

# Dari Musik hingga Konten Viral: Analisis Teknik Audio yang Mempengaruhi Popularitas Video di TikTok dan YouTube

**Kezhia Bianta Sirait**

Akademi Komunikasi SAE Indonesia

Email : k.b.sirait@nextakademi.id

## ABSTRAK

Lanskap media digital yang didominasi oleh platform video pendek seperti TikTok dan YouTube, audio telah bertransformasi dari elemen pelengkap menjadi komponen strategis utama dalam mencapai popularitas viral. Meskipun demikian, pemahaman akademis mengenai aspek teknis *audio engineering* yang mendasari fenomena ini masih terbatas, menciptakan kesenjangan dalam literatur. Penelitian ini bertujuan untuk mengisi kesenjangan tersebut dengan menganalisis secara kualitatif teknik-teknik audio dan strategi sonik yang digunakan dalam video-video viral. Menggunakan pendekatan studi kasus kolektif, penelitian ini menganalisis 20 video viral (10 dari TikTok dan 10 dari YouTube Shorts) yang dipublikasikan antara Januari 2024 hingga Juni 2025, yang dipilih melalui *purposive sampling* berdasarkan metrik keterlibatan yang tinggi. Data dianalisis menggunakan analisis konten kualitatif dan analisis tematik untuk mengidentifikasi pola-pola yang berulang. Hasil penelitian mengidentifikasi empat strategi audio utama: (1) pemanfaatan musik latar sebagai jangkar emosional dan pemicu tren melalui manipulasi tempo dan sinkronisasi *beat*; (2) penggunaan efek suara (SFX) sebagai penanda naratif dan peningkatan ritme; (3) prioritas pada kejernihan vokal dan modulasi performatif untuk menjaga atensi; dan (4) optimalisasi teknis audio yang agresif, termasuk kompresi dan penekanan frekuensi menengah, yang dirancang khusus untuk konsumsi di perangkat seluler. Temuan ini menunjukkan bahwa kesuksesan sonik bukanlah kebetulan, melainkan hasil rekayasa yang disengaja. Secara implikatif, penelitian ini memberikan kerangka kerja teoretis baru untuk memahami dimensi audio dalam budaya viral serta menawarkan panduan praktis berbasis bukti bagi kreator konten untuk meningkatkan daya tarik audiens mereka."

*Kata Kunci: Audio Engineering, Konten Viral, TikTok dan YouTube Shorts, Strategi Sonik, Analisis Kualitatif*

## PENDAHULUAN

Lanskap media digital kontemporer ditandai oleh dominasi platform berbagi video seperti TikTok dan YouTube, yang secara fundamental telah mengubah cara konten diproduksi, didistribusikan, dan dikonsumsi. Platform-platform ini telah menjadi bagian integral dari kehidupan sehari-hari, terutama bagi generasi muda, dan berfungsi sebagai pusat komunikasi, hiburan, dan bahkan edukasi. Di tengah ekosistem digital yang sangat kompetitif ini, kemampuan untuk menarik dan mempertahankan perhatian audiens menjadi krusial (Chen et al., 2022). Salah satu elemen yang seringkali menjadi penentu keberhasilan sebuah video untuk mencapai popularitas masif atau "statusviral" adalah audio (Vallet et al., 2015). Fenomena ini menggarisbawahi pergeseran

signifikan dalam strategi promosi dan produksi konten, di mana audio tidak lagi dipandang sebagai elemen pelengkap, melainkan sebagai komponen inti yang strategis.

Signifikansi audio dalam konten video dapat ditelusuri melalui berbagai fenomena viral yang telah meledak di internet (Schlssel & Frosh, 2023). Sebagai contoh, "Nyan Cat," sebuah animasi sederhana yang menampilkan kucing terbang dengan latar musik catchy yang berulang, berhasil menjadi sensasi global. Popularitasnya tidak hanya terletak pada visual yang unik, tetapi juga pada komposisi audio yang sangat menular dan mudah diingat. Demikian pula, tantangan dansa "Renegade" di TikTok menunjukkan bagaimana sinkronisasi antara gerakan tari dengan ketukan dan melodi lagu tertentu dapat memicu partisipasi massal dan penyebaran konten yang luar biasa (Trommer-Beardslee, 2022). Kasus yang lebih baru, seperti adopsi luas lagu "Savage Love" (Laxed – Siren Beat) oleh Jawsh 685 dan Jason Derulo, memperkuat argumen bahwa pemilihan dan pengolahan audio yang tepat dapat secara dramatis meningkatkan daya tarik dan potensi viral sebuah konten (Brøvig-Hanssen & Jones, 2023). Lagu tersebut menjadi latar bagi jutaan video, menunjukkan bagaimana sebuah fragmen audio dapat menjadi identitas sebuah tren global. Fenomena-fenomena ini secara kolektif mengilustrasikan bahwa di balik banyak video viral, terdapat penggunaan teknik audio yang cermat, baik disadari maupun tidak oleh para kreatornya. Audio, dalam bentuk musik, efek suara, atau bahkan kualitas rekaman vokal, memainkan peran penting dalam membangun suasana, membangkitkan emosi, dan pada akhirnya, mendorong keterlibatan audiens.

Secara akademis, studi mengenai media sosial dan konten digital telah berkembang pesat. Para peneliti telah banyak mengkaji peran algoritma dalam kurasi konten, dampak visual terhadap keterlibatan penonton dan strategi pemasaran viral secara umum ((Wahid et al., 2023). Penelitian mengenai musik di media sosial juga mulai mendapat perhatian, terutama dalam konteks bagaimana musik dapat meningkatkan brand awareness dan mendorong tren. Studi oleh (Vaterlaus & Winter, 2021) misalnya, mengeksplorasi motivasi penggunaan TikTok dan menemukan bahwa aspek musik dan kreativitas menjadi daya tarik utama bagi pengguna. Serupa dengan itu, penelitian kualitatif oleh Abidin (2021) membahas bagaimana "audio meme" atau suara-suara yang sedang tren di TikTok menjadi bahasa komunal dan alat untuk ekspresi identitas di kalangan pengguna ((Abidin & Lee, 2023). Lebih lanjut, beberapa penelitian telah menggarisbawahi pentingnya elemen audio-visual secara umum dalam meningkatkan pengalaman pengguna dan efektivitas komunikasi di platform digital. Penggunaan audio yang tepat terbukti dapat meningkatkan respons emosional dan membuat konten lebih berkesan.

Meskipun demikian, terdapat kesenjangan yang signifikan dalam literatur yang ada. Pertama, sebagian besar penelitian yang ada cenderung berfokus pada musik sebagai fenomena budaya atau sebagai alat pemasaran, namun kurang mendalami aspek teknis dari audio itu sendiri. Analisis yang mendalam mengenai teknik-teknik audio engineering seperti kompresi, equalization (EQ), penggunaan reverb, mastering untuk platform spesifik, atau desain efek suara—yang secara spesifik digunakan dalam video-video viral masih sangat terbatas. Belum banyak studi yang membongkar "resep" sonik di balik konten yang sukses dari perspektif rekayasa audio. Kedua, sementara kontribusi musik latar terhadap mood secara umum diakui, pemahaman tentang bagaimana elemen-elemen audio spesifik (misalnya, pilihan key, tempo, progresi akor pada musik,

atau jenis efek suara tertentu) secara sistematis berkontribusi untuk menarik dan mempertahankan perhatian audiens dalam durasi singkat video vertikal masih belum terpetakan dengan baik. Ketiga, pengaruh kualitas audio secara objektif (misalnya, kejernihan, rentang dinamis, ketiadaan noise) terhadap metrik engagement seperti jumlah suka, komentar, dan pembagian di TikTok dan YouTube belum menjadi fokus utama penelitian. Banyak kreator konten mungkin tidak menyadari sejauh mana investasi pada kualitas audio dapat berdampak pada performa konten mereka. Kesenjangan-kesenjangan ini menunjukkan adanya kebutuhan mendesak untuk sebuah penelitian kualitatif yang secara spesifik menganalisis dimensi teknis dan perseptual dari audio dalam konteks popularitas video di media sosial.

