



ACCADEMIA NAZIONALE DEI LINCEI

CENTRO LINCEO INTERDISCIPLINARE «BENIAMINO SEGRE»

Lunedì 14 settembre 2015

Tavola Rotonda

Speech audio archives: preservation, restoration, annotation, aimed at supporting the linguistic analysis.

As is well-known, there are many audio corpora out there in Europe and USA, witnessing an extraordinary cultural and linguistic richness awaiting research and cultural heritage exploitation. One needs not spend a lot of words to underline the extreme importance of this diverse repository in terms of cultural heritage preservation. They are an endless mine of linguistic variability even within one and the same politically homogeneous territory. It is however equally well-known that these different corpora are based on a multiplicity of formats and stored in databases using incompatible standards, which make their interoperability impossible at the present stage. It is thus very important and urgent to develop intelligent tools to foster the attainment of such a crucial goal.

During the past 50 years many resources have been collected and recorded by several national research centers (mostly dedicated to dialectological and sociolinguistic studies). They witness language usage by speakers differing in geographical origin, age and educational status. In quite a few cases, these materials constitute the living evidence of a language state that has now disappeared.

The European scientific community has in recent times turned more sensitive to problems relating to the maintenance and exploitation of oral archives witnessing important cultural layers of language diversity. But if we do not act now this treasure may be lost forever. The reason is simple: a great deal of the most valuable materials have been collected time back, and are still preserved on obsolete types of support, deemed to perish forever before long. This endangers their preservation and cramps their consultation.

In order to make these cultural assets available to a trans-national circulation, the speech materials have to be annotated, supplied with metadata for on-line searching queries, and – if necessary – restored. We need thus an interdisciplinary approach developing strong synergies among many disciplines.

14,20 Saluto del Direttore del Centro Linceo Interdisciplinare “Beniamino Segre”, Tito Orlandi

14,30 Amedeo De Dominicis (Centro Linceo Interdisciplinare): *Opening*

14,50 Sergio Canazza - Giovanni De Poli (Dept. of Information Engineering, University of Padova): *A Systemic Approach to the Preservation of Speech Audio Documents: Methodology and Software Tools*

15,40 Maristella Agosti - Nicola Ferro (Dept. of Information Engineering, University of Padova): *Digital Annotations: Why, What, and How*

16,30 Coffee break

17,00 Brigitte Bigi - Grégoire de Montcheuil (Laboratoire Parole et Langage, Aix-en-Provence): *Methodology and software for Semi-Automatic multi-domain annotations: annotating, exploring and sharing data*

17,50 Bernard Bel (Laboratoire Parole et Langage, Aix-en-Provence): *The OAIS model in Digital Humanities: feedback from archival practice in linguistics*

18,40 Closing