

The curation of sound archives: the Dutch Dialect Database

Douwe Zeldenrust

Meertens Institute (Royal Netherlands Academy of Arts and Sciences)

The Netherlands

douwe.zeldenrust@meertens.knaw.nl

Introduction

At the first DH Benelux conference, papers and presentations concerning the digitization, curation and dissemination of digital objects were in a minority. Within this domain the interests in sound archives and the use of sound recordings as a resource were even more limited and restricted to for instance phonologists and musicologists. But in a recent article in *Digital Humanities Quarterly*, Anne Murray and Jared Wiercinsky placed this subject in a broader perspective. They state that the creation of new knowledge in the humanities 'depends not only on better understanding the role of sound in the work of humanities scholars, but also incorporating this knowledge into the design of sound archives' (Murray).

Use case: the Dutch Dialect Database

This paper focuses on the digitization, curation and design of sound archives. A use case will be presented: the project of the Dutch Dialect Database (NDB) of the Meertens Institute, Royal Netherlands Academy of Arts and Sciences (Oostendorp). The project started in 2009 and was presented at the 2014 DHBenelux conference as the 'Soundbites collection' (Zeldenrust). Since then, the digital sound archive has been renewed significantly. The latest design principles for web-based spoken word archives have been incorporated and extra information has been added. Currently the collection contains Dutch spoken in the Netherlands and also Dutch spoken outside the Netherlands, such as Dutch spoken in the USA and Dutch spoken in France (Archives). The datasets are CLARIN (Common Language Resources and Technology Infrastructure) compatible and are available via the CLARIN interface.

Conclusion

To conclude with, the paper will reflect on the creation, curation and dissemination of humanities digital resources. The combination of sound and written resources with specially designed interfaces and tools will bring new research possibilities. For instance, the availability the collection Dutch Spoken in the USA will open the way for pattern recognition within this dataset (Sijs).

References

- Murray and Jared Wiercinski (2014) 'A Design Methodology for Web-based Sound Archives', *Digital Humanities Quarterly*, Volume 8 Number 2.
- Oostendorp, M. van (2014), 'Phonological and phonetic databases at the Meertens Institute', *The Oxford Handbook of Corpus Phonology*. Eds. J. Durand & G. Kristoffersen. Oxford: OUP. 546-551.
- Sijs, N. van der. (2013) 'Op de schouders van onze voorgangers. Jo Daan en het Amerikaans-Nederlands', *Nieuw Letterkundig Magazijn* 31.1: 30-34.
- Zeldenrust, D.A. (2014) 'Access to data: the Soundbites collection of the Meertens Institute and a flexible approach to the curation and dissemination of humanities digital resources', Digital Humanities Benelux Conference (The Hague).

Archives

- The Archives of the Meertens Institute, Nederlands in Frankrijk, (1964 - 1966).
- The Archives of the Meertens Institute, Nederlands in de USA, collection Jo Daan, (1966).
- The Archives of the Meertens Institute, Nederlands in de USA, collection van Marle, (1994 -1998).

Websites

- www.clarin.eu (Accessed February 5, 2015).
- www.meertens.knaw.nl/ndb (Accessed February 5, 2015).