

MEDIASI ALGORITMIK DAN KONSTRUKSI MULTIMODAL OTORITAS DAN PERSUASI MASKULIN “LEARN ARABIC WITH KHASU” DI PLATFORM YOUTUBE

Dini Arifa Ayunisa

Universitas Airlangga

dini.arifa.ayunisa-2024@fib.unair.ac.id

ABSTRAK

Penelitian ini bertujuan untuk menganalisis konstruksi otoritas maskulin dalam konten pembelajaran bahasa Arab pada saluran YouTube Learn Arabic with Khasu melalui pendekatan analisis multimodal. Kajian ini berangkat dari asumsi bahwa otoritas dalam ruang digital tidak hanya dibangun melalui bahasa verbal, melainkan melalui orkestrasi berbagai mode semiotik yang bekerja secara simultan, meliputi mode verbal, visual, audio, dan komposisional. Data penelitian berupa 12 video terpilih dengan total 186 ujaran yang ditranskripsi dan dianalisis secara kualitatif-kuantitatif. Hasil penelitian menunjukkan bahwa konstruksi otoritas maskulin direproduksi secara sistematis melalui pilihan leksikal bahasa Arab, khususnya penggunaan bentuk gramatikal maskulin jamak dan sapaan kolektif maskulin. Dari total data, 74 ujaran (39,7%) secara eksplisit menggunakan penanda maskulin seperti *yā ikhwan* dan *antum*. Selain itu, struktur imperatif maskulin mendominasi pola komunikasi, memperlihatkan relasi kuasa yang jelas antara pembicara dan audiens. Repetisi ujaran seperti *hādzā muhim jiddan* (27 kali) menunjukkan strategi pedagogis berbasis penegasan dan internalisasi makna.

Secara visual, otoritas dikonstruksi melalui penggunaan *direct gaze* dalam 81% segmen, gestur direktif seperti telunjuk menunjuk kamera, serta *framing medium close-up* yang memberikan nilai *salience* tinggi pada wajah dan gestur tangan. Elemen-elemen ini membangun relasi dominasi yang dikemas dalam kedekatan pedagogis. Dari sisi audio, kontrol intonasi, penekanan dramatis, serta manipulasi volume suara berfungsi mengarahkan perhatian dan memperkuat legitimasi instruksional. Pada level komposisional, desain studio yang menyerupai ruang akademik formal, pencahayaan frontal terang, serta thumbnail kontras tinggi berperan sebagai sinyal profesionalitas sekaligus strategi visibilitas algoritmik. Konsistensi durasi video (8–12 menit) dan pola unggahan rutin menunjukkan integrasi antara performa pedagogis dan logika platform digital.

Secara keseluruhan, penelitian ini menyimpulkan bahwa otoritas maskulin dalam konten pembelajaran bahasa Arab tersebut merupakan hasil orkestrasi multimodal yang terencana dan koheren. Maskulinitas tidak tampil sebagai atribut biologis semata, melainkan sebagai performa pedagogis yang diproduksi, direproduksi, dan diperkuat melalui kombinasi bahasa, gestur, tatapan, prosodi, serta strategi algoritmik. Temuan ini memperkaya kajian analisis wacana multimodal dengan menunjukkan bagaimana identitas dan otoritas dinegosiasikan dalam ekosistem media digital.

Kata kunci: multimodalitas, otoritas maskulin, pembelajaran bahasa Arab, analisis wacana digital.

ABSTRACT

This study aims to analyze the construction of masculine authority in Arabic language learning content on the YouTube channel Learn Arabic with Khasu through a multimodal discourse analysis approach. The study is based on the assumption that authority in digital spaces is not constructed solely through verbal language, but through the orchestration of various semiotic modes operating simultaneously, including verbal, visual, audio, and compositional modes. The research data consist of 12 selected videos containing a total of 186 utterances that were transcribed and analyzed using a qualitative-quantitative approach.

The findings reveal that the construction of masculine authority is systematically reproduced through lexical choices in Arabic, particularly through the use of masculine plural grammatical forms and collective masculine address terms. Out of the total data, 74 utterances (39.7%) explicitly employed masculine markers such as *yā ikhwan* and *antum*. In addition, masculine imperative structures dominate the communication pattern, indicating a clear power relation between the speaker and the audience. The repetition of utterances such as *hādzā muhim jiddan* (27 occurrences) demonstrates a pedagogical strategy based on emphasis and meaning internalization.

Visually, authority is constructed through the use of *direct gaze* in 81% of the analyzed segments, directive gestures such as pointing directly at the camera, and *medium close-up framing* that provides high *salience* to facial expressions and hand gestures. These elements create a relationship of dominance framed within pedagogical intimacy. From the audio perspective, intonation control, dramatic emphasis, and volume manipulation function to guide audience attention and reinforce instructional legitimacy. At the compositional level, the studio design resembling a formal academic setting, bright frontal lighting, and high-contrast thumbnails serve as signals of professionalism as well as strategies for algorithmic visibility. The consistent video duration (8–12 minutes) and regular upload schedule indicate the integration of pedagogical performance with platform logic.

Overall, this study concludes that masculine authority in this Arabic language learning content is the result of a planned and coherent multimodal orchestration. Masculinity does not merely appear as a biological attribute, but as a pedagogical performance produced, reproduced, and reinforced through the combination of language, gesture, gaze, prosody, and algorithmic strategies. These findings contribute to multimodal discourse analysis by demonstrating how identity and authority are negotiated within the digital media ecosystem.

Keywords: multimodality, masculine authority, Arabic language learning, digital discourse analysis.

PENDAHULUAN

Pertumbuhan teknologi komunikasi digital di abad ke-21 telah menghasilkan pergeseran signifikan dalam pola penciptaan, distribusi, dan konsumsi informasi global. Munculnya media sosial berbasis audiovisual tidak hanya mengubah cara orang terhubung, tetapi juga membentuk kembali cara informasi diciptakan dan didistribusikan di ruang publik digital. YouTube telah menjadi salah satu platform paling berpengaruh dalam transisi ini. Awalnya dikenal sebagai situs berbagi video amatir, platform ini sekarang telah berubah menjadi ekosistem media digital yang sangat terinstitusionalisasi dan profesional yang terhubung dengan logika bisnis media global. Menurut Strangelove (2009), YouTube telah mengaburkan perbedaan antara media konvensional dan partisipatif karena para pembuat konten mulai menggunakan taktik produksi profesional, estetika sinematik, dan model distribusi berbasis algoritma yang mirip dengan yang digunakan dalam bisnis penyiaran modern. Menurut Orduña Malea dkk. (2020), YouTube lebih dari sekadar platform untuk hiburan; ia juga berfungsi sebagai gudang budaya, pengetahuan, pendidikan, dan identitas sosial, yang memengaruhi perilaku konsumsi informasi modern.

Evolusi YouTube sebagai media sosial dan edukatif secara intrinsik terkait dengan keberadaan tata kelola algoritmik, yang secara aktif memengaruhi visibilitas materi di ranah digital. Dalam perspektif ini, algoritma tidak hanya berfungsi sebagai sistem teknologi untuk merekomendasikan film kepada pemirsa, tetapi juga sebagai proses kekuatan digital yang mengatur alokasi perhatian publik. Penyedia konten harus menyesuaikan taktik produksi mereka agar tetap "dikenal" oleh sistem algoritmik dengan menggunakan thumbnail yang menarik, pengaturan durasi video, optimasi judul, pola keterlibatan audiens, dan performa visual yang sesuai dengan preferensi platform. Fenomena ini telah menghasilkan ekonomi perhatian, di mana perhatian audiens merupakan komoditas fundamental dalam persaingan media digital. Menurut Wood (2021), budaya algoritmik YouTube memengaruhi perilaku produsen untuk terus-menerus mengubah identitas dan estetika video mereka guna memperoleh penyebaran yang lebih luas melalui sistem rekomendasi platform. Lisnarti (2025) menekankan bahwa media sosial algoritmik menyebabkan hubungan yang rumit antara tuntutan audiens dan logika platform karena algoritma rekomendasi secara aktif memengaruhi pola konsumsi informasi pengguna. Dengan demikian, komunikasi digital di YouTube tidak pernah sepenuhnya netral; komunikasi tersebut terus-menerus dipengaruhi oleh sistem komputasi yang menjalankan platform tersebut.