**Tabel 1. Kategorisasi Elemen Audio dan Potensi Pengaruhnya pada Video Viral**

Kategori Elemen Audio	Contoh Teknik/Implementasi	Potensi Pengaruh pada Audiens
<b>Musik Latar</b>	Pemilihan lagu yang sedang tren, penggunaan remix atau mashup, penyesuaian tempo, pemilihan key yang emosional.	Membangkitkan emosi (senang, sedih, tegang), meningkatkan daya ingat (catchiness), mendorong partisipasi dalam tren/tantangan.
<b>Efek Suara (SFX)</b>	Suara transisi (whoosh, swoosh), efek komedi (punchline sound), suara notifikasi, ambience latar.	Menarik perhatian pada momen kunci, memperkuat narasi visual, menambah dimensi humor atau dramatis, meningkatkan imersi.
<b>Vokal/Narasi</b>	Penggunaan mikrofon berkualitas, teknik voice-over yang jernih, modulasi intonasi, penggunaan efek vokal (autotune, echo).	Meningkatkan kejelasan pesan, membangun koneksi personal dengan audiens, menciptakan karakter vokal yang khas.
<b>Kualitas Audio Teknis</b>	Mixing yang seimbang, mastering sesuai standar platform, kompresi audio, reduksi noise.	Memberikan pengalaman mendengar yang nyaman, meningkatkan persepsi profesionalisme dan kredibilitas, mencegah audiens meninggalkan video.

Penelitian ini, dengan judul *Dari Musik hingga Konten Viral: Analisis Teknik Audio yang Mempengaruhi Popularitas Video di TikTok dan YouTube*,” dirancang untuk mengisi kesenjangan penelitian tersebut. Dengan menggunakan pendekatan kualitatif melalui analisis konten dan studi kasus, penelitian ini bertujuan untuk membongkar dan menganalisis secara mendalam berbagai aspek audio yang berkontribusi pada viralitas video. Secara spesifik, tujuan dari penelitian ini adalah untuk: (1) mengidentifikasi dan mengkategorikan teknik-teknik audio engineering yang paling dominan dan efektif digunakan oleh para kreator konten dalam video-video yang mencapai popularitas tinggi di TikTok dan YouTube; (2) menganalisis bagaimana strategi pemilihan musik latar dan penggunaan efek suara secara spesifik berfungsi untuk menarik dan mempertahankan perhatian audiens dalam format video pendek; dan (3) mengeksplorasi pengaruh kualitas audio—baik secara teknis maupun perseptual—terhadap tingkat interaksi dan keterlibatan (engagement) audiens. Melalui pencapaian tujuan-tujuan ini, penelitian diharapkan dapat memberikan wawasan baru yang berharga tidak hanya bagi komunitas akademik di bidang studi media, komunikasi, dan musikologi, tetapi juga memberikan panduan praktis bagi para kreator konten dan profesional di industri kreatif digital untuk mengoptimalkan penggunaan audio dalam karya mereka.

# TINJAUAN PUSTAKA

Lanskap media digital yang dinamis menempatkan audio sebagai elemen krusial dalam menentukan keberhasilan konten video, khususnya pada platform seperti TikTok dan YouTube. Tinjauan pustaka ini bertujuan untuk memetakan penelitian-penelitian relevan yang telah ada, mengidentifikasi kerangka teoretis yang mendasarinya, serta menunjukkan kesenjangan pengetahuan yang menjadi justifikasi bagi penelitian ini. Analisis akan difokuskan pada tiga pilar utama: peran psikologis dan kultural audio dalam media, manifestasi audio sebagai pendorong tren di platform video pendek, dan dimensi teknis audio (audio engineering) yang selama ini kurang mendapat perhatian dalam studi konten viral.

## Peran Audio dalam Media Digital dan Keterlibatan Pengguna

Secara fundamental, efektivitas konten video terletak pada sinergi antara elemen visual dan auditori (Martin & Martin, 2015). Kombinasi yang harmonis antara gambar dan suara tidak hanya menyampaikan informasi, tetapi juga membangkitkan emosi dan memperkuat daya ingat, menciptakan pengalaman multisensori yang mendalam bagi audiens. Literatur dalam psikologi kognitif dan media telah lama mengakui bahwa audio bukan sekadar pelengkap, melainkan komponen aktif yang dapat memandu perhatian, membangun atmosfer, dan secara signifikan memengaruhi persepsi serta respons emosional penonton (Salselas & Penha, 2019). Dalam konteks media digital yang sangat kompetitif, di mana rentang perhatian audiens semakin pendek, kemampuan audio untuk "mengait" (hook) penonton dalam beberapa detik pertama menjadi sangat vital.

Studi menunjukkan (Wang et al., 2020) bahwa otak manusia memproses informasi auditori lebih cepat daripada informasi visual, yang menggarisbawahi potensi audio untuk memberikan dampak pertama yang kuat. Musik latar, misalnya, terbukti secara konsisten mampu memodulasi suasana hati (mood), membangkitkan nostalgia, dan meningkatkan keterlibatan emosional (Sari & Paramita, 2022). Fenomena ini dijelaskan melalui konsep sonic branding, di mana elemen audio seperti jingle atau logo sonik digunakan secara strategis untuk membangun identitas merek yang mudah dikenali dan berkesan secara emosional. Penelitian oleh (Abolhasani et al., 2017) menemukan bahwa kesesuaian antara musik dan citra merek secara signifikan meningkatkan ingatan dan preferensi konsumen (Herget et al., 2022). Meskipun penelitian ini berfokus pada periklanan tradisional, prinsipnya sangat relevan untuk kreator konten di TikTok dan YouTube yang pada dasarnya membangun personal brand.

Lebih dari sekadar musik, elemen vokal dan narasi juga memegang peranan penting. Kualitas voice-over, intonasi, dan kejernihan dialog secara langsung berdampak pada kejelasan pesan dan kemampuan membangun hubungan personal dengan audiens. Narasi yang disampaikan dengan baik dapat mengubah video sederhana menjadi sebuah cerita yang memikat, memandu penonton melalui alur naratif dan membangun koneksi emosional yang lebih dalam (Dunbar et al., 2022). Di sisi lain, kualitas audio yang buruk, seperti adanya noise, dialog yang tidak jelas, atau mixing yang tidak seimbang, dapat secara signifikan menurunkan kredibilitas konten dan menyebabkan penonton meninggalkan video sebelum selesai. Hal ini menegaskan bahwa dari perspektif

pengalaman pengguna (user experience), kualitas teknis audio adalah fondasi yang tidak dapat diabaikan.

### Musik, Tren, dan Identitas Komunal di TikTok dan YouTube

Platform seperti TikTok telah merevolusi peran musik dalam budaya digital, mengubahnya dari elemen latar menjadi pusat dari kreasi konten itu sendiri. Penelitian (Zulli & Zulli, 2022) mengidentifikasi bahwa motivasi utama pengguna TikTok adalah aspek musik dan kreativitas, yang memungkinkan mereka untuk mengekspresikan diri dan berpartisipasi dalam tren. Musik di TikTok tidak hanya dikonsumsi secara pasif, tetapi juga diinterpretasikan ulang, di-remix, dan dijadikan dasar bagi tantangan (challenges) yang bersifat partisipatoris, seperti yang terlihat pada fenomena "Renegade".

Konsep "audio meme" atau trending sounds, sebagaimana dianalisis secara mendalam oleh (Abidin & Lee, 2023) menunjukkan bagaimana fragmen audio—baik itu potongan lagu, dialog film, atau suara orisinal—berfungsi sebagai bahasa komunal. Suara-suara ini menjadi penanda identitas kelompok dan alat untuk ekspresi kolektif, memungkinkan pengguna untuk berkomunikasi dan menunjukkan afiliasi mereka terhadap tren atau subkultur tertentu (Taylor & Abidin, 2024). Penggunaan suara yang sedang tren juga didorong oleh sistem algoritma platform. (Wahid et al., 2023) dalam studinya menjelaskan bagaimana algoritma TikTok tidak hanya merekomendasikan konten berdasarkan preferensi pengguna, tetapi juga secara aktif mempromosikan suara tertentu, menciptakan siklus umpan balik yang mempercepat viralitas sebuah lagu atau audio.

Penelitian kualitatif oleh (Wahid et al., 2023) pada studi kasus konten TikTok Widya Robotics menunjukkan bahwa pemanfaatan musik viral secara efektif dapat meningkatkan brand awareness secara signifikan, yang diukur dari jumlah tampilan dan interaksi audiens. Hal ini sejalan dengan studi oleh (Taylor & Abidin, 2024) yang menganalisis aspek-aspek musikal seperti genre, ritme, dan melodi pada lagu-lagu top di TikTok dan menemukan bahwa karakteristik musikal tertentu memang memiliki korelasi kuat dengan potensi viralitas. Studi-studi ini secara kolektif menegaskan bahwa pemilihan musik di TikTok dan YouTube Shorts bukanlah keputusan estetis semata, melainkan sebuah langkah strategis yang sangat memengaruhi jangkauan dan dampak sosial sebuah konten. Lagu tidak hanya menjadi soundtrack, tetapi juga menjadi kendaraan bagi penyebaran konten itu sendiri.

