Changes to Census figures on reported ‘main languages’ in Manchester, 2011–2021

Yaron Matras (Aston Institute for Forensic Linguistics & University of Haifa)
Santiago Chau Príncipe (University of Manchester)

Executive summary

We compared a number of changes in self-reported ‘main language’ in Manchester wards between the census dates 2011 and 2021. A significant increase, both proportionally and in real numbers, is found for Arabic and Urdu, the largest languages after English. Notable increases are also found for Italian, Spanish, Romanian and Portuguese. A number of other languages show proportional growth with lower overall numbers. We try to explain some of the developments by correlating with other datasets, examining geographical spread and co-occurrence with other languages, and direct observations. The increase of Arabic and possibly Romanian is considered due to a combined effect of recent migration and the coming of age of new respondents who identify Arabic as their main language. For the other languages migration appears to be central, though backgrounds tend to be complex: Numbers for Italian and Spanish are partly attributable to a young generation of originally South Asian background, whose parents will have listed Urdu as main language. Portuguese reflects migration from South America as well as Africa. The attempt to compare data

raises fundamental issues of methodology starting with the wording of the census question and on to the format in which results are released with respect to the detail on named languages and present and historical ward boundaries. We conclude that the process would greatly benefit from tighter coordination and consistency of various factors, and from integrating procedures to compile and collate language data through a variety of different sources.

Introduction

In November 2022 the Office for National Statistics (ONS) released data on languages from the national Census 2021 for England and Wales. It showed an increase of 22% over the past decade in the number of people who declared a ‘main language’ other than English: The figure in 2021 was 5.1 million, or 8.9% of the population, up from 4.2 million representing 7.7% of the population in 2011. In the city of Manchester, with an overall resident population of 552,000, around 18% of respondents to the national Census 2021, almost 100,000 people, and more than twice the average for England and Wales, declared a language other than English as their ‘main language’, up from 16.6% in 2011.

Many experts are of the opinion that these figures under-estimate our society’s multilingualism. This has to do with the way in which the Census question was formulated, an issue to which we return below. Nonetheless, the Census clearly shows that Britain is a multilingual society. Languages constitute a major social and cultural indicator, one that needs to be understood.

For Manchester, the most widely reported main languages in Census 2021 were Urdu (ca. 17,500), Arabic (ca. 10,400), Chinese (ca. 8,000), and Polish (ca. 6,500). Panjabi and Spanish each had between 4,000–4,500, Portuguese and Italian around 3,000 each, and Persian, Kurdish, Somali and French between 2,000–2,500 each.

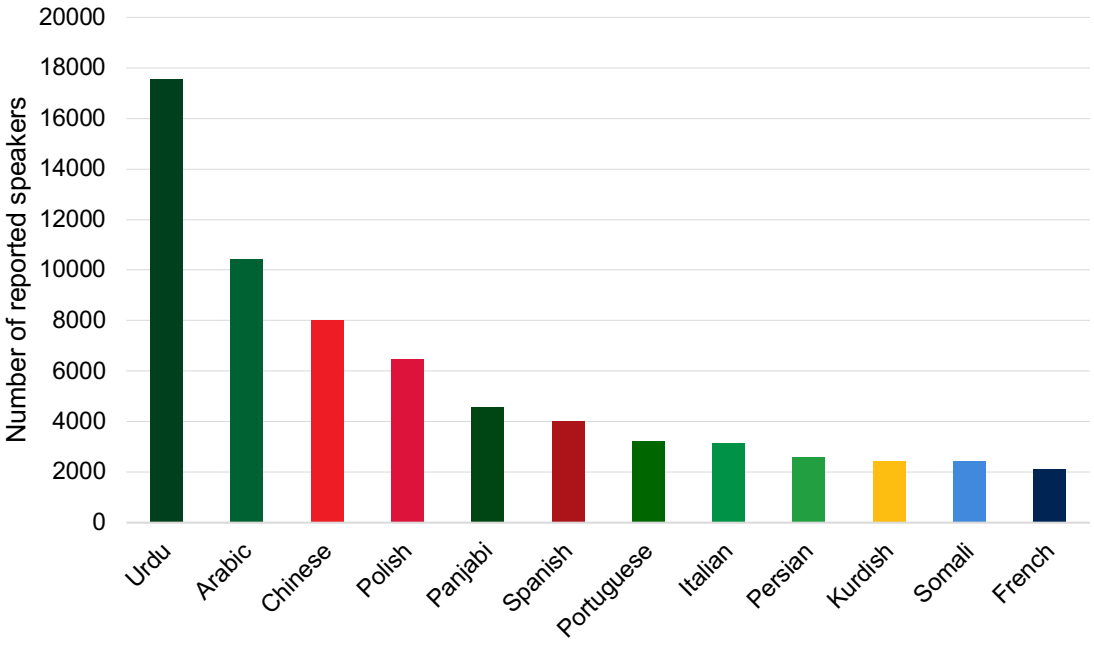

Figure 1: Most widely reported languages in Manchester according to Census 2021.

Why is it important to have information about languages?

Manchester has been celebrating itself as 'Britain's City of Languages' since reports on its multilingualism were first released in 2013 by a research team led by Yaron Matras at the Multilingual Manchester project. Flagging its multilingualism as a distinctive characteristic, the city has since applied successfully for recognition as UNESCO City of Literature. Cultural institutions like the Manchester Poetry Library and Manchester Museum have integrated multilingualism into their outreach programmes. The city has been marking UNESCO International Mother Language Day through annual celebrations since 2017. Multilingualism has become a badge of the city's cultural and social identity and a theme in public events that bring together residents to demonstrate their belonging to the local community.

Information about languages helps understand the composition of our local population. Since all our interactions involve language, recognising patterns of language use is a key component of basic demographics: Languages are an important part of heritage. They are a theme around which individuals rally and join together as communities.

Knowing about languages and giving them public recognition can support confidence building and mutual respect, and help counteract marginalisation. Languages are also important skills. They have the potential to help forge and sustain global links in a variety of domains; understanding languages and their role in shaping identities support cross-cultural understanding and dialogue. Language skills support the career development of individuals. They can create new opportunities for the local work force. Information about preferred languages can help us understand language needs. It can help plan and set up resources in culture and help support links with other localities. It can support wellbeing and communication across generations. Understanding language needs also supports equality by ensuring that key services can make adequate provisions to make them accessible to all.

Problems with the Census question

The question ‘What is your main language?’ was first introduced into the national Census in 2011. We therefore have figures to compare over the past decade. But the wording of the question is ambiguous in several respects. As a result, the data collected by the Census are unlikely to represent the full picture of multilingualism in the country.

First, it is not clear whether ‘main’ refers to proficiency, emotional attachment, order of acquisition, or frequency of use, to name but some.

Next, the Census questionnaire allowed respondents to list just a single ‘main language’. That fails to capture the reality of multilingual repertoires. This was demonstrated by the fact that many respondents in fact tried to list more than one language; however, entries like ‘Berber Arabic’ were either discarded by the Statistics Authority or re-interpreted as referring to (a regional variety or dialect of) just a single language.

The positioning of the question within the Census questionnaire also implies a potential bias. Census takers who chose to specify a language other than English were asked to indicate their proficiency in English. Many will have interpreted that as

stigmatising. Indeed, after Census 2011 results were published, some tried to use them in the political debate to suggest that poor knowledge of English in certain communities raised risks of political radicalisation.

For all those reasons it is likely that many respondents under-reported their use of other languages. Annual School Census data on ‘first languages’ show that pupils from multilingual households typically constitute around 40% of the school population in major cities. In 2020, the number of primary and secondary school pupils in Manchester whose recorded ‘first language’ was not English was 37,551, or 43% of the total school population. The School Census too under-reports ‘first languages’ since it is designed to anticipate learning difficulties and so it does not usually capture the home languages of pupils who appear to be completely fluent in English when they enter school. It is quite likely therefore that the majority of school children in Manchester come from multilingual households. We do not know to what extent the School Census can be regarded as a clear proxy for adult multilingualism: On the one hand, multilingual school children will certainly have multilingual parents. On the other hand, speakers of other languages tend to be a younger population in Manchester and may also have on average more children than English speakers. But the ways data are elicited in the two surveys make a comparison difficult.

