Normalized confusion matrix (mistral-large-latest) action -0.29 0.03 0.00 0.03 0.01 0.01 0.22 0.12 0.00 0.08 0.13 0.03 0.05 aggression -0.05 0.68 0.00 0.00 0.00 0.00 0.00 0.00 0.05 0.23 0.00 0.00 delivery -0.12 0.03 0.02 0.44 0.02 0.02 0.05 0.00 0.11 0.16 0.00 0.00 0.05 interaction -0.00 0.06 0.00 0.00 0.00 0.00 <mark>0.90</mark> 0.00 0.00 0.00 0.03 0.00 0.00 movement -0.06 0.00 0.00 0.00 0.03 0.14 0.08 0.61 0.00 0.00 0.03 0.00 0.06 music $-0.01\ 0.00\ 0.02\ 0.00\ 0.01\ 0.00\ 0.00\ 0.00\ 0.87\ 0.05\ 0.02\ 0.02\ 0.00$ narration $\pm 0.03 \ 0.06 \ 0.00 \ 0.00 \ 0.00 \ 0.03 \ 0.03 \ 0.00 \ 0.61 \ 0.08 \ 0.14 \ 0.00$ setting -0.02 0.00 0.02 0.00 0.02 0.04 0.00 0.12 0.00 0.07 0.04 0.68 0.00

toward $+0.00\,0.00\,0.01\,0.01\,0.00\,0.00\,0.01\,0.00\,0.00\,0.00\,0.00\,0.00\,0.00$ action movement interaction MUSIC apaite delivery

True label

narration Predicted label