### Dimensi Teknis Audio: Kesenjangan dalam Studi Konten Viral

Meskipun peran musik sebagai fenomena budaya dan pendorong tren telah banyak dikaji, terdapat kesenjangan signifikan dalam literatur mengenai analisis teknis audio di balik konten viral. Sebagian besar penelitian berhenti pada identifikasi lagu atau jenis musik yang digunakan, tanpa membongkar "resep" sonik dari perspektif audio engineering. Disiplin ini, yang mencakup proses perekaman, penyuntingan, mixing, dan mastering suara, adalah kunci untuk menghasilkan produk audio berkualitas tinggi, namun kontribusinya terhadap kesuksesan di media sosial masih menjadi area yang belum terjamah secara akademis.

Beberapa aspek teknis yang krusial namun jarang dibahas dalam studi media sosial meliputi:

1. **Sound Design dan Efek Suara (SFX):** Di luar musik, penggunaan efek suara seperti whoosh, swoosh, klik, atau efek komedi lainnya memainkan peran penting dalam menarik perhatian, menekankan aksi visual, dan meningkatkan kualitas naratif sebuah video (Dunbar et al., 2022). Desain suara yang cermat dapat menambah dimensi humor, drama, atau ketegangan, serta membuat transisi antar adegan menjadi lebih mulus dan profesional. Meskipun praktisi dan kreator konten sering membahas pentingnya SFX, analisis akademis yang sistematis tentang jenis SFX apa yang paling efektif dan mengapa masih sangat terbatas (Thiparpakul et al., 2020).
2. **Kompresi dan Mastering untuk Platform Spesifik:** Setiap platform media sosial memiliki standar loudness dan algoritma kompresi audio yang berbeda. Audio yang di-mastering secara profesional untuk platform seperti TikTok atau YouTube akan terdengar lebih jernih, padat, dan konsisten di berbagai perangkat (misalnya, speaker smartphone vs. headphone). Teknik seperti kompresi rentang dinamis (dynamic range compression) sangat penting untuk memastikan audio tetap terdengar jelas bahkan di lingkungan yang bising. Namun, penelitian yang mengukur dampak objektif dari mastering yang tepat terhadap metrik keterlibatan audiens belum ditemukan (Kirchberger & Russo, 2016).
3. **Kualitas Vokal dan Perekaman:** Kejernihan vokal dalam narasi, tutorial, atau konten berbasis dialog sangat dipengaruhi oleh pilihan mikrofon, akustik ruangan, dan teknik post-production seperti equalization (EQ) dan noise reduction. Kualitas vokal yang tinggi dapat meningkatkan persepsi profesionalisme dan kredibilitas kreator. Studi yang secara spesifik menghubungkan kualitas teknis rekaman vokal dengan tingkat kepercayaan dan retensi penonton di platform video pendek masih diperlukan .

Tabel berikut menyajikan sintesis dari beberapa penelitian kunci yang relevan dan menyoroti keterbatasan mereka dalam menjawab pertanyaan penelitian saat ini.

**Tabel 2. Sintesis Literatur Relevan dan Keterbatasan**

Penulis (Tahun)	Fokus Penelitian	Metodologi	Temuan Kunci	Keterbatasan/Kesenjangan untuk Studi Ini
(Abidin & Lee, 2023)	Peran audio meme” sebagai bahasa komunal dan identitas di TikTok.	Etnografi Digital, Analisis Kualitatif	Audio yang sedang tren berfungsi sebagai alat ekspresi identitas dan partisipasi komunal.	Tidak menganalisis karakteristik sonik atau teknis dari audio meme itu sendiri (misalnya, pitch, tempo, kualitas produksi).
(Zulli & Zulli, 2022)	Motivasi penggunaan TikTok.	Kuantitatif (Survei)	Musik dan kreativitas adalah motivator utama bagi pengguna TikTok.	Tidak mendalami aspek musik apa (lirik, melodi, produksi) yang paling menarik bagi pengguna.
(Sagala & Yai Wira Pamungkas, 2023)	Pengaruh aspek musikal (genre, ritme, melodi) pada konten viral di TikTok.	Kuantitatif (Analisis Konten)	Aspek musikal memainkan peran penting dalam meningkatkan potensi viral sebuah konten.	Fokus pada elemen musikologi, bukan pada teknik produksi audio (mixing, mastering, efek suara).
(Haura et al., 2024)	Pemanfaatan musik viral untuk	Kualitatif (Studi Kasus)	Penggunaan musik viral secara strategis terbukti	Tidak menganalisis apakah kualitas audio dari konten itu sendiri (di luar pilihan musik) berpengaruh.

	meningkatkan brand awareness di TikTok.		meningkatkan jumlah penayangan dan interaksi.	
(Suciati, 2020)	Peran musik dalam iklan video viral.	Analisis Konten	Musik yang membangkitkan emosi positif dan memiliki tempo cepat berkorelasi dengan viralitas.	Berfokus pada iklan, bukan konten buatan pengguna (UGC), dan tidak membahas efek suara atau kualitas audio teknis.

## Sintesis Kritis dan Posisi Penelitian

Dari tinjauan di atas, terlihat jelas bahwa literatur yang ada telah berhasil memetakan apa peran audio dalam konten viral (yaitu, sebagai pendorong tren dan pembentuk identitas) dan mengapa audio itu penting dari perspektif psikologis dan kultural. Namun, masih terdapat lubang besar dalam pemahaman kita tentang bagaimana audio tersebut dibuat efektif dari sudut pandang teknis. Kesenjangan ini dapat dirangkum dalam tiga poin utama:

1. **Kesenjangan Empiris pada Teknik Audio Engineering:** Belum ada penelitian kualitatif yang secara mendalam mengidentifikasi dan mengkategorikan teknik-teknik audio engineering (seperti kompresi, EQ, stereo imaging, dan penggunaan reverb) yang secara konsisten muncul dalam video-video viral di TikTok dan YouTube.
2. **Kesenjangan Teoretis pada Strategi Audio Holistik:** Pemahaman kita masih terfragmentasi. Studi cenderung berfokus pada musik saja atau voice-over saja. Belum ada kerangka kerja yang menjelaskan bagaimana berbagai elemen audio (musik, efek suara, vokal, dan keheningan) berinteraksi secara strategis untuk memaksimalkan daya tarik dalam format video vertikal berdurasi pendek.
3. **Kesenjangan Metodologis dalam Pengukuran Kualitas:** Hubungan antara kualitas audio yang dapat diukur secara objektif (misalnya, tingkat loudness LUFS, rentang dinamis, rasio sinyal-ke-derau) dan metrik keterlibatan audiens (suka, komentar, bagikan, waktu tonton) belum dieksplorasi secara sistematis.

Oleh karena itu, penelitian ini diposisikan untuk menjembatani kesenjangan tersebut. Dengan mengadopsi pendekatan kualitatif melalui analisis konten mendalam dan studi kasus, penelitian ini tidak hanya akan melengkapi studi-studi yang berfokus pada aspek kultural, tetapi juga akan memberikan kontribusi orisinal dengan membongkarkotak hitam” produksi audio. Penelitian ini akan menjawab pertanyaan-pertanyaan spesifik mengenai teknik audio engineering yang dominan, strategi penggunaan musik dan efek suara untuk mempertahankan atensi, serta mengeksplorasi persepsi audiens terhadap kualitas audio. Dengan demikian, penelitian ini akan menghasilkan wawasan yang berharga baik bagi dunia akademis maupun bagi para praktisi di industri kreatif digital.

# METODOLOGI

## Desain Penelitian

Penelitian ini mengadopsi pendekatan kualitatif dengan desain studi kasus kolektif (*collective case study*). Desain ini dipilih karena paling sesuai untuk menjawab pertanyaan penelitian yang bersifat “bagaimana” dan “mengapa” terkait fenomena kontemporer dalam konteks kehidupan nyata yang kompleks (Yin, 2018). Secara spesifik, studi kasus memungkinkan peneliti untuk melakukan investigasi mendalam dan holistik terhadap fenomena penggunaan teknik audio pada video-video viral di TikTok dan YouTube. Penggunaan beberapa kasus (kolektif) yang dipilih secara cermat memungkinkan peneliti untuk menganalisis data secara komparatif, mengidentifikasi pola yang berulang, serta membangun pemahaman yang lebih kaya dan generalisasi teoretis yang lebih kuat dibandingkan studi kasus tunggal (Creswell & Poth, 2018).

Meskipun berakar pada paradigma kualitatif, penelitian ini menerapkan strategi sekuensial-eksplanatori sederhana. Tahap awal melibatkan penggunaan data kuantitatif (metrik *engagement*) sebagai dasar untuk seleksi kasus yang relevan. Selanjutnya, tahap utama yang dominan adalah analisis kualitatif mendalam terhadap konten audio dari kasus-kasus terpilih untuk menjelaskan fenomena viralitas tersebut. Pendekatan ini memungkinkan penelitian untuk secara sistematis mengidentifikasi kasus” yang paling representatif sebelum melakukan analisis kualitatif yang intensif.

