



PROGRESS OVERVIEW FOR THE SAM PROJECT

A Fourcin and Sam Partnership

University College London

The multi-lingual ESPRIT Speech Assessment Methodology Project (2589) started in March 1989. Eight countries collaborate in the work; six from the European community, two from EFTA.

This contribution is concerned to provide a first overview of the practical achievements arising from the European collaboration made possible by the coherent adoption of standard sets of protocols, methods, hardware and software in the field of speech assessment methodology. Detailed reports of particular aspects of the work are presented at the conference by members of the SAM consortium.

Databases are generated according to sets of common criteria, ranging from the specification of digital formats, standard media and labelling structures, to the use of common techniques of annotation and linguistic organisation. At the level of annotation, a European computer compatible Standard Phonetic Alphabet, SPA, is now being applied in all of the language environments. Databases, annotated by using the SAM phonetic alphabet, have been generated both at the segmental and the sentence levels of definition. Manual labelling methods have been employed initially as a precursor both to the immediate use of this material for the assessment of recogniser performance and to the development of semi-automatic methods to substantially reduce the labour of future database production.

An especial feature of the work is necessarily cross-linguistic, namely the assessment of recognisers in different language environments. Members of the SAM partnership are employing the databases generated within the group

so that common target recognition equipment can be assessed with the use of language material from all the partner countries. Phonetically and linguistically consistent material makes it possible to gain an insight into the comparative performance of a particular system in different language environments, and its controlled structure provides the basis for a diagnostic assessment of fallibility, which is otherwise difficult to obtain in such a comprehensive fashion.

Output assessment of synthesiser performance is similarly approached on the basis of the multi-lingual use of evaluation techniques, using strictly cross comparable linguistically organised protocols. The use here too of reference material in the different language environments makes it possible to provide a consistent foundation for comparing the output of different synthesizers with the characteristics of natural speech in an objective and quantifiable way.

Basic to the generation of labelled databases, and the development of techniques for input and output assessment is the use not only of common approaches to data and assessment structures but also of common workstation facilities. These make it possible for the same operations of data generation and control associated with largely identical assessment procedures to be employed in different laboratories.