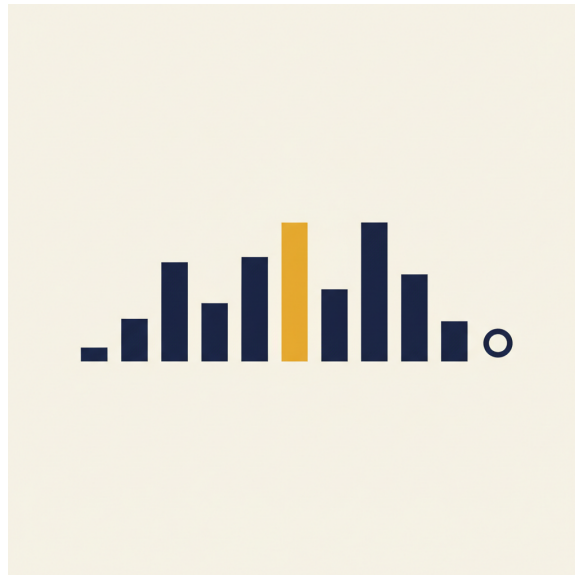




TIER 1 - FOUNDATIONS * V1.0 -- MAY 2026

AI ON THE GO

Voice mode, dictation, and mobile AI for the people who don't sit at desks. When voice is faster than typing, when it isn't, and the etiquette + pitfalls that separate a useful tool from a frustrating one.



BY

Alex Jahn / Agent Logic

v1.0 -- May 2026

Anyone who spends most of their day with their hands full and a phone in their pocket -- trades, sales, parents, drivers, walkers

15-20 minutes

Free. Forever.

EDITION

AUDIENCE

READ TIME

COST

Prepared by Agent Logic / alexanderjahn79@icloud.com / theaiguwyi.com

CONTENTS

What's in here

- 1 The desk-shaped assumption 3**
Read most articles about AI productivity. They assume you're sitting at a desk with two monitors, a quiet room, a cup of coffee, and...
- 2 Three things people call "voice AI" 4**
Get these straight; they're different tools.
- 3 Voice mode compared -- what each one's good at 4**
Snapshot of the major free or near-free tiers as of mid-2026. The space moves fast; the framework outlasts any specific ranking.
- 4 When voice is faster than typing -- the 30-second test 5**
Voice isn't always faster. It feels modern, so people overuse it for things that would have taken less time to type. The test that works:
- 5 Three worked sessions 7**
You're a carpenter walking a kitchen remodel near completion. You need to walk every surface, note every defect, and end up with a clean...
- 6 Etiquette and pitfalls 8**
Be ready for these. They're not bugs you'll fix; they're failure modes you'll learn to spot.
- 7 Mobile-only AI -- patterns for the small screen 10**
Even when you're not using voice, mobile AI is its own animal. Three patterns that work better on a phone than on a desktop:
- 8 When NOT to use voice or mobile 10**
Three situations where voice and mobile AI are the wrong tool:
- 9 Where to go from here 11**
You have voice mode and mobile down. Next in the Tier 1 expansion:

SECTION 1

The desk-shaped assumption

Most AI advice is written for people who sit

Read most articles about AI productivity. They assume you're sitting at a desk with two monitors, a quiet room, a cup of coffee, and twenty minutes to compose the perfect prompt. That's a real audience -- software engineers, marketers, knowledge workers -- and the advice is fine for them.

Then there's everyone else. The carpenter walking a job site. The mom with a toddler on her hip. The salesperson driving between calls. The plumber under a sink. The parent at 6 AM trying to figure out if today's the day school sends the kid home for that thing. The trades, the sales floor, the front line, the family. Most of the world.

Voice mode and mobile AI exist for those people. The advice is different. The use cases are different. The pitfalls are different. This module covers all of it.

What you'll have by page 13

By the end of this primer:

- You'll know the moment.
- You'll have a Siri+ChatGPT integration on iPhone -- what each one's actually good at.
- You'll know
- You'll have between calls -- so you can see what a real session looks like.
- You'll know the yourself with one.

three diff

quick cor

when voi

three wor

etiquette

Plain English, hands-busy use cases, and the honest limits -- voice still mishears trade jargon, names, and numbers.

The keyboard is a desk-shaped assumption. Most of the day, most people aren't at a desk. The fix is in your pocket -- once you know how to use it.

SECTION 2

Three things people call "voice AI"

Get these straight; they're different tools.

1. Voice mode -- you talk to the AI, the AI talks back. Real conversation, no text in between. ChatGPT has Voice Mode (and Advanced Voice). Claude has voice on its mobile app. Gemini has Gemini Live. Siri-with-ChatGPT on iPhone is the entry-level version. This is the one people mean when they say "voice AI."