YouTube telah mengubah perannya menjadi lingkungan pembelajaran daring yang populer seiring dengan evolusi media digital. Konten pendidikan digital berkembang pesat karena budaya modern lebih menyukai media pembelajaran yang lebih mudah beradaptasi, visual, interaktif, dan mudah diakses daripada teknik pembelajaran tradisional. Video pembelajaran di YouTube tidak lagi menampilkan penjelasan lisan dasar, tetapi menggunakan taktik komunikasi multimodal yang mengintegrasikan komponen verbal, visual, audio, gerak tubuh, warna, tipografi, dan desain spasial untuk menghasilkan pengalaman belajar yang lebih meyakinkan dan menarik. Menurut Bhatia (2022), video internet merupakan jenis komunikasi multimodal yang kompleks karena makna dihasilkan melalui interaksi simultan dari banyak modalitas semiotik, bukan hanya teks vokal. Efektivitas konten pembelajaran digital di YouTube sangat ditentukan oleh kemampuan kreator dalam mengelola integrasi yang kohesif antara aspek visual, auditori, dan lisan untuk memaksimalkan keterlibatan audiens. Seiring dengan perkembangan budaya digital, kreator pendidikan telah mengadopsi estetika profesional seperti pencahayaan studio, peminjaman kamera yang stabil, manipulasi intonasi suara, dan thumbnail dengan kontras tinggi untuk mempertahankan retensi audiens dan meningkatkan kinerja algoritma konten mereka (Ramadhan, 2025 ; Silitonga et al., 2025).

Fenomena ini juga telah menyebar di negara-negara Arab dan Teluk, di mana YouTube telah muncul sebagai arena publik digital baru yang memungkinkan orang untuk mengekspresikan identitas budaya, sosial, dan gender mereka secara lebih bebas daripada media tradisional. Ketersediaan platform digital memberikan peluang bagi penulis lokal untuk menciptakan narasi alternatif tentang identitas Arab modern yang menantang dominasi media konvensional dan hegemonik. Menurut penelitian Al-Khajaeman (2022), YouTube di wilayah Arab telah menjadi forum penting untuk menegosiasikan identitas lokal, termasuk penggambaran maskulinitas dan feminitas di ruang digital. Terlepas dari sifat demokratis yang diklaim oleh platform digital, konstruksi gender dalam media Arab tetap sangat dipengaruhi oleh masyarakat patriarki yang menempatkan laki-laki sebagai peran dominan dalam lingkungan publik, intelektual, dan pendidikan. Representasi laki-laki lebih erat kaitannya dengan otoritas, profesionalisme, dan validitas pengetahuan daripada representasi perempuan.

Dalam konteks budaya Arab, bahasa memiliki posisi yang sangat penting dalam pembentukan identitas gender karena struktur tata bahasa Arab secara inheren mengandung penandaan maskulin dan feminin. Bahasa Arab bukan hanya alat komunikasi, tetapi juga instrumen ideologis yang membentuk

konstruksi sosial mengenai relasi gender. Muassomah et al. (2025) menjelaskan bahwa tata bahasa Arab merupakan sistem linguistik yang sangat bergender (*gendered grammar*), sehingga penggunaan pronomina, bentuk verba, dan sapaan tertentu dapat mereproduksi identitas maskulin maupun feminin secara eksplisit. Penggunaan bentuk sapaan seperti *yā ikhwan*, *antum*, atau *yā ustadz* dalam komunikasi digital tidak hanya berfungsi sebagai unsur linguistik biasa, tetapi juga menjadi penanda relasi kuasa dan legitimasi sosial tertentu. Dalam praktik komunikasi edukatif, penggunaan bentuk-bentuk maskulin tersebut secara tidak langsung membangun posisi pengajar laki-laki sebagai figur otoritatif yang memiliki kontrol terhadap proses penyampaian pengetahuan. Dengan demikian, bahasa Arab menjadi medium penting dalam reproduksi identitas maskulin di ruang digital.

Dominasi representasi maskulinitas dalam media digital juga terlihat dalam berbagai penelitian mengenai konten edukasi di YouTube. Gezici dan Saygin (2022) menemukan bahwa video edukasi di YouTube menunjukkan bias gender yang cukup signifikan karena laki-laki lebih sering tampil sebagai narator utama dan diasosiasikan dengan otoritas intelektual dibandingkan perempuan. Penelitian Rafi'atussyifa dan Nor (2023) juga menunjukkan bahwa laki-laki cenderung menggunakan bahasa yang lebih direktif, dominan, dan persuasif dalam vlog YouTube dibandingkan perempuan. Fenomena ini memperlihatkan bahwa konstruksi maskulinitas di ruang digital tidak hanya dibangun melalui tampilan visual, tetapi juga melalui pola komunikasi verbal, strategi persuasi, gestur tubuh, intonasi suara, dan pengelolaan hubungan dengan audiens. Ketika figur perempuan jarang muncul dalam ruang edukasi tertentu, maka terjadi apa yang disebut sebagai *symbolic annihilation*, yaitu penghilangan simbolik yang secara tidak langsung membangun persepsi bahwa domain tersebut merupakan wilayah eksklusif laki-laki. Dalam konteks pembelajaran bahasa Arab, kondisi ini menjadi semakin menarik karena dominasi maskulinitas tidak hanya tampil secara visual, tetapi juga direproduksi melalui struktur gramatikal bahasa Arab itu sendiri.

Untuk memahami kompleksitas komunikasi digital tersebut, diperlukan pendekatan teoritis yang mampu menganalisis hubungan antara bahasa, visual, audio, dan desain media secara simultan. Oleh karena itu, penelitian ini menggunakan pendekatan Analisis Wacana Multimodal yang dikembangkan oleh Kress dan van Leeuwen. Pendekatan ini memandang bahwa makna tidak hanya dibangun melalui bahasa verbal, tetapi juga melalui integrasi berbagai mode semiotik seperti gambar, warna, gestur, tipografi, audio, dan tata ruang visual. Kress dan van Leeuwen (2006) melalui konsep “tata bahasa visual” menjelaskan bahwa komunikasi visual memiliki struktur dan aturan tersendiri yang dapat dianalisis melalui tiga metafungsi utama, yaitu makna representasional, makna interaktif, dan makna komposisional. Makna representasional berkaitan dengan bagaimana identitas dan realitas sosial direpresentasikan dalam media; makna interaktif berkaitan dengan hubungan sosial antara pembicara dan audiens; sedangkan makna komposisional berkaitan dengan tata letak visual, *salience*, dan integrasi elemen dalam satu bingkai komunikasi. Discourse Analyzer (2024) menjelaskan bahwa teori Kress dan van Leeuwen memungkinkan peneliti memahami bagaimana elemen visual bekerja sebagai sistem semiotik yang memengaruhi interpretasi audiens terhadap pesan media. Selain itu, Dontcheva-Navratilova (2023) menegaskan bahwa persuasi digital modern sangat bergantung pada kredibilitas multimodal yang dibangun melalui integrasi elemen verbal, visual, dan audio secara simultan.

Meskipun penelitian mengenai multimodalitas, media digital, dan representasi gender telah berkembang cukup pesat, masih terdapat celah penelitian (*research gap*) yang signifikan. Sebagian besar penelitian terdahulu lebih banyak membahas bias gender dalam media secara umum atau menganalisis multimodalitas dalam iklan, film, dan konten promosi. Penelitian Atma dan Awad (2023), misalnya, lebih berfokus pada stereotip gender dalam iklan televisi melalui pendekatan multimodal, sedangkan Dewi dan Tawami (2022) meneliti representasi keterampilan menulis dalam video iklan edukatif. Andersson (2024) mengkaji ekspresi multimodal dalam video reaksi YouTube terhadap aktivisme transgender, sementara Lim dan Ang (2024) membahas performa identitas dalam vlog mahasiswa YouTuber. Namun, penelitian-penelitian tersebut belum secara spesifik menghubungkan konstruksi maskulinitas dengan mediasi algoritma dalam konteks pembelajaran bahasa Arab berbasis YouTube. Selain itu, kajian mengenai konten edukasi bahasa Arab masih cenderung berfokus pada aspek linguistik murni tanpa melihat bagaimana otoritas pedagogis dibangun melalui interaksi kompleks antara bahasa Arab, visual, audio, dan logika platform digital. Dengan demikian, masih sangat terbatas penelitian yang membahas bagaimana maskulinitas diproduksi sebagai performa pedagogis multimodal yang sekaligus dioptimalkan untuk kebutuhan visibilitas algoritmik di YouTube.

