



PHÚC BANI

Truyền nghề YouTube · Xây kênh view ngoại

QUÀ TẶNG MIỄN PHÍ

GIỌNG ĐỌC MIỄN PHÍ CHO KÊNH KHÔNG LỘ MẶT

Ba công cụ miễn phí, cách chọn giọng theo tệp khán giả
— kèm bộ lệnh bấm chạy là ra mp3 và checklist trước khi ghép

CÔNG CỤ: EDGE TTS · GOOGLE AI STUDIO · ELEVENLABS

3

công cụ

7

bước

3

prompt

5

lỗi hay gặp

"Kênh không lộ mặt thì giọng đọc chính là gương mặt của kênh."



TRONG TÀI LIỆU NÀY CÓ GÌ

Đây là cẩm nang để bạn có giọng đọc cho video mà không tốn một đồng nào — chọn đúng công cụ, chọn đúng chất giọng theo tệp khán giả, rồi biến cả kịch bản thành file audio sạch trong vài phút. Dành cho người làm kênh không lộ mặt, đặc biệt là kênh view ngoại.

PHẦN 1	Ba công cụ tạo giọng miễn phí — dùng cái nào, khi nào	Bảng so sánh thẳng thật
PHẦN 2	Chọn chất giọng theo insight tệp khán giả	Chỗ 90% người mới làm sai
PHẦN 3	Quy trình 7 bước: từ kịch bản tới file audio sạch	Phần thực thi chính
PHẦN 4	Bộ script bấm chạy là ra mp3 — cài trong 5 phút	Kèm mã copy được
PHẦN 5	3 prompt viết kịch bản hợp giọng đọc máy	Copy dán là dùng
PHẦN 6	5 lỗi giọng đọc làm tụt giữ chân + checklist	In ra dán cạnh màn hình

Đọc tài liệu này thế nào cho đúng

- **Đừng đọc hết rồi mới làm.** Mở song song một cửa sổ công cụ, đọc tới đâu bấm tới đó.
- **Làm thử một đoạn 30 giây trước.** Nghe lại bằng tai nghe. Ứng rồi mới chạy cả kịch bản.
- **Chốt một giọng và giữ nguyên cho cả kênh.** Đổi giọng giữa chừng là khán giả quen tai bỏ đi — giọng đọc chính là gương mặt của kênh không lộ mặt.

Một câu cần nhớ trước khi lật trang

Giọng hay không cứu được kịch bản dở. Nhưng giọng sai tệp thì kịch bản hay cỡ nào người ta cũng tắt trong mười giây đầu. Chọn giọng là việc của người làm kênh, không phải việc của công cụ.



PHẦN 1

BA CÔNG CỤ TẠO GIỌNG MIỄN PHÍ

Cả ba đều miễn phí ở mức nào đó, nhưng miễn phí của mỗi bên một kiểu. Chọn sai là làm được vài video rồi tắt, hoặc audio dính watermark không gỡ được.

Bảng so sánh nhanh

CÔNG CỤ	ĐƯỢC GÌ	VƯỚNG GÌ
EDGE TTS	Miễn phí thật, không giới hạn số chữ, không cần tài khoản. Hàng trăm giọng, đủ tiếng Việt, Anh, Ấn, Tây Ban Nha.	Chạy bằng dòng lệnh trên máy
GOOGLE AI STUDIO	Miễn phí có hạn mức. Điều khiển được cảm xúc, tốc độ, cách nhấn nhá bằng câu lệnh. Ghép được hội thoại hai người.	Cần tài khoản Google
ELEVENLABS	Chất giọng tự nhiên nhất thị trường, nhưng bản miễn phí chỉ vài nghìn ký tự mỗi tháng và yêu cầu ghi nguồn.	Hợp làm mẫu, không hợp chạy sản lượng

1. Edge TTS — chọn mặc định cho người chạy sản lượng

Đây là bộ giọng của Microsoft, thứ đang đọc trong tính năng đọc to của trình duyệt Edge. Có một gói mã nguồn mở cho phép gọi thẳng bộ giọng đó từ máy bạn. Không đăng ký, không thẻ, không hạn mức ký tự. Một kịch bản mười sáu phút đọc xong trong khoảng một phút.

