

Estimation of the number of Welsh speakers in England

Introduction

The number of Welsh speakers in England is a topic of interest as they must represent the major part of the Welsh-speaking diaspora. Their numbers have been the matter of speculation. To try to produce a reliable estimate of their numbers we have used the ONS Longitudinal Study¹ (LS), a record linkage study containing census information on a 1 per cent sample of the population of Wales and England, to throw light on the issue.

Using the LS offers a means of finding out about the former Welsh-speaking ability of those LS members now resident in England who lived in Wales at the time of the 1971, 1981 or 1991 Census. It does not give an indication of their ability at the time of the 2001 Census. They can be split too into 2 groups: those born in Wales, and others, and this provides another categorical split of interest.

Methodology

The 2001 Census reported that 610,000² of the population of England was born in Wales. The ONS Longitudinal Study (LS) contained records of 6,313 people resident in England at the time of the 2001 Census who were born in Wales.

The LS contains details, where available, from the 1971, 1981, 1991 and 2001 Censuses. All of those Censuses contained a question, asked in Wales only, about the Welsh-speaking ability of respondents. If the members of the LS sample were resident in Wales at the time of one of the Censuses held in 1971, 1981 or 1991 therefore, and had answered the question, it was possible to say whether at that time the sample members could speak Welsh. Where a record could be matched with more than one of the 1971-1991 Censuses, the most recent recording of Welsh-speaking status was used in the subsequent analysis.

As the earliest Census linked was the 1971 Census, no information about their Welsh-speaking ability would be available for any sample members who although they were resident in Wales at some time moved to England before 1971, not returning at the time of the intervening Censuses. Nor would information be available for anyone who, although born in 1968³ or later, had left Wales before the next Census. The information for the sample is therefore truncated, with no direct information available for those under 13 years of age, and information for only a relatively small proportion of some of the older age groups of those born in Wales, particularly those aged over 45 by 2001.

¹ <http://www.statistics.gov.uk/about/data/methodology/specific/population/LS/default.asp>

² Source: 2001 Census, table KS05.

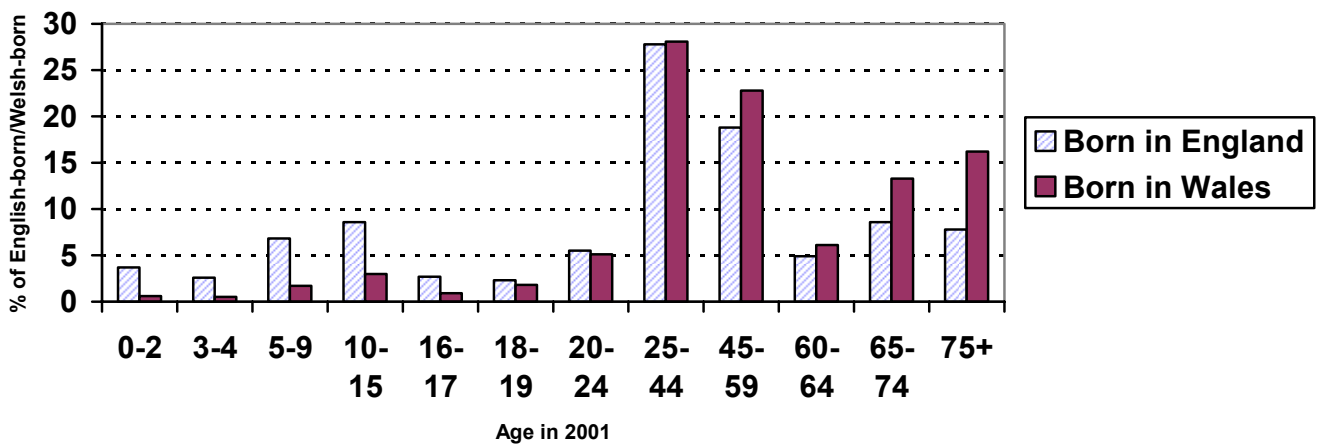
³ Although the question about Welsh is asked of all ages, Census tabulations of the variable are usually limited to those aged 3 or over as children below three may not be considered by their parents (who generally make the Census return) as able to speak any language.