Berdasarkan kondisi tersebut, saluran YouTube *Learn Arabic with Khasu* menjadi objek penelitian yang relevan dan penting untuk dikaji lebih mendalam. Kanal ini tidak hanya menghadirkan pembelajaran bahasa Arab, tetapi juga menunjukkan penggunaan strategi komunikasi multimodal yang

sangat terstruktur dalam membangun citra pengajar laki-laki yang otoritatif, profesional, dan persuasif. Penggunaan sapaan maskulin, struktur imperatif bahasa Arab, tatapan langsung ke kamera (*direct gaze*), manipulasi intonasi suara, gestur direktif, desain studio akademik, serta thumbnail kontras tinggi memperlihatkan bahwa otoritas pedagogis dalam kanal tersebut tidak hadir secara natural, melainkan diproduksi melalui orkestrasi multimodal yang sengaja dirancang. Selain itu, pola produksi konten yang konsisten juga menunjukkan adanya keselarasan dengan logika algoritma YouTube, sehingga maskulinitas tidak hanya menjadi identitas sosial, tetapi juga strategi performatif untuk memperoleh legitimasi dan visibilitas digital.

Penelitian ini menjadi penting dilakukan karena memiliki kontribusi teoretis maupun praktis dalam kajian linguistik, media digital, gender, dan analisis wacana multimodal. Secara teoretis, penelitian ini memperluas penerapan teori Multimodal Discourse Analysis ke dalam konteks pembelajaran bahasa Arab berbasis platform digital yang dimediasi algoritma. Penelitian ini juga memperkaya kajian sosiolinguistik Arab dan studi maskulinitas dengan menunjukkan bahwa identitas gender dalam ruang digital dibangun melalui interaksi kompleks antara bahasa, visual, audio, dan desain algoritmik platform. Secara praktis, penelitian ini diharapkan mampu meningkatkan literasi media masyarakat mengenai bagaimana otoritas dan persuasi dibangun secara halus dalam konten edukasi digital. Selain itu, penelitian ini juga dapat menjadi referensi bagi kreator konten pendidikan untuk memahami pentingnya strategi multimodal dalam meningkatkan efektivitas komunikasi, keterlibatan audiens, dan legitimasi pedagogis di ruang digital.

Berdasarkan uraian tersebut, penelitian ini bertujuan untuk menganalisis bagaimana otoritas dan persuasi maskulin dikonstruksi secara multimodal dalam saluran YouTube *Learn Arabic with Khasu*. Fokus penelitian diarahkan pada analisis interaksi antara mode verbal bahasa Arab, mode visual seperti gestur dan orientasi tubuh, mode audio seperti intonasi dan penekanan suara, serta mode komposisional seperti framing kamera dan desain studio dalam membangun legitimasi pedagogis maskulin. Selain itu, penelitian ini juga bertujuan untuk menganalisis konstruksi makna representasional (*representational meaning*) maskulinitas melalui interaksi antara wacana verbal bahasa Arab, khususnya penggunaan bentuk sapaan maskulin dan pilihan leksikal yang merepresentasikan identitas "ustadz" atau "ahli" serta elemen visual seperti gestur instruktif dan orientasi tubuh dalam saluran "Learn Arabic with Khasu". Mendeskripsikan proses pembangunan makna interaktif (*interactive meaning*) yang digunakan oleh host untuk menegosiasikan relasi otoritas dan membangun keterlibatan emosional dengan audiens melalui manipulasi teknis seperti kontak mata (*gaze*), penekanan audio (intonasi), dan pengaturan jarak kamera (*proximity*) untuk menciptakan keintiman sekaligus dominasi instruksional, dan Menjelaskan kontribusi makna komposisional (*compositional meaning*) dalam membentuk pesan yang koheren melalui tata letak studio, pencahayaan, dan pembingkai kamera (*framing*), serta bagaimana mekanisme mediasi algoritma YouTube berfungsi sebagai sinyal estetika yang memperkuat otoritas pengajar. Melalui pendekatan Analisis Wacana Multimodal, penelitian ini diharapkan mampu menunjukkan bahwa maskulinitas dalam ruang edukasi digital bukan sekadar identitas biologis, melainkan performa sosial dan semiotik yang diproduksi, direproduksi, dan dinegosiasikan secara strategis dalam lanskap media digital kontemporer.

RUMUSAN MASALAH

Berdasarkan latar belakang penelitian yang telah dipaparkan, masalah utama dalam penelitian ini dirumuskan sebagai berikut:

1. Bagaimana makna representasional (*representational meaning*) maskulinitas dikonstruksi melalui interaksi antara wacana verbal bahasa Arab (pilihan leksikal dan bentuk sapaan seperti *ya ikhwan*) serta elemen visual (gestur dan orientasi tubuh) dalam saluran "Learn Arabic with Khasu"?
2. Bagaimana makna interaktif (*interactive meaning*) host dibangun untuk menciptakan relasi otoritas dan strategi persuasi terhadap audiens melalui penggunaan kontak mata (*gaze*), manipulasi audio (intonasi dan penekanan), serta pengaturan jarak kamera (*proximity*)?
3. Bagaimana makna komposisional (*compositional meaning*) yang mencakup tata letak studio, pencahayaan, dan pembingkai kamera (*framing*) membentuk pesan yang koheren sekaligus selaras dengan mekanisme mediasi algoritma YouTube untuk mencapai visibilitas konten?

TUJUAN PENELITIAN

Tujuan utama dari penelitian ini adalah untuk membedah bagaimana otoritas maskulin dan strategi persuasi dibangun secara sengaja melalui desain komunikasi digital yang kompleks. Secara spesifik, penelitian ini bertujuan untuk:

1. Menganalisis konstruksi makna representasional (*representational meaning*) maskulinitas melalui interaksi antara wacana verbal bahasa Arab, khususnya penggunaan bentuk sapaan maskulin dan pilihan leksikal yang merepresentasikan identitas "ustadz" atau "ahli" serta elemen visual seperti gestur instruktif dan orientasi tubuh dalam saluran "Learn Arabic with Khasu".
2. Mendeskripsikan proses pembangunan makna interaktif (*interactive meaning*) yang digunakan oleh host untuk menegosiasikan relasi otoritas dan membangun keterlibatan emosional dengan audiens melalui manipulasi teknis seperti kontak mata (*gaze*), penekanan audio (intonasi), dan pengaturan jarak kamera (*proximity*) untuk menciptakan keintiman sekaligus dominasi instruksional.
3. Menjelaskan kontribusi makna komposisional (*compositional meaning*) dalam membentuk pesan yang koheren melalui tata letak studio, pencahayaan, dan pembingkaihan kamera (*framing*), serta bagaimana mekanisme mediasi algoritma YouTube berfungsi sebagai sinyal estetika yang memperkuat otoritas pengajar.

MANFAAT PENELITIAN

Penelitian mengenai konstruksi otoritas maskulin dan persuasi dalam konten edukasi bahasa Arab di YouTube ini diharapkan dapat memberikan kontribusi signifikan sebagai berikut:

1. Manfaat Teoretis

- a. Penelitian ini memberikan kontribusi pada pengembangan teori Multimodal Discourse Analysis (MDA) dengan memperluas penerapan kerangka kerja Kress dan van Leeuwen ke dalam domain edukasi digital yang dimediasi oleh algoritma. Hal ini memperdalam pemahaman tentang bagaimana berbagai mode semiotik berinteraksi untuk menciptakan koherensi pesan dalam platform video global.
- b. Memperkaya literatur dalam studi maskulinitas dan sosiolinguistik Arab dengan membedah bagaimana otoritas pria dikonstruksi secara tekstual dan visual melalui penggunaan tata bahasa Arab yang bergender. Penelitian ini memperkuat argumen bahwa gender bukan hanya realitas sosial, tetapi juga realitas tekstual yang memengaruhi dinamika hubungan antara pria dan wanita.
- c. Memberikan perspektif baru mengenai hubungan antara desain semiotik dan tata kelola algoritma (*algorithmic governance*) di media sosial. Studi ini menunjukkan bagaimana strategi komunikasi disusun secara sengaja agar tetap "dikenali secara algoritmik" (*algorithmically recognizable*) guna memenangkan perhatian dalam ekonomi perhatian digital.