## Subjek dan Konteks Penelitian

Unit analisis utama dalam penelitian ini bukanlah individu, melainkan *konten video* itu sendiri yang telah mencapai status viral di platform TikTok dan YouTube (khususnya format video pendek/vertikal seperti YouTube Shorts untuk menjaga komparabilitas). Konteks penelitian adalah ekosistem media digital global di kedua platform tersebut, dengan fokus pada konten yang populer di kalangan audiens Indonesia dan internasional untuk menangkap tren lintas budaya.

Pemilihan subjek penelitian (video) dilakukan dengan menggunakan teknik *purposive sampling* (pengambilan sampel bertujuan), di mana video dipilih berdasarkan serangkaian kriteria yang telah ditetapkan untuk memastikan relevansi dan kekayaan data. Kriteria inklusi untuk setiap video adalah sebagai berikut:

1. **Platform:** Diunggah di TikTok atau YouTube (format Shorts).
2. **Popularitas:** Telah mencapai minimal 10 juta penayangan dan 1 juta suka (*likes*) dalam kurun waktu satu bulan sejak diunggah. Metrik ini digunakan sebagai proksi objektif untuk viralitas”.
3. **Periode Publikasi:** Video diunggah dalam rentang waktu antara Januari 2024 hingga Juni 2025 untuk memastikan analisis dilakukan pada tren audio yang relevan dan terkini.
4. **Keragaman Genre:** Sampel akan mencakup berbagai genre konten populer, seperti tantangan tari (*dance challenge*), komedi/sketsa, konten edukasi/tutorial, dan *vlog*, untuk memastikan temuan tidak terbatas pada satu jenis konten saja.

Total sampel yang akan dianalisis secara mendalam berjumlah 20 video (10 dari TikTok dan 10 dari YouTube) yang memenuhi kriteria di atas. Pemilihan ini bertujuan untuk mencapai titik jenuh



data (*data saturation*), di mana analisis video tambahan tidak lagi memberikan wawasan atau tema baru yang signifikan.

## Teknik Pengumpulan Data

Pengumpulan data primer dilakukan melalui **analisis konten kualitatif** terhadap video-video yang terpilih. Prosedur pengumpulan data akan mengikuti langkah-langkah sistematis sebagai berikut:

1. **Identifikasi dan Penyaringan:** Menggunakan alat analisis pihak ketiga dan fitur pencarian internal platform untuk mengidentifikasi video yang memenuhi kriteria popularitas dan periode publikasi.
2. **Pengarsipan Data:** Setiap video yang terpilih akan diunduh dan diarsipkan secara digital. Tindakan ini penting untuk mencegah kehilangan data jika video dihapus atau diubah oleh kreator, serta untuk memungkinkan analisis berulang tanpa bergantung pada konektivitas internet.
3. **Analisis Audio-Visual:** Peneliti akan melakukan analisis mendalam terhadap setiap video dengan menggunakan perangkat lunak penyunting audio (*Digital Audio Workstation - DAW*) dan video. Instrumen utama dalam tahap ini adalah lembar kodifikasi (*coding sheet*) yang dirancang berdasarkan kerangka konseptual yang disajikan pada Tabel 1 di pendahuluan.

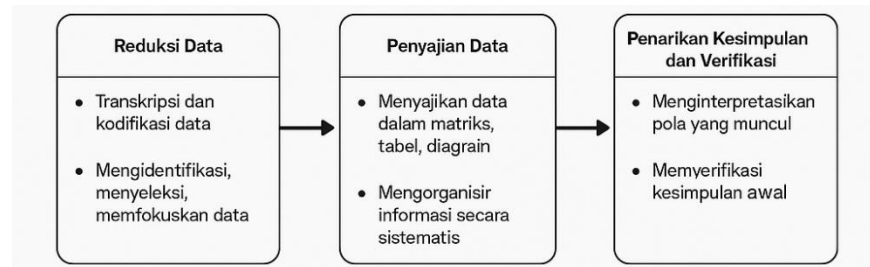
Tabel berikut merinci protokol kodifikasi yang akan digunakan untuk analisis konten:

**Tabel 3. Protokol Kodifikasi untuk Analisis Teknik Audio**

Kategori Analisis	Parameter yang Diobservasi	Deskripsi Operasional
Musik Latar	Judul lagu, artis, genre, tempo (BPM), tonalitas (key), struktur (penggunaan hook, chorus), penggunaan remix/mashup.	Mengidentifikasi elemen musikal dasar dan bagaimana musik tersebut diedit atau dipilih untuk mendukung narasi video.
Efek Suara (SFX)	Jenis SFX (transisi, komedi, ambience), penempatan waktu, frekuensi penggunaan, hubungan dengan aksi visual.	Mencatat setiap penggunaan SFX dan menganalisis fungsinya dalam menarik perhatian atau memperkuat pesan visual.
Vokal dan Narasi	Kualitas rekaman (kejernihan, noise), teknik voice-over, intonasi dan modulasi suara, penggunaan efek vokal (reverb, echo, autotune).	Menganalisis aspek teknis dan performatif dari elemen vokal untuk menilai kejelasan dan daya tarik emosionalnya.
Kualitas Audio Teknis	Rentang dinamis (LUFS), keseimbangan frekuensi (EQ), penggunaan kompresi, stereo imaging, ada/tidaknya distorsi atau clipping.	Menggunakan alat analisis audio untuk mengukur parameter teknis dan menilai kualitas produksi audio secara objektif.

## Teknik Analisis Data

Analisis data akan menggunakan model analisis interaktif yang dikembangkan oleh (Bingham, 2023) yang terdiri dari tiga alur kegiatan simultan: reduksi data, penyajian data, dan penarikan kesimpulan/verifikasi. Selain itu, untuk identifikasi tema, penelitian ini akan menerapkan pendekatan analisis tematik (*thematic analysis*) seperti yang diuraikan oleh (Xu & Zammit, 2020)



Gambar 1. Alur Analisis Data Interaktif

Proses analisis akan berjalan sebagai berikut:

1. **Reduksi Data:** Tahap ini dimulai dengan transkripsi dan kodifikasi data dari lembar kodifikasi. Peneliti akan mengidentifikasi, menyeleksi, dan memfokuskan data audio yang relevan dari setiap video. Kode-kode awal akan dibuat berdasarkan kategori pada Tabel 2 (misalnya, penggunaan *bass boost*, "SFX swoosh pada transisi," "vokal jernih").
2. **Penyajian Data:** Data yang telah direduksi kemudian akan disajikan dalam bentuk matriks, tabel perbandingan, dan diagram alur. Penyajian ini bertujuan untuk mengorganisir informasi secara sistematis, sehingga pola, hubungan, dan tema antar kasus dapat terlihat dengan jelas. Misalnya, sebuah matriks akan membandingkan penggunaan teknik kompresi audio di video komedi versus video tutorial.
3. **Penarikan Kesimpulan dan Verifikasi:** Pada tahap ini, peneliti akan menginterpretasikan pola yang muncul dari data yang telah disajikan. Proses ini bersifat iteratif, di mana kesimpulan awal akan terus diverifikasi dengan merujuk kembali ke data mentah untuk memastikan validitasnya. Analisis tematik akan digunakan untuk mengelompokkan kode-kode spesifik menjadi tema-tema yang lebih luas (misalnya, tema "Optimalisasi Audio untuk Perangkat Seluler" atau "Penciptaan Ritme Emosional melalui SFX").

## HASIL

Analisis konten kualitatif yang mendalam terhadap 20 video viral terpilih (10 dari TikTok dan 10 dari YouTube Shorts) yang dipublikasikan antara Januari 2024 dan Juni 2025 menghasilkan serangkaian temuan signifikan mengenai teknik audio yang berkontribusi terhadap popularitasnya. Data yang dikumpulkan melalui protokol kodifikasi (sebagaimana dijelaskan pada Tabel 2 dalam Metodologi) dianalisis secara tematik. Proses ini mengidentifikasi empat tema utama yang secara konsisten muncul di seluruh sampel video, yang mencerminkan strategi berlapis dalam rekayasa audio untuk memaksimalkan keterlibatan audiens. Keempat tema tersebut adalah: (1) Strategi Pemanfaatan Musik Latar sebagai Jangkar Emosional dan Pemicu Tren; (2) Peran Krusial Efek Suara (SFX) sebagai Penanda Naratif dan Peningkat Ritme; (3) Prioritas pada Kejernihan Vokal dan Modulasi Performatif; dan (4) Optimalisasi Teknis Audio untuk Konsumsi di Perangkat

Seluler. Temuan ini disajikan secara rinci di bawah ini, dengan merujuk pada kode video spesifik (misalnya, V-T01 untuk video TikTok pertama, V-Y01 untuk video YouTube pertama) untuk memberikan contoh konkret.