Assuming that those born in Wales were also resident in Wales at some point enables us to assess the completeness of the matching process for those born in Wales. Similar calculations are not possible for those not born in Wales. We can identify whether they were resident in Wales in 1971, 1981 or 1991 but if they were not in Wales at those points, we do not know whether or not they were ever resident in Wales. The vast majority of those resident in England in 2001 never resided in Wales of course. The analysis presented here distinguishes therefore between those born in Wales and others.

The age distribution of the Welsh-born residents of England is quite different from the age distribution of the English-born residents of England, as can be seen from Figure 1. Fewer of those born in Wales are aged under 18: 6.7%, compared to 24.4% of those born in England. Although people of all ages migrate, the peak rate of emigration from Wales currently is around age 18⁴ when school-leavers migrate, largely when moving into higher education. Although those aged 25-44 form around 28% of both those born in England and those born in Wales, amongst those aged 20 and over, higher proportions of those born in Wales are aged 45 and over.

Figure 1

Percentage age distribution of the English- and Welsh-born residents of England, 2001



Source: ONS Longitudinal Study, author's analysis

This age distribution means that the truncation in the availability of information on the Welsh-speaking ability of the young is considerably less important for the task of estimation than is the short-fall in the information available for the older age groups.

⁴ Government Actuary's Department: Population projections, 2004-based principal assumption

Estimation for those born in Wales

Some basic statistics of the LS sample *of those born in Wales* are shown in Tables 1 and 2.

Table 3 also shows the percentage found to be Welsh-speaking in the sample, by Census source, and the percentage Welsh-speaking in the Census for the same cohort, as far as it was possible to obtain. For those age groups where more than one Census could be the source of an indication of Welsh-speaking ability, one can see that the Census at which the informants' age-group included age 18 provided more matches than any of the other Censuses. Probable significant differences between the sample and Census percentages are indicated. Of the 16 comparisons made, 2 are significantly different and one of these is not comparing exactly the same age cohorts. This last point is important as in the case of those aged 60 or over in 2001, the choice of the age cohort in earlier years was made subjectively. A greater number of significant differences could have been obtained had different cohorts been used for comparison.

This analysis therefore suggests that generally those born in Wales but resident in England in 2001 were not significantly different, in terms of the proportion of them who spoke Welsh, from others of the same age cohort when they were last in Wales.

Looking at the sample Welsh-speaking percentages with 2001 age-groups across the 1971-1991 Census, one finds that for only one age group – those aged 25 – 34 in 2001 – does the 95% confidence intervals for the difference between the proportion in 2 successive Censuses not include zero. For that cohort, the point estimates and 95% confidence intervals for the percentage speaking Welsh were:

| | |
|---|---------------|
| when aged up to 4 at the 1971 Census: 2.8%: | 0.8% - 9.6% |
| when aged 5 to 14 at the 1981 Census: 14.1%: | 10.2% - 19.2% |
| when aged 15 to 24 at the 1991 Census: 26.1%: | 21.2% - 31.7% |

(The percentage speaking Welsh aged up to four is probably depressed because children under 3 may not be recorded as speaking Welsh simply because they can not speak. As they grow, some of them will come to speak Welsh as their home language or mother tongue).

For this 25 – 34 age group, both the LS figures and the Census figures reflect the effect of schooling in increasing the percentage which spoke Welsh amongst the cohort. In estimating what percentage of those aged up to 10, and 10 to 14, in 2001 could speak Welsh it would seem justifiable to assume that their ability would be based on acquisition in the home, rather than via the school system, as most of them would probably have received most of their education in England. (A percentage of 7% would be reasonable, reflecting the percentage of 3 to 4-year olds at the time of the 2001 Census who in Wales could speak Welsh and lived in a family where both the parents, or the single parent, could speak Welsh).

