

Daily Progress Report – 2025-09-05 (Fri)

Prepared by: **Tiong-Sik Ng**

English

1. Summary of Accomplishments

- Successfully ported everything to the company's AWS account, with testing confirming that behavior, latency, and barge-in are working as intended.
- Verified that permission issues were resolved, allowing end-to-end call handling to proceed smoothly.
- Conducted manual call tests after confirming stability in the pipeline.
- Environment is now stable enough for iterative improvements on audio quality and conversational naturalness.

2. Issues & Risks

- **Muffled speech quality:** Current audio output sounds compressed or dull, likely due to encoding choices or double conversion. Needs fixing before declaring the system production-ready.
- **Robotic tone:** Generated speech still feels synthetic. Improvement will require either fine-tuning or switching to a more natural model.

3. Next Steps

1. Fix the muffled voice output by ensuring PCM encoding and proper preprocessing before Connect playback.
 2. Categorize filler speech clips and finalize the prompt library for production use.
 3. Explore fine-tuning on **Korean datasets** (e.g., KSponspeech) to improve prosody and naturalness.
 4. Continue stability checks in the company environment to ensure no hidden dependencies remain.
-

한국어

1. 금일 수행 사항

- 회사 AWS 계정으로 모든 환경을 성공적으로 이관하였으며, 동작, 지연(latency), 바지인 기능이 정상 작동함을 테스트로 확인.

- 권한 문제를 해결하여 전체 콜 처리 플로우가 안정적으로 진행됨을 확인.
- 파이프라인 안정성 확보 후, 직접 콜 테스트까지 수행 완료.
- 현재 환경은 오디오 품질 및 대화 자연스러움 개선을 위한 반복 작업이 가능한 상태임.

2. 이슈 & 리스크

- **음성이 다소 먹먹하게 들림:** 인코딩 방식 또는 이중 변환 문제로 추정. Connect 재생 전 PCM 인코딩 확인 필요.
- **로봇 같은 억양:** 발화가 여전히 인위적으로 들림. 한국어 데이터셋 기반의 파인튜닝 필요.

3. 다음 단계

1. 음성 출력의 먹먹한 문제를 해결하여 PCM 기반으로 재인코딩 및 전처리 과정 개선.
2. Filler speech 를 카테고리별로 정리하고, 프로덕션용 프롬프트 라이브러리를 최종 확정.
3. **KSponspeech** 등 한국어 데이터셋을 활용한 파인튜닝으로 억양과 자연스러움 개선.
4. 회사 계정 환경에서 추가적인 숨은 의존성이 없도록 지속 점검.