2. Manfaat Praktis

- a. Hasil penelitian ini dapat menjadi referensi bagi para pengembang konten edukasi bahasa Arab di YouTube untuk memahami pentingnya aspek multimodal dalam meningkatkan **daya persuasi** dan efektivitas pesan instruksional. Penggunaan strategi komunikasi yang tepat dapat meningkatkan aksesibilitas dan keterlibatan audiens Gen Z yang terbiasa dengan realitas multibahasa.
- b. Meningkatkan **literasi media** dan kesadaran kritis bagi para pembelajar bahasa mengenai bagaimana identitas gender dan otoritas dibangun secara halus dalam materi pembelajaran berbasis video. Hal ini memungkinkan audiens untuk lebih waspada terhadap bias gender sistematis yang mungkin muncul dalam konten edukatif.
- c. Memberikan panduan bagi para kreator konten di wilayah Arab mengenai cara mengelola berbagai sumber daya semiotik untuk membangun **autentisitas dan pengaruh** di luar kendali pengawas pintu (*gatekeepers*) media tradisional. Melalui pendekatan multimodal yang tepat, kreator dapat menegosiasikan peran mereka secara lebih otonom dalam ruang publik digital.

TINJAUAN TEORETIS DAN KAJIAN TERKAIT

Analisis Wacana Multimodal dan Tata Bahasa Visual

Perkembangan media digital kontemporer telah mengubah paradigma komunikasi dari yang sebelumnya berpusat pada bahasa verbal menjadi komunikasi berbasis multimodal yang melibatkan integrasi berbagai sumber daya semiotik secara simultan. Dalam konteks media audiovisual seperti YouTube, makna tidak lagi diproduksi hanya melalui ujaran atau teks tertulis, tetapi melalui kombinasi kompleks antara bahasa, visual, audio, warna, gerak tubuh, tipografi, tata ruang, hingga desain antarmuka digital. Oleh karena itu, pendekatan linguistik tradisional yang hanya menitikberatkan pada aspek verbal dianggap tidak lagi memadai untuk menjelaskan dinamika komunikasi digital modern. Untuk memahami bagaimana pesan, identitas, dan relasi kuasa dibangun dalam ruang media digital, diperlukan pendekatan yang mampu menganalisis interaksi antar-mode semiotik secara simultan dan terintegrasi. Dalam konteks tersebut, Analisis Wacana Multimodal atau *Multimodal Discourse Analysis* (MDA) menjadi pendekatan yang relevan dan penting digunakan.

Analisis Wacana Multimodal berkembang dari tradisi semiotika sosial (*social semiotics*) yang dipelopori oleh Michael Halliday dan kemudian dikembangkan lebih lanjut oleh Gunther Kress dan Theo van Leeuwen. Pendekatan ini memandang bahwa setiap mode komunikasi, baik verbal maupun nonverbal, memiliki potensi semiotik untuk membangun makna sosial. Kress dan van Leeuwen (2006) menjelaskan bahwa komunikasi visual memiliki “tata bahasa” tersendiri sebagaimana bahasa verbal memiliki struktur gramatikal. Oleh karena itu, gambar, warna, framing, sudut kamera, pencahayaan, gestur, dan orientasi tubuh dapat dianalisis secara sistematis sebagai sistem tanda yang merepresentasikan realitas sosial tertentu. Dalam perspektif ini, makna bukanlah sesuatu yang hadir secara alamiah, melainkan hasil konstruksi sosial yang diproduksi melalui pilihan semiotik tertentu dalam konteks budaya dan ideologis tertentu.

Kress dan van Leeuwen melalui konsep *visual grammar* atau tata bahasa visual mengembangkan tiga metafungsi utama dalam analisis multimodal, yaitu makna representasional (*representational meaning*), makna interaktif (*interactive meaning*), dan makna komposisional (*compositional meaning*). Ketiga metafungsi tersebut menjadi kerangka analisis utama dalam memahami bagaimana identitas, relasi sosial, dan ideologi direpresentasikan dalam media visual. Makna representasional berkaitan dengan bagaimana individu, objek, atau peristiwa direpresentasikan dalam teks multimodal. Dalam konteks video pembelajaran digital, representasi dapat terlihat melalui pilihan pakaian, latar studio, ekspresi wajah, posisi tubuh, maupun objek simbolik yang digunakan kreator konten untuk membangun citra tertentu. Kress dan van Leeuwen menegaskan bahwa representasi visual tidak bersifat netral karena selalu membawa ideologi sosial tertentu yang memengaruhi cara audiens memahami realitas.

Makna interaktif berfokus pada hubungan sosial yang dibangun antara produsen media dan audiens melalui berbagai elemen visual dan audio. Dalam komunikasi audiovisual, relasi sosial dapat dibangun melalui tatapan mata (*gaze*), sudut kamera (*camera angle*), jarak sosial (*proximity*), kualitas suara, hingga manipulasi intonasi. Tatapan langsung ke kamera, misalnya, menciptakan apa yang disebut Kress dan van Leeuwen sebagai *demand image*, yaitu kondisi ketika audiens seolah-olah “dituntut” untuk terlibat secara personal dengan pembicara. Sebaliknya, tatapan tidak langsung menciptakan *offer image*, di mana audiens hanya diposisikan sebagai pengamat pasif. Dalam konteks YouTube, penggunaan *direct gaze*, medium close-up, serta intonasi persuasif menjadi strategi penting untuk membangun kedekatan emosional sekaligus mempertahankan otoritas komunikatif kreator konten terhadap audiens.

Makna komposisional berkaitan dengan bagaimana elemen-elemen visual diorganisasikan dalam satu bingkai komunikasi sehingga menghasilkan pesan yang koheren. Tata letak visual, salience, framing, warna, pencahayaan, dan posisi objek memiliki peran penting dalam menentukan fokus perhatian audiens. Kress dan van Leeuwen menjelaskan bahwa elemen yang ditempatkan di pusat visual atau diberi kontras warna yang kuat cenderung memiliki nilai informasi lebih tinggi dibandingkan elemen lain. Dalam konteks media digital seperti YouTube, makna komposisional menjadi sangat penting karena berkaitan langsung dengan strategi estetika dan visibilitas algoritmik. Thumbnail kontras tinggi, framing wajah dominan, pencahayaan terang, dan penggunaan tipografi besar bukan hanya persoalan estetika visual, tetapi juga bagian dari strategi semiotik untuk menarik perhatian audiens di tengah kompetisi ekonomi perhatian digital.

Selain Kress dan van Leeuwen, berbagai penelitian kontemporer juga memperkuat relevansi pendekatan multimodal dalam kajian media digital. Kilby (2021) menjelaskan bahwa komunikasi modern semakin bergantung pada integrasi teks dan visual karena audiens digital cenderung memproses informasi

secara simultan melalui berbagai mode semiotik. Bateman dan Tseng (2023) menambahkan bahwa video digital merupakan bentuk komunikasi multimodal yang sangat kompleks karena makna dibangun secara dinamis melalui hubungan antara suara, gambar bergerak, teks layar, dan desain visual. Dalam konteks media daring, multimodalitas menjadi strategi utama untuk membangun narasi, kredibilitas, dan persuasi secara lebih efektif.

Pendekatan MDA juga banyak digunakan untuk menganalisis representasi identitas gender dan relasi kuasa dalam media digital. Andersson (2024), misalnya, menggunakan analisis multimodal untuk mengkaji ekspresi impoliteness dalam video reaksi YouTube terhadap aktivisme transgender. Hasil penelitian tersebut menunjukkan bahwa gestur, ekspresi wajah, intonasi, dan framing visual memiliki kontribusi besar dalam membangun posisi ideologis pembicara. Penelitian Lennon dan Kilby (2021) juga menunjukkan bahwa komunikasi visual mampu membangun nasionalisme dan identitas politik melalui integrasi simbol visual dan teks verbal. Dengan demikian, pendekatan multimodal tidak hanya relevan untuk memahami struktur komunikasi media, tetapi juga penting dalam membedah konstruksi identitas sosial dan ideologi dalam ruang digital.

Dalam penelitian ini, pendekatan MDA digunakan untuk menganalisis bagaimana otoritas maskulin dikonstruksi dalam saluran YouTube *Learn Arabic with Khusus* melalui integrasi mode verbal bahasa Arab, mode visual seperti gestur dan tatapan, mode audio seperti intonasi dan penekanan suara, serta mode komposisional seperti framing kamera dan desain studio. Pendekatan ini memungkinkan peneliti memahami bahwa otoritas pedagogis dalam media digital bukan sekadar hasil kompetensi linguistik, tetapi merupakan produk dari orkestrasi multimodal yang dirancang secara strategis untuk membangun legitimasi dan persuasi di ruang digital.

