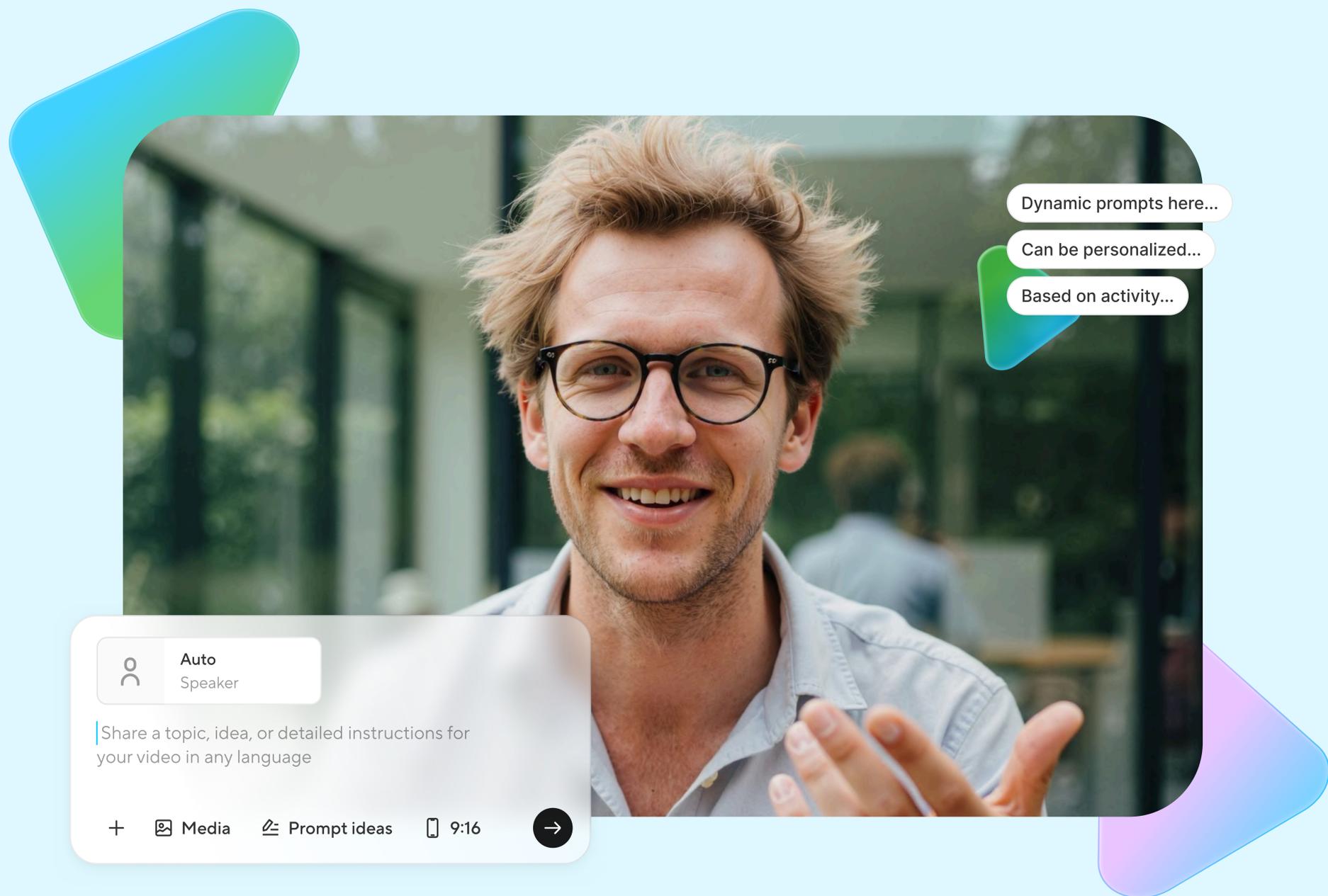


Hands-On Lab Guide

# How to get the most out of HeyGen's Video Agent



Auto  
Speaker

Share a topic, idea, or detailed instructions for your video in any language



Media



Prompt ideas



9:16



Most people prompt Video Agent like they're Googling. That's why their videos look generic.

The difference between a forgettable AI video and one that actually converts? **How you direct.**

We spent months on Video Agent, refining prompts that consistently produce production-quality videos because we know exactly how our Agent "thinks."



