1: The NECTE project

UNESCO’s Convention for the Safeguarding of Intangible Cultural Heritage includes in its definition of the latter ‘oral traditions and expressions, including language as a vehicle of the intangible cultural heritage’. The Newcastle Electronic Corpus of Tyneside English (NECTE) is a project devoted to the preservation of materials relating to the linguistic and cultural heritage of Tyneside. The NECTE project:

1. Preserves interviews with Tyneside people of the late 1960s and the early 1990s that provide fascinating insights not only into how Tynesiders spoke at those times, but also into their lives and attitudes.
2. Uses the most up-to-date information technology to provide ready access to this material on the Web, and to ensure that it will not be lost to future generations.

The NECTE project amalgamates two separate collections of recorded speech. One of these collections was made in the late 1960s and early 1970s as part of the Tyneside Linguistic Survey (TLS) project, based in the Department of English Language at the University of Newcastle upon Tyne, the other collection was made in 1994 as part of the Phonological Variation and Change (PVC) project based in the same university’s Department of Speech. The Tyneside Linguistic Survey was an ambitious and groundbreaking project, in many ways ahead of its time (see Strang et al. 1968, Pellowe et al. 1972). The key personnel included the late Professor Barbara Strang and John Pellowe; Vince McNeany, a locally-born mature student who conducted all the interviews and transcribed them phonetically; Val Jones, who carried out the computational analysis; Joan Beal, Anthea Fraser Shields (now Gupta) and John Local, who worked as graduate students on related projects, and Graham Nixon.

The aims and methods of the TLS were:

• Theoretically, in setting out to gain scientific understanding of Tyneside English by extracting it from actual usage rather than interpreting it in terms of theoretical preconceptions;
• Methodologically, in using (i) advanced mathematical concepts and (ii) computer technology to achieve this understanding.

The TLS project was funded by what was then the Social Science Research Council (now ESRC) and originally set out to conduct 150 loosely-structured interviews with people from Gateshead on the south bank of the Tyne. A stratified random sample was selected, and all interviews were conducted by the same person, Vince McNeany. The interviews were recorded onto reel-to-reel analogue tapes, 103 of which survive, though three of these are damaged. Some interviews were also carried out in Newcastle, but very few of these have survived (as far as we know). One of the ways in which the TLS project was extremely advanced for its time was in attempting to use multivariate analysis to determine similarities between speakers. This involved the use of an elaborate and very detailed hierarchical system of coding. The chart on which this code was recorded survives (on the back of a sheet of wallpaper), so we were able to scan and preserve this. An extract appears in figure 1 below.

Figure 1: extract from the TLS coding chart. Here, the symbol in the box represents a set of words in which the vowel in, e.g. keep, people is found. Each line represents a major phonetic variant, and the symbols on each line very fine distinctions of pronunciation.

The speech was transcribed onto index cards, on which codes for various levels of linguistic analysis were also entered on different lines, as illustrated in figure 2. Here the sets of numbers on the fourth, fifth and sixth lines each correspond to one of the fine phonetic distinctions in figure 1. The line of speech recorded is ‘The wife’s, the wife’s sister says erm…’.
The TLS researchers then had the codes punched onto cards (this was the way in which electronic data was entered in the early 1970s) in order to carry out analysis.

These methods of coding and transcription, and the forms in which records were preserved proved problematic to us in various respects, but we have been able to restore them to their original form and to include them in our corpus as one of the NECTE data representations. Figures 1 and 2 thus represent historical data which may well be important to anybody with an interest in the history of science and/or linguistics. NECTE has made no attempt either to review the TLS phonetic transcriptions relative to the original audio recordings, or to extend the phonetic representation to what the TLS did not cover, but the TLS transcriptions are offered as an historical artefact.

Phonological Variation and Change in Contemporary Spoken English (PVC) project was carried out much more recently, so both the data and the means of coding and transcription are much more accessible. This project was funded by the ESRC, the successor to the body which funded the TLS. The key personnel were Lesley Milroy, Jim Milroy, Gerry Docherty, Paul Foulkes, David Walshaw, Dominic Watt and Penny Oxley. The primary data consists of high quality recordings on DAT tapes. Speakers were recorded in pairs consisting of friends, husbands and wives, or siblings. The fieldworker stayed in the background and encouraged the speakers to converse freely, and the resulting recordings average 60 minutes in length. A fuller account of the methodology of this project can be found in Milroy et al. 1997.

