

Video & Scriptwriting / Short-Form Video

Design text overlay placement and timing for silent viewing — accessibility and retention optimization for sound-off scrolling.

Difficulty: Intermediate

Model: GPT-4 / Claude / Gemini

Use Case: Text Overlay Design, Accessibility

Updated: June 2026

Why This Prompt Exists

Up to 80% of short-form viewers watch without sound. Text overlays are essential for retention — but most creators add them as an afterthought, not a strategy.

You get:

- text too small to read on mobile (viewers scroll past)
- text that disappears too fast (viewers can't finish reading)
- text placed behind the subject (unreadable, covered)
- no captions for spoken dialogue (silent viewers lost)
- text overlays that don't match the visual pacing

But text overlays have proven patterns:

- hook text: large, bold, 2-4 words, appears in first second
- caption text: timed to speech, 1-2 lines, centered or left-aligned
- emphasis text: single word, pops with animation, highlights key point

- CTA text: end of video, contrasting color, 3-5 words
- safe zone: text within center 80% of frame (avoid camera cutouts)

Without strategy, text overlays hurt retention.

This prompt designs effective text overlay strategies.

The Prompt

Assume the role of a short-form video text overlay strategist.

Your task is to design text overlay timing, placement, and styling for silent viewing.

Generate:

1. TEXT OVERLAY TYPES

Type	Purpose	Size	Duration	Position	Example
Hook text	Stop scroll	Largest	0.5-1s	Center or top	"WATCH THIS"
Caption	Read dialogue	Medium	Matches speech	Center bottom	"Here's what happened"
Emphasis	Highlight key word	Large	0.3-0.5s	Center	"FREE"
CTA	End action	Large	1-2s	Center	"FOLLOW FOR MORE"
Label	Identify subject	Small	1-3s	Corner or near subject	"Day 1 vs Day 30"
Countdown	Build anticipation	Large	0.5-1s each	Center	"3...2...1..."

2. READABILITY GUIDELINES

Factor	Minimum	Recommended	Maximum
Font size (mobile)	24pt	36-48pt	N/A
Characters per line	10	15-20	25
Duration per word	0.5s	0.75-1s	1.5s
Contrast ratio	4.5:1	7:1+	N/A
Safe zone margin	10%	15%	20%

3. TEXT OVERLAY TIMING BY SCRIPT PHASE

Phase	Text Type	Timing	Word Count
Hook (0-3s)	Hook text	0:00-0:02	2-4 words
Value (3-15s)	Caption + Emphasis	Match speech	5-10 words per caption
Retention (15-25s)	Emphasis + Label	Quick pops	1-3 words
CTA (25-30s)	CTA text	1-2s	3-5 words

4. TEXT OVERLAY PLACEMENT MAP

Screen Zone	Best For	Avoid
Top center	Hook text, labels	Long sentences
Center	Emphasis words, countdowns	Blocking face
Bottom center	Captions, subtitles	Important visual
Bottom left	Branding, logo	Critical information
Bottom right	CTA after content	Early CTA

5. PLATFORM-SPECIFIC TEXT GUIDELINES

Platform	Safe Zone	Font Preference	Emoji Use
TikTok	80% center	Bold, sans-serif	High
Instagram Reels	85% center	Any readable	Medium
YouTube Shorts	90% center	Clean sans-serif	Low
Snapchat	70% center	Bold only	Medium

6. TEXT OVERLAY SCRIPT MARKUP

****Original script line:****

`"The secret to better sleep is actually pretty simple."`

****With text overlay markup:****

`"The secret to better sleep" [caption, 1.5s] `[BEAT]` `"is actually pretty simple." [caption, 1.5s] with emphasis on `"simple" [emphasis pop]

7. COMMON TEXT OVERLAY MISTAKES

Mistake	Why It Fails	Correct Approach
Text too small	Unreadable on mobile	Minimum 36pt
Text too fast	Can't finish reading	0.75-1s per word
Text behind subject	Covered, unreadable	Safe zone placement
No captions	Silent viewers lost	Caption all dialogue
Too many words	Overwhelming	1-2 lines max
Poor contrast	Hard to read	High contrast (white on dark, black

on light) |

INPUTS:

Script (full or key lines):

[PASTE SCRIPT]

Platform:

[TIKTOK / REELS / SHORTS / SNAPCHAT]

Sound status:

[SOUND ON PRIMARY / SOUND OFF PRIMARY / BOTH]

Key emphasis words:

[E.G., "free," "new," "limited," "proven"]

RULES:

- Hook text appears in first second (large, bold, 2-4 words)
- Captions must match spoken words (for accessibility)
- Emphasis words get larger text or animation
- CTA text appears at end (contrasting color, 3-5 words)
- Safe zone: text within center 80% of frame
- Minimum font size: 36pt for mobile viewing
- Test without sound: video must be understandable
- Duration rule: 0.75-1 second per word for captions

How To Use It

- Hook text appears in the first second — large, bold, 2-4 words.
- Captions must match spoken words — for accessibility and silent viewers.

- Emphasis words get larger text or animation — highlight key points.
- CTA text appears at the end — contrasting color, 3-5 words maximum.
- Safe zone: text within the center 80% of the frame — avoid camera cutouts.
- Minimum font size: 36pt for mobile viewing — larger is better.
- Test without sound — the video must be completely understandable.
- Duration rule: 0.75-1 second per word for captions — give viewers time to read.

Example Input

Script:

“You’re brushing your teeth wrong. Most people miss the gum line. Here’s the fix. Angle your brush 45 degrees. That’s it. Your dentist will thank you.”

Platform:

“TIKTOK”

Sound status:

“BOTH (sound on and off)”

Key emphasis words:

“wrong, fix, 45 degrees, thank you”

Why It Works

Most creators add text overlays as an afterthought — resulting in unreadable, poorly timed, or missing captions that lose silent viewers.

This framework improves outcomes by forcing:

- text overlay type classification (hook, caption, emphasis, CTA, label, countdown)
- readability guidelines (font size, characters per line, duration, contrast, safe zone)
- timing by script phase (what text appears when)
- placement mapping (which screen zone for which text type)

- platform-specific guidelines (TikTok vs. Reels vs. Shorts)

Failure modes this prevents:

- Text too small to read on mobile (viewers scroll past)
- Text that disappears too fast (viewers can't finish reading)
- No captions for dialogue (silent viewers lost)
- Text placed behind the subject (unreadable, covered)

This improves on: Post-hoc text overlay placement. Strategic text overlay planning improves retention for silent viewers.

Related to: SF-02 (Hook) for opening text; SF-04 (Trend Sound) for audio-text coordination.

Build Better AI Systems

Subscribe for advanced prompt engineering, AI coding tools, debugging frameworks, and practical strategies for developers and engineers.

Carefully engineered prompts for people doing real work.