Other countries with an English-speaking majority take a more differentiated approach to the Census question: The US, Australia and Ireland ask respondents whether they speak a language other than English in the home. South Africa asks which two languages were spoken most often in the household. New Zealand asks in which languages respondents could have ‘a conversation about a lot of everyday things’. Canada asks about proficiency in English or French, any other language used in the home, and any additional language used on a regular basis, and which language was learnt first.

Despite requests, the Office for National Statistics (ONS) refused to amend the question in the run up to Census 2021. But in an attempt to alleviate some of the shortcomings, more guidance to the question was provided, advising respondents to “think about the language you use naturally” and suggesting “it could be the language you use at home”. When publishing the results ONS also included for the first time

figures on the reported use of different ‘main languages’ in the same household – across generations, within partnerships, and between other household relationships such as friends. Across the country multiple languages were recorded for altogether 6% of households, or 1.5 million people.

The principal reason behind the UK government’s reluctance to amend the Census question appears to be that it is not intended at all as an instrument to collect data on language use. Rather, Question 18 on ‘main language’ is merely a stepping-stone to question 19 on proficiency in English. In effect, the Census question on language serves as an instrument to support the argument that other languages are a hindrance to integration into British society. Various senior politicians have argued that all UK residents should “speak English as first language”. However, among Census respondents who reported a ‘main language’ other than English only 0.3% or 138,000 people reported not to know English at all. In Manchester, the percentage was the same, with only 3,578 respondents stating that they didn’t know English. That suggests that there is no reality behind the assumption that using other languages hinders the acquisition of English, even when another language is perceived as more important and central than English in a person’s daily life.

The Census is therefore not necessarily an accurate snapshot of the use of languages in the country. Instead it represents that portion of the population who were motivated to declare the importance of a language other than English in their daily lives, for one reason or another.

We have no evidence that local authorities actively rely on Census data to plan provisions. Procurement of interpreting and translation services is usually de-centralised, handled by individual service providers and driven mostly by demand. This was demonstrated by the COVID crisis, where in England government agencies like the NHS and Public Health England made minimal efforts to translate information material. Where efforts were made, the choice of languages appeared to be subject to the availability of translators and demand from officials or activists, not based on a systematic survey of needs. Local cultural services too, such as library acquisitions, are driven primarily by demand.

Nonetheless, the availability of some data on languages allows us to identify trends and learn more about the composition of the population.

Comparing Census data on languages: 2011 and 2021

We now have two datasets, covering a decade during which immigration and cultural diversity were widely contested in UK political debates. In the middle of this period was the Brexit referendum, at its end was the COVID pandemic. Both impacted on the composition of UK society, motivating some to leave the country, and making it difficult for others to return, while push factors in various parts of the world continued to drive migration. We now know that the year 2021 in fact saw the biggest net migration to the UK. It is also quite possible that during this decade public debate around issues of identity and culture made many people of all generations more acutely aware of the importance of their heritage and willing to acknowledge it through declarative acts such as those prompted by the Census question. It is therefore difficult to assess the factors behind changes in response to the Census question on ‘main language’.

Issues in the assessment

Various issues around the availability and consistency of the data have complicated our ability to compare the two datasets.

For a start, data from the 2021 Census were not easily accessible. For this report, we needed the distribution of speakers of various languages in each ward. However, in the Bulk Data download option the ONS did not release information on the distribution of languages by wards, the smallest survey unit being local authorities. We therefore had to create a Custom Data set, which allowed us to select wards as a unit using the option ‘Electoral wards and divisions’ for the local authority of Manchester.

However, this confronted us with a second problem. The Custom Data website lists two variables for language: ‘Main language’ (ML) and ‘Main language (detailed)’ (MLd). At first glance the two data sets appear to be similar, showing identical numbers of

‘main language’ for English, French and Spanish speakers, for instance. However, for some language categories the results do not match. And this brings us to our third problem: Not all languages are listed across both options. While the Bulk Data download for ‘Main language (detailed)’ (code TS024) includes languages like Welsh and Portuguese, as well as Akan and Kurdish, in the Custom Data download (which we needed in order to access wards as a survey unit) both ML and MLd datasets combined many of these languages into higher categories. For example, where TS024 listed Amharic, Igbo and Afrikaans as different languages, ML and MLd combined them as ‘African languages’. The data set for the 2011 census resembles the TS024 data set most closely, but it too combines some languages into higher categories.

The two data sets also combined languages differently. For instance, ML has a category for all EU languages, while MLd has a category for Polish and another for the rest of the EU languages. These differences in categorisation are also reflected in the number for each language category. For example, when we add up the numbers for ‘Polish’ and ‘other EU languages’ in MLd to match the equivalent category in ML, we find that there are slight differences: For Manchester ML lists 18,744 speakers of EU languages, MLd has 18,738, while the TS024 data set falls approximately in the middle with 18,740. The overall number of categories also differed: ML has the option to display either 11 or 23 language categories. MLd offers 23 or 26. Both fell short of the detailed 93 language categories offered by TS024. Here too there were some minor discrepancies, as ML calculated 531,479 Manchester residents, MLd had 531,484 and TS024 had 531,483.

In summary, the data was difficult to access since ONS does not release data by ward, there were issues of categorisation and classification of languages, and there were minor discrepancies among the data sets even for comparable categories.

Changes to ward boundaries

Since 2011, changes have been introduced to ward boundaries. Some ward areas were increased while others were reduced in size, and in some cases new wards were

created and old ones disappeared. This presented us with a problem of comparability across the two census years.

To solve this issue as accurately as possible we calculated the percentage of territory lost or gained by individual wards between the two census dates. We then used this percentage to calculate the hypothetical population change that would have taken place had the ward boundaries remained the same. We can illustrate the procedure by examining the ward of Ardwick in Manchester. Figure 2 shows the territory of Ardwick in 2011, Figure 3 shows the shape of Ardwick today and Figure 4 presents the changes:

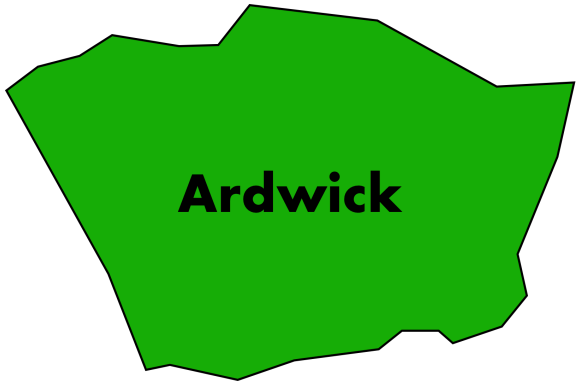

Figure 2: Boundaries of Ardwick in 2011.

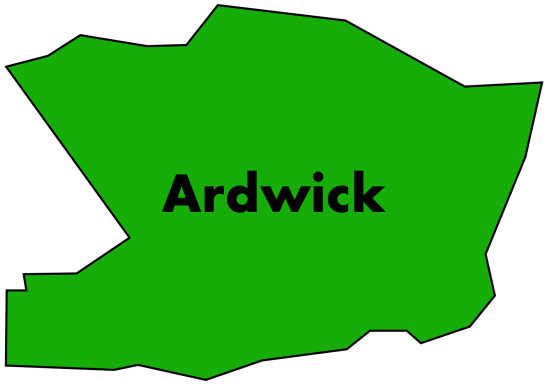

Figure 3: Present boundaries of Ardwick.