- **Được:** miễn phí không giới hạn, xuất thẳng ra mp3, chỉnh được tốc độ và cao độ, hàng trăm giọng cho gần như mọi thị trường bạn định đánh.
- **Vướng:** phải chạy bằng dòng lệnh — nghe đáng sợ nhưng thật ra chỉ là dán một dòng. Phần 4 có sẵn mã, bạn copy là chạy.
- **Giọng Việt:** **vi-VN-NamMinhNeural** (nam) và **vi-VN-HoaiMyNeural** (nữ).

Đây là công cụ chúng tôi dùng để dựng phần lời cho video hướng dẫn nội bộ — nói có sách mách có chứng, không phải giới thiệu cho vui.

2. Google AI Studio — chọn khi cần giọng có cảm xúc

Vào aistudio.google.com, đăng nhập bằng tài khoản Google là dùng được. Điểm mạnh khác hẳn hai công cụ còn lại: bạn ra lệnh cho nó đọc thế nào bằng chính tiếng người — đọc chậm lại, giọng trầm, như đang kể một bí mật — chứ không phải kéo thanh trượt.

- **Hợp với:** kênh kể chuyện, kênh cần cao trào, đoạn hook đầu video cần nhấn nhá.
- **Vướng:** có hạn mức miễn phí theo ngày và chính sách đổi liên tục — kiểm tra lại hạn mức ngay trong trang trước khi định chạy cả loạt video.
- **Mẹo dùng khôn:** để nó đọc **mười lăm giây đầu** (phần quyết định người xem ở lại), phần thân video để Edge TTS đọc. Vừa tiết kiệm hạn mức, vừa được đúng chỗ đáng tiền nhất.



3. ElevenLabs — chọn khi làm mẫu để chốt chất giọng

Chất lượng thì khỏi bàn, nhưng bản miễn phí chỉ đủ vài phút audio mỗi tháng và yêu cầu ghi nguồn khi dùng công khai. Làm kênh thật sự thì hạn mức đó hết trong một video.

- **Dùng đúng cách:** lấy bản miễn phí để **thử và chốt chất giọng** bạn muốn, rồi đi tìm giọng gần giống nhất trong Edge TTS để chạy sản lượng.
- **Đừng:** xây cả kênh trên bản miễn phí rồi tới video thứ ba thì tắc giữa chừng.

Hạn mức và điều khoản của cả ba công cụ đều do bên thứ ba quy định và thay đổi theo thời gian. Con số trong tài liệu này là mức phổ biến tại thời điểm biên soạn — trước khi chạy sản lượng lớn, mở trang giá của họ kiểm tra lại một lượt. Ba công cụ này không phải sản phẩm của Phúc BANI và chúng tôi không hưởng hoa hồng từ chúng.



PHẦN 2

CHỌN GIỌNG THEO INSIGHT TẬP

Người mới chọn giọng theo kiểu nghe hay là lấy. Người làm nghề chọn giọng theo người sẽ nghe nó.

Bốn thứ phải chốt trước khi mở công cụ

GIỚI TÍNH	Tập nữ nội trợ nghe giọng nữ thân thiện. Tập nam công nghệ, tài chính nghe giọng nam chắc tiếng.	Chốt một lần cho cả kênh
TUỔI	Giọng trẻ hợp giải trí, drama. Giọng trung niên hợp kiến thức, tài chính, sức khỏe — nghe là thấy đáng tin.	Sai chỗ này mất uy tín ngay
CHẤT GIỌNG	Ấm và trầm hợp kể chuyện. Sáng và rõ hợp tin tức, hướng dẫn.	Nghe thử tối thiểu năm giọng
NHỊP ĐIỀU	Kể chuyện đọc chậm hơn bình thường một nhịp. Tin tức đọc nhanh, dứt khoát.	Chỉnh bằng tốc độ đọc

