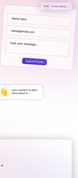


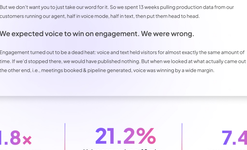
Voice is already how your buyers communicate. Everywhere except your website.

They talk to Siri, they talk to Alexa, they send voice notes to their friends throughout the day. Then they land on your website, and the conversation goes quiet — it's back to scrolling or typing.

At Docket, we're trying to change that.

SCROLL TO READ

We gave our agent a voice, then put it to test.



We built an AI Marketing Agent that works the way a good rep would: it listens, it answers, it moves the conversation forward instead of parking a visitor in front of a form. That agent has run in text for a while now. This year we gave it a voice, same agent, same conversation, just spoken instead of typed, so a visitor can talk to your website the same way they already talk to Alexa in the kitchen. Nothing to dial, nothing to download, the moment someone lands, they can just start talking.

But we don't want you to just take our word for it. So we spent 13 weeks pulling production data from our customers running our agent, half in text, then put them head to head.

We expected voice to win on engagement. We were wrong.

Engagement turned out to be a dead heat: voice and text held visitors for almost exactly the same amount of time. If we'd stopped there, we would have published nothing. But when we looked at what actually came out the other end, i.e., meetings booked & pipeline generated, voice was winning by a wide margin.

1.8×
Higher email capture rate
Voice-over text, matched traffic

21.2%
Voice conversion at 10 mins+
vs 14.9% for text at the same depth

7.4×
More meetings booked
Same traffic, same 13 week window

THE RESEARCH
CHAPTER 01 OF 04

What really happens when your website speaks your buyers' language

We ran the same agent two ways for thirteen weeks. The results rewrote how we think about voice.

CHAPTER 1 The Problem | CHAPTER 2 The Experiment | CHAPTER 3 The Results | CHAPTER 4 The Commitment

CHAPTER 01 THE PROBLEM

Your buyers are talking. Your website is not listening.

Voice is not a feature people are wishing on. It is the default mode most already live in — at home, in the car, in the kitchen. Voice activated elevators, airport maps, and more are turning up in public. It is no longer a new idea; it is becoming table stakes.

7B
WhatsApp Voice Messages
sent last year(META)

1.46M
Slack Voice Notes
Sent every week(Slack)

8.4B
Voice assistants in use
up from 4.2B in 2020(Juniper)

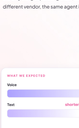
Every number to the right is voice turning up somewhere new. Not one of them is your website. The one place buyers come to actually do business with you is still the one place they can't just talk. They get served a form, or a chatbot.

CHAPTER 02 THE EXPERIMENT

We studied our own agent, two ways, for 13 weeks.

HOW WE RAN IT

We took Docket's agent, already live on customer sites, and compared it running in voice mode against the same agent running in text mode: comparable traffic, the same features switched on, the same 13-week window. That left modality as the only thing separating them, not a different product, not a different vendor, the same agent in two modes.



WHAT WE EXPECTED

We expected voice to win on engagement, keeping people around longer, and assumed that extra time was where the extra pipeline would come from. The obvious bet.

In reality, engagement was almost identical: 195 seconds for voice & 185 for text.

For a moment it looked like a dead end. If both modes held people equally long, the obvious conclusion was that they would convert at the same rate. Then we checked what each conversation actually produced and the gap we saw rewrote the whole voice versus text debate.

The gap we saw rewrote the whole voice versus text debate.

Voice captured emails 1.8 times as often as text, and booked meetings 7.4 times as often, on the same traffic, in the same window, with the mode of conversation as the only thing that changed.

Ask for an email, and voice pulls ahead at 1.8 times the rate. Ask for a meeting, and voice is not even in the same conversation. 7.4 times the rate, on identical traffic, with the same agent.

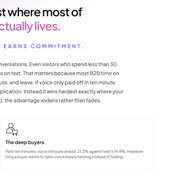
Metric	Voice	Text	Voice edge
Deep engagement	64.2%	66.4%	Equal
Email captured	6.0%	3.4%	1.8X
Meeting booked	8.1%	1.1%	7.4X

CHAPTER 03 THE RESULT

Voice and text engage buyers equally. Only one of them converts them.

THEY WERE THE SAME CONVERSATION, RIGHT UP UNTIL IT WAS TIME TO ACT.

We took Docket's agent, already live on customer sites, and compared it running in voice mode against the same agent running in text mode: comparable traffic, the same features switched on, the same 13-week window. That left modality as the only thing separating them, not a different product, not a different vendor, the same from.



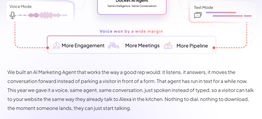
WHY TEXT LOSES BUYERS THE MOMENT IT ASKS THEM TO ACT.

Here's the part that isn't obvious: both modalities load the same booking step. Voice and text both ask the visitor to commit, the same way, at the end. So the gap isn't about one having an extra step and the other not.

It's about what happens before the ask. A voice conversation builds trust the way talking to a person does: tone, pacing, back-and-forth that feels responsive instead of scripted. By the time the booking step comes up, the visitor already trusts who they've been talking to enough to follow through.

A text agent, even a good one, leaves room for interpretation. A visitor has to guess whether a reply is warm or curt, genuine or scripted. That uncertainty is enough to make fewer visitors carry the conviction to finish.

Voice does not win by removing friction. It wins by building enough trust, before the ask, that the friction does not stop them.



Voice wins hardest where most of your traffic actually lives.

TEXT CAPTURES CONTACTS. VOICE EARNS COMMITMENT.

Voice's biggest wins did not come from long, deep conversations. Even visitors who spend less than 30 seconds convert at more than twice the rate on voice as on text. That matters because most B2B time on site is short. People show up, size you up in under a minute, and leave. If voice only paid off in ten-minute conversations it would be a neat finding with limited application. Instead it wins hardest exactly where your traffic actually lives. And for the buyers who do go deep, the advantage widens rather than fades.

The fast buyers

Under 30 seconds, voice converts at 7.6% against text's 3.3%, almost all of them booked meetings. These were your highest intent buyers, they already knew what they wanted and needed a fast way to say yes.

The deep buyers.

Past two minutes, voice still pulls ahead. 21.2% against text's 14.9%. However long a buyer wants to take, voice keeps helping instead of talking.

Voice wins at both ends, the buyer in a hurry and the buyer going deep, and never loses the middle. Length was never the advantage. Readiness was, and voice meets it the moment it shows up.

CHAPTER 04 THE COMMITMENT

Voice won the modality question.

Everything in this study came from traffic these agents were always going to get. Nothing changed but how the conversation happened, and on that alone, voice booked 7.4 times as many meetings as text. The buyers who become pipeline are already on your site. They reach the moment of action and stop, and never we hand them off to a form. That is what the Docket AI Marketing Agent is built to do. It holds the real conversation, books the meeting the moment a buyer is ready, and hands your rep the booking and the reasoning together — while intent is still warm.

The modality question is settled. Your pipeline wasn't leaking from a broken step. It was leaking from trust your website never earned.