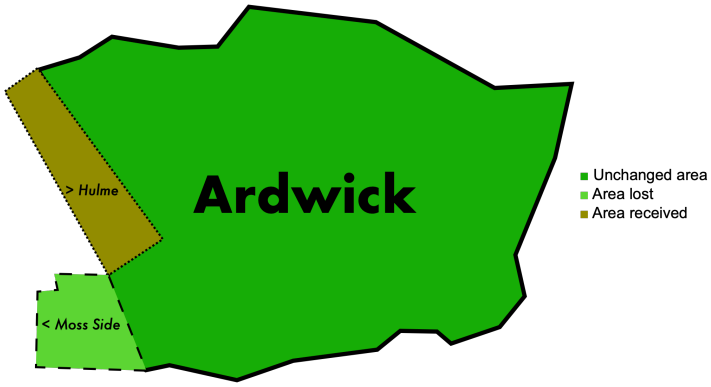

Figure 4: Boundary changes for Ardwick, 2011–present.

As seen in Figure 4, Ardwick lost territory to Hulme and received territory from Moss Side. Ardwick's lost territory represents 0.29 km² or 7% of its earlier area, while the gain from Moss Side represents 0.25 km² or 13.5% of Moss Side's original area.

We used the same percentages to calculate what would have been Ardwick's population in 2011 if it had the area of Ardwick in 2021. In 2011, the number of respondents in Ardwick who declared Spanish as 'main language' was 118. Taking away 7% representing the area lost to Hulme we are left with 93%, or 109.74. That accounts for the loss of area. To account for the gain of area we then add 13.5% of respondents who in 2011 recorded Spanish as 'main language' in Moss Side, or 11.06. That gives us a total of around 120.41, which we round up to 120. This total representing the number of respondents who recorded 'Spanish' in Ardwick in 2011 if the ward had had its present day (2021) boundaries.

We can then compare that projected figure to the one recorded in Ardwick in the 2021 Census. In this case the difference between 118 actually recorded in 2011, and our projection of 120 is rather inisgnificant. Other boundary changes were more comprehensive (see Figure 5 and Figure 6) and the implementation of the procedure there was crucial.

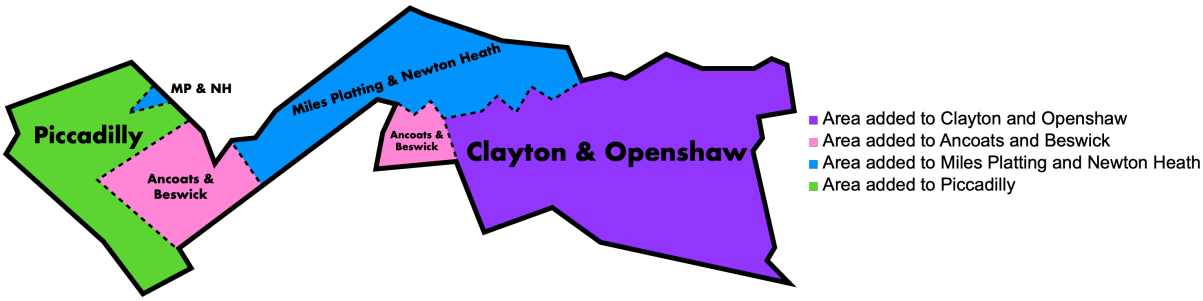

Figure 5: Disintegration of Ancoats and Clayton, showing its modern successor wards.

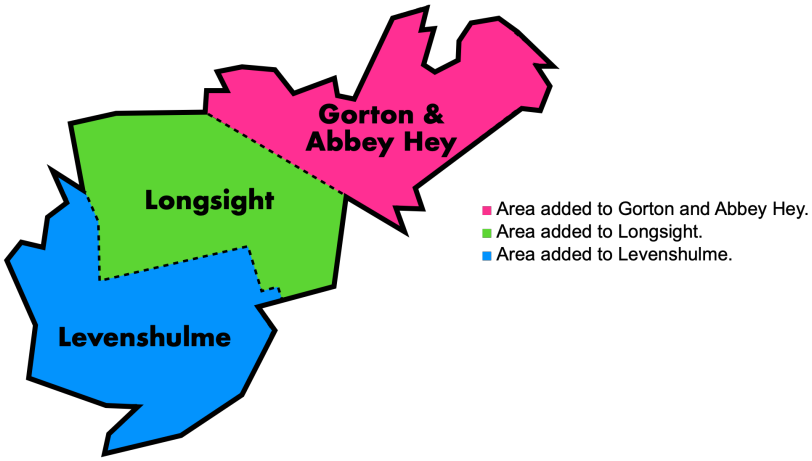

Figure 6: Disintegration of Gorton South, showing its modern successor wards.

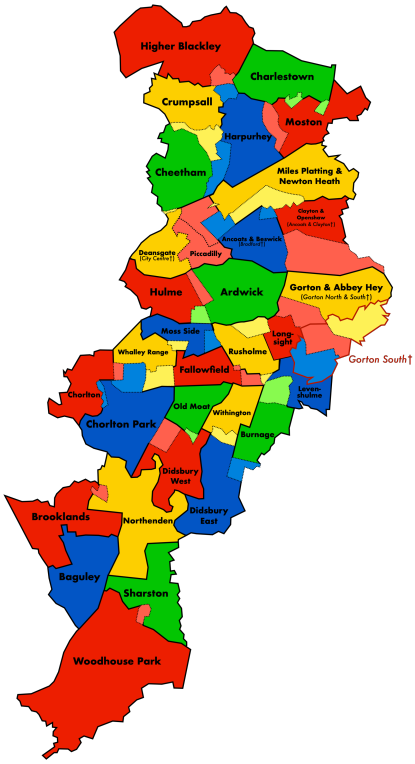

Figure 7: Boundary changes to Manchester's wards, 2011–present

Methodology

We proceeded to select the variable 'Main language' (rather than 'Main language (detailed)') with the option of 23 categories. We downloaded the data set into an Excel (XLSX) spread sheet. In order to compare the two census datasets we made modifications to the figures for each of the wards represented in the 2011 census, as described above. We set a minimum threshold of 5 speakers per ward, with all counts below 5 speakers being amalgamated under a 'Remainder' category. We then compared the number of speakers for each ward listed in the 2011 and 2021 census datasets and calculated their percentage difference.

For the analysis we selected a number of languages based on the following criteria: a) The language should have more than 1000 'main language' entries overall for the city. b) The difference between the two census dates should represent at least around 25±%. We also added c) languages with a change of 100±% in either direction, irrespective of the overall number of respondents. For these selected languages, we

tried to localise the changes in individual wards, examining the five wards with the largest reported number of ‘main language’ entries for the language. Where we were unable to compare data for specific languages, we had no choice but to apply the comparison for the higher category, e.g. ‘South Asian languages’, and deduce possible implications for the language we were interested in.

Principal changes between the two census dates

Overall, Manchester has experienced a growth in the reported numbers of many languages. Some languages, like Italian, Romanian, Igbo and Tigrinya, have doubled or even tripled or quadrupled their reported number of speakers. Table 1 summarises the changes between the 2011 and 2021. The following Figures (Figure 8–Figure 10) show the five wards which have seen a significant increase of the reported speaker populations.

Language	2011 total	2021 total	Change	New reported speakers	2021 Rank (after English)
Urdu	13,095	17,572	+34%	4,477	1
Arabic	7,037	10,425	+48%	3,388	2
Spanish	1,869	4,032	+116%	2,163	6
Portuguese	1,458	3,239	+122%	1,781	8
Italian	753	3,142	+317%	2,389	9
Romanian	720	2,834	+294%	2,114	10
Kurdish	1,886	2,425	+29%	539	12
Malayalam	849	1,947	+129%	1,098	16
Pashto	1,147	1,580	+38%	433	17
Tigrinya	417	1,297	+211%	880	19
Hungarian	353	742	110%	389	24
Akan	306	690	+125%	384	26
Igbo	181	578	+219%	397	30

Table 1: Comparison of total reported speakers between 2011–2021 Censuses and the ranking of the languages (following English) according to the number of their reported speakers in Manchester in 2021.

