**Encoder Output - Layer 1** 20.0 120 -17.5 100 -Frequency dimension 15.0 80 -12.5 10.0 60 -7.5 40 -5.0 20 -2.5 0 50 150 200 0 100 250 300 **Time dimension**