FILE 1 – `family_voice_bridge_product_spec.md`

Document Meta

- **Product**: Family Voice Bridge
- **Owner**: Paloma (Builder)
- **Version**: v0.2
- **Status**: Draft – MVP spec
- **Last updated**: 2025-12-05 (started on 2nd)

---

1. TL;DR

Family Voice Bridge is a **mobile-first web app** that:

- Lets **Mom (Chinese-speaking)** press one big button to speak Chinese → plays **natural English** audio to the English speaker.
- Listens to the **English reply** and plays back **Chinese** to Mom.
- Quietly logs the conversation turns, auto-tags them by "situation / 场景" **(deprioritized)**, and generates a **daily recap + key phrases** so Mom can review and learn over time.

MVP is a **simple, two-turn session** (Mom → reply → done), plus a **Daily Vocabulary and a Daily Learning** screen.

---

2. Background & Problem

Context

- Mom (60), native Chinese, limited English, visiting US for 5.5 months to help with your newborn.
- Home environment: kitchen, baby care, food, postpartum support.
- Devices: mainly Mom's Samsung Galaxy S21 (Android), stable home Wi-Fi.

Core problems

1. **Friction in conversations**
    - Mom has to juggle Google Translate, Perplexity, etc.
    - Conversations stall; she often says less than she wants to.
2. **Lost learning opportunities**
    - Useful daily phrases are not captured or revisited.
    - Her English exposure doesn't compound into learning.
3. **Emotional distance**
    - She can feel like an outsider or "less helpful" in the household.
    - This discourages her from initiating conversations.

Product vision

Build a household “supermind”: a single web app that acts as **translator, memory, and tutor** for the family’s real-life multilingual conversations (can teach husband Chinese as well).

---

3. Goals, Non-Goals & Success Metrics

3.1 Goals (MVP)

1. **Low-friction communication**
    - Mom can comfortably start and finish short bilingual conversations without switching apps or typing.
2. **Daily language exposure & reflection**
    - Each day’s conversations are captured and distilled into a few meaningful phrases and topics for Mom to revisit.
3. **Extensible foundation**
    - Codebase and data model make it easy to add:
        - More languages (Bangla),
        - Simple quizzes,
        - Additional learner profiles (dad, husband).

3.2 Non-Goals (MVP)

- Adaptive/spaced-repetition learning paths.
- Rich user management, auth, or multi-tenant SaaS polish.
- High-latency streaming ASR/TTS; we focus on **discrete, short turns**.
- Complex analytics dashboards or admin console.

3.3 OKRs (MVP)

Objective 1 – Enable smooth, low-friction daily communication

- KR1.1: **≥ 80%** of translation turns finish (speech played back) within **≤ 4 seconds** of speaker finishing.
- KR1.2: Mom uses the app for **≥ 5 sessions/day** on **≥ 5 days/week**.

Objective 2 – Turn daily life into a gentle learning loop

- KR2.1: Daily review shows **≥ 1 summary** + **5 phrases** for any day with ≥ 3 sessions.
- KR2.2: Mom listens to at least **3 of 5 daily phrases** on **≥ 60%** of days.
- KR2.3: Over 5.5 months, the system surfaces **≥ 50 distinct useful phrases** tied to meaningful situations (baby, kitchen, postpartum, etc.).

Objective 3 – Build a reliable, extensible “household supermind”

- KR3.1: **≥ 99% uptime** during waking hours (7am–11pm local).
- KR3.2: **≥ 95%** of turns are logged with a valid situation tag (auto-classification succeeds).
- KR3.3: Adding Bangla support or a basic quiz mode, or a second learner profile is feasible in **≤ 1 week** of focused dev.

---

4. Users & Personas

Primary User – Mom

- 60, native Chinese speaker, limited English.
- Uses Samsung Android phone on home Wi-Fi.
- Loves learning, but lacks companionship and guidance.
- Core needs:
    - Talk to son-in-law, mother-in-law about baby and home.
    - Collaborate with English-speakers in care of her daughter and granddaughter, chores.
    - Learn practical English phrases over time.