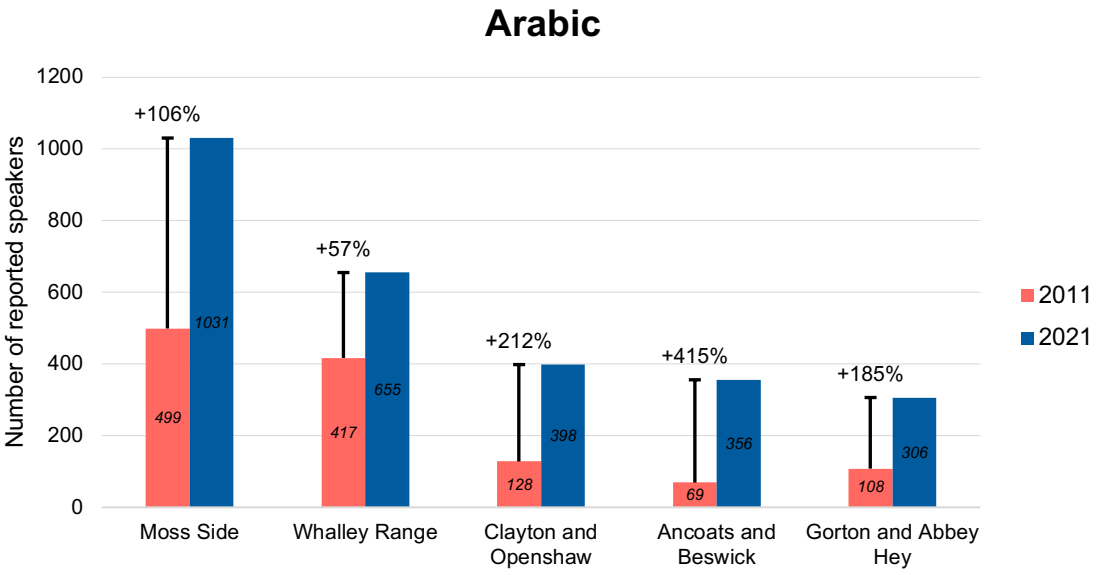

Figure 8: Five wards with a significant change in reported Arabic speakers, 2011–2021.

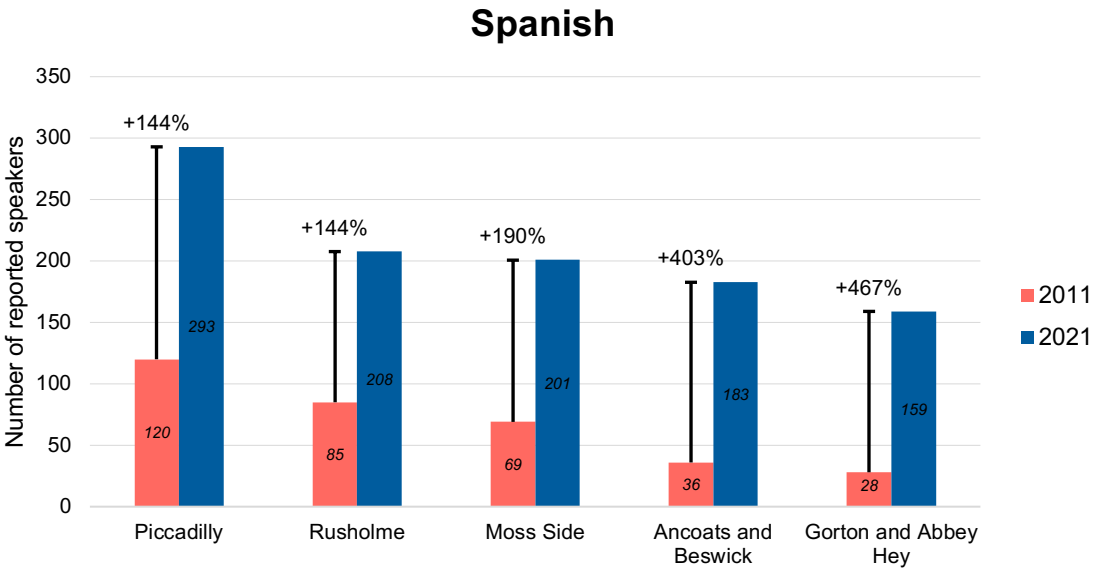

Figure 9: Five wards with a significant change in reported Spanish speakers, 2011–2021.

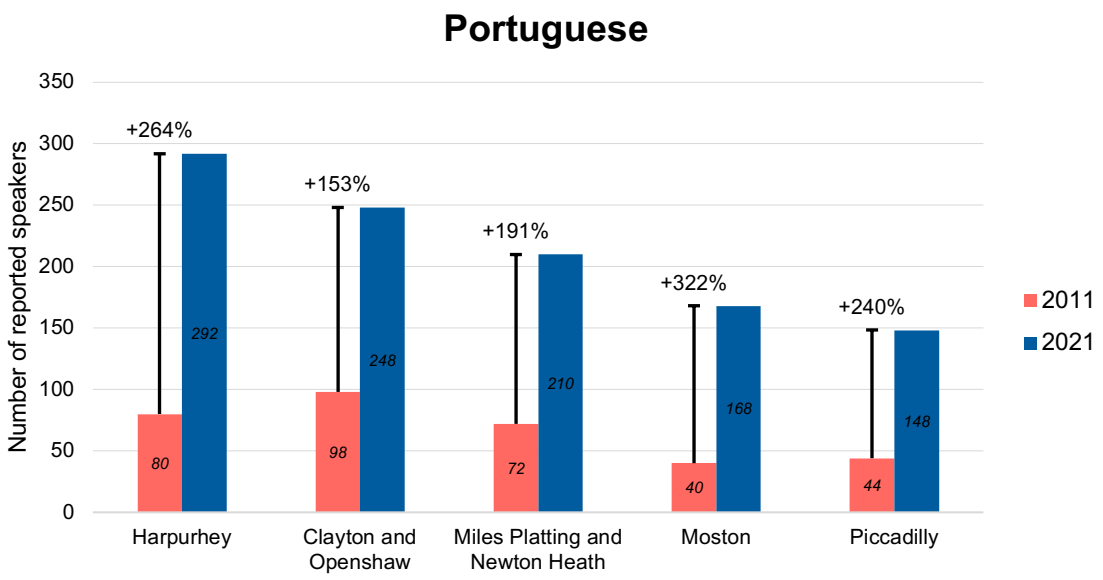

Figure 10: Five wards with a significant change in reported Portuguese speakers, 2011–2021.

As explained above, the 2021 Census groups a number of languages under higher categories at the ward level: Urdu and Malayalam appear under ‘South Asian languages’; Italian, Romanian and Hungarian under ‘EU languages’; Kurdish and Pashto under ‘West or Central Asian languages’; and Tigrinya, Igbo and Akan under ‘African languages’. This impedes a direct comparison between the census dates at ward level. Below (Figure 11–Figure 14) we compared these higher categories.

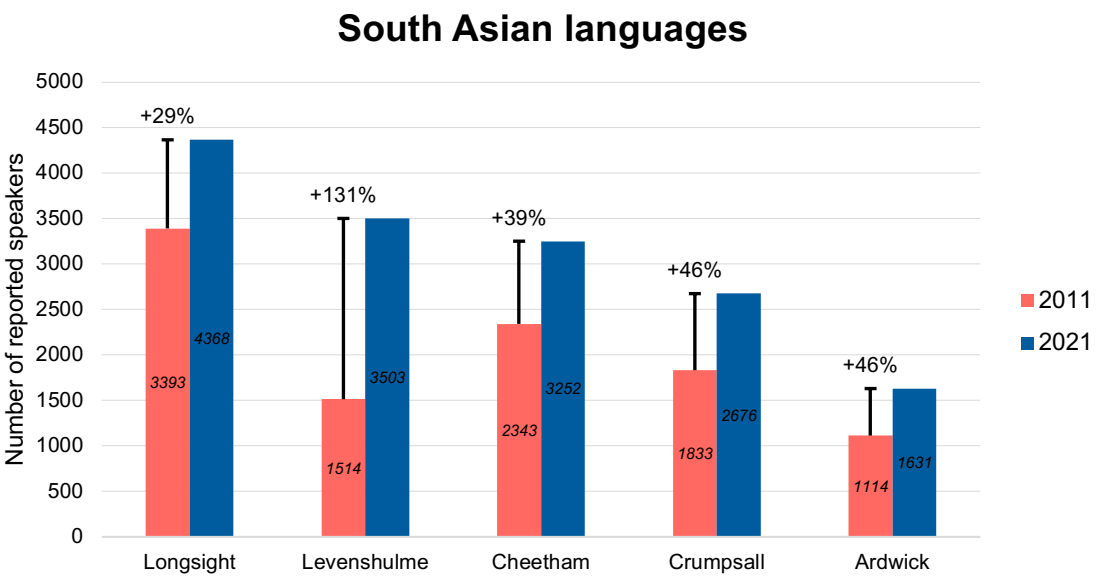

Figure 11: Five wards with a significant change in reported speakers of South Asian languages, 2011–2021.