In 2001 we were awarded a substantial research grant to produce an enhanced electronic corpus resource from a combination of the TLS and the PVC collections which is available to the community in a variety of formats: digitized sound, standard orthographic transcription and phonetic transcription along with a substantial amount of
background information about the project, and the Tyneside dialect. This can be accessed at www.ncl.ac.uk/necte.

2: NECTE as a resource for social and cultural history

There is little point to creating a resource like NECTE unless it provides information of interest not only to academics but also to the community at large, so in the remainder of this paper, we look at ways in which the corpus can be used to gain insight into Tyneside linguistic and cultural heritage. Because speakers were encouraged to talk about their lives, work and leisure, the NECTE materials provide an important resource for social historians. The TLS interviews provide a record of life on Tyneside just before the decline of industry. All the working-class male informants had left school at 14 or 15, mainly to take up jobs in industry. Unemployment was rarely mentioned, though several were on long-term sick leave or had retired due to injuries sustained in industrial accidents. Men had jobs with titles such as chipper, shunter, wire-drawer, moulder and tool-hardener, none of which would be familiar today. The value of the TLS as a social history resource is best illustrated by letting the participants speak for themselves, as in the following extract. Here, G044, a thirty-eight year old, working class man, is the participant, and 01 is the interviewer.

[TLS/G044] they're one of the one of the dirtiest pla-all foundries is dirty have you been in a foundry?
[TLS/01] I haven't actually no
[TLS/G044] and you've got the fumes and-all to put up with you know... it was one of the dirtiest places I've been in .. and the method they had of working was one of the worst places
[TLS/01] aye
[TLS/G044] they worked in gangs you see eh and you know what it is if you give one of them fellows a job on their own they couldn't start it
[TLS/01] aye
[TLS/G044] that's the way it was they used to do a little bit of a job and then their charge-man used to do the rest give them a job on their own they were beat they couldn't do it

Here, G044 describes the working conditions in the foundry in a way which conveys graphically the unpleasantness of the job.

The NECTE corpus provides very valuable information on the changing lives of women in the second half of the 20th century. The surviving TLS tapes include forty interviews with married women. Twenty-six of these describe themselves as housewives, four as retired, and only ten (25%) were in employment. Of these, four were in what we might call ‘white collar’ jobs. Thus the expectation, especially amongst working-class women, was that marriage would lead to the end of paid employment. This expectation led to thwarted ambitions, as illustrated in the next extract, in which G214 is a working-class woman in her thirties and 01 is the interviewer again.

[TLS/G214] I wanted you-see I had a ehmm .. I had a chance to go to art .. an art college in Shields
[TLS/01] mm
[TLS/G214] and my grandparents wouldn't let me because my mother had gone through college and she never finished
[TLS/01] aye
she got married you see ehm because I was brought up by my grandparents and they said ridiculous you'll be married before

you see so they wouldn't eh .. let it so I went into I took a a course at night-schools because I really wanted to go to eh

this college I could have had the scholarship you know

I only wish I could have sort-of done more than what I did you know but I was married quite young

The third extract illustrates the social changes that had taken place between the period in which the TLS interviews were recorded (1969/70) and the 1990’s: indeed, one of the participants here explicitly refers to changes over the last twenty or thirty years. K1 and L1 were both seventeen years old at the time of recording, and K1, as an apprentice motor mechanic, was very aware of being a ‘woman in a man’s job’. As such, her experience testifies both to the expansion of opportunities for women in the second half of the twentieth century, and the extent to which attitudes are/ were still entrenched.

well that's what I get and people’ll say come into the garage and-that and say .. “do you get any stick working here?” and I'll say