2. Dictation -- you talk, your phone or computer types. Apple's built-in dictation. Google's Voice Typing. Otter, Whisper-flavored apps. The output is text; the AI part is hidden in transcription, no conversation back. You can pipe dictated text into ChatGPT or Claude as a prompt -- and that's a workflow worth knowing -- but dictation by itself isn't an AI conversation.

3. AI phone calls -- the AI calls a phone number, or a phone number reaches an AI. Twilio + LLM stacks. Restaurants with AI booking lines. Some businesses have already deployed this. From the user side, it's just a phone call where the other end is a robot that's actually pretty good. From the builder side, that's Tier 3 territory.

This module is about #1 and #2. Both have their moment.

SECTION 3

Voice mode compared -- what each one's good

Snapshot of the major free or near-free tiers as of mid-2026. The space moves fast; the framework outlasts any specific ranking.

ChatGPT Voice / Advanced Voice (Plus tier \$20/mo) -- the most natural-feeling conversation. Genuine back-and-forth, can interrupt mid-sentence, handles light-hearted talk and serious questions equally well. Strong on general knowledge, OK on technical specifics. The closest thing to talking to a competent human assistant. Best for: open-ended thinking, quick questions, talking

through an idea. Weak when it has to handle uncommon names, specific brand-name products, or trade jargon.

Claude Voice (mobile app, Pro tier \$20/mo) -- Claude is the strongest tool in this primer's catalog for actual writing and reasoning. The voice version inherits that. Less natural-sounding than ChatGPT's voice, more deliberate. Good for: working through a hard question, dictating draft content while walking, anything where you want depth over snappy. Weak when you want a quick chatty answer -- it'll over-explain.

Gemini Live (free tier) -- Google's, free, integrates with Calendar/Gmail/Maps if you give it the permissions. Good for: anything involving Google's ecosystem (asking about your schedule, about an email thread, about a trip). Weak when the question is outside Google's data -- it doesn't pull from the broader web with the same fluency.

Siri + ChatGPT (iPhone, free) -- built into iOS. Hold the side button, ask Siri a question, Siri kicks the hard ones to ChatGPT (with your permission). Lowest-friction entry point. Best for: people who don't want to install another app and just want their phone to be smarter. Weak: limited conversation length, can't do extended thinking, no memory across sessions.

The honest read: ChatGPT Voice is the best general-purpose hands-free assistant right now. Claude Voice is the best for actual work. Gemini Live is the best free option if you live in Google. Siri+ChatGPT is the best for "I don't want to think about it" use.

Quick decision tree:

- "I just want a quick answer while I'm doing something else." -> ChatGPT Voice or Siri+ChatGPT.
- "I want to talk through a real question and have the model help me think." -> Claude Voice or ChatGPT Advanced Voice.
- "I want it tied to my calendar / email." -> Gemini Live.
- "I don't want to pay for anything yet." -> Siri+ChatGPT (free) or Gemini Live (free).

SECTION 4

When voice is faster than typing -- the 30-second

Voice isn't always faster. It feels modern, so people overuse it for things that would have taken less time to type. The test that works:

The 30-second test:

Could you type the question -- or the rough idea of the question -- in under 30 seconds without stopping what you're doing?

- **Yes** -> type. Voice is slower for short things because you have to wait for the model to finish talking.
- **No** -> use voice. Either your hands are busy, or the question is long and rambling enough that talking is genuinely faster.

That's it. Stop importing "I want to feel modern" into the decision and you'll halve the moments you waste on voice that should have been a sentence in a chat window.

The four moments where voice clearly wins

1. **Hands literally occupied.** Driving, carrying a toddler, holding a piece of trim, washing dishes, on a ladder. Voice is the only option. Use it.
2. **Walking around a job site or a property.** Narrating what you see is faster than stopping to type it. The model can take the rambling and structure it later.
3. **Long, half-formed thinking.** "I'm trying to decide whether to take on this kitchen remodel -- they want it done in three weeks, I have the crew, but the timing crowds the porch project. Walk me through how to think about it." Talking through 90 seconds of context produces a richer response than the typed version, where most people would compress to one line and lose the nuance.
4. **The "I'm tired and I need to dump this somewhere" moment.** End of the day, you have a tangle in your head -- what to do tomorrow, who to call back, what the kid said about the teacher. Voice it out. The model will help you sort.

The three moments where voice clearly loses

1. **Anything you'll need to copy/paste.** Quotes, names, addresses, links, code. Voice is bad at all of this. Use text.
2. **Anything in a public space where you'd be embarrassed.** Don't dictate your divorce confession on a bus. People can hear you.
3. **Quick lookups.** "Convert 18'-6" to inches." Type that. Voice will take longer to start talking than the answer takes to read.