Mediasi Algoritma dan Tata Kelola Platform

Perkembangan platform digital modern tidak dapat dipisahkan dari keberadaan algoritma sebagai sistem pengatur distribusi informasi di ruang media. Dalam konteks media sosial kontemporer, algoritma tidak lagi sekadar perangkat teknis yang membantu pengguna menemukan konten, tetapi telah berkembang menjadi mekanisme kuasa yang menentukan visibilitas, popularitas, dan legitimasi suatu narasi di ruang publik digital. YouTube sebagai salah satu platform video terbesar di dunia menggunakan sistem rekomendasi berbasis algoritma untuk mengatur video mana yang muncul di halaman utama, kolom rekomendasi, maupun hasil pencarian pengguna. Sistem tersebut bekerja berdasarkan berbagai indikator seperti retensi penonton, jumlah klik, durasi menonton, pola interaksi, dan konsistensi unggahan konten. Akibatnya, kreator konten harus menyesuaikan strategi komunikasi mereka agar tetap dikenali dan diprioritaskan oleh sistem algoritmik platform.

Fenomena tersebut melahirkan konsep *algorithmic governance* atau tata kelola algoritmik, yaitu kondisi ketika algoritma berfungsi sebagai pengatur utama dalam menentukan arus informasi dan partisipasi publik di media digital. Wood (2021) menjelaskan bahwa budaya algoritmik mendorong kreator untuk terus menyesuaikan produksi konten mereka sesuai dengan preferensi platform demi mempertahankan visibilitas digital. Dalam kondisi ini, algoritma bekerja layaknya “penjaga gerbang” (*digital gatekeeper*) yang menentukan konten mana yang dianggap relevan dan layak mendapatkan perhatian publik. Dengan demikian, relasi kuasa dalam media digital tidak lagi hanya dimiliki institusi media tradisional, tetapi juga berada dalam sistem algoritma yang tidak sepenuhnya transparan bagi pengguna.

Dalam ekonomi perhatian (*attention economy*), perhatian audiens menjadi komoditas utama yang diperebutkan kreator konten. Oleh karena itu, produksi konten digital modern tidak hanya berorientasi pada penyampaian informasi, tetapi juga pada kemampuan menarik dan mempertahankan perhatian audiens selama mungkin. Thumbnail kontras tinggi, judul provokatif, durasi video ideal, penggunaan ekspresi wajah dramatis, hingga manipulasi intonasi suara merupakan bagian dari strategi optimasi algoritmik yang sengaja dirancang untuk meningkatkan performa distribusi konten. Fitriansyah dan Lubis (2023) menjelaskan bahwa kreator media digital, termasuk pendakwah dan pengajar daring, secara aktif memanfaatkan logika algoritma media sosial untuk memperluas jangkauan audiens mereka. Fenomena ini menunjukkan bahwa komunikasi digital modern tidak dapat dilepaskan dari strategi performatif yang selaras dengan kebutuhan algoritmik platform.

Konsep *algorithmic imaginary* juga menjadi penting dalam memahami perilaku kreator konten di media digital. Istilah ini merujuk pada asumsi, spekulasi, dan pemahaman kreator mengenai cara kerja algoritma platform, meskipun mekanisme internal algoritma sebenarnya bersifat tertutup (*black box*). Kreator kemudian mengembangkan berbagai strategi berdasarkan persepsi tersebut, seperti menentukan

waktu unggahan, memilih thumbnail tertentu, mempertahankan durasi video yang ideal, atau menciptakan pola interaksi yang dianggap disukai algoritma. Dengan kata lain, kreator tidak hanya memproduksi konten untuk audiens manusia, tetapi juga untuk “audiens algoritmik” yang menentukan distribusi dan visibilitas konten mereka di platform digital.

Dalam konteks YouTube, strategi algoritmik sangat berkaitan dengan pembentukan identitas dan otoritas digital. Lim dan Ang (2024) menjelaskan bahwa kreator YouTube modern melakukan *commodifying the self*, yaitu proses menjadikan identitas diri sebagai komoditas digital yang dikemas secara visual dan performatif untuk menarik perhatian audiens. Otoritas dan kredibilitas tidak lagi hanya bergantung pada kapasitas intelektual, tetapi juga pada kemampuan membangun citra autentik melalui desain visual, gaya bicara, dan interaksi dengan audiens. Dontcheva-Navratilova (2023) menegaskan bahwa persuasi digital modern sangat bergantung pada kredibilitas multimodal yang dibangun melalui kombinasi verbal, visual, dan audio secara simultan.

Dalam penelitian ini, konsep mediasi algoritma digunakan untuk memahami bagaimana saluran *Learn Arabic with Khasu* membangun otoritas maskulin yang tidak hanya bersifat linguistik dan visual, tetapi juga terintegrasi dengan logika visibilitas platform YouTube. Penggunaan thumbnail kontras tinggi, framing wajah dominan, desain studio profesional, konsistensi unggahan, dan durasi video tertentu menunjukkan bahwa strategi pedagogis dalam kanal tersebut telah disesuaikan dengan kebutuhan distribusi algoritmik platform digital. Dengan demikian, otoritas maskulin dalam media digital dipahami bukan sekadar identitas sosial, tetapi juga strategi performatif yang dirancang agar tetap “dikenali secara algoritmik” dalam kompetisi ekonomi perhatian digital.

Konstruksi Linguistik Gender dalam Bahasa Arab

Bahasa Arab merupakan salah satu bahasa yang memiliki sistem gramatikal sangat bergender (*gendered grammar*). Dalam bahasa Arab, gender tidak hanya muncul sebagai identitas biologis atau sosial, tetapi tertanam secara struktural dalam tata bahasa melalui penggunaan pronomina, bentuk verba, nomina, dan sistem sapaan tertentu. Setiap pilihan linguistik dalam bahasa Arab dapat merepresentasikan identitas maskulin maupun feminin secara eksplisit. Oleh karena itu, bahasa Arab tidak hanya berfungsi sebagai alat komunikasi, tetapi juga sebagai medium ideologis yang mereproduksi relasi gender dalam masyarakat.

Struktur gramatikal bahasa Arab secara inheren merepresentasikan relasi gender karena hampir seluruh elemen linguistik memiliki penanda maskulin dan feminin. Dalam praktik komunikasi sehari-hari, bentuk maskulin sering digunakan sebagai bentuk default atau universal, termasuk ketika merujuk pada kelompok campuran laki-laki dan perempuan. Kondisi tersebut menyebabkan maskulinitas memperoleh posisi dominan dalam representasi linguistik dan sosial. Penggunaan pronomina seperti *antum* (kalian laki-laki), bentuk imperatif maskulin, atau sapaan seperti *yā ikhwan* tidak hanya memiliki fungsi gramatikal, tetapi juga membangun relasi kuasa tertentu dalam komunikasi sosial (Muassomah et al., 2025).

Dalam konteks media digital, penggunaan bentuk linguistik maskulin sering dikaitkan dengan representasi otoritas, profesionalisme, dan legitimasi intelektual. Penelitian Rafi’atussyifa dan Nor (2023) menunjukkan bahwa laki-laki dalam vlog YouTube cenderung menggunakan bahasa yang lebih direktif, dominan, dan persuasif dibandingkan perempuan. Penggunaan kalimat imperatif, nada tegas, dan bentuk sapaan otoritatif menjadi strategi linguistik untuk mempertahankan posisi dominan dalam interaksi digital. Dalam konteks pembelajaran bahasa Arab, pola komunikasi tersebut menjadi semakin signifikan karena struktur gramatikal bahasa Arab sendiri mendukung reproduksi maskulinitas sebagai norma dominan.

Selain aspek verbal, konstruksi maskulinitas dalam media juga diperkuat melalui representasi sosial dan visual. Leeder (2024) menjelaskan bahwa maskulinitas hegemonik sering diasosiasikan dengan kemandirian, kontrol, profesionalisme, dan dominasi intelektual. Dalam media digital, laki-laki lebih sering direpresentasikan sebagai figur ahli, pengambil keputusan, atau pemegang otoritas pengetahuan. Ketika figur perempuan jarang hadir dalam ruang edukasi tertentu, maka terjadi apa yang disebut *symbolic annihilation*, yaitu penghilangan simbolik yang membangun persepsi bahwa domain tersebut merupakan ruang eksklusif laki-laki.