Bảng giọng gợi ý theo thị trường

Việt Nam	vi-VN-NamMinhNeural (nam, chắc tiếng) · vi-VN-HoaiMyNeural (nữ, sáng rõ)
Mỹ	en-US-GuyNeural , en-US-ChristopherNeural (nam) · en-US-JennyNeural , en-US-AriaNeural (nữ)
Anh	en-GB-RyanNeural (nam) · en-GB-SoniaNeural (nữ) — chất giọng Anh hợp ngách lịch sử, tài liệu
Ấn Độ	en-IN-PrabhatNeural (nam) · en-IN-NeerjaNeural (nữ) — thị trường dễ lên view, RPM thấp, hợp giai đoạn nuôi kênh
Tây Ban Nha	es-MX-JorgeNeural , es-ES-AlvaroNeural (nam) · es-MX-DaliaNeural (nữ)
Xem toàn bộ	Chạy lệnh edge-tts --list-voices để xem đủ danh sách hàng trăm giọng của mọi thứ tiếng

Luật một giọng

Một kênh chỉ dùng **một giọng chính** xuyên suốt. Khán giả nhớ kênh không lộ mặt bằng tai chứ không bằng mắt. Đổi giọng giữa chừng chẳng khác gì đổi mặt người dẫn — kênh mất nhận diện, giữ chân tụt theo.



PHẦN 3

BẢY BƯỚC TỪ KỊCH BẢN TỚI AUDIO SẠCH

Làm đúng thứ tự này thì một video mười sáu phút mất chừng mười lăm phút để ra file audio hoàn chỉnh, kể cả thời gian nghe soát.

Bước 1 — Làm sạch kịch bản trước khi đưa cho máy đọc

Đây là bước ăn tiền nhất và cũng là bước hay bị bỏ nhất. Máy đọc y nguyên những gì bạn viết, kể cả những thứ bạn không định cho ai nghe.

- Xoá hết chú thích trong ngoặc vuông, ghi chú cảnh, mốc thời gian, tên file.
- Viết số thành chữ ở chỗ dễ đọc sai: 2026 nên viết là hai nghìn không trăm hai sáu nếu bạn muốn chắc chắn; 15% viết là mười lăm phần trăm.
- Bỏ ký tự lạ: dấu sao, dấu thăng, gạch đầu dòng — máy đọc sẽ ngắc ngứ hoặc đọc thành tiếng.
- Viết tắt phải bung ra: TP.HCM đọc thành gì thì viết đúng như thế.
- Câu dài quá hai dòng thì cắt đôi. Máy không biết lấy hơi, người nghe sẽ thấy ngột.

Bước 2 — Chia kịch bản thành khối

Đừng đưa cả kịch bản mười sáu phút vào một lần. Chia theo đoạn nội dung, mỗi khối tương ứng một phần trong dàn ý. Lý do: sửa một câu thì chỉ phải đọc lại một khối, không phải đọc lại cả video. Đặt tên file theo thứ tự để lát ghép cho đúng.

Bước 3 — Chọn giọng và chốt tốc độ

Kể chuyện	Tốc độ giảm nhẹ, khoảng chậm hơn bình thường một chút. Cao độ để nguyên.
Tin tức, hướng dẫn	Tốc độ tăng nhẹ, dứt khoát, không kéo dài cuối câu.
Ngách trẻ, giải trí	Tốc độ nhanh hơn rõ rệt — nhưng thử một đoạn rồi nghe lại, nhanh quá thành nghe mệt.

Bước 4 — Chạy thử ba mươi giây

Lấy đúng đoạn mở đầu, chạy thử, nghe bằng tai nghe chứ không nghe bằng loa laptop. Nghe xem có chỗ nào máy đọc sai trọng âm, đọc dính chữ, hay ngắt câu sai không. Sửa kịch bản rồi chạy lại. Chỉ khi đoạn ba mươi giây nghe lọt tai mới chạy toàn bộ.