Secondary Users

- **You (bilingual Chinese/English)**
    - Orchestrates conversations, occasionally checks logs and daily reviews.
    - May use Daily Review to help Mom practice.
- **Husband & Mother-in-law (English/Bangla)**
    - Participate by replying in English.
    - Want low mental load: they just talk; app handles translation.

Future User

- **Child (6–12 months +)**
    - Later: use real conversations as a language-learning hub in Chinese/English/Bangla.

---

5. Scope & Features (MVP)

5.1 Feature 1: One-Turn Voice Translation (Mom ↔ English Speaker)

- Mobile-first web app with a **single large "Speak"** button for Mom.
- **Session flow (fixed 2 turns)**:
    1. Mom taps Speak → app records Chinese audio.
    2. Backend:
        - `/v1/audio/transcriptions` → Chinese text.
        - `/v1/chat/completions` → Chinese → English.
    3. Frontend uses Web Speech API to speak English.
    4. App enters **reply listening mode**:
        - Listens up to **10 seconds** of English reply.
        - Ends early if **3 seconds of silence** after speech has begun **(deprioritized)**.
    5. Backend:
        - Transcribes English → text.
        - Translates English → Chinese.
    6. Frontend speaks Chinese back to Mom.
- Each pair (Mom + reply) = **one session** with **two turns**.

5.2 Feature 2: Daily Conversation Log

- For each **turn**, backend stores:
  - `session_id`, timestamps
  - role (Mom, partner)
  - `source_lang`, `source_text`
  - `target_lang`, `translated_text`
  - `situation_tag` (populated asynchronously)
- "Today's Conversations" screen:
  - Lists sessions (chronological).
  - Each row shows:
    - Time
    - Short label (e.g., "Kitchen – Dinner")
    - Quick preview of Mom sentence and reply.
  - Tapping a session shows the two turns with:
    - Original + translation text
    - Optional replay of TTS

5.3 Feature 3: Daily Vocabulary & Key Phrases

- For each day with enough data:
  - A **short topic summary** (Chinese + English).
  - A **"What's new today"** snippet (Chinese + English).
  - A list of **5 English phrases**:
    - English phrase
    - Chinese translation
    - Play button for English audio
- Simple "mark as listened" / basic interaction tracking **(deprioritized).**

5.4 Feature 4: Infrastructure & Tech (MVP Decisions)

- **Frontend**
  - Mobile-friendly web app (React / Next.js or similar).
  - Mic input: `getUserMedia` + `MediaRecorder` or equivalent.
  - TTS: Web Speech API (`speechSynthesis`) for English & Chinese.
- **Backend**
  - App server (e.g., FastAPI / Node / Next.js API routes) plus database.
  - Uses Student Portal API:
    - `/v1/audio/transcriptions` – ASR (Chinese & English).
    - `/v1/chat/completions` – translation & daily summaries.
    - `/v1/embeddings` – optional later for better "What's new".
- **Latency strategy**
  - **Hot path**: ASR → translation → TTS only.
  - **Background**: daily digest, embeddings, situation classification.

---

6. High-Level Architecture

- **Client (Browser)**
  - Mobile web app
  - Responsibilities:
    - Capture audio for Mom and reply speaker.
    - Play back TTS in appropriate language.
    - Display sessions, daily review, key phrases.
- **Application Backend**
  - Session orchestration and DB persistence.
  - Calls Student Portal's `/v1/audio/transcriptions`, `/v1/chat/completions`, and optionally `/v1/embeddings`.
  - Implements:
    - Session API (start session, store turns).
    - Daily summary generator job.
    - API for Daily Review & Conversation Log.
- **Database**
  - Core tables:
    - `households`, `users`
    - `conversation_sessions`, `conversation_turns`
    - `daily_summaries`, `daily_key_phrases`
  - Store minimal, non-PII text where possible (no raw audio).