

Exercise 5: Custom Vocabulary for Your Domain

Intermediate

All Modes

15–20 minutes

LEARNING OBJECTIVE

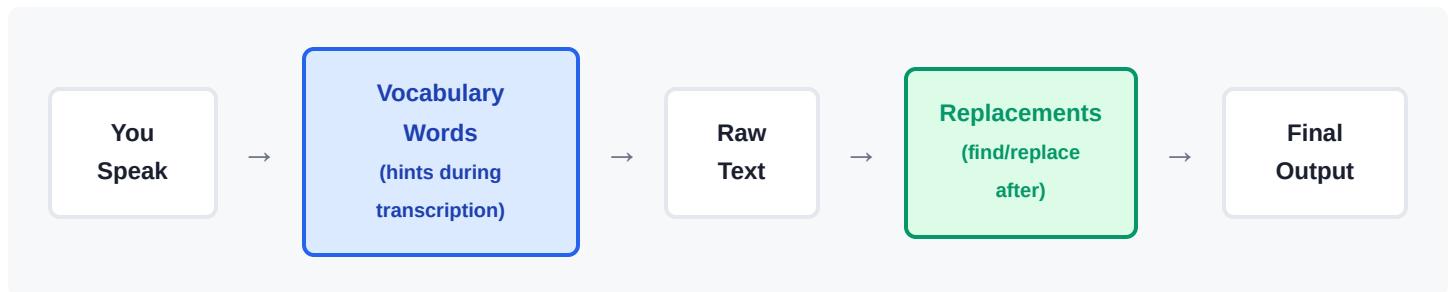
Configure SuperWhisper to correctly recognize your organization's terminology, acronyms, and jargon. You'll learn the critical difference between Vocabulary Words and Replacements—and why using the wrong one can actually make transcription worse.

Why This Matters

Every field has its own language. Public health has PHI, HIPAA, DSHS, and epidemiology. Your IT team has OAuth, PostgreSQL, and Kubernetes. SuperWhisper's general-purpose models don't know these terms exist—so they'll mishear them as similar-sounding common words. Custom vocabulary fixes this, but there's a catch: **too much vocabulary makes things worse**. This exercise teaches you the right approach.

Two Different Tools for Two Different Problems

SuperWhisper gives you two ways to improve recognition. They work at different stages of the process, and understanding when to use each is essential.





Vocabulary Words

When: DURING transcription

How: Hints to the AI about terms to listen for

Best for: Uncommon proper nouns, technical acronyms

Reliability: ~80-90% (AI-dependent)

Limit: 15-20 words maximum



Replacements

When: AFTER transcription

How: Programmatic find-and-replace

Best for: Fixing consistent mishearings, formatting

Reliability: 100% (not AI-dependent)

Limit: No practical limit

PRO TIP: WHEN IN DOUBT, USE REPLACEMENTS

Replacements are 100% reliable because they happen *after* transcription—they're just programmatic text substitution. Vocabulary words help the AI during transcription, but adding too many actually **confuses the model** and degrades accuracy. Think of vocabulary as a "top 15" list of your most important terms, and use replacements for everything else.

Step 1: Access Vocabulary Settings

1

Open Settings → Vocabulary

Click the SuperWhisper menu bar icon → Settings → Vocabulary tab

You'll see two sections: **Vocabulary Words** (top) and **Replacements** (bottom). We'll configure both.

Step 2: Add Vocabulary Words (Sparingly!)

Add only your MOST IMPORTANT terms—the ones the AI frequently mishears. These are hints to help the model recognize uncommon words.

Category	Example Words to Add
Organization	APH, Austin Public Health, HIPAA, DSHS, HHSC
Medical Terms	PHI, EHR, ICD-10, CPT, epidemiology, comorbidity
Technical	API, JSON, OAuth, PostgreSQL, Kubernetes
Names	[Add 3-5 frequently-mentioned colleague names]
Local Terms	[Project names, system names, department acronyms]

⚠ Critical: The 15-20 Word Limit

This is not a soft suggestion—it's a hard constraint. Adding more than ~15-20 vocabulary words **actively degrades transcription quality**. The AI becomes overwhelmed with hints and starts making worse guesses. If you need to teach SuperWhisper more terms, use Replacements instead.

