

## Introducing the Diachronic Electronic Corpus of Tyneside English (DECTE): Methods, Data and Applications

Karen Corrigan & Isabelle Buchstaller (*Newcastle University*)

The last decade has seen a move towards the use of ever larger corpora in linguistics, prompted by and, in turn, sparking the development of a number of publicly available on-line data-bases (i.e. COCA, NECTE, SCOTS). However, none of these provide access to longitudinal data permitting analyses of real time language change. The ONZE corpus (Gordon *et al.* 2007) remains the only dialectal English language dataset that provides access to recordings spanning more than 3 decades.

This presentation showcases a monitor corpus of dialectal Tyneside speech developed along similar lines to ONZE. Since 2006, researchers at Newcastle University, UK have been collecting an extensive data-base of local speech from informants with various age, gender and class profiles. Importantly, Newcastle University has already digitized and enhanced two older corpora of Tyneside speech, the 1960's *Tyneside Linguistic Survey* (TLS), and materials from the 1994 *Phonological Variation and Change in Contemporary Spoken English* (PVC), which have been consolidated as NECTE (<http://www.ncl.ac.uk/necte/> & Allen *et al.* 2007). The new monitor corpus (NECTE2), which augments NECTE, constitutes the *Diachronic Electronic Corpus of Tyneside English* (DECTE). As Table 1 shows, expanding NECTE thus makes DECTE a very rare example of regional and social linguistic data spanning five decades and offering a wealth of contemporary and diachronic interview data (transcripts and recordings) created using identical protocols and all publicly available on-line.

Table 1: Birthdates for the speakers in DECTE

<b>NECTE2</b> (2007-2009)	<b>Younger</b> 1967 - 1990	<b>Older</b> 1923 - 1966		
<b>PVC</b> (1994/5)	<b>NECTE</b>	<b>Younger</b> 1954 - 1977	<b>Older</b> 1911 - 1953	
<b>TLS</b> (1960s)			<b>Younger</b> 1925 - 1968	<b>Older</b> 1895 - 1934

In this presentation we will:

- Showcase DECTE, explaining its functionality and user-friendly interface;
- Report on two pilot studies of linguistic change in Tyneside English that take advantage of the full range of diachronic evidence now available.

### References

- Allen, W.H.A., Beal, J.C., Corrigan, K.P., Maguire, W. & Moisl, H.L. 'A Linguistic 'Time Capsule': The Newcastle Electronic Corpus of Tyneside English' in Beal, J.C., Corrigan, K.P. & Moisl, H.L. (eds.) *Creating and Digitizing Language Corpora: Vol.2, Diachronic Databases*, pp. 16-48. Houndmills: Palgrave.
- Gordon, E., Maclagan, M. & Hay, J. (2007) 'The ONZE Corpus', in Beal, J.C., Corrigan, K.P. & Moisl, H.L. (eds.), pp. 82-104.