Bước 5 — Chạy toàn bộ và ghép

Chạy lần lượt từng khối, ghép lại theo thứ tự. Chừa khoảng lặng ngắn giữa các phần lớn — đây là chỗ người nghe thở, cũng là chỗ bạn cắt cảnh cho tự nhiên.

Bước 6 — Nghe soát toàn bộ một lượt

#Phúc BANI



- Nghe hết, không tua. Bạn đang thay mặt khán giả nghe lần đầu.
- Đánh dấu chỗ nào thấy chán — đó là chỗ kịch bản có vấn đề, không phải giọng có vấn đề.
- Chỗ nào máy đọc sai tên riêng thì sửa cách viết trong kịch bản rồi chạy lại riêng câu đó.

Bước 7 — Ghép vào video theo nhịp

Đọc tới đâu ghép hình tới đó. Đừng dựng hình xong mới nhét tiếng vào — làm ngược thì hình và tiếng lệch nhau, người xem cảm nhận được ngay dù không chỉ ra được tại sao.

Chỗ hay hỏng nhất

Chín trên mười trường hợp audio nghe chán không phải do giọng máy, mà do **kịch bản viết cho mắt đọc chứ không viết cho tai nghe**. Viết xong hãy đọc to lên một lượt — chỗ nào bạn hụt hơi thì máy cũng sẽ đọc ngộp chỗ đó.



PHẦN 4

BỘ SCRIPT BẮM CHẠY LÀ RA MP3

Phần này dành cho ai chọn Edge TTS. Làm một lần, dùng mãi. Không cần biết lập trình, chỉ cần copy đúng dòng lệnh.

Cài đặt — làm một lần, khoảng năm phút

BƯỚC 1	Cài Python bản mới nhất từ trang python.org. Khi cài nhớ tích ô Add Python to PATH.	Máy đã có Python thì bỏ qua
BƯỚC 2	Mở cửa sổ dòng lệnh, gõ: pip install edge-tts	Chờ khoảng một phút
BƯỚC 3	Gõ: edge-tts --list-voices để kiểm tra đã cài xong	Hiện ra danh sách giọng là thành công

Lệnh cơ bản — đọc một đoạn text thành mp3

edge-tts --voice vi-VN-NamMinhNeural --file kichban.txt --write-media audio.mp3

Đặt file kịch bản tên kichban.txt cùng thư mục, chạy lệnh trên, xong là có audio.mp3. Muốn giọng nữ thì đổi vi-VN-NamMinhNeural thành vi-VN-HoaiMyNeural. Muốn tiếng Anh Mỹ thì đổi thành en-US-GuyNeural.

Lệnh có chỉnh tốc độ và cao độ

edge-tts --voice en-US-GuyNeural --rate=-8% --pitch=-2Hz --file script.txt --write-media voice.mp3

- **--rate** là tốc độ đọc. Số âm là chậm lại, số dương là nhanh lên. Kể chuyện thường dùng khoảng âm năm tới âm mười phần trăm.
- **--pitch** là cao độ. Số âm làm giọng trầm xuống, nghe chắc và đáng tin hơn.
- **--write-subtitles phu-de.vtt** thêm vào cuối lệnh sẽ xuất luôn file phụ đề khớp tiếng.

Chạy hàng loạt nhiều khối một lúc

Để tất cả file kịch bản trong một thư mục, đặt tên 01.txt, 02.txt, 03.txt. Trên Windows, mở PowerShell tại thư mục đó và chạy đúng một dòng dưới đây, nó sẽ đọc lần lượt từng file và xuất ra mp3 cùng tên.

```
Get-ChildItem *.txt | ForEach-Object { edge-tts --voice vi-VN-NamMinhNeural --file $_.Name --write-media ($_.BaseName + ".mp3") }
```

