IETF 124, 6 November 2025

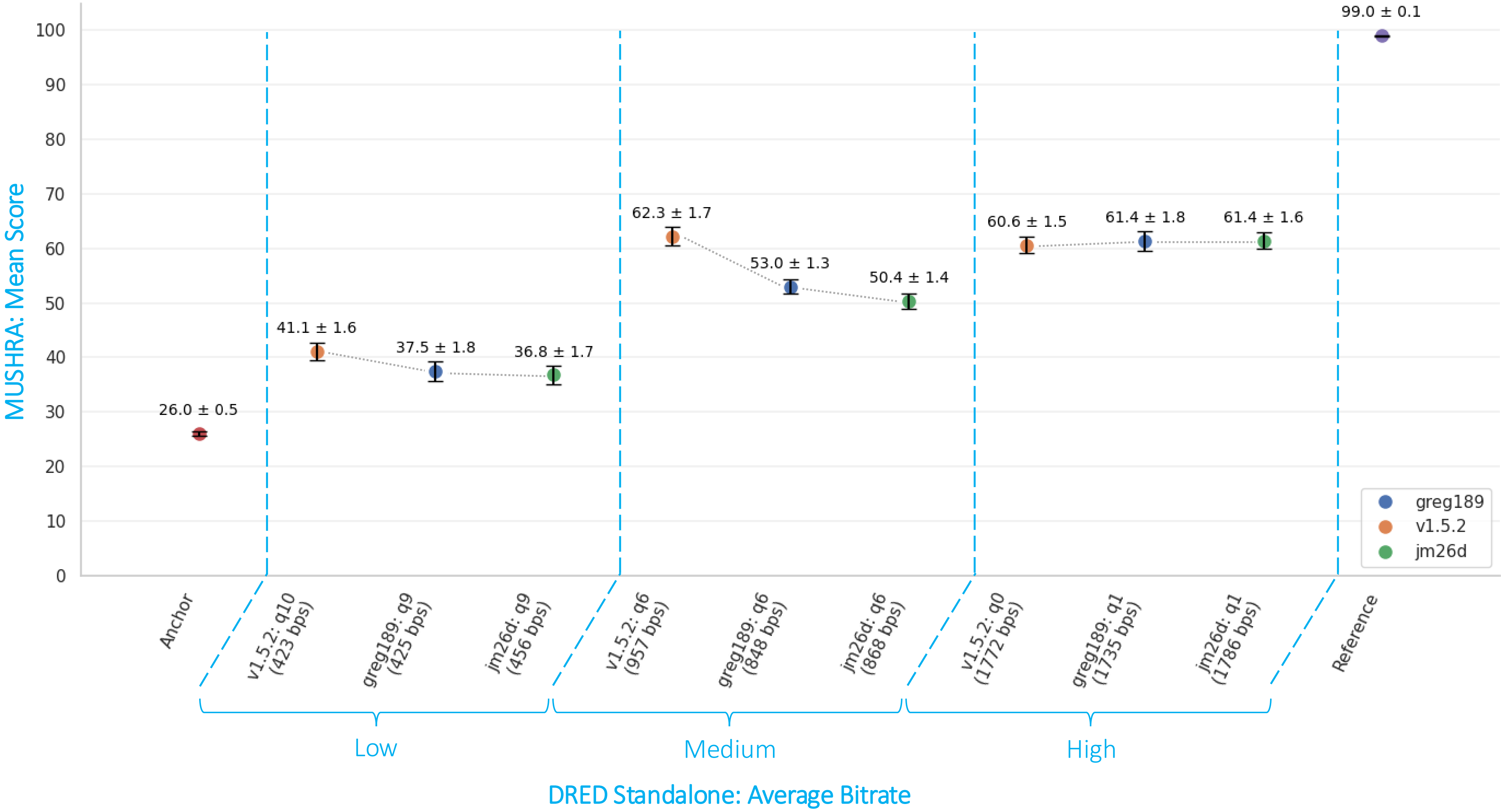
draft-lechler-mlcodec-test-battery-01

A Comprehensive Listening Test Battery
for ML Audio Codec Evaluation

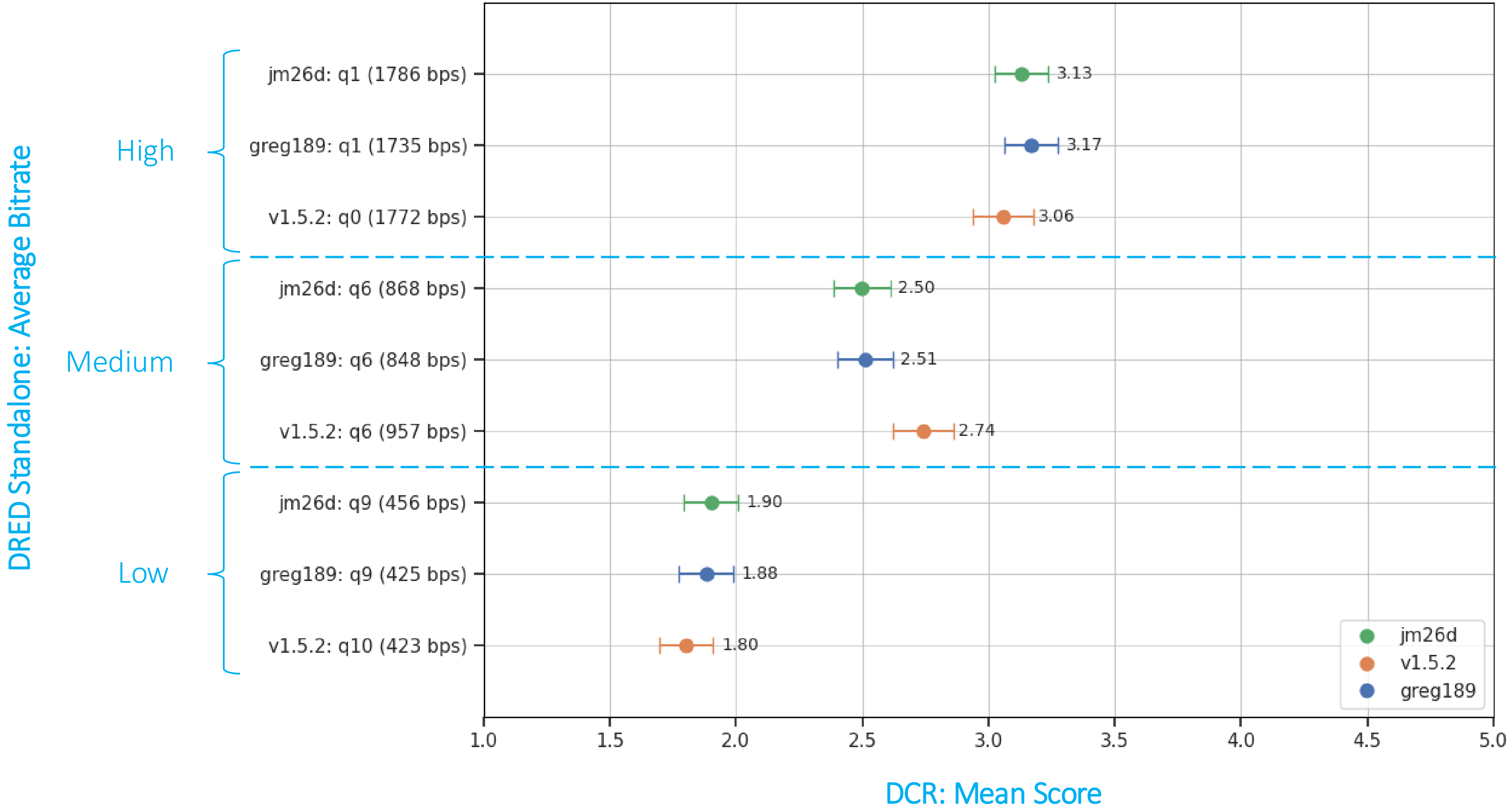
Laura Lechler & Kamil Wojcicki (Cisco Systems)