### The basics

## What Video Agent actually needs

Before you type a single word, understand the three controls at your fingertips:

**Avatar selection** — Pick a specific avatar or let Auto mode find one that fits your content. Pro tip: You can also go avatar-free with voice-over only, but you must explicitly say "no avatar" in your prompt.

**Duration** — Let Auto mode decide based on your content or choose 30s, 1 min, 2 min, etc.

Note: Video Agent follows your script more than this duration number, e.g. if you provide a short clip but selected 2 min, it will likely only create a short video.

**Aspect ratio** — Create videos in portrait or landscape, or simply leave as Auto.

These are your baseline. The magic happens in the prompt itself.

### The prompt

## More context = better videos

At minimum, your prompt should describe the content you're trying to deliver. Here's what a basic prompt looks like:

```
"Introduce HeyGen to knowledge workers, talk about its Talking Avatar models, how people use it, and mention Video Agent at the end."
```

That works. But the more context and intent you provide, the better Video Agent can structure scenes, pacing, and visuals.



## The pro move

# Use your script directly

This is the single biggest upgrade most people miss.

Paste your full video script into the prompt. Video Agent will follow it scene-by-scene while improving flow, timing, and visuals automatically.

Here's a script-driven prompt in action:

**Intro** (A-roll, motion graphics overlay): "If your work is mostly explaining things – updates, ideas, decisions – video usually helps, but making it takes too much time ."

**What is HeyGen** (motion graphics B-roll): "HeyGen helps introverts turn ideas into production-ready videos – without cameras, editing, or studios."

**Talking Avatars** (A-roll + demo cut): "Our talking avatar models offer realistic and natural delivery using your own digital identity."

**Use cases** (motion graphics list): "Teams use it for internal training, online education, product explanations, and knowledge sharing."

**Introducing Video Agent** (end beat): "And with our new Video Agent – one prompt becomes a structured, animated video, end to end."

**End card:** HeyGen · Empower knowledge-based creators

The agent may make small grammar and pacing edits. That's a feature, not a bug

## Attachments

# Give Video Agent reference material

You can upload files to help Video Agent understand your content:

**Images and videos:** Product screenshots, existing assets, diagrams, or any media you want included. Pro tip: Upload your own photo and ask the agent to use it as your talking avatar.

**PDFs and documents:** Training materials, research papers, or product docs. The agent will extract key information.

**Critical:** When uploading references, add context about how you want them used. For example:

"Use the attached product screenshots as B-roll when discussing features"

"Reference the attached PDF for accurate technical specifications"



# Advanced prompting

Here's my personal favorite prompt addition. I add this to almost everything:

```
"Use minimal, clean styled visuals. Blue, black, and white are the main colors. Leverage motion graphics as B-rolls and A-roll overlays. Use AI videos when necessary. When real-world footage is needed, use Stock Media. Include an intro sequence, outro sequence, and chapter breaks using Motion Graphics."
```

Try it. You'll immediately notice the difference. But why does this work? Let's break it down.

## Define your visual style and colors

Video Agent executes style requirements consistently across your entire video. Without a defined style, scenes can look disconnected.

Style	Best for	Prompt addition
Minimalistic	Corporate, tech, SaaS	"Use minimalistic, clean visuals with lots of white space"
Cartoon/animated	Education, kids content	"Use cartoon-style illustrated visuals"
Bold and vibrant	Marketing, social	"Use bold, vibrant colors and dynamic visuals"
Cinematic	Brand films, high-end	"Use cinematic quality visuals with dramatic lighting"
Flat design	Modern, app demos	"Use flat design style with geometric shapes"
Gradient modern	Tech, startup	"Use modern gradient backgrounds and sleek transitions"
Retro/vintage	Nostalgia, creative	"Use retro-inspired visuals with warm tones"
<your style>	<your brand>	<your prompt>

You can even specify exact colors and fonts:

```
"Use #1E40AF as primary blue, #F8FAFC as background white, and #0F172A for text. Use Inter font family throughout."
```



## When to use what

This is where most people get stuck. Match the media type to the content type.

### Motion graphics

Animated graphic elements: text animations, icons, charts, shapes, transitions.  
Captions are auto generated.