SECTION 5

Three worked sessions

Worked session 1 -- Walking a punch list at a job site

You're a carpenter walking a kitchen remodel near completion. You need to walk every surface, note every defect, and end up with a clean punch list to send to the client.

The session:

You, walking through the room: "Hey ChatGPT, I'm walking a punch list on a kitchen remodel. I'm going to dictate everything I notice, room corner to room corner. At the end, organize it by surface -- cabinets, countertop, plumbing, paint -- and mark the four most critical items as priority. Just listen until I say I'm done."

strike side. Right cabinet door catches on the spice pull-out -- needs a hinge adjustment. Toe-kick on the run under the sink has a gap at the seam, needs caulk. The crown molding return on the left wall has a hairline at the joint."

7/8ths instead of 1 inch -- the client may notice or not, flag it. The sink lip has caulk that's pulling at the back left corner. The faucet hole is centered to the sink but not to the window -- that's not on us, but the client may ask."

end: "Done. Organize and prioritize."

You, walki

You, walki

(Continue

The model gives you back a clean four-section punch list with priorities flagged. You email it to the client from your phone. You did in five minutes what would have taken twenty in a notebook plus another fifteen typing it up.

The honest part: the model will mistype "1/8ths" as "1 8th" or "an eighth" -- go through the output once before sending. It will also miss when you change subjects without warning, so use phrases like "moving to the counter now" to give it segmentation hooks.

Worked session 2 -- Parenting at 6 AM

You're up early with the kid, half-functional. You need to figure out three things before the day starts: lunchbox contents, the calendar status, and whether you remembered to RSVP to a thing.

You, half-awake: "Hey ChatGPT. I'm checking on my kid's lunchbox needs. Yesterday I packed turkey and cheese, the day before that PB&J. The school sent a note last week about a peanut-allergic kid in class -- what was that situation again? I want to swap to an alternative for the rest of the week. What goes in a lunchbox that's protein-decent and not peanut?"

The model rolls with the mostly-coherent prompt. You get four lunch ideas, the reminder about the

allergy note. It takes 90 seconds. You start the day already ahead.

The honest part: if the model claims to remember the school note, verify. The school's actual note has the actual rules. The model is reconstructing from your previous conversation -- could be right, could be off.

Worked session 3 -- Driving between calls

You're a salesperson or contractor driving from one call to the next. You have ten minutes. You want to do something useful with it.

You at a stoplight: "Voice mode. I just left a call with a couple in Oshkosh -- they want a deck rebuild, mid-summer install, budget range \$18 to \$22K. They asked me to follow up with two references. I have references from a similar job last year. I want to draft a follow-up email -- short, warm, naming the two reference projects, attaching nothing yet, scheduling a site walk. Walk me through it."

By the next stoplight you have the email draft in your head. You pull over for two minutes to dictate it into your email app, send it, and you're back on the road. That's a 12-minute block where most people would have just driven.

The honest part: dictating an actual email through voice mode requires you to either (a) copy the text from the chat to your email, which means stopping the car, or (b) ask the model to email it for you, which most consumer voice modes can't do directly. You're going to have to stop briefly to actually send. Voice is the thinking layer; typing is still the sending layer most days.

Voice is for the dictation, the rambling, and the thinking-out-loud. Typing is still better for the sending, the copying, and anything where the exact words matter.

SECTION 6

Etiquette and pitfalls

What voice mode regularly gets wrong

Be ready for these. They're not bugs you'll fix; they're failure modes you'll learn to spot.

- **Names.** First names, last names, brand names, place names. Voice mode will hear "Boenning" as "boning," "Schmidt" as "Smith," "Fond du Lac" as "fund du lock." Always re-read names in the output.
- **Numbers with units.** "Three eighths" gets rendered "3/8" sometimes, "three eights" other times, "0.375" never (because most people don't say it that way). If a measurement matters, type it.
- **Trade jargon.** "Soffit" becomes "sofa." "Joist" becomes "joys." "Romex" becomes "rumex" or "remix." If you work in a vocabulary the model wasn't well-trained on, voice will struggle. Slowing down helps; substituting common words helps; verifying the output helps.
- **Punctuation.** Voice modes guess at punctuation. Long thoughts come out as run-on sentences. If you need clean punctuation, edit before sending.

Public-space etiquette

People nearby can hear you. This sounds obvious until you realize you've narrated your tax situation in line at the grocery store. Some rules:

- **Public transit, restaurants, waiting rooms** -- text mode only. Don't put your business in the air.
- **A car with passengers who aren't your intimate family** -- same rule, text mode.
- **Outdoors in private spaces** (your yard, walking the dog) -- voice is fine.
- **A jobsite with a crew or other trades** -- be aware that everything you say is now part of the work-site conversation. If you're saying anything sensitive about a client or a coworker, save it for the truck.