Evaluation of DRED improvement candidates

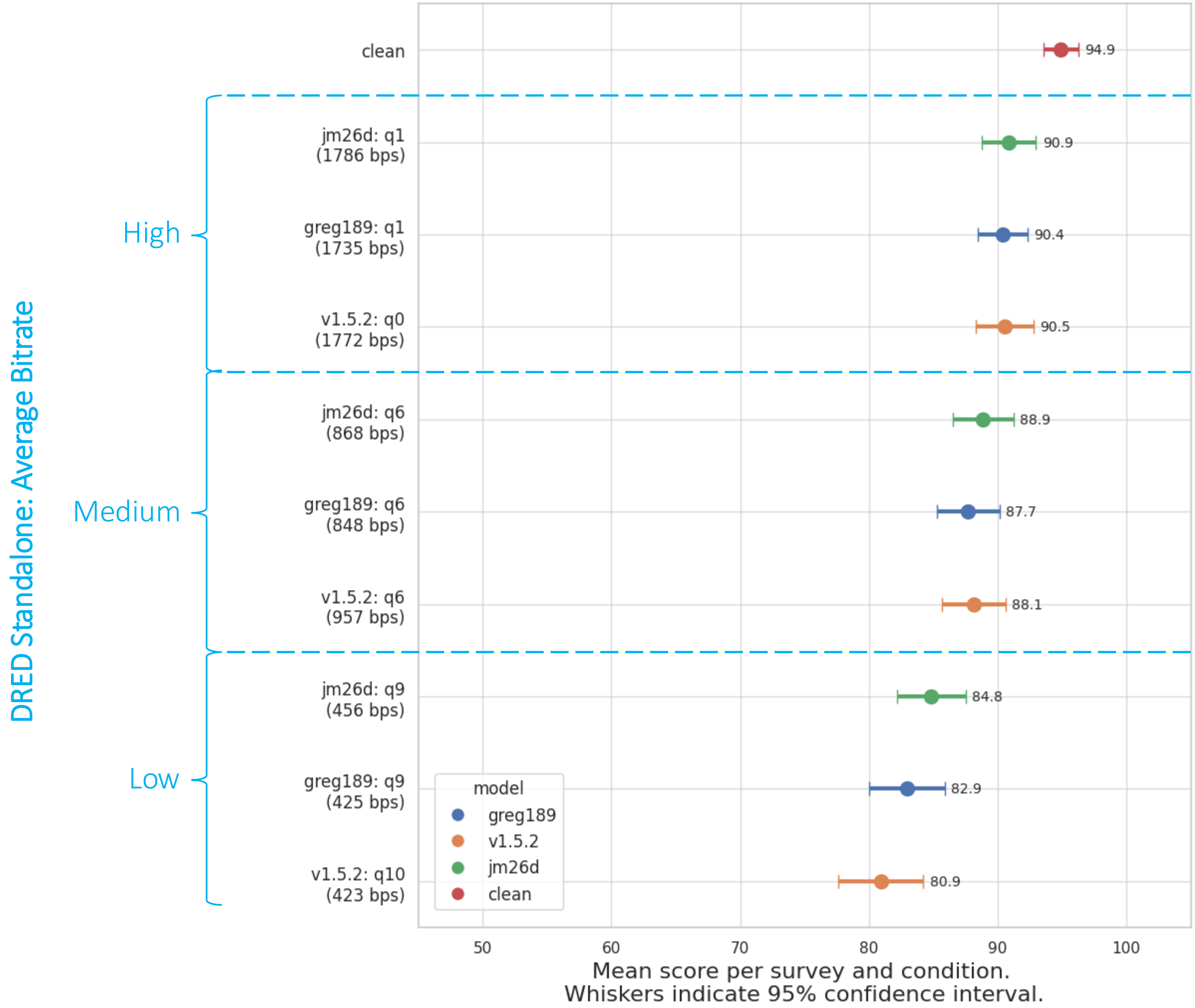
Quality in clean conditions



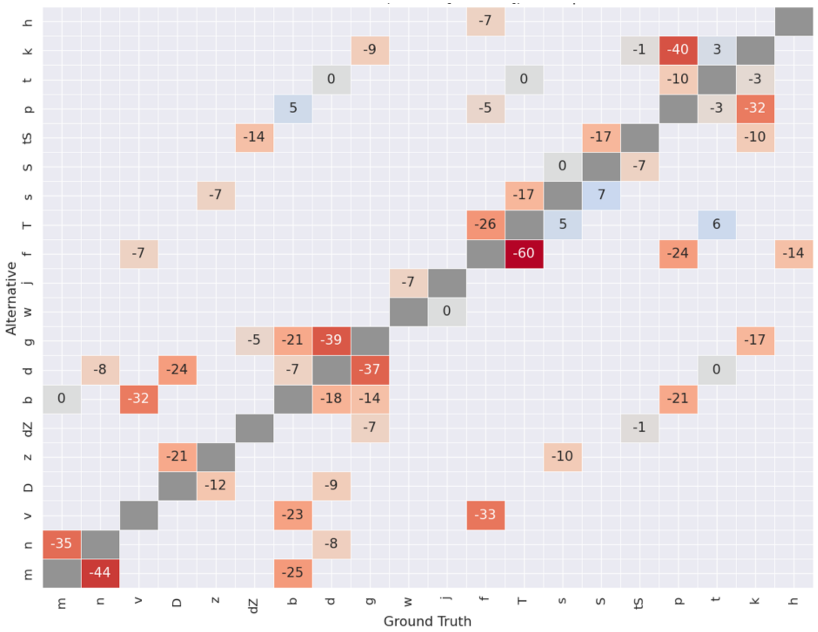
Quality in light real-world noise and reverb