Estimation for all the other age groups can reasonably be based on the observed percentage of Welsh speakers within the entire sample of the age group because of the lack of significance in the percentages observed across the source Censuses (and the general absence of significant differences between the sample estimates and Census estimates for cohorts). Using the lower and upper limits from the Wilson score 95% confidence intervals⁵ produces corresponding lower and upper limits on the overall percentage who could at one time speak Welsh amongst those Welsh-born residents of England aged 10 and over. The central estimate is 16.0%, with a lower estimate of 14.1% and a higher estimate of 19.2%. Applying these percentages to the *total* numbers born in Wales produces estimates of the number born in Wales resident in England in 2001 who could speak Welsh: the central estimate is 98,000, the lower estimate is 86,000, the higher 117,000. Despite the complication of this examination, the central estimate barely differs from that obtained by using the following simpler approach.

Table S15 of the 2001 Census provides a broad age breakdown of 0 – 15, 16 – pensionable age, and over pensionable age, by sex, of the number of those born in Wales, resident in England. Applying the relevant percentages of Welsh-speaking in Wales, calculated from table CAS146 of the 2001 Census, for the 16 and over age groups, but using an estimate of 7% for those under 16 (as suggested above but not incorporated in the subsequent estimation), produces an estimate of 16.8% Welsh-speaking, representing 102,000 speakers.

(Using a similar approach but using data from table 6.3 of Focus on Wales: Its People⁶ which tabulates the numbers of Welsh-born living in all other parts of the UK, produces an estimate of another 1,000 Welsh-speakers living in Scotland and Northern Ireland).

⁵ Macros written by Robert Newcombe of Cardiff University were used to calculate confidence intervals.

http://www.cardiff.ac.uk/medicine/epidemiology_statistics/research/statistics/newcombe/proportions/index.htm

⁶ <http://www.statistics.gov.uk/focuson/wales/default.asp>

Table 1 LS Residents of England, 2001, born in Wales: basic LS statistics

| Age in 2001 | Can speak Welsh | Cannot speak Welsh | Total | Recorded in Wales 1971-1991; missing response | No census records 1971-1991 | Census records available; Not recorded in Wales 1971-1991 | Total LS sample |
|-------------|-----------------|--------------------|-------|---|-----------------------------|---|-----------------|
| Under 10 | 0 | 0 | 0 | 0 | 174 | 0 | 174 |
| 10 to 14 | 0 | 14 | 14 | 43 | 17 | 75 | 149 |
| 15 to 19 | 23 | 53 | 76 | 0 | 31 | 102 | 209 |
| 20 to 24 | 49 | 140 | 189 | 52 | 17 | 65 | 323 |
| 25 to 34 | 104 | 463 | 567 | 0 | 41 | 206 | 814 |
| 35 to 44 | 87 | 502 | 589 | 6 | 32 | 331 | 958 |
| 45 to 59 | 74 | 333 | 407 | 36 | 27 | 968 | 1438 |
| 60 or over | 25 | 157 | 182 | 4 | 82 | 1,980 | 2,248 |
| All ages | 362 | 1,662 | 2,024 | 141 | 421 | 3,727 | 6,313 |

Source: ONS Longitudinal Study, author's analysis

Table 2 LS Residents of England, 2001, born in Wales: estimation

| Age in 2001 | Can speak Welsh | Cannot speak Welsh | Total | % able to speak Welsh | 95% confidence interval (based on Wilson score) | |
|----------------|-----------------|--------------------|-------|-----------------------|---|-------------|
| | | | | | Lower limit | Upper limit |
| Under 10 | - | - | - | - | - | - |
| 10 to 14 | 0 | 14 | 14 | 0 | 0 | 21.5 |
| 15 to 19 | 23 | 53 | 76 | 30.3 | 21.1 | 41.3 |
| 20 to 24 | 49 | 140 | 189 | 25.9 | 20.2 | 32.6 |
| 25 to 34 | 104 | 463 | 567 | 18.3 | 15.4 | 21.7 |
| 35 to 44 | 87 | 502 | 589 | 14.8 | 12.1 | 17.9 |
| 45 to 59 | 74 | 333 | 407 | 18.2 | 14.7 | 22.2 |
| 60 or over | 25 | 157 | 182 | 13.7 | 9.5 | 19.5 |
| All 3 and over | 362 | 1,662 | 2,024 | 16.0* | 14.1* | 19.2* |