Bảng tra nhanh khi quên



Xem tất cả giọng	edge-tts --list-voices
Lọc giọng tiếng Việt	edge-tts --list-voices findstr vi-VN
Đọc một câu ngắn để thử	edge-tts --voice vi-VN-NamMinhNeural --text "Xin chào" --write-media thu.mp3
Xuất kèm phụ đề	thêm --write-subtitles ten-file.vtt vào cuối lệnh
Đọc chậm lại	thêm --rate=-10%
Giọng trầm hơn	thêm --pitch=-3Hz

Nếu máy báo không tìm thấy lệnh edge-tts, đóng cửa sổ dòng lệnh rồi mở lại một lần nữa. Chín mươi phần trăm trường hợp là do cửa sổ cũ chưa nhận đường dẫn mới sau khi cài.



PHẦN 5

BA PROMPT VIẾT KỊCH BẢN HỢP GIỌNG MÁY

Kịch bản viết cho người đọc và kịch bản viết cho máy đọc là hai thứ khác nhau. Ba prompt dưới đây bắt AI viết ngay từ đầu theo kiểu tai nghe được.

Prompt 1 — Viết kịch bản cho tai, không cho mắt

Bạn là người viết kịch bản lời bình cho video YouTube không lộ mặt, ngách [NGÁCH], thị trường [THỊ TRƯỜNG]. Kịch bản này sẽ được một giọng đọc máy đọc lên, nên hãy viết theo luật sau: câu ngắn, mỗi câu một ý; không dùng ký hiệu, gạch đầu dòng hay chữ viết tắt; số viết thành chữ ở chỗ dễ đọc sai; không viết chú thích trong ngoặc. Trước khi viết, hãy hỏi tôi cho đến khi bạn tự tin chín mươi lăm phần trăm là hiểu đúng ý. Nếu không chắc thông tin, hãy để trống, đừng bịa. Độ dài [SỐ] từ, sai số cộng trừ năm mươi.

Prompt 2 — Sửa kịch bản có sẵn thành bản đọc được

Đây là kịch bản của tôi. Hãy sửa lại để một giọng đọc máy đọc lên nghe tự nhiên: cắt câu dài quá hai dòng, bỏ ký tự lạ, bung chữ viết tắt, đổi số thành chữ ở chỗ dễ đọc sai, thêm chỗ ngắt nghỉ hợp lý. Giữ nguyên ý và giữ nguyên độ dài. Chỉ ra cho tôi những câu bạn đã sửa và vì sao.

Prompt 3 — Viết chỉ dẫn giọng cho Google AI Studio

Tôi sắp đưa đoạn mở đầu này cho công cụ đọc có điều khiển cảm xúc. Hãy viết cho tôi một câu chỉ dẫn giọng đọc bằng tiếng Anh, mô tả rõ tốc độ, cao độ, cảm xúc và chỗ cần nhấn, sao cho mười lăm giây đầu giữ được người xem trong ngách [NGÁCH]. Đưa ba phương án khác nhau để tôi thử.

Vì sao phải có prompt riêng cho phần này

Người mới thường viết kịch bản xong mới phát hiện máy đọc trúc trắc, rồi ngồi sửa tay từng câu. Ba prompt trên đẩy việc đó về **trước khi viết** — rẻ hơn nhiều so với sửa sau.



PHẦN 6

NĂM LỖI GIỌNG ĐỌC LÀM TỤT GIỮ CHÂN

Đây là những lỗi khiến người xem tắt video mà họ không giải thích được lý do — chỉ thấy nghe mệt.

01 Đọc đều một nhịp từ đầu đến cuối

Bạn thấy: Người xem trụ được ba phút rồi rời, biểu đồ giữ chân tụt thẳng.

Gốc rễ: Cả kịch bản một tốc độ, không có chỗ nhanh chỗ chậm, không có khoảng lặng.

Chữa ở: Chia kịch bản thành khối và chạy mỗi khối một tốc độ hơi khác nhau. Đoạn cao trào chậm lại, đoạn liệt kê nhanh lên. Chừa khoảng lặng nửa giây trước câu chốt.