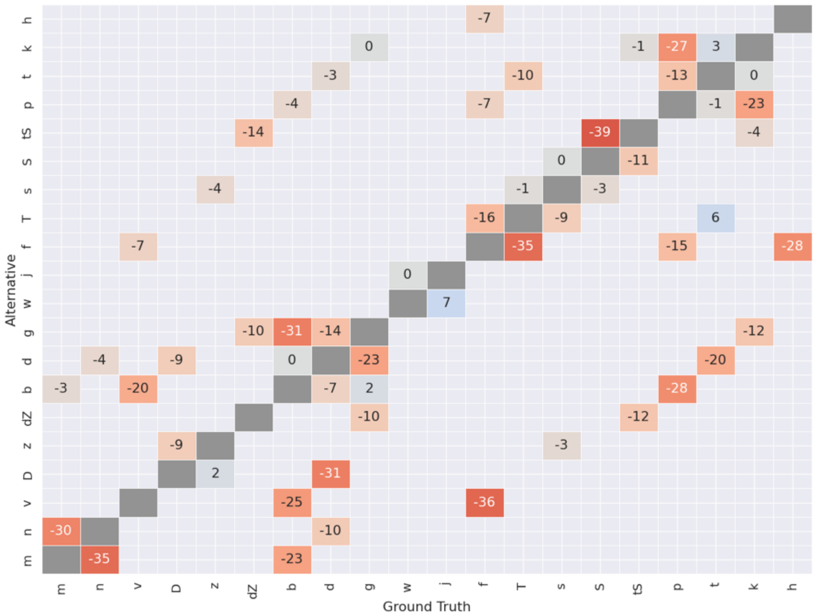
Intelligibility in clean conditions



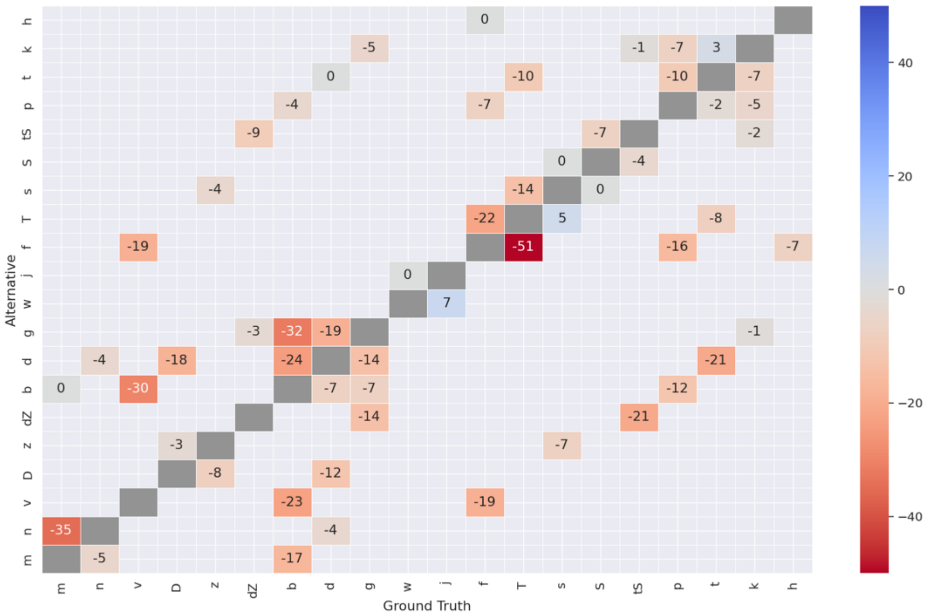
Intelligibility Deep Dive:
Delta phoneme-score matrices (low bitrates)