*: weighted by total sample size (last column in Table 1)

Source: ONS Longitudinal Study, author's analysis

Table 3 Residents of England, 2001, born in Wales: comparison of Census and LS

| Age group in 2001 Census | Source of Welsh-speaking status | | | Total LS sample size including records not matched with 1971-91 Census in Wales |
|----------------------------|---------------------------------|---------------------|--------------------|---|
| | 1971 Census | 1981 Census | 1991 Census | |
| Under 10 | | | | |
| Age at Census: | Not possible | Not possible | Not possible | 174 |
| 10 - 14 | | | | |
| Age at Census: | Not possible | Not possible | 0 – 4 | |
| % Welsh speaking in Census | | | 16.1 (age 3-4) | |
| % Welsh-speaking in LS | - | - | 0 | |
| Total in sample | - | - | 14 | 149 |
| 15 – 19 | | | | |
| Age at Census: | Not possible | Not possible | 5 – 9 | |
| % Welsh speaking in Census | | | 24.7 | |
| % Welsh-speaking in LS | - | - | 30.3 | |
| Total in sample | - | - | 76 | 209 |
| 20 – 24 | | | | |
| Age at Census: | Not possible | 0 – 4 | 10 – 14 | |
| % Welsh speaking in Census | | 13.3 (age 3-4) | 26.9 | |
| % Welsh-speaking in LS | - | <15 | 27.3 | |
| Total in sample | - | 17 | 172 | 323 |
| 25 – 34 | | | | |
| Age at Census: | Up to 4 | 5 -14 | 15 – 24 | |
| % Welsh speaking in Census | 11.3*** (age 3-4) | 18.4 | 17.1*** | |
| % Welsh-speaking in LS | <4 | 14.1 | 26.1 | |
| Total in sample | 72 | 227 | 268 | 814 |
| 35 – 44 | | | | |
| Age at Census: | 5 -14 | 15 - 24 | 25 – 34 | |
| % Welsh speaking in Census | 15.7 | 14.9*** | 14.1 | |
| % Welsh-speaking in LS | 14.6 | 16.3 | 9.6 | |
| Total in sample | 212 | 294 | 83 | 958 |
| 45 – 59 | | | | |
| Age at Census: | 15 - 29 | 25 - 39 | 35 – 49 | |
| % Welsh speaking in Census | 15.9 (age 15-24) | 14.5 (age 25-34) | 15.4 | |
| % Welsh-speaking in LS | 17.7 | 17.7 | 20.6 | |
| Total in sample | 248 | 96 | 63 | 1,438 |
| 60 and over | | | | |
| Age at Census: | 30+ | 40+ | 50+ | |
| % Welsh speaking in Census | 18.3 (age 25-44) | 16.8 (age 35-44) | 17.9 (age 5-65) | |
| % Welsh-speaking in LS | 10.9 | 8.0 | 20.6 | |
| Total in sample | 64 | 50 | 68 | 2,248 |

***: value does not lie within Wilson score 95% confidence interval calculated for the sample percentage.

Source: Censuses of Population and ONS Longitudinal Study, author's analysis

Estimation for those not born in Wales

The 2001 Census reported that 48,529,000⁷ of the population of England was born outside Wales, in England or elsewhere. 89% of these non-Welsh born residents of England were born in England. The ONS Longitudinal Study (LS) contained records of 502,281 people resident in England at the time of the 2001 Census who were born outside Wales. 87% of those were born in England. Whereas the 6,313 LS sample of Welsh-born English residents yielded 2,024 with a previous Census record in Wales and information concerning their ability to speak Welsh, the sample of those born outside Wales yielded 2,165, a higher absolute number but a much smaller percentage of the total sample. Table 4 presents a breakdown of them by age. Table 5 provides a comparison with the Welsh-born sample.

In Table 5 it can be seen that amongst those aged 24 or under there was no significant difference between the percentages able to speak Welsh amongst those in the LS sample born in Wales and born outside Wales. For the 25 to 34-year-old age groups the percentage able to speak Welsh amongst those born outside Wales was significantly lower than amongst the Welsh-born: 10.0% compared to 18.3%. The sample sizes amongst the older age groups prevent exact comparisons but it is clear that the differences are much larger.

