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, conceptione to the speaker-adaptive this appro

4. EXPERIMENTAL RESULTS

4.1 CMU Alph Data

Multi-Speaker. The 3 male and 3 female speakers listed in Table 1 were used to train and test the four different architectures. The results (%correct, ex words, averaged over all 6 sp table 2. I

2. IMPROVED CONTINUOUS RESULTS

Speaker Dependent (CMU Alph Data)			
600/3000 train, 400/2000 test sentences/words			
speaker	SPHINX[2]	MS-TDNN[2]	our MS-TDNN
mjmt	96.0	97.5	**98.5**
mdbs	83.9	89.7	**91.1**
maem	–	–	**94.6**
fcaw	–	–	**98.8**
flgt	–	–	**86.9**
fee	–	–	**91.0**

Speaker Independent (Res. Manag. Sp	
109/11000 train, 11/900)	
SPHINX[6]	
	+ Sen
88.7	9

Table

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) for speaker-
independent, connected letter recognition which out
performs an HMM based system (SPH
vious MS-TDNNs [2], and
tectures