v1.5.2: q10 (423 bps)
vs.
clean

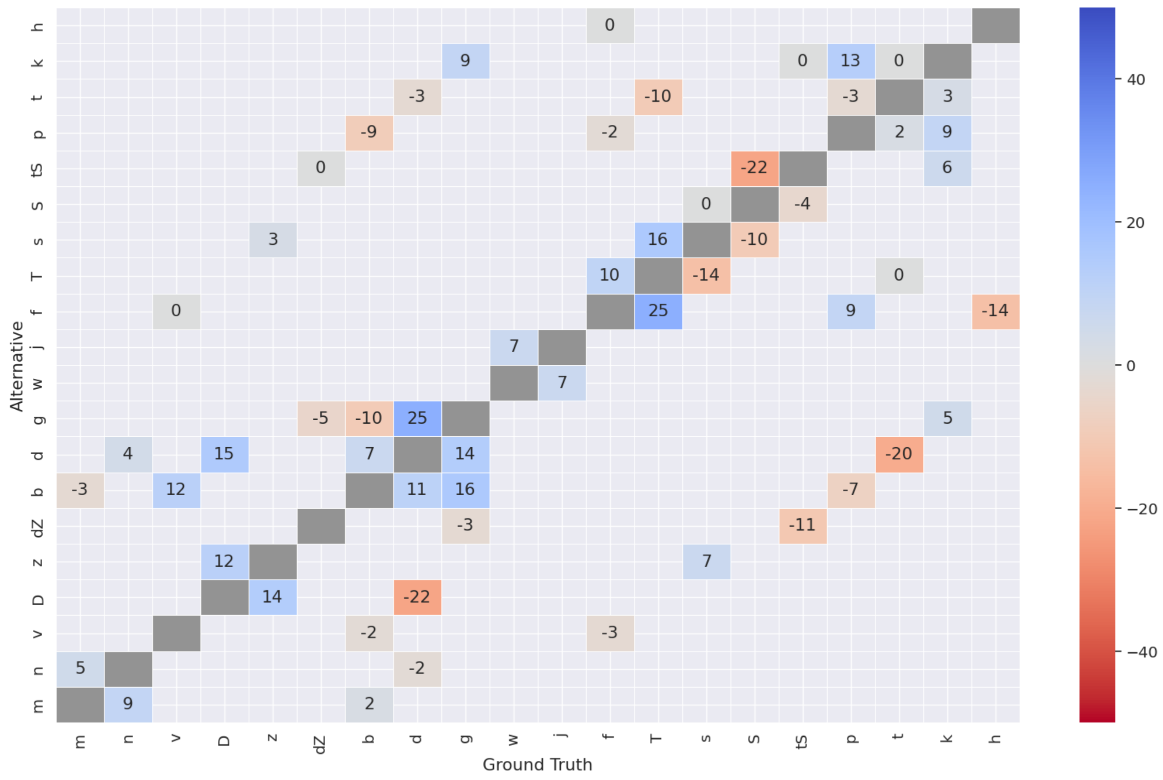


greg189: q9 (425 bps)
vs.
clean

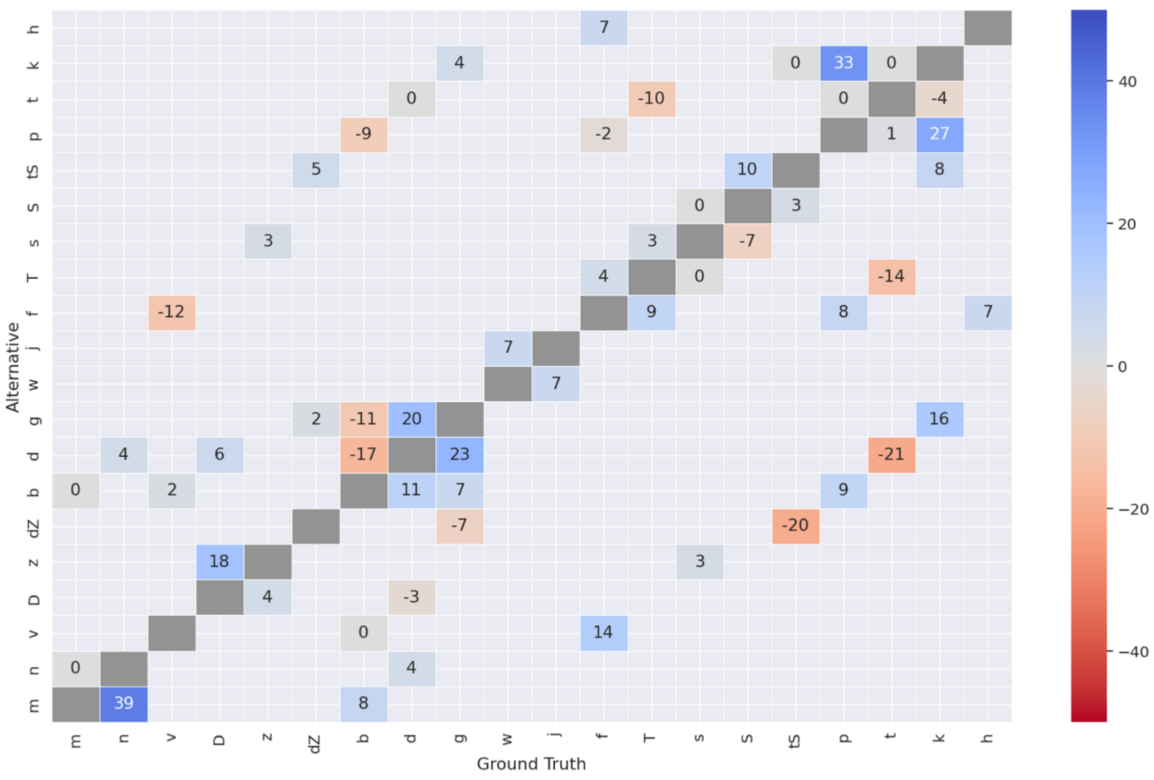


jm26d: q9 (456 bps)
vs.
clean

Intelligibility Deep Dive:
Delta phoneme-score matrices (low bitrates)



