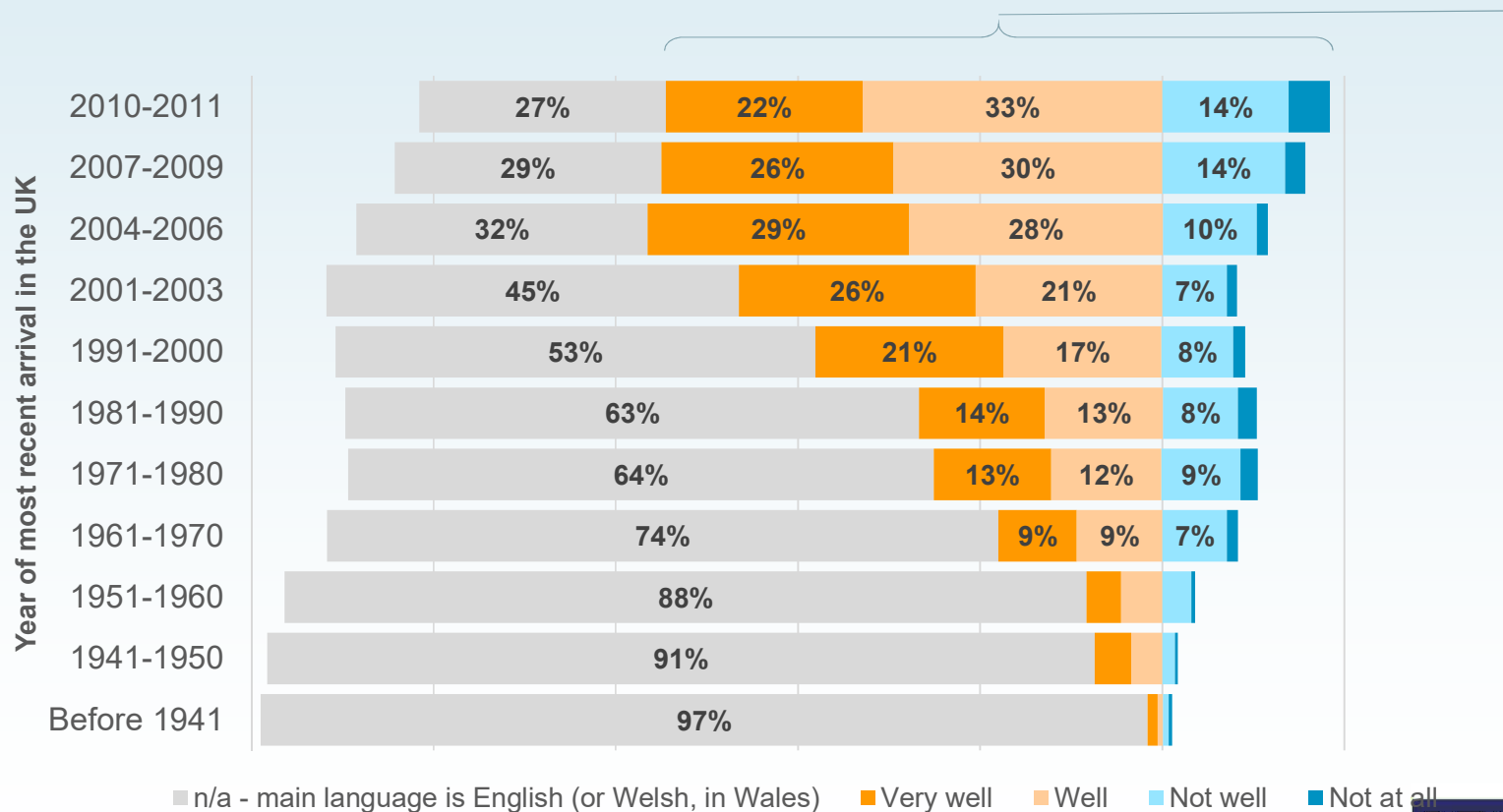
- Safeguarded data via UK Data Service
  - <https://census.ukdataservice.ac.uk/get-data/microdata.aspx>
- Secure data
  - Accreditation required; access via statistical agency secure settings

- For almost all languages, the majority of people who use that main language also report good English proficiency
  - Exceptions: very small numbers of speakers; signed languages
  - Also true at regional level

# "How well can you speak English?"

(amongst 2011 Census (England and Wales) respondents not born in UK)

Main language is not English (or Welsh, in Wales)



Source: 2011 Census, safeguarded microdata



---

# Questions

Contact details:

Oliver Duke-Williams: [o.duke-williams@ucl.ac.uk](mailto:o.duke-williams@ucl.ac.uk)

Vassilis Routsis: [v.routsis@ucl.ac.uk](mailto:v.routsis@ucl.ac.uk)