RQ1: Channel Decomposition — Text-Image vs Text-Audio Contributions