#### Best for:

- A-roll overlays: Lower thirds, bullet points alongside avatar, animated callouts
- B-roll scenes: Full-screen animated explanations, data visualizations
- Chapter cards: Section breaks, intros, outros
- Information display: Statistics, comparisons, timelines

Example:

"Use motion graphics to display the 5 key benefits as animated bullet points appearing one by one while the avatar speaks."

### AI-generated images and videos

Created by generative AI based on your descriptions.

#### Best for:

- Conceptual illustrations
- Custom scenarios that stock footage won't cover
- Stylized visuals in a particular artistic style
- Product mockups in various contexts

Example:

"Generate an AI image showing a futuristic office where humans and AI work together: use this as B-roll for the 'future of work' section."

### Stock media

Real-world footage from stock libraries.

#### Best for:

- Authentic scenes (real offices, cities, people)
- Industry-specific content (medical, manufacturing, retail)
- Emotional moments
- Establishing shots

Example:

"Use stock footage of a busy corporate office for B-roll when discussing workplace productivity."



## Quick reference

# The media type matrix

Content type	Motion graphics	AI generated	Stock media
Data/statistics	✅ Best	❌	❌
Abstract concepts	✅ Good	✅ Best	❌
Real environments	❌	⚠️ Can work	✅ Best
Brand elements	✅ Best	❌	❌
Human emotions	❌	⚠️ Uncanny	✅ Best
Custom scenarios	⚠️ Limited	✅ Best	⚠️ May not exist
Technical diagrams	✅ Best	❌	❌

## Scene-by-scene prompting

# Maximum control

When you need precise output, prompt each scene individually.

Basic structure:

```
Scene 1: [Scene Type]
Visual: [Describe exact visual]
VO/script: "[What the avatar says]"
Duration: [Approximate length]
```

Here's a detailed product launch video example:

```
Scene 1: Intro (motion graphics)
Visual: Animated logo reveal with particle effects, brand colors sweep
Duration: 3 seconds
```

```
Scene 2: Hook (A-roll with overlay)
Visual: Avatar on branded background, text overlay "The Future is Here"
VO: "What if I told you that creating professional videos just got 10x easier?"
Duration: 5 seconds
```



Scene 3: Problem statement (stock media B-roll)

Visual: Stock footage of frustrated person at computer, then clock ticking

VO: "Traditional video production takes weeks. Coordinating schedules, booking studios, endless editing rounds..."

Duration: 8 seconds

Scene 4: Solution introduction (A-roll + motion graphics overlay)

Visual: Avatar speaking, animated product logo appears beside them

VO: "Introducing HeyGen Video Agent – your AI-powered video creation partner"

Duration: 6 seconds

Scene 5: Feature showcase (motion graphics B-roll)

Visual: Animated screen recording style, showing interface with callouts

VO: "Simply describe what you want, and watch your video come to life"

Duration: 10 seconds

Scene 6: Benefits (motion graphics list)

Visual: 3 benefits animate in one by one with icons

VO: "Save time. Maintain consistency. Scale your content."

Duration: 8 seconds

Scene 7: CTA (A-roll)

Visual: Avatar, confident pose, CTA text overlay

VO: "Try HeyGen Video Agent today and transform how you create videos"

Duration: 5 seconds

Scene 8: End card (motion graphics)

Visual: Logo, tagline, website URL, social handles

Duration: 4 seconds



# Real prompts you can steal

## Compliance training:

Use a professional female avatar. Make a compliance training video explaining phishing in detail. Use examples and list top watch-outs. Leverage motion graphics as A-roll overlay and B-roll to help explain core concepts.

## Educational explainer (VO only):

Create a 1-minute video about camera aperture. Use minimal science diagrams and visualizations. No avatar needed, only voice-over. Cool neutrals (navy, cyan), thin-line diagrams, and slow elegant motion. B-roll is abstract scientific illustrations. Sequencing: definition → diagram expansion → conceptual layering, with fade-through transitions.

## Brand story (animated):

Make a video telling the story of how Twitch got started. Use cartoon-style animations and overlays. I want Twitch's iconic colors and fonts. Use motion graphics overlays and AI-generated B-roll.

## The bottom line

Video Agent isn't magic; it's a production partner that executes your creative direction.

The more specific you are about content, style, media types, and scene structure, the closer you'll get to exactly what you envision.

Start with a script. Define your visual style. Match media types to content types. Prompt scene-by-scene when precision matters.

You had the message. Now you own the production.

**Try these prompts with your next video.**

**Then tell me what worked.**

