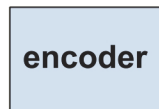


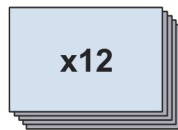


w2v

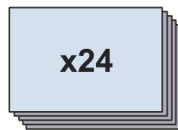


optimal layer selection

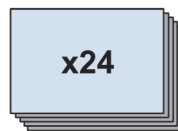
vqw2v



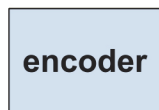
w2v2



XLSR

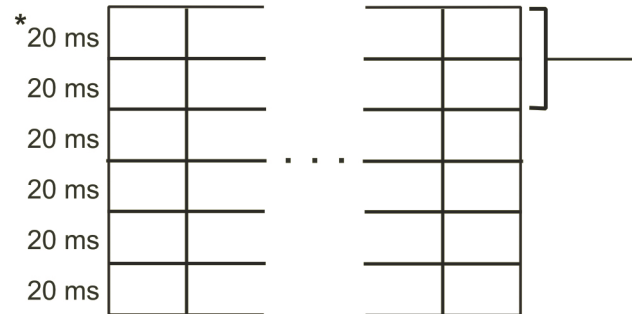


DeCoAR



MFCC

Methods

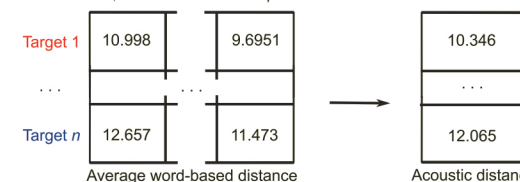
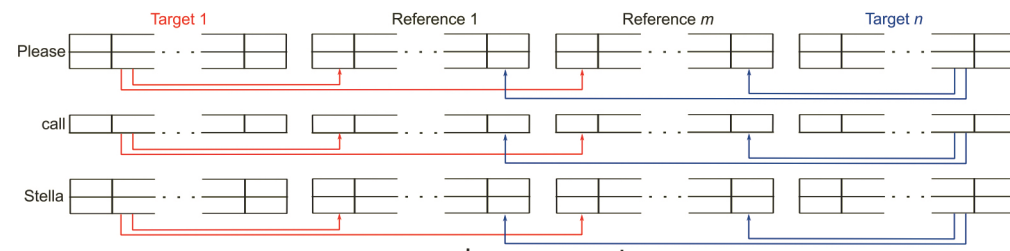


Dimensions dependent on algorithm

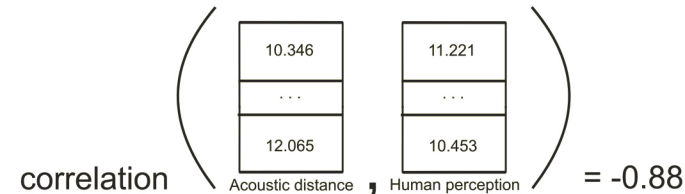
Forced alignment

	Start	End
PLEASE	0.4800	0.8800
CALL	0.9802	1.0888
STELLA	1.2890	1.4899
ASK	2.1670	2.4868
HER	2.4868	2.5965
TO	2.5965	...
...		

Feature extraction



Dynamic time warping



Correlation