Privacy traps

Voice goes through cloud servers to be transcribed and processed. The platforms collect what you say. In practice, that means:

- **Don't say specific Social Security numbers, account numbers, passwords.** Same rule as text mode, more important here because you're more likely to forget.
- **Don't dictate someone else's private information.** A client's diagnosis, a coworker's conflict, anything you wouldn't text into a public group chat.
- **Be aware of who's listening on your end.** If you're using voice mode on speakerphone in a shop with a TV nearby, the TV mic might pick up the AI's reply. The AI's reply is fine; what's on your end is yours to manage.

The pattern: treat voice mode like a phone call you can't take back. If you wouldn't say it on a phone call, don't say it to the AI.

SECTION 7

Mobile-only AI -- patterns for the small screen

Even when you're not using voice, mobile AI is its own animal. Three patterns that work better on a phone than on a desktop:

1. Photo + question. Take a picture of the thing, ask the model. "What is this part?" "Is this safe to eat past the date?" "What does this dashboard light mean?" "Read the warranty card and tell me how to register." The vision capability of the major models is now solid; this is one of the highest-utility daily uses on mobile.

2. Quick capture, edit later. Voice in a long rambling thought, save the chat. Don't try to refine it on the phone -- small screen, small keyboard, hard to edit. When you're back at a desk, open the chat, refine. Voice was for capture; the desk is for shaping.

3. Voice + share-sheet. iPhone share sheet (or Android equivalent) often has a way to pipe a captured photo or piece of text directly into ChatGPT or Claude. Set this up once, and the moment you see something -- receipt, sign, document, manual page -- you can capture-and-ask in two taps.

3

Three failure modes

make voice mode harder than it looks: trade jargon mishearing, public-space privacy slips, and over-using voice for things that would have been faster to type. Get those right and the rest is just practice.

SECTION 8

When NOT to use voice or mobile

Three situations where voice and mobile AI are the wrong tool:

- **Long-form writing.** Drafting an actual proposal, an actual report, an actual letter that has to be precise. Voice is for sketching; the keyboard is for finishing. A two-page proposal voiced out loud will need 30 minutes of cleanup; the same proposal typed will need 5 minutes of polish.
- **Anything that needs exact reproduction.** Code, exact quotes, recipes (hands-busy moments where someone asked the model to read a recipe out loud are a major exception, but writing a recipe down via voice gets you "1 cup flour" rendered as "won cup flour"). Type it.
- **Decision moments where you need to slow down.** Hard family conversations, important professional decisions, anything where the wrong answer is expensive. Voice goes fast and surfaces fast. Some questions benefit from going slow.

Within those limits, voice and mobile AI together cover roughly half of the ways most people would actually want to use AI day-to-day. The desk-bound advice covers the other half. Most people only know the second half. Now you know both.

SECTION 9

Where to go from here

You have voice mode and mobile down. Next in the Tier 1 expansion:

- **Free tools, real differences** -- ChatGPT vs Claude vs Gemini vs Copilot for personal use. When each one wins, when to switch, the privacy difference between them.

After that, two more on big decisions, parenting alongside AI, and reading walls of text -- then Tier 1 is full at 12 modules.

Get the next module the day it drops: theaiguywi.com/training

One email per release. No drip. No spam. Opt out anytime.

If you want voice mode, mobile patterns, and AI etiquette installed across an entire crew or sales team -- the rules trained in once, the privacy red lines drawn before someone embarrasses the company -- that's the consulting offer. We do it the way I run it in my own carpentry business.

Reach out: alexanderjahn79@icloud.com

A short call. Honest scope. We figure out together if it's a fit.

Closing -- the lock-in line

The phone in your pocket is more capable than the desktop computer you grew up with. Most people use 10% of what it can do because all the AI advice they read assumes they'll be at a desk later. This module is the on-the-go-shaped version of the rest of the curriculum.

You have voice. You have mobile patterns. You have the etiquette. The only thing left is the practice -- say the words out loud the first dozen times even when it feels weird, and by the second week you'll wonder how you used to do it any other way.

30

The 30-second test:

if you can type the question in under 30 seconds without stopping what you're doing, type. Otherwise, voice. That single rule disqualifies most of the bad voice-mode habits before they form.

You have the on-the-go ramp. Two more modules in this batch.

Agent Logic --

Fond du Lac, WI. This is module 8 of 12 in Tier 1 (Personal).

© 2026 Agent Logic. Share freely.

theaiguyn