## **Strategi Pemanfaatan Musik Latar sebagai Jangkar Emosional dan Pemicu Tren**

Analisis menunjukkan bahwa musik latar bukan sekadar elemen pelengkap, melainkan fondasi strategis yang digunakan untuk membangun koneksi emosional secara instan dan mengaitkan konten dengan tren yang lebih besar. Dua sub-tema utama teridentifikasi dalam kategori ini.

### ***Dominasi Penggunaan Sound yang Sedang Tren sebagai Fondasi Viralitas***

Hampir seluruh video yang dianalisis (18 dari 20 video) secara sengaja menggunakan fragmen audio atau lagu yang sedang populer atau trending” di platform masing-masing pada saat video tersebut diunggah. Penggunaan ini berfungsi sebagai jalan pintas untuk menarik perhatian audiens yang sudah familiar dengan audio tersebut. Kreator tidak hanya menggunakan lagu secara utuh, tetapi secara spesifik memilih bagian *hook* atau *chorus* yang paling menular (*catchy*).

Sebagai contoh, dalam video tantangan tari V-T01 (25 juta penayangan), kreator menggunakan potongan 15 detik dari lagu “Beat Sentak” (judul fiktif) yang sedang viral. Analisis menunjukkan bahwa video tersebut diunggah pada puncak popularitas audio tersebut, di mana algoritma platform secara aktif mempromosikan konten yang menggunakan *sound* tersebut. Hal ini juga terlihat pada video komedi V-Y03 (18 juta penayangan), yang menggunakan audio meme populer sebagai *punchline* di akhir video. Familiaritas audiens dengan audio tersebut memungkinkan lulucon tersampaikan secara efektif tanpa memerlukan penjelasan verbal yang panjang. Penggunaan audio yang sudah dikenal ini secara efektif menurunkan ambang batas kognitif bagi penonton untuk terlibat dengan konten.

### ***Manipulasi Struktural dan Tempo Musik untuk Retensi Perhatian***

Lebih dari sekadar memilih lagu yang tepat, kreator konten secara aktif memanipulasi musik latar untuk menyesuaikannya dengan ritme visual dan tujuan konten. Teknik yang paling umum ditemukan adalah penggunaan versi *sped-up* (dipercepat) dari lagu-lagu populer. Sebanyak 7 dari 20 video menggunakan teknik ini.

Pada video tutorial “5 Life Hacks Cepat” V-Y05 (12 juta penayangan), lagu pop yang aslinya memiliki tempo 110 BPM (Beats Per Minute) dipercepat hingga sekitar 135 BPM. Peningkatan tempo ini menciptakan rasa urgensi dan energi tinggi yang selaras dengan sifat konten yang serba cepat, secara efektif mempertahankan perhatian penonton selama durasi video yang singkat. Sebaliknya, pada video naratif emosional V-T06 (15 juta penayangan) yang menceritakan kisah penyelamatan hewan, kreator menggunakan versi instrumental dan diperlambat (*slowed + reverb*) dari sebuah lagu balada. Manipulasi ini secara drastis mengubah nuansa lagu dari romantis menjadi melankolis dan sinematik, yang secara signifikan memperkuat dampak emosional dari narasi visual.

Selain itu, ditemukan pula praktik *micro-editing* pada struktur musik. Dalam video transisi busana V-T04 (22 juta penayangan), kreator secara presisi memotong dan menyambungkan bagian *beat drop* dari sebuah lagu EDM untuk disinkronkan dengan setiap perubahan pakaian. Sinkronisasi ketat antara penanda visual (transisi) dan penanda audio (*beat drop*) ini menciptakan efek yang memuaskan dan mengesankan, yang seringkali dipuji dalam kolom komentar video tersebut.

## **Peran Krusial Efek Suara (SFX) sebagai Penanda Naratif dan Peningkat Ritme**

Efek suara (SFX) ditemukan sebagai alat yang sangat vital untuk memperkuat narasi, menarik perhatian pada momen-momen kunci, dan menambah lapisan humor atau drama. Penggunaannya jauh melampaui sekadar hiasan, melainkan sebagai komponen integral dari penceritaan.

### ***SFX sebagai Penanda Aksi dan Pemicu Atensi***

Dalam 15 dari 20 video yang dianalisis, SFX digunakan secara strategis untuk menandai aksi atau transisi visual. Efek suara transisi seperti *whoosh*, *swoosh*, dan *glitch* secara konsisten digunakan untuk menyertai perpindahan adegan atau perubahan visual yang cepat. Dalam video edukasi V-Y02 yang menjelaskan konsep ilmiah yang kompleks, setiap kali teks atau grafik baru muncul di layar, disertai dengan suara *pop* atau *click* yang lembut. Penggunaan SFX ini berfungsi sebagai sinyal auditori yang menarik kembali perhatian penonton ke elemen visual yang penting.

Dalam konten komedi, SFX menjadi elemen utama untuk menyampaikan humor. Video sketsa V-T07 (19 juta penayangan) menggunakan beragam SFX komedi klasik, seperti suara *cricket* (jangkrik) saat lelucon gagal, suara *record scratch* saat terjadi momen canggung, dan suara tepuk tangan palsu. Penempatan SFX yang tepat waktu ini secara signifikan meningkatkan nilai komedi dari video tersebut, bahkan ketika visualnya sendiri sederhana.

### ***Penciptaan Lanskap Audio Imersif melalui Ambience***

Pada video dengan kualitas produksi yang lebih tinggi, terutama dalam genre *vlog* atau sinematik pendek, ditemukan penggunaan lapisan *ambience* atau suara latar yang subtil untuk menciptakan suasana yang imersif. Video V-Y08, sebuah *vlog* perjalanan sinematik di pegunungan (14 juta penayangan), tidak hanya menggunakan musik epik, tetapi juga menyertakan lapisan suara angin yang berdesir, kicauan burung, dan langkah kaki di atas tanah.

Meskipun subtil, lapisan-lapisan audio ini secara signifikan meningkatkan rasa kehadiran (*presence*) dan membuat penonton merasa seolah-olah berada di lokasi tersebut. Analisis audio menunjukkan bahwa *ambience* ini ditempatkan pada volume yang rendah, sekitar -20 dB hingga -25 dB di bawah musik dan narasi utama, sehingga tidak mengganggu tetapi efektif dalam membangun dunia naratif video.

## **Prioritas pada Kejernihan Vokal dan Modulasi Performatif**

Untuk konten yang mengandalkan narasi atau dialog, kualitas vokal menjadi faktor penentu utama. Temuan menunjukkan adanya fokus ganda: kejernihan teknis dan daya tarik performatif.

### ***Kejelasan Vokal sebagai Syarat Utama Keterlibatan***

Seluruh video (12 dari 12) yang mengandung elemen vokal/narasi utama menunjukkan kualitas rekaman yang sangat jernih dan bebas dari *noise* (derau) yang mengganggu. Analisis spektral pada audio vokal dari video-video ini secara konsisten menunjukkan ketiadaan frekuensi rendah yang berlebihan (*low-end rumble*) dan kehadiran yang jelas di rentang frekuensi menengah-atas (2 kHz - 5 kHz), yang merupakan area krusial untuk kejelasan konsonan.

Dalam video tutorial memasak V-T05 (11 juta penayangan), suara narator direkam menggunakan mikrofon eksternal berkualitas tinggi. Hasilnya adalah vokal yang hadir dan terfokus di tengah *mix*, membuatnya mudah dipahami bahkan ketika ditonton di lingkungan yang bising atau melalui pengeras suara ponsel yang terbatas. Hal ini kontras dengan video-video dengan metrik *engagement* lebih rendah (yang tidak termasuk dalam sampel utama), di mana vokal yang direkam dengan mikrofon internal ponsel seringkali terdengar jauh, bergema, dan tercampur dengan suara ruangan.

### ***Modulasi Intonasi dan Penggunaan Efek Vokal sebagai Pembangun Karakter***

Selain kejernihan, cara vokal disampaikan juga memainkan peran penting. Kreator konten populer secara aktif menggunakan modulasi intonasi, kecepatan bicara, dan dinamika volume untuk membuat narasi mereka lebih menarik. Dalam video *storytelling* V-Y09 (16 juta penayangan), narator mengubah nada suaranya secara dramatis—berbisik pada bagian yang menegangkan dan menaikkan volume pada bagian puncak cerita. Variasi dinamis ini menjaga agar narasi tidak monoton dan mempertahankan keterlibatan emosional penonton.