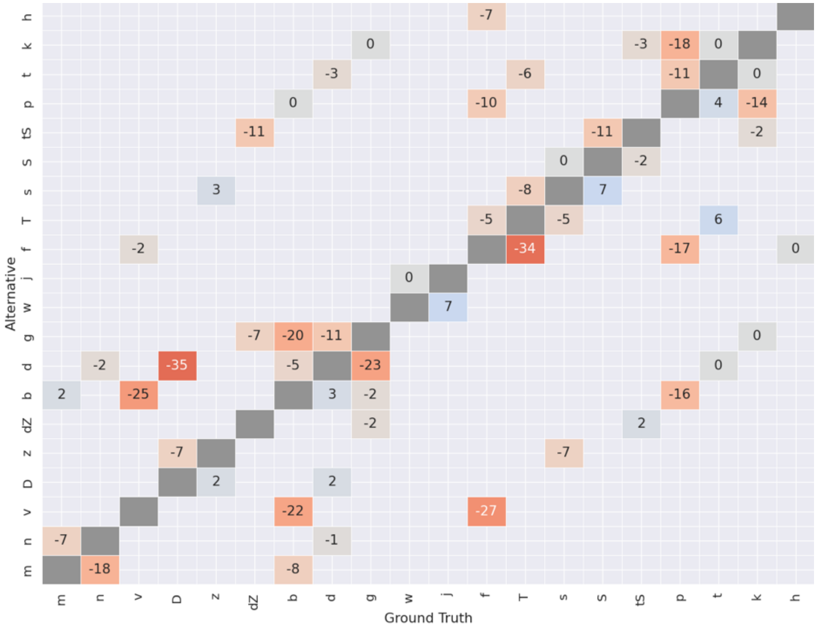
greg189: q9 (425 bps)
vs.
v1.5.2: q10 (423 bps)



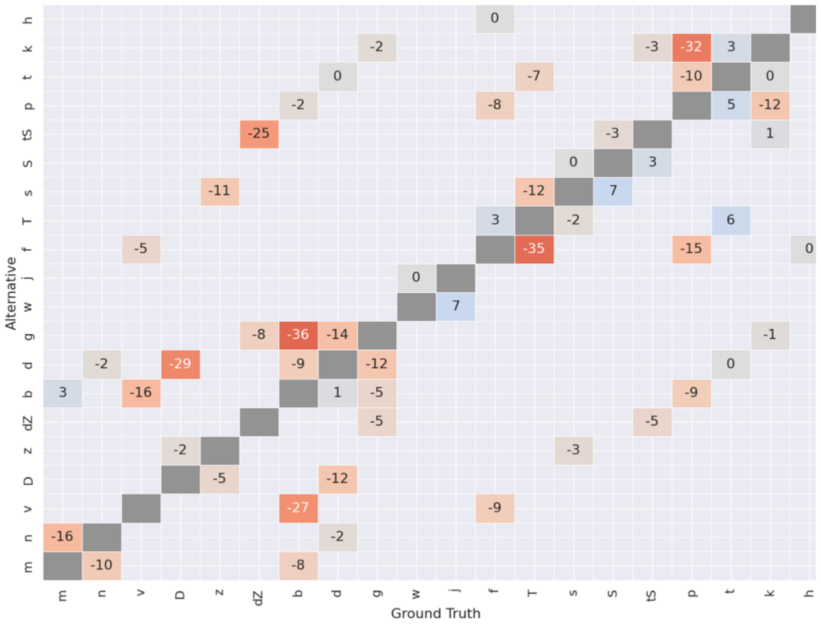
jm26d: q9 (456 bps)
vs.
v1.5.2: q10 (423 bps)

Appendix

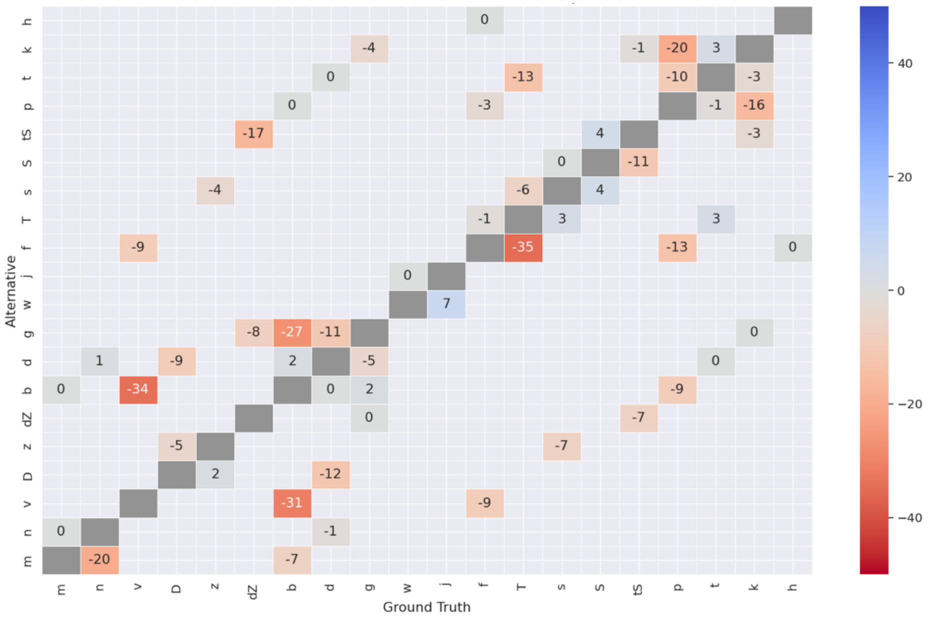
Intelligibility Deep Dive:
Delta phoneme-score matrices (medium bitrates)