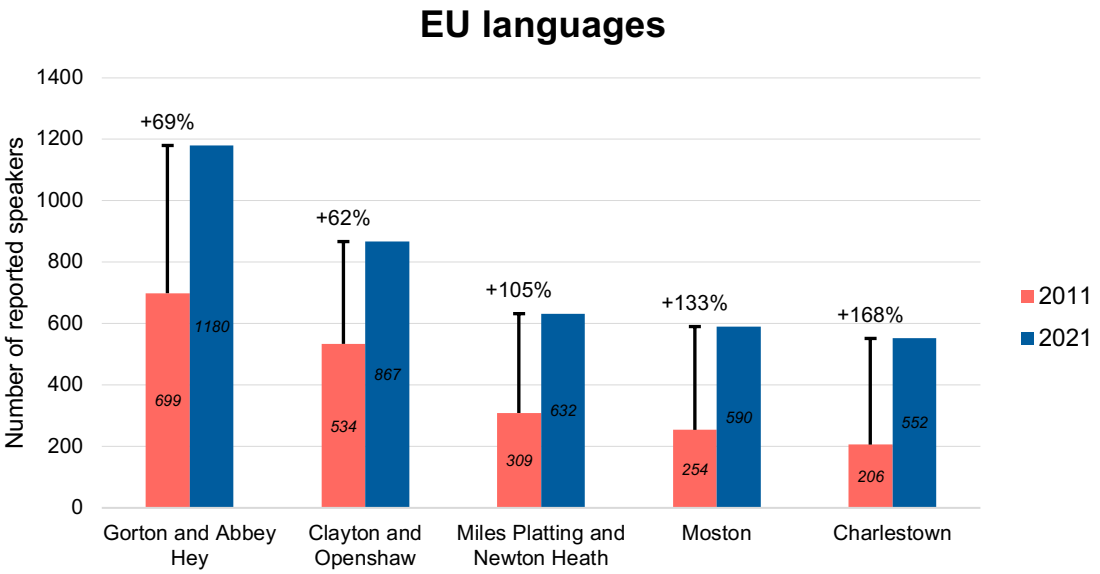

Figure 12: Five wards with a significant change in reported speakers of EU languages, 2011–2021.

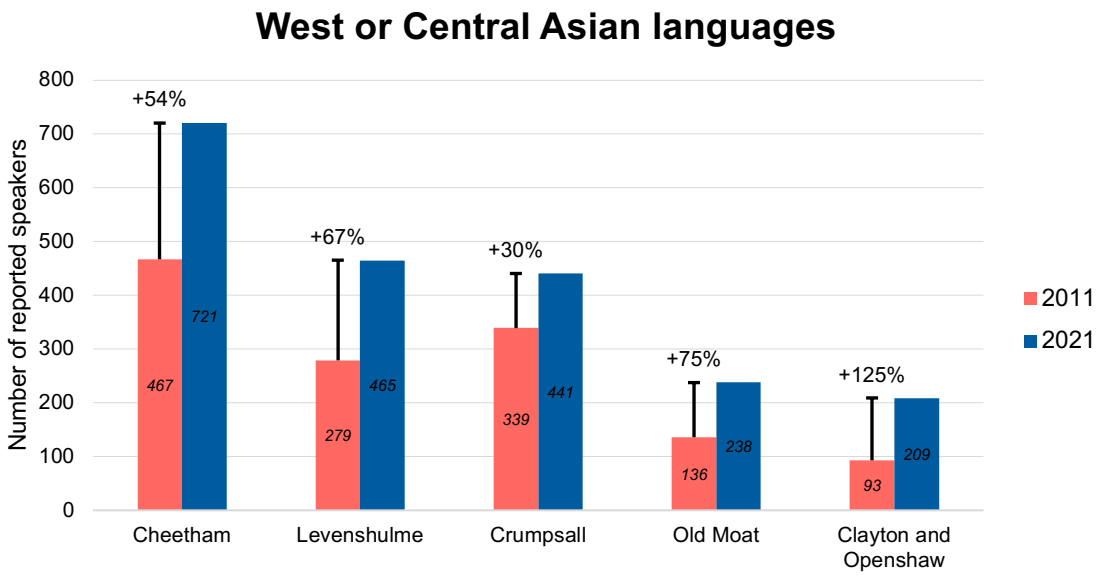

Figure 13: Five wards with a significant change in reported speakers of West or Central Asian languages, 2011–2021.

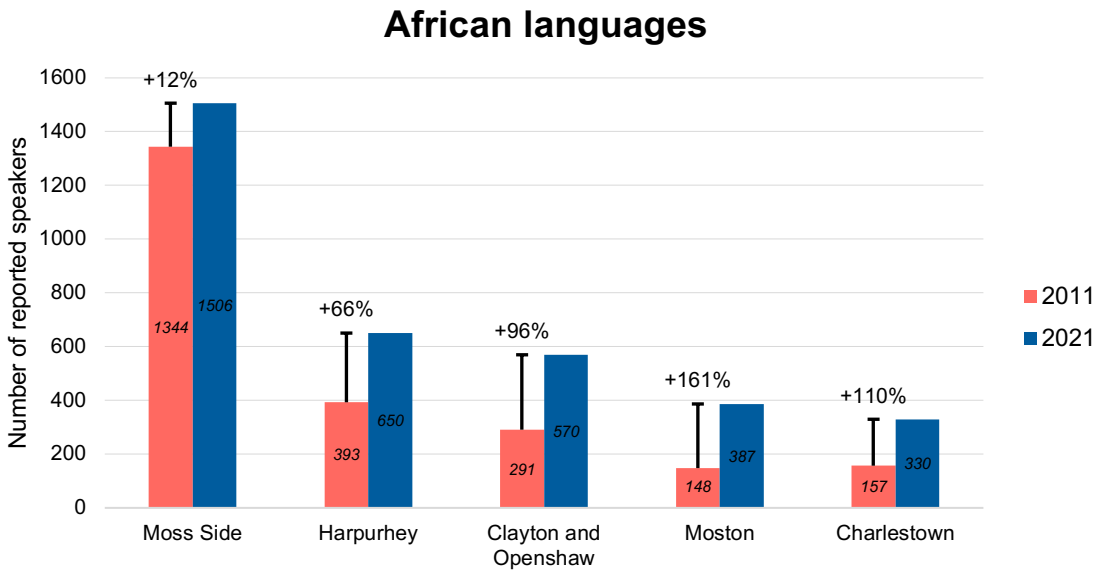

Figure 14: Five wards with a significant change in reported speakers of African languages, 2011–2021.

We have School Census data for Manchester from 2010 and 2020, which we can compare with the two national Census data sets from 2011 and 2021 respectively (Table 2).