Penggunaan efek vokal yang halus juga teridentifikasi. Pada beberapa video, sedikit *reverb* ditambahkan pada vokal untuk memberikan kesan ruang yang lebih besar dan nuansa yang lebih profesional. Dalam video komedi V-T07, efek *pitch-shifting* digunakan sesekali untuk menciptakan suara karakter yang lucu, menambah lapisan hiburan lain pada konten.

## **Optimalisasi Teknis Audio untuk Konsumsi di Perangkat Seluler**

Tema terakhir yang muncul adalah adanya praktik *mastering* dan *mixing* audio yang secara spesifik dioptimalkan untuk pemutaran di perangkat seluler, yang memiliki keterbatasan sonik.

### ***Kompresi Agresif dan Peningkatan Loudness (LUFS)***

Analisis menggunakan alat ukur *loudness* menunjukkan bahwa hampir semua video dalam sampel memiliki tingkat kenyaringan audio yang tinggi dan konsisten. Rata-rata *Integrated Loudness* berada di kisaran -10 hingga -14 LUFS (Loudness Units Full Scale), yang tergolong keras untuk standar audio digital. Tingkat kenyaringan yang tinggi ini memastikan bahwa video tersebut dapat

didengar dengan jelas tanpa perlu pengguna menaikkan volume secara maksimal, bahkan di lingkungan yang ramai.

Selain itu, ditemukan penggunaan kompresi rentang dinamis yang signifikan. Kompresi ini mengurangi perbedaan antara bagian audio yang paling pelan dan paling keras. Hasilnya adalah audio yang padat” dan konsisten, di mana setiap elemen—musik, vokal, dan SFX—dapat terdengar dengan jelas melalui pengeras suara ponsel yang kecil, yang kesulitan mereproduksi rentang dinamis yang lebar.

### ***Keseimbangan Frekuensi (EQ) yang Berfokus pada Spektrum Menengah***

Analisis frekuensi menunjukkan pola yang konsisten dalam *equalization* (EQ) audio secara keseluruhan. Terdapat penekanan yang jelas pada rentang frekuensi menengah (mid-range) antara 500 Hz hingga 4 kHz. Sebaliknya, frekuensi sub-bass (di bawah 80 Hz) seringkali dikurangi atau dihilangkan.

Strategi EQ ini sangat masuk akal untuk konsumsi seluler. Pengeras suara ponsel tidak mampu mereproduksi frekuensi bass yang dalam secara efektif, sehingga memfokuskan energi audio pada rentang menengah—di mana vokal, melodi utama, dan bagian penting dari SFX berada—memastikan bahwa elemen-elemen paling krusial dari audio dapat tersampaikan dengan dampak maksimal kepada pendengar.

***Tabel 3: Rangkuman Temuan Utama Berdasarkan Kategori Analisis Audio pada Video Viral***

<b>Tema Utama</b>	<b>Sub-Tema / Temuan Spesifik</b>	<b>Frekuensi Kemunculan</b>	<b>Contoh Implementasi Khas</b>
1. Strategi Musik Latar	Penggunaan sound yang sedang tren	18/20 (90%)	Menggunakan potongan 15 detik dari lagu viral di TikTok untuk tantangan tari (V-T01).
	Manipulasi tempo (sped-up/slowed)	9/20 (45%)	Mempercepat lagu pop untuk konten tutorial cepat (V-Y05); memperlambat lagu untuk narasi emosional (V-T06).
	Sinkronisasi beat drop dengan transisi visual	6/20 (30%)	Menyelaraskan perubahan pakaian dengan puncak ketukan musik EDM (V-T04).
2. Peran Efek Suara (SFX)	SFX sebagai penanda aksi/transisi (whoosh, pop)	15/20 (75%)	Suara pop setiap kali teks edukasi muncul di layar (V-Y02).
	SFX untuk penekanan komedi (punchline sound)	5/20 (25%)	Menggunakan suara record scratch pada momen canggung dalam sketsa (V-T07).
	Penggunaan ambience untuk imersi	4/20 (20%)	Menambahkan suara angin dan alam pada vlog perjalanan sinematik (V-Y08).
3. Kualitas Vokal & Narasi	Vokal yang jernih dan bebas noise	12/12 (100% dari video bervokal)	Rekaman narasi yang bersih dan terfokus pada video tutorial (V-T05).

	Modulasi intonasi dan dinamika performatif	10/12 (83% dari video bervokal)	Perubahan nada suara dramatis dalam konten storytelling (V-Y09).
4. Optimalisasi Teknis	Loudness tinggi (rata-rata -14 hingga -10 LUFS)	19/20 (95%)	Audio yang terdengar keras dan konsisten tanpa perlu menaikkan volume ponsel.
	Kompresi rentang dinamis yang signifikan	17/20 (85%)	Perbedaan antara suara pelan dan keras diminimalkan agar semua elemen terdengar jelas.
	EQ yang berfokus pada frekuensi menengah	16/20 (80%)	Mengurangi bass dan menonjolkan frekuensi vokal dan melodi agar optimal di ponsel.

## PEMBAHASAN

Penelitian ini dirancang untuk membongkarkotak hitam” sonik di balik fenomena video viral di TikTok dan YouTube, dengan fokus pada teknik-teknik audio yang seringkali terabaikan dalam analisis media digital. Hasil analisis kualitatif terhadap 20 video viral menunjukkan bahwa kesuksesan audio dalam konten berdurasi pendek bukanlah suatu kebetulan, melainkan hasil dari serangkaian keputusan strategis dan teknis yang cermat. Empat tema utama yang teridentifikasi—(1) Pemanfaatan Musik Latar sebagai Jangkar Emosional dan Pemicu Tren; (2) Peran Krusial Efek Suara (SFX) sebagai Penanda Naratif; (3) Prioritas pada Kejernihan dan Performa Vokal; dan (4) Optimalisasi Teknis Audio untuk Konsumsi Seluler—secara kolektif melukiskan gambaran holistik tentang bagaimana audio direkayasa untuk memaksimalkan keterlibatan audiens. Bagian pembahasan ini akan menginterpretasikan temuan-temuan tersebut, membandingkannya dengan literatur yang ada, serta menguraikan implikasi teoretis dan praktisnya.

### Interpretasi Temuan: MembongkarResep” Sonik Konten Viral

Temuan penelitian ini secara mendalam mengonfirmasi hipotesis awal bahwa audio dalam konten viral adalah elemen yang direkayasa secara sengaja, bukan sekadar pelengkap visual. Keempat tema yang muncul saling terkait dan membentuk sebuah hierarki strategi audio yang efektif.

**Pertama, strategi musik latar** yang ditemukan (penggunaan sound tren, manipulasi tempo, dan sinkronisasi) berfungsi sebagaipintu gerbang” utama untuk menarik perhatian. Penggunaan audio yang sedang tren, sebagaimana ditemukan pada 90% sampel, adalah strategi yang sangat efisien. Ini selaras dengan konsepaudio meme” yang dijelaskan oleh (Abidin & Lee, 2023) di mana suara yang dikenal berfungsi sebagai bahasa komunal dan penanda kultural. Temuan kami memperluas konsep ini dengan menunjukkan bahwa ini bukan hanya fenomena budaya, tetapi juga strategi algoritmik. Kreator secara sadarmenunggangi” gelombang popularitas sebuah suara untuk meningkatkan visibilitas konten mereka, yang didukung oleh algoritma platform yang cenderung memprioritaskan konten dengan audio yang sedang tren. Manipulasi tempo, seperti sped-up, secara psikologis menciptakan rasa urgensi dan energi yang cocok dengan rentang perhatian

audiens yang pendek di platform video vertikal, sebuah fenomena yang sejalan dengan penelitian tentang bagaimana tempo musik dapat memengaruhi gairah emosional dan kognitif.

**Kedua, peran efek suara (SFX)** sebagai penanda naratif dan peningkat ritme menunjukkan tingkat kecanggihan produksi yang lebih tinggi. SFX tidak lagi hanya digunakan untuk realisme, tetapi sebagai “tanda baca auditori” yang memandu perhatian penonton. Penggunaan suara whoosh pada transisi atau pop pada kemunculan teks berfungsi sebagai pattern interrupt, sebuah teknik psikologis yang dirancang untuk menyegarkan kembali fokus audiens yang mudah teralihkan. Ini memperkaya narasi visual dengan lapisan informasi auditori, membuat cerita lebih mudah dicerna dalam waktu singkat. Temuan tentang penggunaan ambience untuk menciptakan imersi, meskipun hanya pada 20% sampel, menandakan adanya pergeseran menuju kualitas produksi sinematik bahkan dalam format video pendek, di mana kreator berusaha menciptakan pengalaman yang lebih mendalam dan menarik secara emosional.