“aye” and they’ll say “do you get any when you first worked here” “aye” I think I’m going to get it all the through do you know what I mean? It’s never going to stop

well so?
do you know what I mean

I mean it’s .. I cannot understand it me because people still like that I-mean they should eh .. they should know by now that .. I-mean man can do a woman’s job and a woman can do a [ii] man’s job

exactly it’s different now from what it was say

twenty thirty year ago whatever it is it’s a lot different all equal rights

The NECTE corpus also provides a record of cultural heritage. The TLS interviews provide a record of customs and pastimes which are disappearing. They included a question about childhood games, and the middle-aged and older speakers in particular provide names for and items used in games which are very regionally specific, but unknown to 21st-century youth. They recall childhood games such as hitchy-dabbers, Jack-shine-the-lantern, multikitty, relievo, spannies and tiggy-on-high, along with names for particular types of marbles, such as allies, liggies, muggers and spuggies.

3: NECTE as a record of dialect heritage

Of course, the main reason for creating the NECTE corpus was to preserve and make accessible records of the changing dialect of Tyneside in the second half of the twentieth century. The NECTE corpus provides a record of speakers of Tyneside English from different social groups and from two different periods. The TLS did not set
out to select ‘broad’ or ‘pure’ dialect speakers, aiming instead at a balanced sample which would reflect the variability of urban speech. However, the TLS tapes do preserve a record of traditional linguistic features which are much rarer in the 1994 PVC recordings. The extract below is from an interview with a forty-five year old working-class woman who had a minimal education, never moved away from the same area of Gateshead, rarely even visits Newcastle and finds the idea of going away for holidays laughable (‘we divn’t gan on holidays, man!’). This lack of exposure to other speech communities has resulted in her using a form of dialect which was conservative even in 1969.

This extract contains examples of conservative Tyneside dialect at all linguistic levels. In terms of pronunciation:

- The words about and house are pronounced with the monophthong (single vowel) [u:] as in food. In modern Tyneside, this is rarer, except in the use of Toon (‘town’) to refer to the City of Newcastle, or, more usually, Newcastle United Football Club.

- The pronunciation of first is identical to forced. This can still be heard on Tyneside, but it is rarer than it used to be.

In terms of grammar, one dialect feature to note is the lack of the plural marker –s in:

- and eh I lived in with my mother for not quite two year; but I lived in there for about oh .. eighteen or nineteen year

This is still typical of Tyneside and many other northern varieties of English, where the plural of nouns is not marked after numerals. After all, we already know that two and nineteen are more than one, so the plural marker would be redundant. In terms of dialect vocabulary, a short extract can only yield a few items, but varnigh meaning ‘very nearly’ is quite an archaic Tyneside word, and, whilst shifted is a Standard English word, the sense of ‘moved house’ is more local to Tyneside. In the course of preparing NECTE, we built up a glossary of local terms, initially as a guide to spelling for transcribers, but this constitutes an important resource in itself, and is available on the website. An extract from the glossary appears as Figure 3, below.

| Ally | marble |
| Ally oukin, - ookin | truce word in children's games |
Figure 3: extract from the NECTE glossary

We have so far concentrated on the TLS component of the NECTE corpus, since this is older and therefore more likely to contain the more archaic language normally associated with dialect heritage. The more recently-collected PVC corpus allows us to capture innovations in the language of Tyneside. One feature which has been widely discussed both in the media and in scholarly journals (Tagliamonti & Hudson 1999, Dailey-O’Cain 2000, Buchstaller 2006) is the use of *like* as a verb to introduce a quotation or reported speech, as in the following, from a seventeen-year-old, middle-class boy:

People says ‘Right, you’re buying us twenty tabs’. I was like ‘No’

We found that the only speakers in the whole corpus who used *like* in this way were the sixteen- and seventeen-year-olds in the PVC corpus, suggesting that this feature was a fairly recent innovation in Tyneside speech in 1994.

4. Conclusion

In this paper, we have attempted to demonstrate the ways in which the NECTE corpus preserves the non-tangible heritage of Tyneside, recording as it does the social history, culture and dialect of the region. Since the corpus became available it has attracted a good deal of interest from scholars in a number of disciplines as well as interested members of the public. We hope that readers will access the website (www.ncl.ac.uk/necte) and make use of this rich and valuable resource.
References


---

¹ This project was funded by the Arts and Humanities Research Council grant number RE11776