Language	2010 total	2020 total	Change	Census change ‘main language’ city-wide
Urdu	5644	9387	+66%	+34%
Arabic	1984	4343	+219%	+48%
Spanish	103	892	+866%	+116%
Portuguese	271	674	+248%	+122%
Italian	71	1134	+1591%	+317%
Romanian	59	454	+769%	+294%
Kurdish	201	865	+430%	+29%
Malayalam	156	274	+175%	+129%
Pashto	214	676	+315%	+38%
Tigrinya	66	284	+430%	+211%
Igbo	41	223	+543%	+219%
Akan	93	311	+334%	+125%
Hungarian	6	198	+3300%	+110%
Romani	24	46	+191%	–97%

Table 2: Comparison of School Census data on ‘first language’, 2010–2020.

Consistently, increases between the two dates are of a considerably higher measure in the School Census compared to the national census. We find a number of cases where the ratio is 1:2 or 1:3, in other words, the increase in schools for the same period was 2 or 3 times that of the increase on the national Census. We find this pattern for Urdu, Portuguese, Romanian, Tigrinya, Igbo and Akan. For Malayalam, many speakers of which are medical professionals recruited to work in the NHS, the ratio of 1:1.3 in the increase in the two datasets over the time period would seem to reflect the ratio of children to adults given the likely age and social position of respondents.

On average, the school age population of Manchester constitutes around 16% of the total population. In some ethnic minorities, with a younger overall population and larger nuclear families, that percentage is likely to be higher. If we add, for each individual language in our sample, the number of reported school age children in 2020 and the number of self-reported adults in 2021 to give the total population of reported speakers for each language, we find that the school age children represent between 12% (Malayalam), 16–18% (Portuguese, Spanish, Romanian, Tigrinya), 21–26% (Hungarian, Italian, Kurdish) and 29–38% (Arabic, Urdu, Pashto). Given that many of

the Hungarian speakers belong to the Romani minority (with higher average birth rates) and that speakers of Italian are often of Pakistani heritage (with family sizes typical of Western Asia), these proportions seem to make sense. They serve in that way as partial verification of the reliability of both datasets, or else their similar likely level of inaccuracy. The significant increase in the School Census data compared with the national Census could then be explained by one or several factors: A natural growth of a school age population among a population of adults who were too young to have school aged children in 2011; immigration over the past decade of a young population with school aged children; and rising awareness of school staff with regard to pupils' home languages leading to greater diligence in recording them.

Discussion

For a small number of languages, the data clearly point to an increase over the past decade. Some of this can be attributed to general population growth. But the overall proportion of people declaring a 'main language' other than English in the national Census has also grown. That may be a reflection of different patterns of family size and a younger population in certain language groups. In a small part the growth may reflect greater awareness of, and self-confidence around home and heritage languages, leading more respondents to declare a 'main language' other than English. Such a process may have been supported or even driven by the favourable public discourse in the city around languages that has emerged since 2013; it may also be a defiant reaction to the negativity associated with multilingualism in the national public discourse since the Brexit referendum debate in 2016.

There is little doubt, however, that there has also been a growth due to new people settling in the city. The rise of Arabic to the position of second largest language other than English, after Urdu, reflects the pull factor of Manchester as a city with an established Arabic speaking community where multiple institutions operate in Arabic, from retail to services and on to several dozen community based supplementary schools that teach Arabic to children of migrants of Arab background. Arabic also shows one of the highest rates of language vitality among young people and it is certainly possible that many young people who grew up in Manchester and are first

time respondents to the census identified with Arabic as ‘main language’. But some evidence on rising numbers of interpreting and translation requests for Arabic (collected by the Multilingual Manchester project) seem to point to an increase of new arrivals, many of them refugees from Syria, Sudan, Yemen, and the Bidooni community of Kuwait. For Tigrinya, Pashto, and Kurdish the increase likely to reflect the arrival of refugees from Eritrea, Afghanistan, Syria, and Iran.

The growth of Urdu especially cannot in our opinion be attributed to natural demographic growth owing to the declining vitality of the language among the established population of South Asian background; in other words, the increase is unlikely to reflect the coming of age (as adult census respondents) of young people who spent most of their childhood or were born in Manchester and must instead be to the arrival of a new population.

The increase for Italian and to some extent Spanish partly reflects the immigration of families of South Asian background coming from EU countries and making use of the re-settlement opportunities that existed until Brexit regulations came into force in 2021. Many are attracted to the city owing to its dense South Asian population. That could be behind the increase of Spanish in Rusholme and Moss Side (Italian is subsumed under EU Languages for the ward-based data set and is therefore difficult to localise).

For Malayalam, the increase can be attributed to continuing professional recruitment into the local health care sector.

The distribution by wards partly helps us understand the picture. For West and Central Asian languages the growth can be localised in Cheetham, Levenshulme, and Crumpsall, while for South Asian languages it can be localised in these wards as well as in Longsight. This points to existing concentrations of speakers acting as a pull factor for the settlement of new arrivals. The localisation of Arabic is different, with concentrations in Moss Side and Whalley Range.

For Portuguese we find a different pattern, with a strong preference for Harpurhey and Moston, strengthening the impression that a substantial proportion of those who reported Portuguese as ‘main language’ are of African background. African languages,

too, recorded a significant growth in Harpurhey, Moston and Charlestown, contiguous areas in North Manchester, while much of the growth for EU languages is attributable to Gorton and Abbey Hey, where much of the Romanian and Romanian Roma community resides (indicating in turn that Roma reported the national language 'Romanian' rather than their ethnic minority language 'Romani' – the similarity in labels being completely coincidental – as 'main language'), and to Moston, where there is a large Hungarian speaking community (many of them too are Roma).

Spanish and Portuguese also show concentrations and growth in Piccadilly. We interpret those as representing recent arrivals who are young professionals of Latin American background.

Finally, East Manchester, represented by the wards of Clayton and Openshaw, Miles Platting and Newton Heath, and Ancoats and Beswick, appears to have been a substantial settlement target for speakers of a variety of languages include all those in our sample except for Spanish and South Asian languages. We attribute this to residential opportunities that satisfy budget considerations and are not counteracted by, or simply supplement the pull factor of established concentrations of residents of shared backgrounds.

Numbers for some of the city's major languages remained at similar levels between the two census dates. This is the case for Polish (with 6,471 declared 'main language' in 2021, compared to 6,447 in 2011), Panjabi (4,591 in 2021, compared to 4,719 in 2011), Bengali (3,367 in 2021, compared to 3,114 in 2011), and Persian (2,580 in 2021, compared to 2,660 in 2011). For Polish, the reported process of return to Poland, both before and after the Brexit referendum, may have offset the demographic growth of the local population. For Panjabi and Bengali, immigration may have offset demographic changes in the opposite direction among established communities that include an elderly population.

Numbers for Bulgarian increased from 381 in 2011 to 605 in 2021, a growth of 58%, but within an overall small population. For Somali, one the city's larger languages, the increase (2,415 in 2021, compared to 2,958 in 2011) is with 22% just below the threshold we set for our selected sample languages. Here too the changes might be

attributed to a combination of population growth and immigration. For Somali in particular, Manchester provides a strong pull factor given the presence of a large community in Moss Side and Rusholme and the preference among Somalis to reside in densely localised communities.

For a number of European languages that are equally below our selected sample's threshold for population size, the comparison shows a decline. Declines of 25% or more are identified for Slovak (353 in 2021, compared to 685 in 2011), German (589 in 2021, compared to 936 in 2011), Dutch (283 in 2021, compared to 439 in 2011). The numbers are relatively small but anecdotal evidence points to a connection with emigration arising from post-Brexit discomfort.

Final observations pertain to languages with much smaller speaker populations but which are nonetheless noteworthy. The number of those declaring Yiddish as a 'main language' increased in Manchester from 5 to 62 (a more than tenfold increase, or 1140%, technically the largest increase across the two census dates). This might be attributable to changing attitudes or even to natural demographic growth in the community that is known to have the highest birth rate in Britain, though it is more likely to also reflect the re-settlement of Hasidic Jews from Israel ahead of Brexit, who were able to claim ancestry in one of the EU countries of central and eastern Europe. The slight increase of Hebrew from 50 to 67 (34%) is likely to be a part of the same process. For other languages such as Romani, Caribbean Creoles, Cornish, Scottish Gaelic or Oceanic languages the differences in percentages are seemingly high but the actual numbers are of little significance, whether interpreted in real terms or in terms of changing attitudes to the act of self-declaration, which is in effect what the census captures most accurately.