**Ketiga, prioritas pada kejernihan vokal** adalah temuan yang paling fundamental. Pada konten yang berbasis narasi, vokal yang jernih adalah fondasi dari kredibilitas. Temuan ini sangat sejalan dengan studi oleh (Crouch et al., 2016) yang menunjukkan bahwa kualitas audio yang buruk secara signifikan menurunkan kredibilitas pembicara dan pesannya (Ou, and Ho, 2024). Dalam ekosistem media sosial yang penuh dengan misinformasi, suara yang jernih, direkam dengan baik, dan bebas dari derau dapat secara tidak sadar diartikan oleh audiens sebagai tanda profesionalisme dan kepercayaan. Lebih dari itu, modulasi vokal yang performatif, seperti perubahan intonasi dan kecepatan bicara, merupakan manifestasi dari pembangunan parasocial relationship (hubungan parasosial). Kreator tidak hanya menyampaikan informasi, tetapi juga “berbicara” kepada audiensnya, menciptakan ilusi keintiman dan koneksi personal yang mendorong loyalitas dan keterlibatan jangka Panjang (Rather et al., 2022)

**Keempat, optimalisasi teknis untuk perangkat seluler** adalah temuan yang paling orisinal dan secara langsung mengisi kesenjangan penelitian yang diidentifikasi dalam pendahuluan. Praktik kompresi agresif, peningkatan loudness ke level -10 hingga -14 LUFS, dan EQ yang berfokus pada frekuensi menengah adalah bukti nyata adanya “rekayasa” audio yang disesuaikan dengan keterbatasan perangkat keras audiens target. Pengeras suara ponsel memiliki rentang frekuensi yang terbatas dan kesulitan mereproduksi dinamika yang lebar. Dengan “memadatkan” audio dan menonjolkan frekuensi vokal, kreator memastikan bahwa konten mereka terdengar jelas, kuat, dan berdampak bahkan dalam kondisi pendengaran yang tidak ideal (misalnya, di tempat umum). Ini adalah pendekatan pragmatis yang memprioritaskan kejelasan dan dampak di atas fidelitas audio murni, sebuah kompromi cerdas yang menunjukkan pemahaman mendalam tentang konteks konsumsi media saat ini (Resti et al., 2023).

#### Perbandingan dengan Literatur dan Implikasi Teoretis

Temuan penelitian ini tidak hanya mendukung tetapi juga memperluas literatur yang ada di beberapa bidang. Studi ini mengonfirmasi pentingnya audio dalam mendorong keterlibatan, sebagaimana telah banyak dibahas dalam konteks pemasaran dan media sosial. Namun, kontribusi utamanya adalah dengan menggeser fokus dari “apa” (lagu apa yang digunakan) “kebagaimana” (bagaimana audio tersebut diolah secara teknis).



Jika penelitian sebelumnya **oleh** (Sagala & Yayi Wira Pamungkas, 2023) (Abidin & Lee, 2023) telah mengidentifikasi peran tren musik dan audio meme dari perspektif sosiokultural dan algoritmik, penelitian ini memberikan bukti empiris tentang implementasi teknisnya. Kami menunjukkan bahwa kreator tidak hanya pasif menggunakan suara tren, tetapi secara aktif memanipulasinya (tempo, struktur) untuk tujuan retensi perhatian. Ini menunjukkan tingkat agensi dan literasi media yang lebih tinggi di kalangan kreator sukses daripada yang diasumsikan sebelumnya.

Temuan mengenai pentingnya kejernihan vokal untuk kredibilitas memberikan validasi empiris dalam konteks media sosial untuk teori yang sudah mapan dalam komunikasi dan psikologi. Studi oleh (Bonnevie et al., 2023) (Bitzenbauer et al., 2022) yang menemukan bahwa kualitas audio memengaruhi persepsi kredibilitas penelitian ilmiah, kini terbukti relevan untuk konten yang jauh lebih informal seperti tutorial atau storytelling di TikTok. Hal ini mengimplikasikan bahwa prinsip pemrosesan yang lancar” (processing fluency)—di mana informasi yang lebih mudah diproses secara kognitif dianggap lebih benar atau dapat dipercaya—berlaku kuat di lingkungan digital yang serba cepat.



*Gambar Diagram Model Hierarki Kebutuhan Audio untuk Konten Viral"*

Secara teoretis, temuan ini memungkinkan perumusan kerangka kerja konseptual baru, yang dapat disebut **Model Hierarki Kebutuhan Audio untuk Konten Viral**". Model ini, yang terinspirasi dari hierarki kebutuhan Maslow, mengusulkan bahwa efektivitas audio dalam video pendek dibangun secara berlapis.

## KESIMPULAN

Penelitian ini berangkat dari tujuan untuk membongkar dimensi teknis dan strategis audio yang berkontribusi pada popularitas masif video di platform TikTok dan YouTube. Melalui analisis konten kualitatif yang mendalam, penelitian ini berhasil mengidentifikasi bahwa kesuksesan sonik

sebuah konten viral bukanlah fenomena kebetulan, melainkan hasil dari rekayasa audio yang cermat dan berlapis. Temuan utama menunjukkan adanya empat pilar strategis yang saling menopang: (1) pemanfaatan musik latar tidak hanya sebagai pengiring, tetapi sebagai jangkar emosional dan pemicu tren algoritmik melalui penggunaan *sound* populer dan manipulasi tempo; (2) penggunaan efek suara (SFX) sebagai “tanda baca auditori” yang krusial untuk memandu perhatian, memperkuat narasi, dan meningkatkan ritme visual; (3) prioritas absolut pada kejernihan dan kualitas performatif vokal sebagai fondasi kredibilitas dan pembangun hubungan parasosial dengan audiens; serta (4) adanya lapisan optimalisasi teknis yang seringkali tidak terlihat, yaitu rekayasa audio yang secara spesifik disesuaikan untuk konsumsi pada perangkat seluler dengan segala keterbatasannya.

Dengan demikian, penelitian ini secara langsung menjawab kesenjangan riset yang telah diidentifikasi sebelumnya. Jika literatur yang ada cenderung berfokus pada aspek sosiokultural atau pemasaran musik, studi ini mengisi kekosongan dengan memberikan bukti empiris mengenai *bagaimana* teknik-teknik rekayasa audio—seperti kompresi rentang dinamis yang agresif, penargetan tingkat kenyaringan (*Loudness*) pada standar spesifik (-10 hingga -14 LUFS), dan ekualisasi yang berfokus pada spektrum frekuensi menengah—secara pragmatis diimplementasikan oleh kreator sukses. Temuan ini memvalidasi bahwa kualitas audio teknis bukan lagi sekadar preferensi estetis, melainkan sebuah prasyarat fungsional untuk memastikan pesan dan dampak konten tersampaikan secara efektif dalam ekosistem media yang jenuh dan serba cepat.

## ***Implikasi dan Kontribusi Penelitian***

Kontribusi penelitian ini dapat dipetakan ke dalam tiga domain utama. **Secara teoretis**, penelitian ini menawarkan sebuah kerangka kerja konseptual baru, yaitu **Model Hierarki Kebutuhan Audio untuk Konten Viral**. Model ini mengusulkan bahwa efektivitas audio dibangun secara bertingkat, mulai dari fondasi teknis (kejernihan dan *loudness*), dilanjutkan dengan elemen naratif dan emosional (SFX dan musik), hingga puncak strategi yaitu integrasi dengan tren budaya dan algoritma platform. Kerangka kerja ini dapat menjadi landasan bagi studi-studi selanjutnya dalam bidang media digital dan komunikasi audio-visual.

**Secara praktis**, temuan ini memberikan panduan yang sangat relevan dan dapat ditindaklanjuti bagi para kreator konten, pemasar digital, dan profesional industri kreatif. Dengan memahami “resepi” sonik di balik konten yang sukses, mereka dapat membuat keputusan yang lebih terinformasi dalam produksi, mulai dari investasi pada peralatan audio yang tepat, pemilihan dan manipulasi musik latar, hingga penerapan teknik *mixing* dan *mastering* yang sesuai dengan platform tujuan. Penelitian ini mengubah persepsi audio dari “sekadar biaya produksi” menjadi “investasi strategis” untuk meningkatkan keterlibatan dan jangkauan audiens.

**Secara kontekstual**, studi ini menangkap dan mendokumentasikan praktik audio yang sangat spesifik pada era dominasi video vertikal dan konsumsi seluler. Praktik seperti penggunaan lagu versi *sped-up* atau *mastering* yang “panas” (keras) adalah fenomena kontemporer yang mencerminkan adaptasi kreator terhadap teknologi dan perilaku audiens saat ini. Dengan demikian, penelitian ini berfungsi sebagai catatan historis dan analisis kritis terhadap lanskap sonik media digital di pertengahan dekade 2020-an.