The crucial difference, of course, between those residents of England who were born in Wales and those born outside Wales is that for the former it is known that they were at some point in Wales whereas for the latter group, unless they were recorded by a Census in Wales, we do not know whether they ever resided in Wales (although we can be sure that the vast majority of them never did).

Of the total 6.8% could speak Welsh. A 95% confidence interval is 5.8% to 7.9% but this really only relates to those non-Welsh-born residents of England who had been resident in Wales at one of the 1971-1991 Censuses. We will assume that the Welsh speaking ability of these can be used to calculate an estimate for the whole population. The 147 who were recorded as Welsh speaking in those Censuses represented just 0.029% of the total 502,253 LS sample⁸ of non-Welsh-born residents of England. Applying these percentages to the Census total of 42,358,885 non-Welsh-born residents of England produces very approximately a central estimate of 12,000 with a lower estimate of 8,000 and a higher of 13,000.

⁷ Source: 2001 Census, table KS05.

⁸ Sample with relevant full coding of age, source etc.

Table 4 Residents of England, 2001 in LS: basic LS statistics

| Age in 2001 | Can speak Welsh | Cannot speak Welsh | Total | Recorded in Wales 1971-1991; missing response | No census record 1971-1991 | Census records available; Not recorded in Wales 1971-1991 | Total |
|-------------|-----------------|--------------------|-------|---|----------------------------|---|---------|
| 0 to 14 | 3 | 21 | 24 | 21 | 63,557 | 29,155 | 92,757 |
| 15 to 24 | 46 | 111 | 157 | 22 | 7,589 | 50,991 | 58,759 |
| 25 to 34 | 41 | 369 | 410 | 21 | 9,133 | 60,073 | 69,637 |
| 35 to 54 | 49 | 915 | 964 | 47 | 8,490 | 134,688 | 144,189 |
| 55+ | 8 | 602 | 610 | 24 | 5,289 | 130,988 | 136,911 |
| All ages | 147 | 2018 | 2,165 | 135 | 94,058 | 405,895 | 502,253 |

Source: ONS Longitudinal Study, author's analysis

Table 5 Residents of England, 2001 in LS: comparison of those born outside Wales with those born in Wales

| Age in 2001 | Born outside Wales % able to speak Welsh | Age in 2001 | Born in Wales % able to speak Welsh | 95% CI for difference in percentage |
|----------------|--|-------------|-------------------------------------|-------------------------------------|
| 0 to 14 | 12.5 | 3 to 14 | 0 | -10.52 to 31.00 |
| 15 to 24 | 29.3 | 15 to 24 | 27.2 | -6.53 to 11.18 |
| 25 to 34 | 10.0 | 25 to 34 | 18.3 | -12.58 to -3.91 |
| 35 to 54 | 5.1 | | | |
| 55 and over | 1.3 | | | |
| | | 35 to 44 | 14.8 | |
| | | 45 to 59 | 18.2 | |
| | | 60 and over | 13.7 | |
| All 3 and over | 6.8 | | 17.9 | -13.09 to -9.13 |

Source: ONS Longitudinal Study, author's analysis

Discussion of the basis of the estimates

The total central estimate of the number of people resident in England in 2001 who could at an earlier date speak Welsh is 110,000, 98,000 of whom were born in Wales.

There are a number of qualifications to this estimate, some of which have already been noted:

1. This is an estimate of those who could at an earlier time speak Welsh. Not all of them would still consider themselves able to speak Welsh. Other analyses of the LS⁹ have attempted to quantify the scale of loss. This is likely to be significant. For example, on the basis of the previous analyses one might expect 16% of those 25 – 34 years old in 2001 who spoke Welsh in 1991 to consider themselves as unable to speak Welsh by 2001.