Step 3: Add Replacements (Use Liberally)

Replacements are your workhorse. Use them for common mishearings, capitalization fixes, and formatting preferences. There's no practical limit.

You Say / AI Hears	Replace With	Why
a p h	APH	Correct capitalization for acronym
hipaa	HIPAA	Correct capitalization for acronym
p h i	PHI	Convert spaced letters to acronym
cloud code	Claude Code	Fix common mishearing (sounds similar)
v s code	VS Code	Convert spaced letters to product name
github	GitHub	Correct brand capitalization
okay	OK	Personal formatting preference
todo	TODO	Developer convention (comments)

Why Replacements Are So Reliable

Replacements happen *after* the AI has finished transcribing. They're simple text find-and-replace operations—no AI guessing involved. If the AI consistently hears "cloud code" when you say "Claude Code," a replacement fixes it 100% of the time.

Step 4: Test and Iterate

2

Test Your Vocabulary

Try dictating sentences with your added terms. Speak naturally—don't over-enunciate.

Test sentence: "We need to update the PHI documentation for HIPAA compliance before sending it to DSHS."

All three acronyms should render correctly.

3

Note Persistent Errors

If a term is still wrong after adding it to vocabulary, note exactly what the AI heard.

Example: You say "epidemiology" → AI hears "epidemia ology"

Add a replacement: "epidemia ology" → "epidemiology"

4

Add Replacements for Persistent Errors

Replacements are your fallback when vocabulary hints aren't enough.

Experiment: Discover Your Personal Mishearings

Everyone's voice is different, and the AI will mishear different words for different people. Spend 5 minutes dictating your common work terms and note what the AI gets wrong:

- Dictate your most-used acronyms (even common ones like "API")
- Say colleague names you mention frequently
- Use technical jargon from your specific projects
- Try speaking faster vs. slower—does accuracy change?

Your Turn: Build Your Personal List

✍ Vocabulary Words (Your Top 10-15)

List the terms you use most frequently that a general-purpose model won't know. Remember: quality over quantity.

Replacements (As Many as Needed)

For each term the AI consistently mishears, add a replacement. You can add to this list over time.

Heard As	Replace With

PRO TIP: THIS IS AN ONGOING PROCESS

Your vocabulary list isn't something you set once and forget. Keep a running note (physical or digital) of terms that get misheard during your daily work. Every few weeks, add new replacements for persistent errors. Over time, your transcription accuracy will steadily improve as your custom vocabulary grows.

Troubleshooting

Accuracy Got Worse After Adding Vocabulary?

You probably added too many words. The AI is now confused by conflicting hints. **Solution:** Remove all vocabulary words and start fresh with only your top 5. Add more slowly and test after each addition.

Replacement Not Working?

Replacements are case-sensitive and exact-match. If you added "APH" → "Austin Public Health" but the AI outputs "aph" (lowercase), it won't match. **Solution:** Add both "APH" and "aph" as separate replacements.

Names Still Getting Mangled?

Proper names are the hardest for AI to get right. If "Dr. Ramirez" consistently becomes "Dr. Romero," add it as both a vocabulary word AND a replacement. The combination often works better than either alone.

✓ Completion Checklist

- Opened Settings → Vocabulary
- Added 5-15 key vocabulary words (NOT more than 20)
- Added replacements for common acronyms and formatting
- Tested dictation with domain-specific terms
- Added replacements for any persistent errors discovered
- Created a note/document to track future mishearings

What's Next?

You've now mastered the fundamentals: modes, prompts, context, and vocabulary. In the final exercise, **Exercise 6: Agentic Handoff Mode**, you'll learn an advanced pattern for developers—using SuperWhisper to transform rambling spoken descriptions into structured task specifications that AI coding agents can execute effectively.