Konstruksi gender dalam media digital juga berkaitan erat dengan proses *authenticity labor*, yaitu upaya kreator membangun kesan autentik dan kredibel di hadapan audiens. Dalam konteks YouTube, host laki-laki tidak hanya menggunakan bahasa maskulin, tetapi juga mengombinasikannya dengan gestur direktif, tatapan dominan, intonasi tegas, dan desain visual profesional untuk membangun citra otoritatif.

Dengan demikian, maskulinitas dalam media digital bukan sekadar identitas biologis, tetapi performa sosial dan semiotik yang diproduksi secara strategis melalui integrasi berbagai mode komunikasi.

Dalam penelitian ini, teori konstruksi linguistik gender digunakan untuk memahami bagaimana bentuk-bentuk gramatikal maskulin bahasa Arab dalam saluran *Learn Arabic with Khasu* berfungsi sebagai mekanisme produksi otoritas pedagogis. Penggunaan sapaan kolektif maskulin, struktur imperatif, dan pronomina jamak maskulin dipahami sebagai strategi linguistik yang tidak hanya mengatur komunikasi pembelajaran, tetapi juga mereproduksi relasi kuasa dan legitimasi sosial tertentu dalam ruang digital.

Bias Gender dalam Konten Edukasi Digital

Media digital sering dipersepsikan sebagai ruang demokratis yang memungkinkan siapa pun berpartisipasi secara bebas dalam produksi dan distribusi informasi. Namun, berbagai penelitian menunjukkan bahwa media digital tetap mereproduksi berbagai bentuk bias sosial, termasuk bias gender. Dalam konteks YouTube, representasi laki-laki dan perempuan tidak selalu seimbang karena algoritma dan budaya media digital cenderung memprioritaskan figur tertentu yang dianggap lebih sesuai dengan ekspektasi audiens dan logika platform.

Gezici dan Saygin (2022) menemukan bahwa video edukasi di YouTube menunjukkan bias gender yang signifikan karena laki-laki lebih sering tampil sebagai narator utama dan diasosiasikan dengan otoritas intelektual. Dominasi tersebut terutama terlihat pada bidang-bidang yang dianggap memiliki prestise sosial tinggi, termasuk pendidikan, teknologi, dan agama. Dalam konteks ini, laki-laki lebih sering diposisikan sebagai sumber pengetahuan yang sah, sedangkan perempuan cenderung direpresentasikan secara subordinatif atau bahkan tidak hadir sama sekali dalam ruang visual edukatif (Gezici dan Saygin, 2022).

Fenomena tersebut berkaitan dengan konsep *symbolic annihilation*, yaitu kondisi ketika kelompok tertentu dihilangkan atau direduksi representasinya dalam media sehingga memperkuat marginalisasi sosial mereka. Ketika saluran edukasi secara konsisten hanya menampilkan figur laki-laki sebagai pengajar, maka ruang pendidikan tersebut secara tidak langsung dikonstruksi sebagai domain eksklusif maskulin. Kondisi ini dapat memengaruhi persepsi audiens mengenai siapa yang dianggap memiliki otoritas dan legitimasi dalam bidang tertentu.

Bias gender dalam media digital tidak hanya muncul melalui representasi visual, tetapi juga melalui strategi multimodal yang lebih kompleks. Konstruksi gender dalam media modern dibangun melalui interaksi antara teks verbal, visual, warna, gerak tubuh, dan audio (Asif et al., 2025). Dalam konteks video pembelajaran, laki-laki sering menggunakan gaya komunikasi yang lebih direktif dan dominan, sementara desain visual dan audio mendukung citra profesional serta otoritatif mereka. Bias gender dalam media digital bekerja secara halus melalui integrasi berbagai sumber daya semiotik yang membentuk persepsi audiens secara terus-menerus.

Dalam penelitian ini, kajian bias gender digunakan untuk memahami bagaimana otoritas maskulin direproduksi dalam konten pembelajaran bahasa Arab di YouTube melalui kombinasi strategi verbal, visual, audio, dan algoritmik. Penelitian ini memandang bahwa representasi maskulin dalam media edukasi digital bukan sekadar hasil pilihan individu kreator, tetapi juga dipengaruhi oleh struktur budaya patriarkal dan logika algoritmik platform digital yang lebih luas.

METODE PENELITIAN

Penelitian ini menggunakan pendekatan kualitatif interpretif dengan orientasi analisis multimodal untuk mengkaji secara mendalam bagaimana otoritas maskulin dan strategi persuasi pedagogis dikonstruksi dalam konten pembelajaran bahasa Arab pada saluran YouTube *Learn Arabic with Khasu*. Pendekatan kualitatif interpretif dipilih karena penelitian ini tidak hanya berupaya mengidentifikasi unsur-unsur linguistik secara deskriptif, tetapi juga menafsirkan bagaimana makna sosial, relasi kuasa, identitas gender, dan legitimasi pedagogis diproduksi melalui interaksi berbagai mode semiotik dalam media digital audiovisual. Dalam paradigma interpretif, realitas sosial dipahami sebagai hasil konstruksi makna yang dibentuk melalui praktik komunikasi dan pengalaman sosial, sehingga analisis terhadap teks digital tidak cukup dilakukan melalui penghitungan statistik semata, melainkan memerlukan interpretasi kontekstual terhadap simbol, gestur, visual, dan strategi performatif yang digunakan subjek komunikasi.

Pendekatan ini relevan digunakan karena objek penelitian berupa video pembelajaran YouTube merupakan teks multimodal yang memadukan unsur verbal, visual, audio, dan komposisional secara simultan. Dalam konteks media digital, makna tidak hadir secara tunggal melalui bahasa verbal, tetapi dikonstruksi melalui orkestrasi berbagai sumber daya semiotik yang bekerja secara bersamaan. Oleh

sebab itu, penelitian ini memandang komunikasi digital sebagai praktik sosial yang dipengaruhi oleh konteks budaya, ideologi gender, dan logika platform algoritmik. Dengan menggunakan pendekatan interpretif, peneliti dapat memahami bagaimana identitas maskulin tidak hanya ditampilkan secara visual, tetapi juga dinegosiasikan melalui pilihan bahasa Arab, intonasi suara, framing kamera, gestur tubuh, hingga desain studio yang merepresentasikan profesionalisme dan otoritas edukatif.

Penelitian ini berfokus pada saluran YouTube *Learn Arabic with Khasu* sebagai sumber data utama. Kanal tersebut dipilih karena memperlihatkan karakteristik konten edukasi digital yang sangat kuat secara multimodal, baik dari sisi verbal, visual, maupun audio. Selain itu, kanal ini menunjukkan pola produksi konten yang profesional dan konsisten sehingga dapat dikategorikan sebagai *Professionally Generated Content* (PGC). Konten dalam kanal tersebut tidak hanya menampilkan pengajaran bahasa Arab secara linguistik, tetapi juga memperlihatkan strategi komunikasi visual dan performatif yang secara sistematis membangun citra pengajar laki-laki sebagai figur otoritatif, profesional, dan persuasif. Pemilihan kanal ini juga didasarkan pada relevansinya dengan fokus penelitian mengenai konstruksi otoritas maskulin dalam media edukasi digital berbasis algoritma.

Teknik penentuan data dilakukan melalui metode *purposive sampling*. Teknik ini digunakan karena penelitian tidak membutuhkan jumlah data yang besar secara kuantitatif, melainkan membutuhkan data yang kaya akan elemen multimodal dan relevan dengan fokus penelitian. Peneliti memilih dua belas video dari saluran *Learn Arabic with Khasu* berdasarkan beberapa kriteria utama, yaitu: (1) video memiliki jumlah penonton dan tingkat interaksi tinggi; (2) video menampilkan host laki-laki secara dominan dalam keseluruhan segmen; (3) video memuat interaksi verbal bahasa Arab yang intensif; (4) video memperlihatkan penggunaan elemen visual dan audio secara jelas; dan (5) video memiliki durasi yang representatif untuk dianalisis secara mendalam. Dari dua belas video tersebut diperoleh total 186 ujaran utama yang kemudian ditranskripsi dan dianalisis secara multimodal.