02 Câu quá dài, máy đọc không kịp lấy hơi

Bạn thấy: Nghe ngộp, người xem phải căng tai theo, mệt rồi bỏ.

Gốc rễ: Kịch bản viết theo văn viết, câu ba bốn dòng mới chấm.

Chữa ở: Luật cứng: câu nào dài hơn hai dòng thì cắt đôi. Đọc to lên một lượt, chỗ nào bạn hụt hơi thì cắt chỗ đó.

03 Sai trọng âm tên riêng và số liệu

Bạn thấy: Khán giả bản địa nhận ra ngay là giọng máy, mất tin tưởng.

Gốc rễ: Để nguyên tên riêng nước ngoài và con số cho máy tự đoán.

Chữa ở: Viết lại theo cách phát âm trong kịch bản, hoặc chạy riêng câu đó rồi ghép. Với thị trường nước ngoài, nhờ AI kiểm tra cách đọc tên riêng trước.

04 Giọng không khớp nội dung

Bạn thấy: Video kiến thức nghiêm túc mà giọng trẻ tươi tắn — người xem thấy sai sai, không tin nội dung.

Gốc rễ: Chọn giọng theo cảm tính của người làm kênh, không theo tệp khán giả.

Chữa ở: Quay lại Phần 2, chốt bốn thứ: giới tính, tuổi, chất giọng, nhịp điệu. Chốt xong thì giữ nguyên cho cả kênh.

05 Audio to nhỏ không đều giữa các đoạn

Bạn thấy: Người xem phải chỉnh âm lượng giữa chừng — lý do rất phổ biến khiến họ tắt.

Gốc rễ: Ghép nhiều khối audio chạy ở các lần khác nhau, có khối to có khối nhỏ.

Chữa ở: Chạy tất cả các khối trong cùng một lượt với cùng một giọng và cùng thông số. Trước khi xuất, chuẩn hoá âm lượng một lần cho cả file.



CHECKLIST TRƯỚC KHI GHÉP AUDIO VÀO VIDEO

In trang này ra, dán cạnh màn hình. Tick đủ mới ghép.

- ☐ Kịch bản đã xoá hết chú thích, ký tự lạ, chữ viết tắt chưa bung.
- ☐ Câu dài hơn hai dòng đã cắt đôi hết.
- ☐ Số và tên riêng đã viết theo cách đọc đúng.
- ☐ Đã chốt một giọng chính và giữ đúng giọng đó như các video trước.
- ☐ Đã chạy thử ba mươi giây đầu và nghe lại bằng tai nghe.
- ☐ Tốc độ và cao độ giống hết các video trước của kênh.
- ☐ Tất cả các khối chạy cùng một lượt, cùng thông số.
- ☐ Đã nghe soát toàn bộ một lượt không tua.
- ☐ Âm lượng đều từ đầu đến cuối.
- ☐ Có khoảng lặng ngắn giữa các phần lớn.
- ☐ Đã xuất kèm file phụ đề nếu định đăng đa ngôn ngữ.
- ☐ Đã lưu lại thông số giọng vào hồ sơ kênh để video sau dùng lại.

BƯỚC TIẾP THEO

Có giọng đọc rồi thì phần khó nhất của kênh không lộ mặt vẫn còn nguyên: chọn ngách, chọn từ khoá, dựng kịch bản giữ chân và đăng đều tay cho tới lúc kênh chạy. Đó là thứ không có công cụ nào làm hộ được — phải có lộ trình và có người soi.

Đi tiếp cùng chúng tôi

Thử thách xây kênh của Phúc BANI hướng dẫn từng bước theo ngày, làm thật chứ không học suông: **ymm.vn/challenge**. Kho bài viết, prompt và quy trình cập nhật liên tục tại **phuc.vn**. Trợ lý viết kịch bản dựng sẵn: **phucgpt.com**.

Tài liệu này miễn phí, bạn được phép chia sẻ nguyên bản cho bạn bè. Đừng cắt logo rồi bán lại — làm thế thì lần sau chúng tôi không có động lực soạn tiếp.