v1.5.2: q6 (957 bps)
vs.
clean

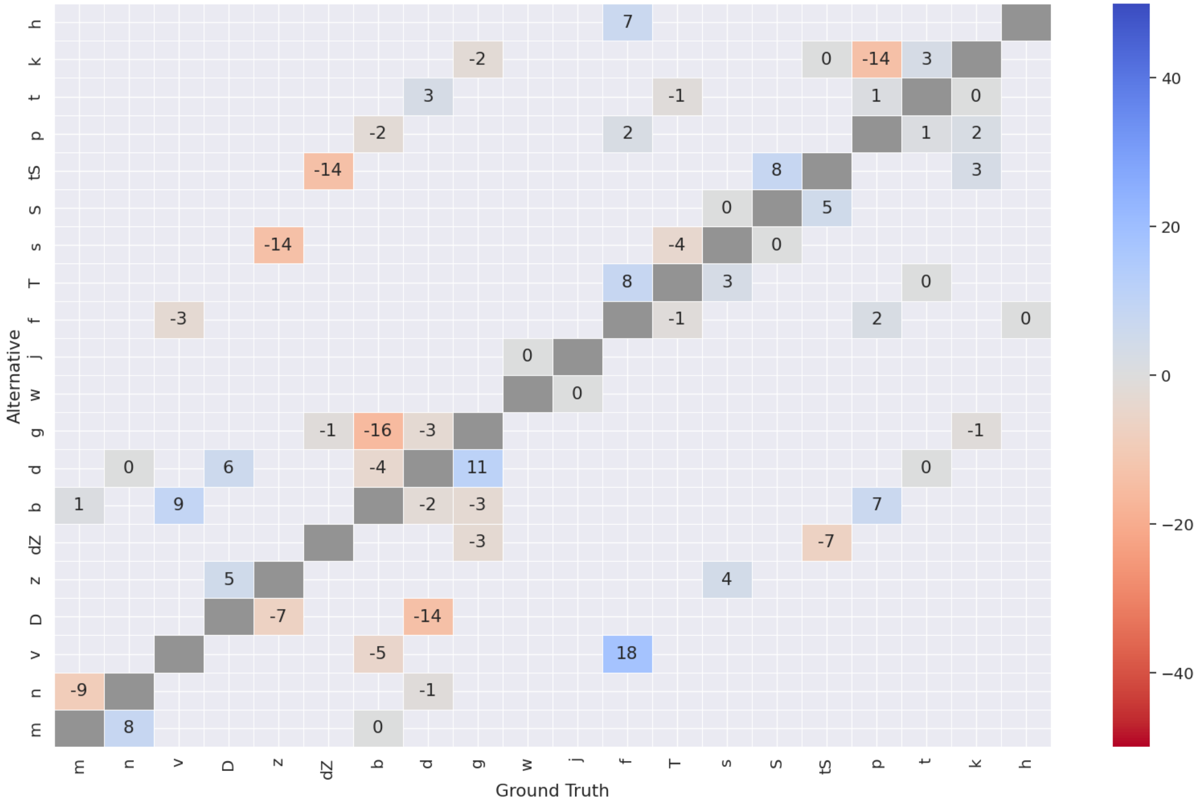


greg189: q6 (848 bps)
vs.
clean

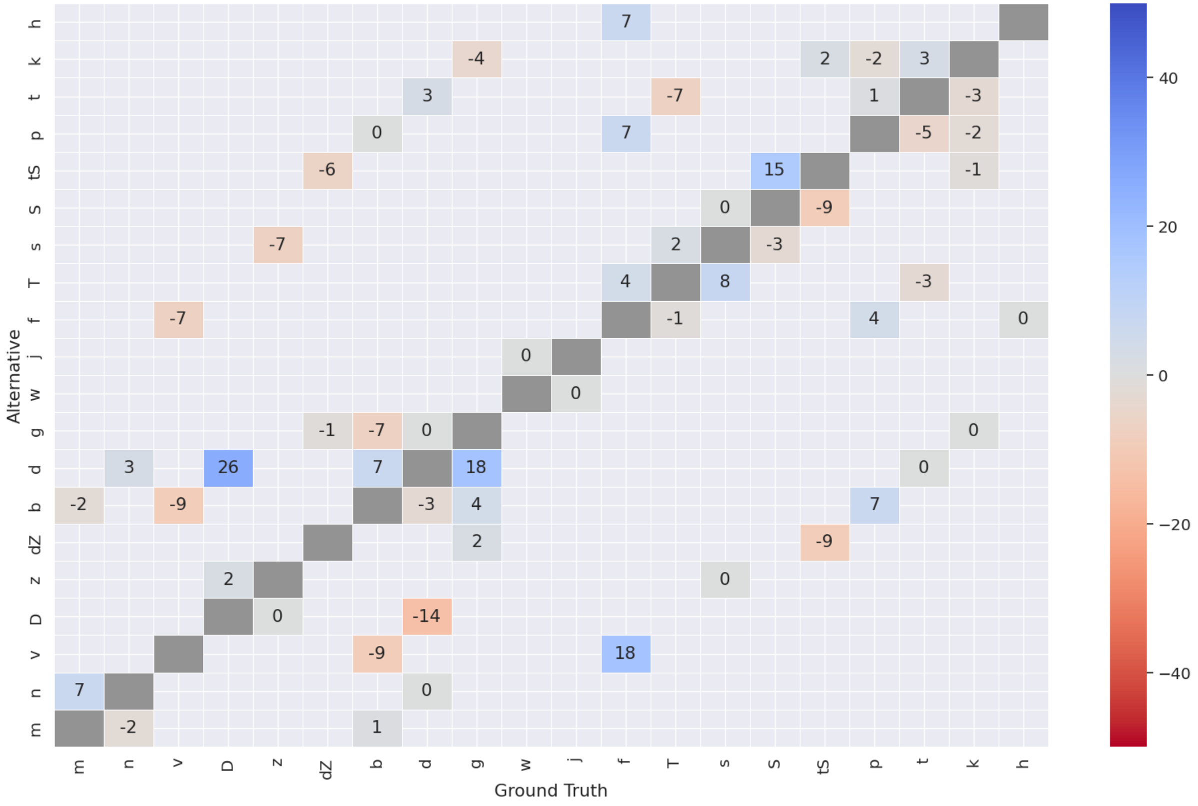


jm26d: q6 (868 bps)
vs.
clean

Intelligibility Deep Dive:
Delta phoneme-score matrices (medium bitrates)