Unit analisis dalam penelitian ini mencakup seluruh tindakan komunikatif yang diproduksi host selama proses pembelajaran berlangsung. Unit analisis tersebut meliputi mode verbal berupa ujaran bahasa Arab, pilihan leksikal, struktur imperatif, bentuk sapaan maskulin, dan pronomina jamak maskulin; mode visual berupa ekspresi wajah, gerakan tangan, arah tatapan mata (*gaze*), orientasi tubuh, posisi kamera, serta desain latar studio; mode audio berupa intonasi, volume suara, tempo bicara, penekanan prosodi, dan jeda dramatik; serta mode komposisional berupa tata letak visual, pencahayaan, framing, thumbnail, warna dominan, dan tipografi Arab yang digunakan dalam video. Seluruh mode tersebut dianalisis sebagai sistem semiotik yang saling berinteraksi dalam membangun makna otoritas dan persuasi maskulin.

Landasan teoritis penelitian ini menggunakan pendekatan *Multimodal Discourse Analysis* (MDA) yang dikembangkan oleh Kress dan van Leeuwen melalui konsep *visual grammar* atau tata bahasa visual. Pendekatan ini digunakan karena mampu menganalisis hubungan antara bahasa, visual, dan audio secara simultan dalam teks audiovisual digital. Kress dan van Leeuwen (2006) menjelaskan bahwa komunikasi visual memiliki struktur semiotik tersendiri sebagaimana bahasa verbal memiliki sistem gramatikal. Dalam kerangka ini, penelitian menganalisis tiga metafungsi utama, yaitu makna representasional (*representational meaning*), makna interaktif (*interactive meaning*), dan makna komposisional (*compositional meaning*). Makna representasional digunakan untuk melihat bagaimana identitas maskulin dan posisi otoritatif direpresentasikan melalui bahasa Arab, pakaian, gestur, dan simbol visual tertentu. Makna interaktif digunakan untuk mengkaji bagaimana hubungan sosial antara host dan audiens dibangun melalui tatapan mata, intonasi, dan pengaturan jarak kamera. Sementara itu, makna komposisional digunakan untuk menganalisis bagaimana tata letak visual, pencahayaan, framing, dan desain studio membentuk koherensi pesan dan strategi persuasi digital.

Selain pendekatan MDA, penelitian ini juga menggunakan kerangka *Multiplicity* sebagai perangkat konseptual tambahan untuk memahami kompleksitas komunikasi digital. Kerangka *Multiplicity* mencakup empat dimensi utama, yaitu *Modes*, *Mediations*, *Varieties*, dan *Purposes*. Dimensi *Modes* berkaitan dengan penggunaan berbagai sumber daya semiotik seperti suara, teks, gerakan tubuh, dan visual dalam membangun makna. Dimensi *Mediations* berkaitan dengan medium komunikasi yang digunakan, yaitu YouTube sebagai platform digital berbasis algoritma yang memengaruhi distribusi dan visibilitas konten. Dimensi *Varieties* menganalisis variasi bahasa Arab yang digunakan, termasuk penggunaan *Modern Standard Arabic* (MSA) dan bentuk-bentuk linguistik maskulin. Sementara itu, dimensi *Purposes* digunakan untuk memahami tujuan komunikatif konten, seperti membangun legitimasi pedagogis, persuasi audiens, dan keterlibatan emosional. Keempat dimensi tersebut dianalisis secara terintegrasi untuk memahami bagaimana performa maskulinitas dibangun dalam konteks media digital kontemporer.

Teknik pengumpulan data dilakukan melalui beberapa tahapan sistematis. Tahap pertama adalah dokumentasi audiovisual dengan mengunduh seluruh video yang telah dipilih dari saluran YouTube *Learn Arabic with Khasu*. Tahap kedua adalah proses transkripsi verbatim terhadap seluruh ujaran verbal bahasa Arab untuk memastikan akurasi analisis linguistik dan pragmatik. Dalam tahap ini, peneliti menuliskan seluruh ujaran sesuai bentuk aslinya dalam bahasa Arab, kemudian menerjemahkannya ke dalam bahasa Indonesia untuk kepentingan interpretasi makna. Tahap ketiga adalah pengambilan *screenshot* pada segmen-segmen penting yang memperlihatkan interaksi antara mode verbal, visual, dan audio, seperti penggunaan direct gaze, gestur menunjuk kamera, atau perubahan intonasi dramatik. Tahap keempat adalah proses pengkodean (*coding*) terhadap seluruh data berdasarkan kategori metafungsi visual Kress dan van Leeuwen serta dimensi Multiplicity.

Proses analisis data dilakukan secara bertahap dan berulang (*iterative analysis*). Pada tahap awal, peneliti melakukan reduksi data dengan mengidentifikasi ujaran, gestur, dan elemen visual yang paling relevan dengan konstruksi otoritas maskulin. Tahap berikutnya adalah kategorisasi data berdasarkan mode semiotik, seperti mode verbal, visual, audio, dan komposisional. Setelah itu, peneliti melakukan interpretasi mendalam terhadap hubungan antar-mode semiotik untuk memahami bagaimana makna otoritas dan persuasi dibangun secara simultan. Analisis dilakukan dengan memperhatikan konteks budaya Arab, struktur gramatikal bahasa Arab yang bergender, serta logika algoritmik YouTube sebagai medium distribusi konten digital. Dengan demikian, penelitian ini tidak hanya memeriksa apa yang dikatakan host, tetapi juga bagaimana pesan tersebut diproduksi dan diperkuat melalui orkestrasi multimodal.

Dalam analisis verbal, peneliti memfokuskan perhatian pada penggunaan bentuk imperatif maskulin, pronomina jamak maskulin seperti *antum*, serta sapaan kolektif seperti *yā ikhwan*. Bentuk-bentuk linguistik tersebut dianalisis sebagai strategi reproduksi relasi kuasa dan legitimasi pedagogis. Dalam analisis visual, peneliti mengkaji penggunaan direct gaze, gestur instruktif, posisi tubuh, dan framing kamera sebagai mekanisme pembentukan dominasi dan kedekatan sosial. Dalam analisis audio, peneliti menelaah bagaimana penekanan intonasi, volume suara, tempo bicara, dan jeda dramatik digunakan untuk mengarahkan perhatian audiens dan memperkuat urgensi pesan. Sementara itu, analisis komposisional berfokus pada desain studio, pencahayaan, thumbnail, warna dominan, dan tata letak visual yang berfungsi sebagai strategi estetika sekaligus strategi visibilitas algoritmik.

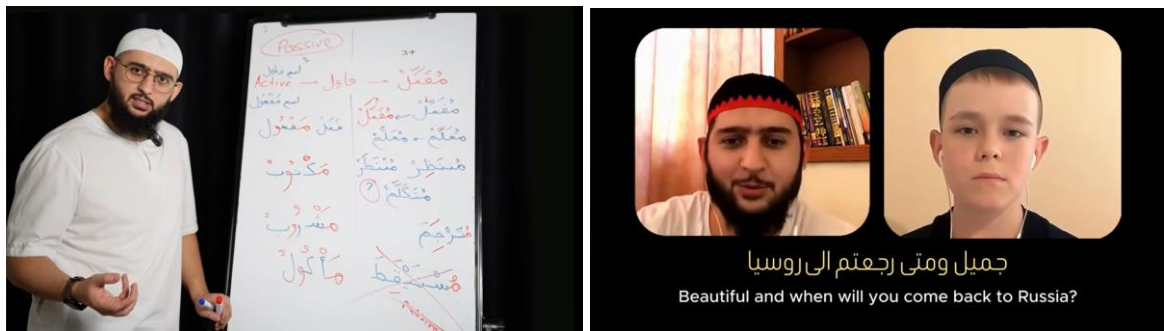
Untuk menjamin validitas dan reliabilitas penelitian, peneliti menggunakan teknik triangulasi teori dan triangulasi mode semiotik. Triangulasi teori dilakukan dengan mengintegrasikan perspektif dari analisis wacana multimodal, sociolinguistik Arab, kajian media digital, dan studi gender untuk menghasilkan interpretasi yang lebih komprehensif. Sementara itu, triangulasi mode semiotik dilakukan dengan membandingkan dan memverifikasi hubungan antara data verbal, visual, dan audio sehingga hasil analisis tidak hanya bergantung pada satu jenis data tertentu. Peneliti juga melakukan pengecekan berulang terhadap hasil transkripsi dan pengkodean untuk meminimalkan bias interpretatif dan menjaga konsistensi kategorisasi data.