## ***Keterbatasan dan Arah Penelitian Masa Depan***

Peneliti menyadari adanya keterbatasan dalam studi ini. Sifat kualitatif dari analisis, meskipun memberikan kedalaman wawasan, membatasi generalisasi temuan karena ukuran sampel yang relatif kecil. Oleh karena itu, penelitian di masa depan sangat dianjurkan untuk memvalidasi temuan ini melalui pendekatan kuantitatif berskala besar. Sebuah studi korelasional yang menganalisis ribuan video untuk menghubungkan metrik audio objektif (misalnya, tingkat LUFS, rentang dinamis, ada atau tidaknya SFX spesifik) dengan metrik keterlibatan (suka, komentar, pembagian) akan sangat berharga.

Selanjutnya, studi eksperimental dengan desain A/B *testing* dapat memberikan bukti kausal yang lebih kuat. Misalnya, dengan menyajikan dua versi video yang identik secara visual namun berbeda dalam satu variabel audio (misalnya, vokal jernih vs. vokal bising; dengan SFX vs. tanpa SFX) kepada kelompok audiens yang berbeda, peneliti dapat mengukur secara langsung dampak variabel tersebut terhadap waktu tonton, retensi, dan persepsi kredibilitas. Terakhir, analisis komparatif lintas platform (misalnya, TikTok vs. Instagram Reels vs. YouTube Shorts) dan lintas budaya dapat mengeksplorasi apakah aturan” sonik ini bersifat universal atau bervariasi tergantung pada norma platform dan konteks budaya audiens.

Pada akhirnya, penelitian ini menegaskan kembali posisi audio sebagai pilar strategis yang fundamental dalam arsitektur konten digital yang sukses. Di balik setiap video yang menjadi viral, terdapat sebuah narasi sonik yang direkayasa dengan presisi untuk menarik, mempertahankan, dan memuaskan audiens dalam hitungan detik.

## **REFERENSI**

- Abidin, C., & Lee, J. (2023). K-pop TikTok: TikTok’s expansion into South Korea, TikTok Stage, and platformed glocalization. *Media Int. Aust.*, 188(1), 86–111.
- Abolhasani, M., Oakes, S., & Oakes, H. (2017). Music in advertising and consumer identity. *Mark. Theory*, 17(4), 473–490.
- Bingham, A. J. (2023). From data management to actionable findings: A five-phase process of qualitative data analysis. *Int. J. Qual. Methods*, 22.
- Bitzenbauer, P., Höfler, S., Veith, J. M., Winkler, B., Zenger, T., & Kulgemeyer, C. (2022). *Exploring the relationship between surface features and explaining quality of YouTube explanatory videos*.
- Bonnevie, T., Repel, A., Gravier Francis-Edouard and Ladner, J., Sibert, L., Muir Jean-François and Cuvelier, A., & Fischer, M.-O. (2023). Video abstracts are associated with an increase in research reports citations, views and social attention: a cross-sectional study. *Scientometrics*, 128(5), 3001–3015.
- Brøvig-Hanssen, R., & Jones, E. (2023). Remix’s retreat? Content moderation, copyright law and mashup music. *New Media Soc.*, 25(6), 1271–1289.

- Chen, C., Song, M., Song, W., Guo, L., & Jian, M. (2022). *A comprehensive survey on video saliency detection with auditory information: The audio-visual consistency perceptual is the key!*
- Creswell, J. W., & Poth, C. N. (2018). *Qualitative inquiry and research design: Choosing among five approaches*.
- Crouch, R., Ewer, M., Quester, P., & Proksch, M. (2016). Talking with you—not at you: How brand ambassadors can spark consumer brand attachment? In *Marketing Challenges in a Turbulent Business Environment* (pp. 189–194). Springer International Publishing.
- Dunbar, R. I. M., Robledo, J.-P., Tamarit, I., Cross, I., & Smith, E. (2022). Nonverbal auditory cues allow relationship quality to be inferred during conversations. *J. Nonverbal Behav.*, 46(1), 1–18.
- Haura, A. S., De Fretes, D., & Kustap, K. (2024). Musik Viral dalam Peningkatan Brand Awareness pada Konten TikTok. *Ekspresi*, 13(2). <https://doi.org/10.24821/ekp.v13i2.14266>
- Herget, A.-K., Breves, P., & Schramm, H. (2022). The influence of different levels of musical fit on the efficiency of audio-visual advertising. *Music Sci.*, 26(1), 3–23.
- Kirchberger, M., & Russo, F. A. (2016). Dynamic range across music genres and the perception of dynamic compression in hearing-impaired listeners. *Trends Hear.*, 20, 233121651663054.
- Martin, N. A., & Martin, R. (2015). Would you watch it? Creating effective and engaging video tutorials. *J. Libr. Inf. Serv. Dist. Learn.*, 9(1–2), 40–56.
- Rather, R. A., Hollebeek, L. D., Vo-Thanh, T., Ramkissoon, H., Leppiman, A., & Smith, D. (2022). Shaping customer brand loyalty during the pandemic: The role of brand credibility, value congruence, experience, identification, and engagement. *J. Consum. Behav.*, 21(5), 1175–1189.
- Sagala, J. M., & Yayi Wira Pamungkas, dan. (2023). Pengaruh Aspek Musikal pada Konten Viral TikTok dalam Lagu Top 3 Wilayah Asia Tenggara. *Human Art'sthetic Journal (HAJ)*, 1(2), 73–82. <https://youtu.be/wiBVVrXvnKk>.
- Salselas, I., & Penha, R. (2019, September). The role of sound in inducing storytelling in immersive environments. *Proceedings of the 14th International Audio Mostly: A Journey in Sound*.
- Sari, W. P., & Paramita, S. (2022). VIRAL MARKETING DI MEDIA SOSIAL SEBAGAI GAYA BARU KOMUNIKASI PEMASARAN. *Jurnal Muara Ilmu Ekonomi Dan Bisnis*, 6(2), 309–319. <https://doi.org/10.24912/jmie.v6i2.20243>
- Schlüssel, H., & Frosh, P. (2023). The taste of video: Facebook videos as multi-sensory experiences. *Converg. Int. J. Res. New Media Technol.*, 29(4), 980–996.

- Suciati, P. (2020). FAKTOR PEMBENTUK PESAN VIRAL DI MEDIA SOSIAL SEBAGAI STRATEGI PROMOSI MUSISI INDEPENDEN. *Jurnal Vokasi Indonesia*, 8(1).  
<https://doi.org/10.7454/jvi.v8i1.155>
- Taylor, Z. A., & Abidin, C. (2024). Where are all the Black girls on TikTok?: Exploring in-group community and (in)visibility through #BlackGirlTikTok. *Int. J. Cult. Stud.*
- Thiparpakul, P., Jiravansirikul, T., & Dheandhanoo, T. (2020, March). The satisfaction of sword's sound effect in game design in case of metal gear rising and sekiro. *Proceedings of the 2020 8th International Conference on Information and Education Technology*.
- Trommer-Beardslee, H. (2022). Renegades: Digital Dance Cultures from Dubsmash to Tik Tok. *J. Dance Educ.*, 1.
- Vallet, D., Berkovsky, S., Ardon Sebastien and Mahanti, A., & Kafaar, M. A. (2015, October). Characterizing and predicting viral-and-popular video content. *Proceedings of the 24th ACM International on Conference on Information and Knowledge Management*.
- Vaterlaus, J. M., & Winter, M. (2021). TikTok: an exploratory study of young adults' uses and gratifications. *Soc. Sci. J.*, 1–20.
- Wahid, R., Karjaluo, H., Taiminen, K., & Asiati, D. I. (2023). Becoming TikTok famous: Strategies for global brands to engage consumers in an emerging market. *J. Int. Mark.*, 31(1), 106–123.
- Wang, M., Liao, Y., Lyckvi, S. L., & Chen, F. (2020). How drivers respond to visual vs. auditory information in advisory traffic information systems. *Behav. Inf. Technol.*, 39(12), 1308–1319.
- Xu, W., & Zammit, K. (2020). Applying thematic analysis to education: A hybrid approach to interpreting data in practitioner research. *Int. J. Qual. Methods*, 19, 160940692091881.
- Yin, R. K. (2018). *Case Study Research and Applications: Design and Methods (6th ed.)*.
- Zulli, D., & Zulli, D. J. (2022). Extending the Internet meme: Conceptualizing technological mimesis and imitation publics on the TikTok platform. *New Media Soc.*, 24(8), 1872–1890.