**Phonological levelling, diffusion and divergence in Liverpool and its hinterland**

There are three cohorts of speakers from each of three geographical localities: Liverpool, St Helens & Skelmersdale. The three cohorts are (1) archive speakers (born between 1890-1943); (2) older speakers (born between 1918-1952); teen speakers (born between 1992-1994).

The methodology for data collection was a loosely structured interview. For the archive cohort the speakers are giving oral history interviews and for the older and teen cohorts the data collection method is based on William Labov’s sociolinguistic interview design.

In the archive cohort, the topics in the interviews focus on growing up and working in Liverpool, St Helens and Skelmersdale, mostly from the period around the turn of the 20th century and around the period of the First World War. The audio data, although not the transcript, are stored at the North West Sound Archive, based in Clitheroe Castle. We thank the North West Sound Archive for making these data available for this project.

The data in the older and teen cohorts were specifically recorded as part of this project. The interview topics centred on speakers’ impressions of and beliefs about regions of north-west England and the accents spoken there (specifically, Liverpool, St Helens and Skelmersdale). This often includes a discussion of the linguistic features the participants believe are part of the corresponding accents, and the stereotypes which surround them.

The submitted dataset, which consists of tab-delimited text files, is sorted first by geographical location (Liverpool / Skelmersdale / St Helens) and then by speaker cohort (Archive / Older / Teen). The number of files in each dataset is as follows:

/Liverpool/Archive: 22 files

/Liverpool/Older: 6 files

/Liverpool/Teen: 11 files

/Skelmersdale/Archive: 6 files

/Skelmersdale/Older: 6 files

/Skelmersdale/Teen: 7 files

/St Helens/Archive: 10 items

/St Helens/Older: 7 files

/St Helens/Teen: 9 files

Total number of datafiles: 84

This corresponds to data from the followings numbers of speakers (since some speakers were recorded in pairs or groups of three, the same file may represent the interview data from two or three speakers):

/Liverpool/Archive: 22 speakers

/Liverpool/Older: 12 speakers

/Liverpool/Teen: 20 speakers

/Skelmersdale/Archive: 11 speakers

/Skelmersdale/Older: 10 speakers

/Skelmersdale/Teen: 14 speakers

/St Helens/Archive: 12 speakers

/St Helens/Older: 11 speakers

/St Helens/Teen: 16 speakers

Total number of speakers: 128

File names are structured as follows:

LOCALITY CODE - COHORT - GENDER – NUMBER

* Locality code: Liv=Liverpool; SK=Skelmersdale; STH=St Helens
* Cohort: Archive; Older; Teen
* Gender: F=female; M=male

Examples

* Liv\_ArchiveF02 (Liverpool Archive Female speaker 2)
* STH\_TeenF03+M03 (St Helens Teen Female speaker 3 and Male speaker 3)