

was no way to use (say) the vowels of one ISM and the  
consonants of another ISM. In a more flexible tuning-in  
scheme, an individual speaker-mixture can be selected  
for each phoneme independently, concept one  
to the speaker-adaptive  
this approach

#### 4. EXPERIMENTAL RESULTS

##### Speech Data

The 3 male and 3 female speakers listed in Table 1 were used to train and test the four different architectures. There

## 2 INTRODUCTION

Dependent (CMU Alph Data) |  
n 40/200 test sentences/words  
or  
ON

t tasks. For ~~speaker~~  
, we used the CMU "Alph - Data",  
with 1000 sentences (i. e. continuously spelled strings  
of letters in our context) fr  
fema

# MULTI-SPEAKER/SPEAKER-INDEPENDENT ARCHITECTURES FOR THE MULTI-STATE TIME DELAY NEURAL NETWORK

*Hermann Hild and Alex Waibel*

School of Computer Science  
Carnegie Mellon University  
Pittsburgh, PA 15213-3891, USA

## **ABSTRACT**

In this paper we present an improved Multi-State  
Time Delay Neural Network (MS-TDNN) f  
i n d e p e n d e n t ,