⁹ Jones, H (2005) Ability to speak Welsh in the censuses of population: a longitudinal analysis. *Population Trends No. 122 Winter 2005*. http://www.statistics.gov.uk/downloads/theme_population/PopTrends122v1.pdf

2. In the case of the Welsh-born, the estimate relates to the first-generation of emigrants. Some of them may have raised their children to speak Welsh and unless they lived in Wales at the time of one of the 1971-1991 Censuses their Welsh speaking ability will not have been recorded and will not have entered into the estimation process in any way. It is difficult to assess the extent of this problem¹⁰.
3. It is possible others will have learnt Welsh without being either born in Wales or ever resident in the country. The numbers involved are unlikely to be significant.
4. As noted for those born outside Wales the estimate calculated rests on the assumption that the characteristics of sample for which information about Welsh speaking ability is available, i.e. those non-Welsh-born residents of England who had been resident in Wales at one of the 1971-1991 Censuses, may be applied to the total population. The sample actually provides no information about the Welsh speaking ability of any non-Welsh-born who may have resided in Wales before 1971 so the estimation approach used might be expected to *underestimate* the number of Welsh speakers among the non-Welsh-born. On the other hand, those non-Welsh-born people not represented in the sample would be aged at least 28 in 2001. For people over that age who moved to Wales after 1971 the analysis has shown (Table 5) that the percentage of Welsh speakers among them was significantly lower than amongst the younger age groups. So using the percentages obtained from the sample might *overestimate* the percentage Welsh speaking amongst the longer established non-Welsh born emigrants from Wales. In addition there were fewer non-Welsh born living in Wales. (Until 1971 fewer than 18% of Wales' residents were born outside Wales¹¹. Only by 2001 had the percentage grown to 25%). Thus it is probable that the percentage of older non-Welsh-born residents of England who ever lived in Wales is smaller than that which would be found among the younger age groups. To conclude this complicated consideration. There are difficulties with the estimation approach used; some aspects could lead to an overestimate, others to an underestimate. The central estimate of 12,000 still seems reasonable in the circumstances.

¹⁰ Though in relation to both this point and the third point it may be worth recording evidence of an in-flow to Wales of Welsh speakers to higher education. The National Assembly for Wales' Statistical Bulletin 52/2006 *Welsh in Higher Education Institutions, 2004/05* reported "300 students with some teaching through the medium of Welsh were not Welsh domiciled. 195 of these were English domiciled and a further 80 were from the rest of the European Union." <http://new.wales.gov.uk/docrepos/40382/40382313/403824/post16/post16-2006/sb52-2006-e.pdf?lang=en>

¹¹ Office for National Statistics (2004). Focus on Wales: its people. http://www.statistics.gov.uk/downloads/theme_compendia/fow/Wales.pdf

Conclusion and policy implications

The estimate that 110,000 people resident in England can speak Welsh is reasonable. Such a number – together with another thousand in Scotland and Northern Ireland – would mean that the total number of people able to speak Welsh in the UK exceeds 690,000 and that some 17% of them live outside Wales elsewhere in the UK.

What is the relevance of such numbers to language planners and policy makers in Wales? (One may assume that they are of no or hardly any interest to policy makers in England). Possible interest to Welsh planners depends on the “mobility” of the speakers, i.e. how likely they are to return to Wales at some point for an extended period or periods of time, at what ages and for what purpose – of working age? With or without children? Close to retirement¹²? The ethno-linguistic vitality of the Welsh speakers outside Wales matters. To what extent do they represent a market of consumers – for Welsh language broadcasting and other cultural outputs?

Beyond an appreciation of the size and nature of the Welsh speaking diaspora, a key consideration may be the accessibility of the Welsh speaking diaspora to language planners. If they can not be identified or contacted, how are they to be considered?

These questions suggest possible further lines of research.

Acknowledgements

Thanks go to Bola Akinwale, Salah Merad and Karl Ashworth of the ONS for their assistance in using the LS data.

Hywel Jones BSc, DMS, FRSS, CStat
Welsh Language Board
22 January 2007

¹² 2001 Census Standard tables 134, 135 and 146 provide some indications of the extent of “mobility” but only over a 12 month period. All of them tabulate numbers moving into Wales in the 12 months before the 2001 Census.