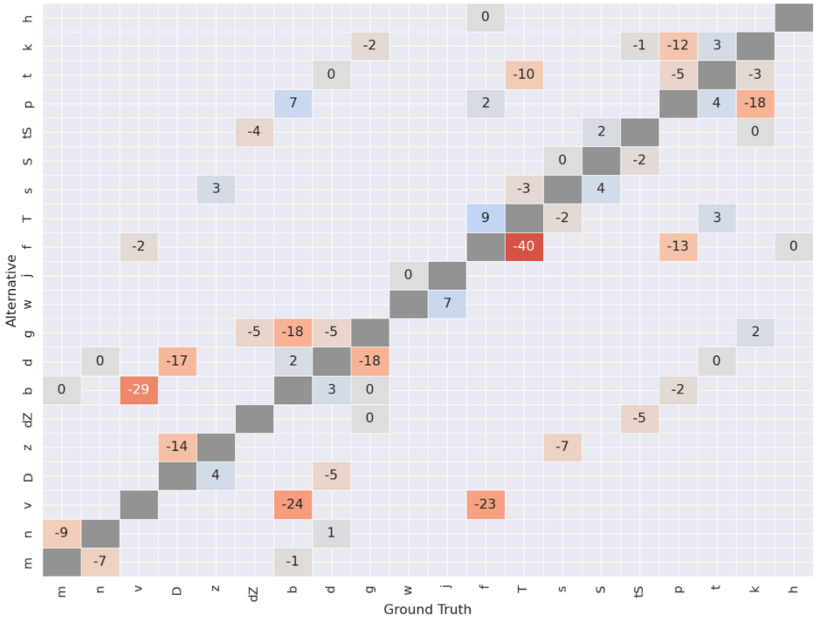


greg189: q6 (848 bps)
vs.
v1.5.2: q6 (957 bps)

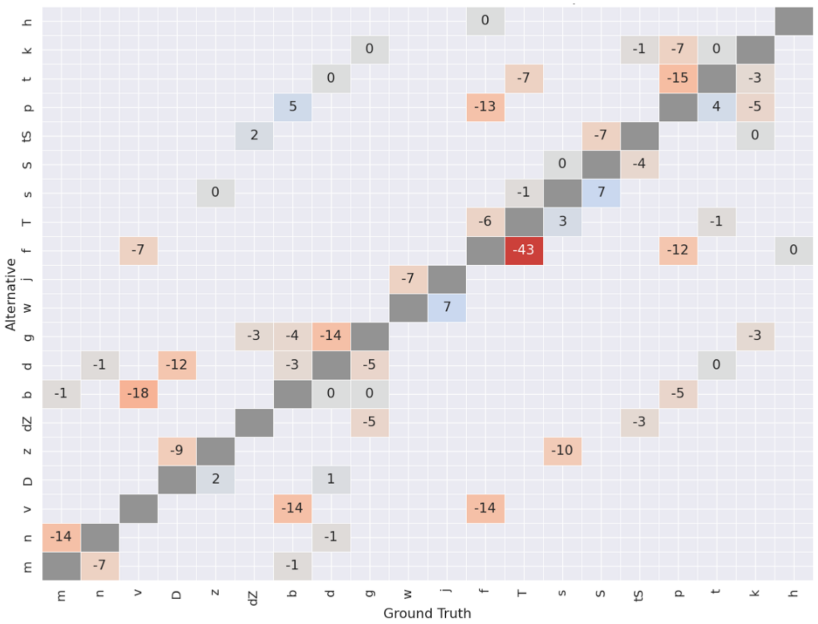


jm26d: q6 (868 bps)
vs.
v1.5.2: q6 (957 bps)

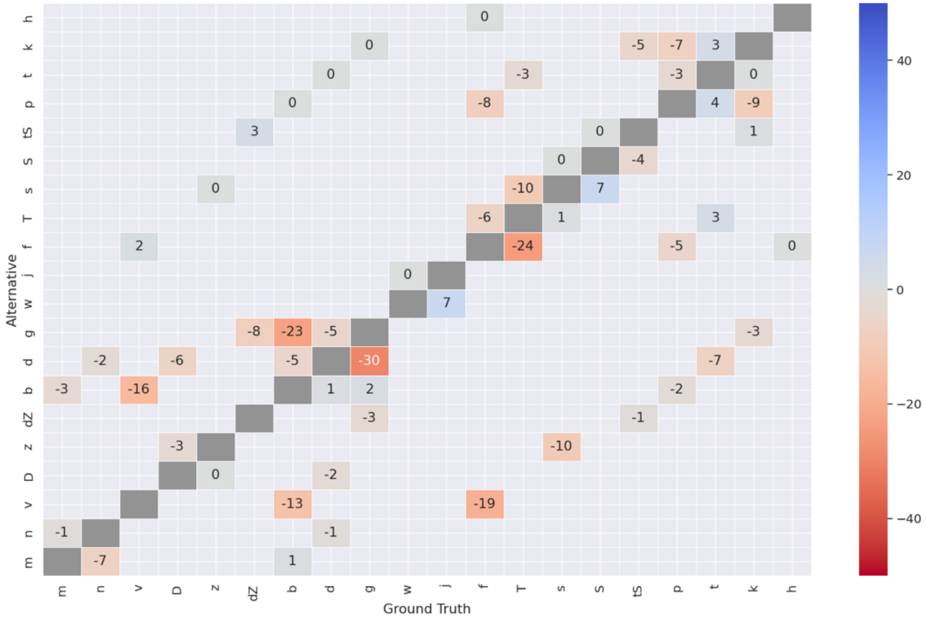
Intelligibility Deep Dive:
Delta phoneme-score matrices (high bitrates)