Conclusions

As we try to make sense of Census data on languages we face a number of obstacles: It is not entirely clear how respondents choose to answer the question given its potential ambiguity. Data are most certainly lost because respondents can only indicate on single 'main language' introducing an artificial hierarchy that does not

reflect the reality of many households, as the ONS's own reporting indicates. We also do not know what data are lost because respondents shy away from declaring regional and non-standard languages (as in the case of Romani, or Caribbean Creoles), what data are lost because respondents shy away from linking home language(s) to proficiency in English.

One of the arguments given by officials for not wanting to amend the question was to ensure comparability between the two census dates. But that has now been impeded due to the lack of consistency in the formats in which data have been released. A further impediment has been the changes to ward boundaries. Since this is a routine process, one would hope that changes ward boundaries should be taken into consideration and search formulae might be developed similar to the one that we've piloted here and built in as algorithms into the download options for Census data.

A more comprehensive picture can be provided by triangulating data sets on language. Above we mentioned the School Census. But data on languages can also be drawn from statistics on library loans, school exams in languages, and requests for interpreting and translation across various public sectors, to name but some. In 2018 the Multilingual Manchester Data Tool was launched to provide proof of concept for a rudimentary method of triangulating such data sets. It encountered numerous impediments, among them the difficulty to persuade local agencies to compile and share data on a regular basis, and the inability of a university based research project to sustain a team of co-workers (academic and technical) to import, update and maintain data on a regular basis, including updates to the Application Programming Interface required to keep track of changes to ward boundaries. The project delivered its proof of concept, and it now represents a snapshot in past time, but it also has become partly dysfunctional due to the lack of technical updates. This demonstrates the risks of relying on an institutional environment that is driven by considerations of short-term reputation and income generation and lacks the infrastructure and holistic, civic commitment that is required in order to provide a reliable and sustainable service to the community. Alternatives need to be found under the auspices of local and national government agencies in order to implement such ventures in the future. They should include the creation of protocols for data compilation and sharing as well as expert consultancy, technical solutions and the modest infrastructure needed to

accommodate them. Manchester has made a move in the right direction when in 2019 it published a report on the city's language diversity having set in motion a consultation process with various city institutions and public services, in response to a draft City Language Strategy. The process has since stalled. The opportunity to study data, and the realisation that the numbers and proportions of the city's multilingual residents continue to rise should give the process further impetus.

References and links consulted

Office for National Statistics, Census 2021 Data on Languages:
https://www.ons.gov.uk/peoplepopulationandcommunity/culturalidentity/language/bulletins/languageenglandandwales/census2021

Census Bulk Data Download website:
https://www.nomisweb.co.uk/census/2021/bulk

Census Custom Data website:
https://www.ons.gov.uk/datasets/create

Multilingual Manchester Data Tool
https://sites.manchester.ac.uk/mlm-datatool/

Matras, Yaron, Rebecca Tipton and Daniele Viktor Leggio. Proposal for a City Language Strategy, July 2018
https://mcrcityoflanguages.org/developing-a-city-language-strategy-for-manchester/

Manchester City Council
Report for Information, Manchester's Language Diversity, 5 September 2019
https://democracy.manchester.gov.uk/documents/s9767/Manchesters%20Language%20Diversity.pdf

Matras, Yaron, and Deepthi Gopal. 2013. What languages are spoken in England and Wales? CoDE Census Briefings
https://hummedia.manchester.ac.uk/institutes/code/briefingsupdated/what-languages-are-spoken-in-england-and-wales.pdf

Matras, Yaron, and Deepthi Gopal. 2013. Who can and cannot speak English? CoDE Census Briefings
http://mlm.humanities.manchester.ac.uk/wp-content/uploads/2015/12/who-can-and-cannot-speak-english.pdf

Matras, Yaron. 'What is your Main Language?' In order to engage with policy, Modern Linguists need a vision for society. Cross-Language Dynamics Research Blog, Aug 6, 2018
http://projects.alc.manchester.ac.uk/cross-language-dynamics/what-is-your-main-language-in-order-to-engage-with-policy-modern-linguists-need-a-vision-for-society/

Matras, Yaron. Improving the census question on 'language' could help repair community relations and Britain's international image post-Brexit, Manchester Policy Blogs, April 1, 2019
https://blog.policy.manchester.ac.uk/growth_inclusion/2019/04/improving-the-census-question-on-language-could-help-repair-community-relations-and-britains-international-image-post-brexit/

Matras, Yaron. 2023. Decolonising language in the city: Multilingual repertoires, institutional practice and civic engagement in a UK urban setting. Working Papers in Urban Language & Literacies 307.
https://wpull.org/wp-content/uploads/2023/01/WP307-Matras-2023-Decolonising-language-in-the-city.pdf

Matras, Yaron, and Alex Robertson. 2015. Multilingualism in a post-industrial city: policy and practice in Manchester. *Current Issues in Language Planning*. 16, 296-314.
https://yaronmatras.files.wordpress.com/2021/10/matrasrobertson2015-copy.pdf

Matras, Yaron, and Leonie Gaiser. 2020. Re-visiting ‘community language’: Arabic in a Western global city. In: Mar-Molinero, C*lare. ed. Researching Language in Superdiverse Urban Contexts: Exploring Methodological and Theoretical Concepts.* Celevdon: Multilingual Matters. 52-78. https://yaronmatras.files.wordpress.com/2021/10/revisiting-community-languages.pdf

Appendix

Heat maps (2021)

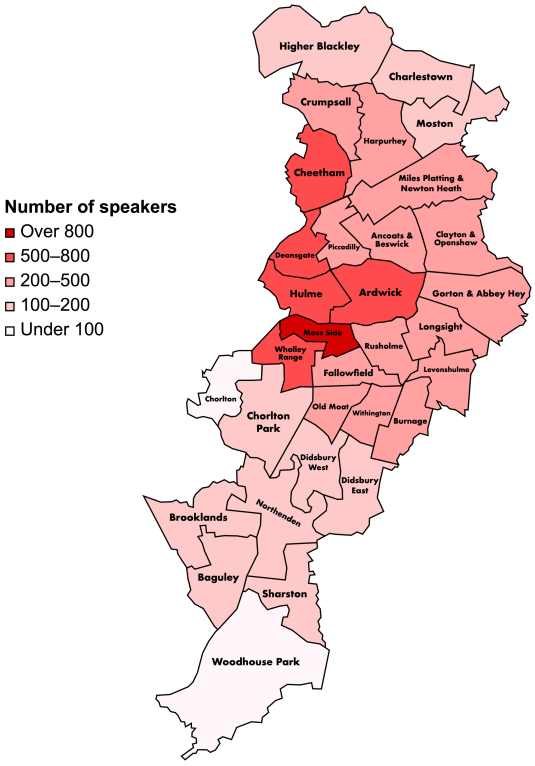

Figure 15: Distribution of Arabic speakers across Manchester’s wards.

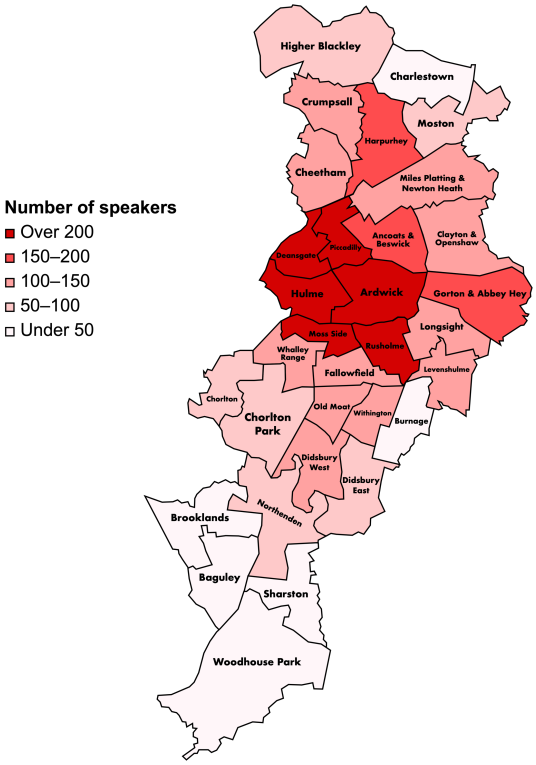

Figure 16: Distribution of Spanish speakers across Manchester's wards.