v1.5.2: q0 (1772 bps)
vs.
clean

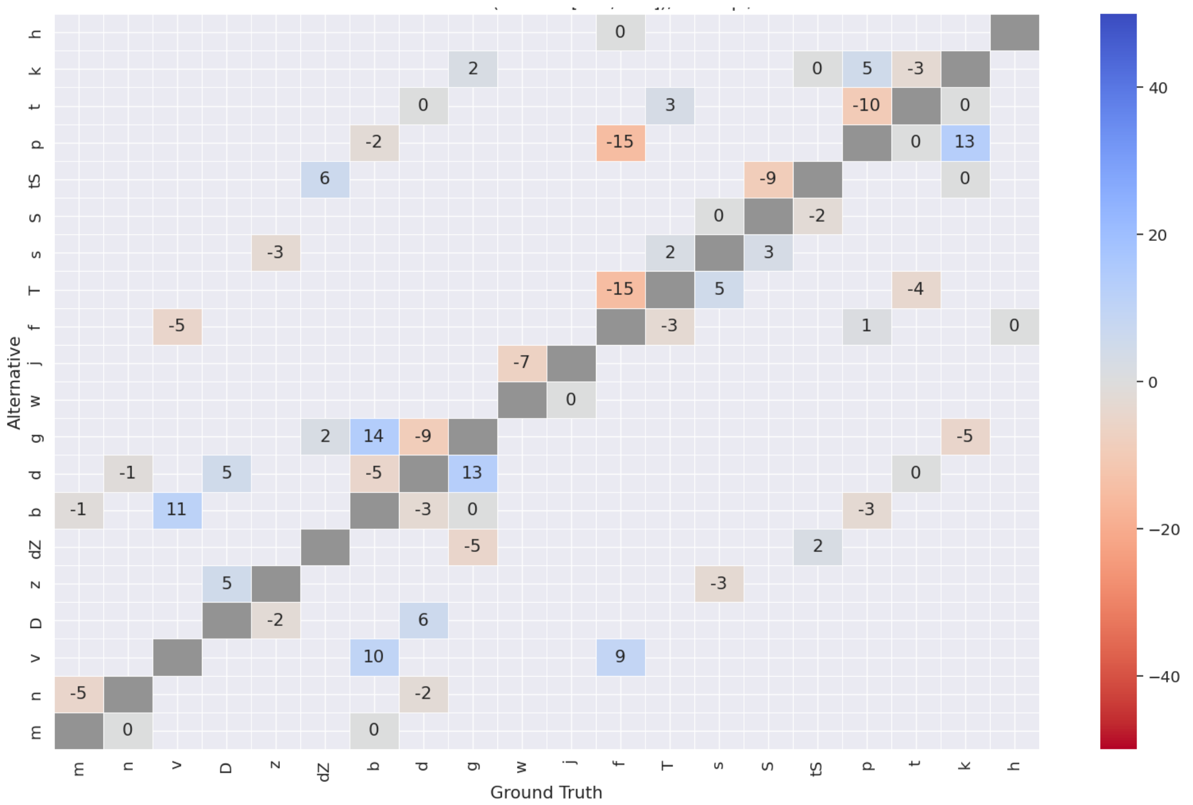


greg189: q1 (1735 bps)
vs.
clean

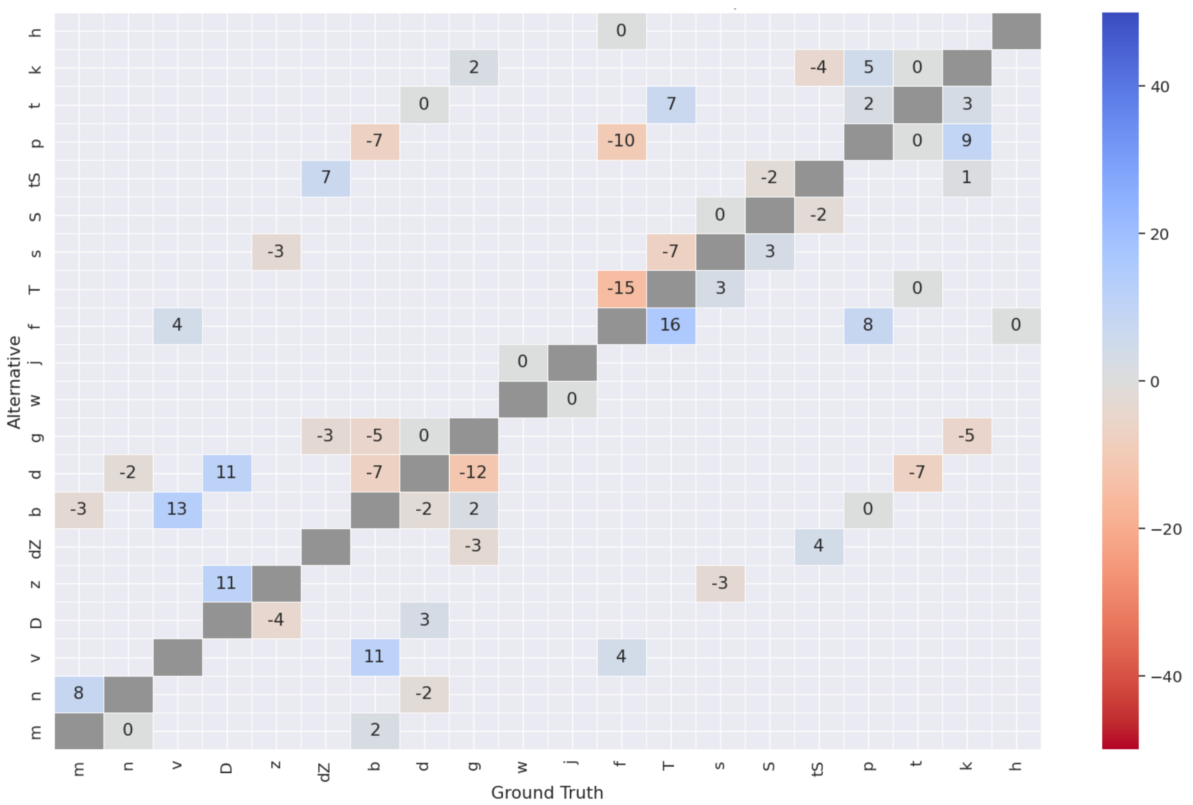


jm26d: q1 (1786 bps)
vs.
clean

Intelligibility Deep Dive: Delta phoneme-score matrices (high bitrates)



greg189: q1 (1735 bps)
vs.
v1.5.2: q0 (1772 bps)



jm26d: q1 (1786 bps)
vs.
v1.5.2: q0 (1772 bps)