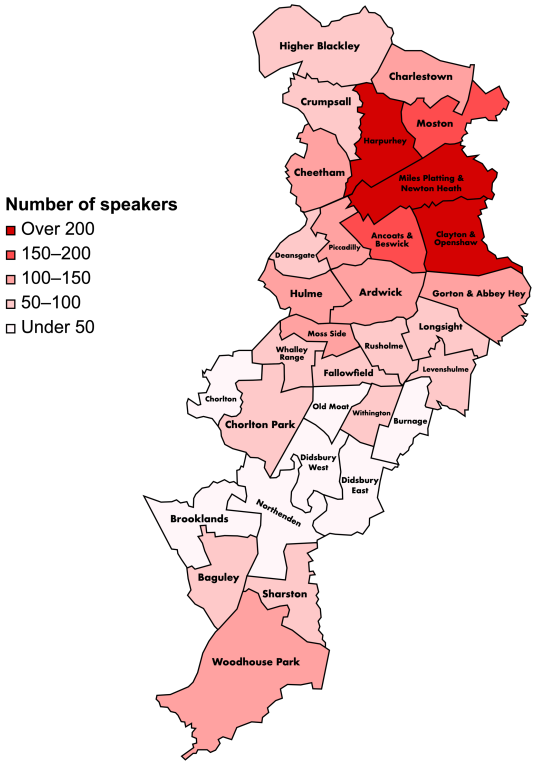

Figure 17: Distribution of Portuguese speakers across Manchester's wards.

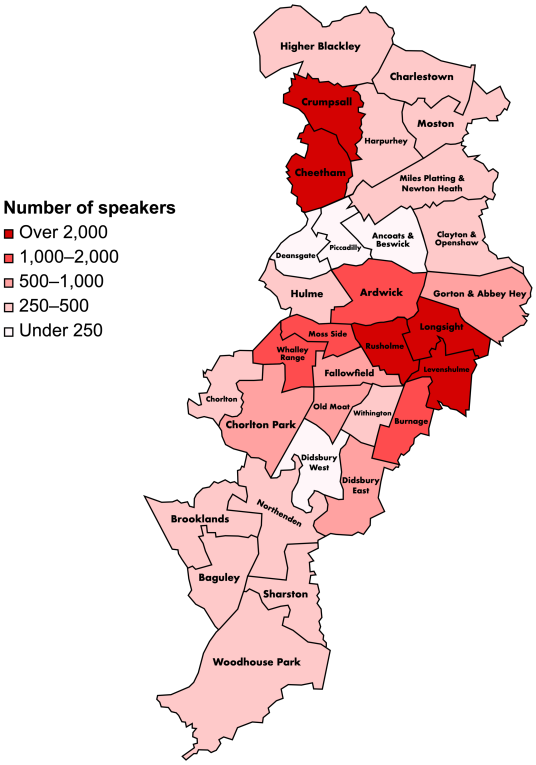

Figure 18: Distribution of South Asian languages across Manchester's wards.

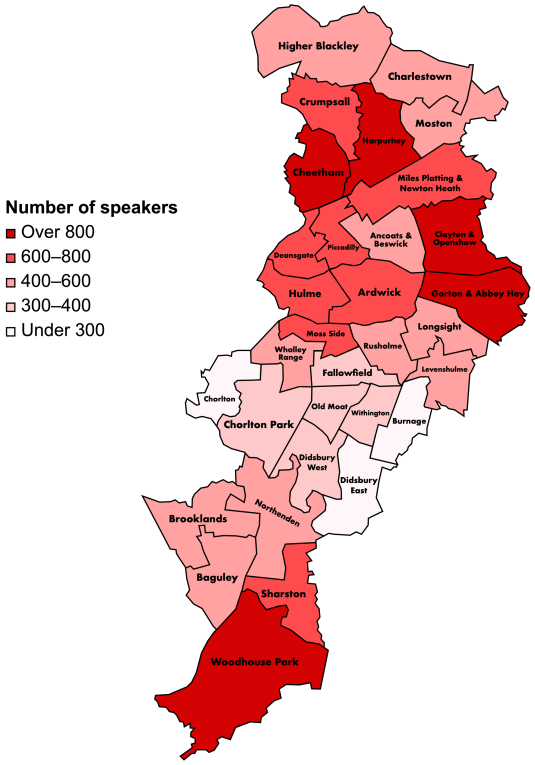

Figure 19: Distribution of EU languages, excluding Spanish, Portuguese and French, across Manchester's wards.

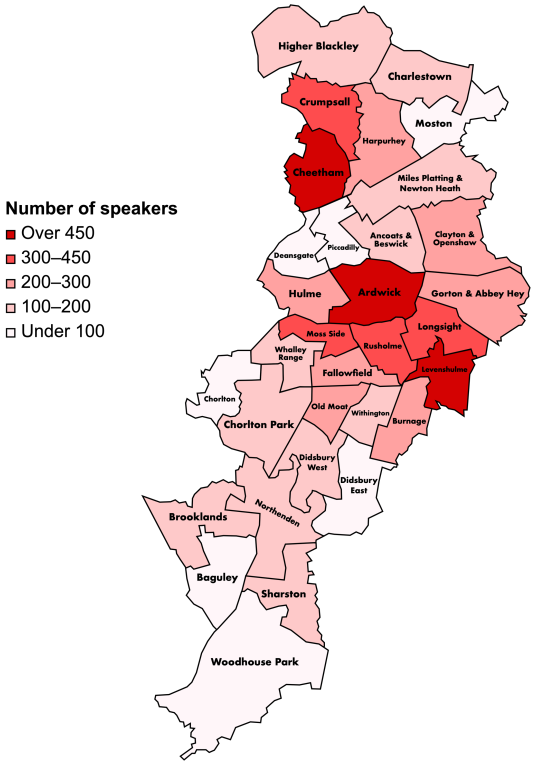

Figure 20: Distribution of West and Central Asian languages across Manchester's wards.

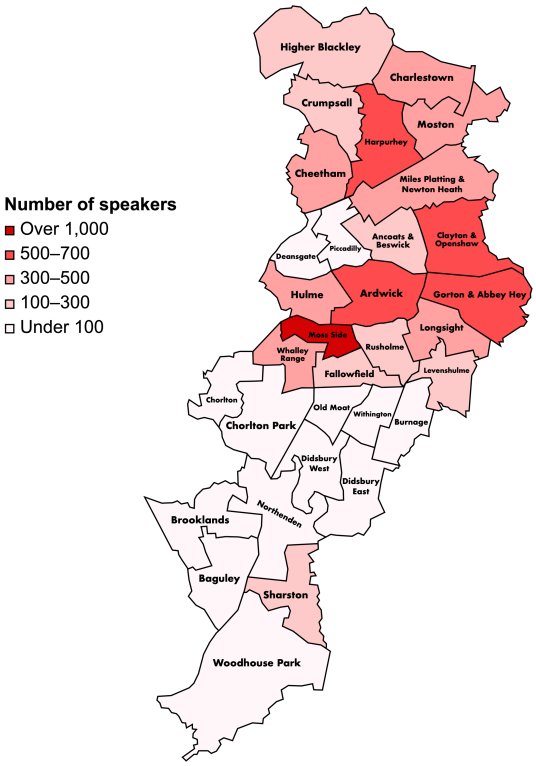

Figure 21: Distribution of African languages across Manchester's wards.