Dalam aspek etika penelitian, seluruh data yang digunakan berasal dari konten publik yang tersedia secara terbuka di platform YouTube. Namun demikian, penelitian tetap memperhatikan prinsip etika akademik dengan tidak memanipulasi data, tidak mengubah konteks ujaran, serta menggunakan data hanya untuk kepentingan ilmiah. Identitas personal di luar nama kanal publik tidak diungkapkan secara berlebihan, dan seluruh proses analisis dilakukan secara objektif sesuai kaidah penelitian ilmiah. Dengan prosedur metodologis tersebut, penelitian ini diharapkan mampu menghasilkan pemahaman yang mendalam mengenai bagaimana otoritas maskulin dikonstruksi secara multimodal dalam lanskap media edukasi digital kontemporer.

HASIL DAN PEMBAHASAN

Gambaran Umum Pola Multimodal dalam Konstruksi Otoritas Maskulin

Hasil analisis terhadap 12 video terpilih dari saluran *Learn Arabic with Khasu* menunjukkan bahwa konstruksi otoritas maskulin tidak dibangun melalui satu mode semiotik tunggal, melainkan melalui orkestrasi simultan antara pilihan leksikal bahasa Arab, gestur instruktif, kontrol intonasi, framing kamera, serta desain studio yang menyerupai ruang akademik formal. Dari total 186 ujaran yang ditranskripsi, ditemukan 74 ujaran (39,7%) yang secara eksplisit menggunakan bentuk gramatikal maskulin atau sapaan kolektif maskulin seperti *يا إخوان* (*yā ikhwan*) dan *أنتم* (*antum*). Pola ini memperlihatkan bahwa maskulinitas bukan sekadar identitas biologis host, melainkan performa linguistik yang secara konsisten direproduksi.

Gambar 4.1 Tangkapan layar video *Learn Arabic with Khasu*

Secara representasional, host diposisikan sebagai figur ahli melalui kombinasi pakaian formal (kemeja polos atau gamis sederhana), latar belakang rak buku, serta penggunaan papan tulis digital. Secara interaktif, penggunaan tatapan langsung ke kamera (direct gaze) dalam 81% segmen analisis membangun relasi dominasi sekaligus keintiman pedagogis. Sementara itu, secara komposisional, pencahayaan frontal terang dan framing medium close-up memberikan nilai salience tinggi pada wajah dan gestur tangan, memperkuat legitimasi instruksional.

Analisis leksikal menunjukkan kecenderungan penggunaan struktur imperatif maskulin serta bentuk fi'il mudhāri' berpenanda jamak maskulin. Berikut matriks data utama yang mengilustrasikan temuan tersebut:

No	Ujaran Bahasa Arab	Terjemahan	Jumlah Kemunculan	Mode Verbal	Mode Visual	Mode Audio
1	يا إخوان انتبهوا جيداً	Wahai saudara-saudara, perhatikan baik-baik	18x	Sapaan maskulin kolektif	Telunjuk menunjuk kamera	Intonasi naik-tegas
2	ركزوا معي	Fokuslah bersama saya	21x	Imperatif maskulin	Gestur tangan terbuka	Penekanan pada "معي"
3	أنتم تستطيعون	Kalian (maskulin) bisa	14x	Pronoun jamak maskulin	Anggukan kepala	Nada afirmatif
4	هذا مهم جداً	Ini sangat penting	27x	Deklaratif otoritatif	Tangan mengepal ringan	Tekanan kuat pada "مهم"
5	اسمعوا النطق الصحيح	Dengarkan pelafalan yang benar	12x	Imperatif edukatif	Mendekat ke kamera	Volume diturunkan (dramatic pause)

Tabel 4.1 Data Ujaran Verbal dan Mode Multimodal

Tabel tersebut memuat lima ujaran bahasa Arab yang diproduksi secara berulang dalam konteks penyampaian materi. Setiap data dianalisis berdasarkan ujaran Bahasa Arab, terjemahan, jumlah kemunculan, mode verbal, mode visual, dan mode audio. Struktur ini menunjukkan bahwa makna tidak hanya dibangun melalui bahasa (verbal), tetapi juga melalui gestur (visual) dan kualitas suara (audio). Dengan kata lain, otoritas, persuasi, dan konstruksi posisi pembicara dimediasi secara multimodal.

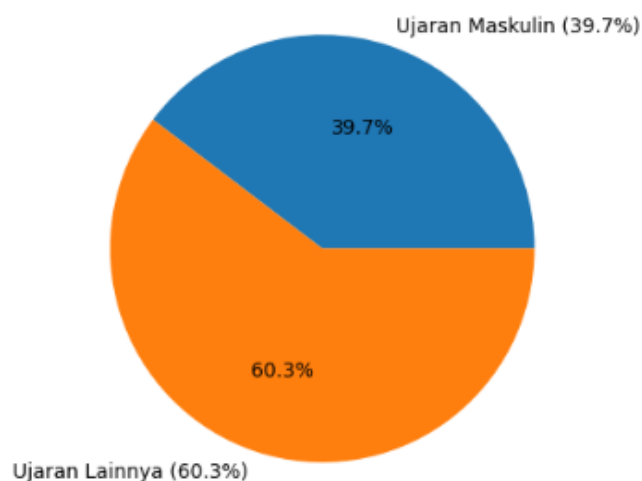


Diagram 4.2 Distribusi Ujaran Maskulin dari total data

Diagram pertama menunjukkan bahwa dari total 186 ujaran yang ditranskripsi, sebanyak 39,7% secara eksplisit menggunakan bentuk gramatikal maskulin atau sapaan kolektif maskulin, sedangkan 60,3% lainnya berupa ujaran netral atau bentuk lain yang tidak secara langsung menandai maskulinitas gramatikal.

Proporsi 39,7% ini bukan angka kecil. Hampir separuh dari keseluruhan produksi verbal mengandung penandaan maskulin yang jelas. Hal ini menunjukkan bahwa maskulinitas dalam konteks ini bukan sekadar identitas biologis pembicara, melainkan performa linguistik yang diproduksi secara konsisten melalui pilihan morfologi dan pronomina bahasa Arab. Dalam sistem gramatika Arab, bentuk maskulin jamak sering menjadi default, tetapi frekuensi tinggi dalam konteks pedagogis ini memperlihatkan adanya reproduksi posisi otoritatif yang diasosiasikan dengan bentuk maskulin.

Sementara itu, bagian 60,3% menunjukkan bahwa tidak semua ujaran dimuati penanda maskulin eksplisit. Namun, keberadaan hampir 40% ujaran maskulin sudah cukup kuat untuk membangun framing identitas audiens sebagai kelompok yang diasumsikan maskulin secara gramatikal.

1. يا إخوان انتبهوا جيداً (18x)

Makna: “Wahai saudara-saudara, perhatikan baik-baik.”

Secara verbal, ujaran ini menggunakan sapaan kolektif maskulin *يا إخوان* yang mengonstruksi audiens sebagai kelompok laki-laki atau diasumsikan maskulin secara gramatikal. Bentuk imperatif *انتبهوا* mempertegas relasi hierarkis: pembicara berada pada posisi memberi instruksi. Secara visual, telunjuk yang menunjuk kamera menciptakan efek deiksis langsung, seolah-olah audiens sedang “ditunjuk” secara personal. Ini memperkuat daya tekan perintah. Dari sisi audio, intonasi naik dan tegas memperjelas fungsi direktifnya.

2. ركزوا معي (21x)

Makna: “Fokuslah bersama saya.”

Ini merupakan data dengan frekuensi cukup tinggi (21 kali), menunjukkan strategi repetisi sebagai teknik pedagogis. Secara verbal, bentuk imperatif maskulin *ركزوا* menegaskan ajakan yang bersifat kolektif. Frasa *معي* (bersama saya) secara implisit membangun relasi kedekatan, namun tetap dalam struktur komando. Secara visual, gestur tangan terbuka memberi kesan inklusif dan mengundang partisipasi. Dari aspek audio, penekanan pada kata *معي* memperkuat sentralitas pembicara sebagai sumber otoritas pengetahuan.

3. أنتم تستطيعون (14x)

Makna: “Kalian bisa.”

Secara verbal, penggunaan pronomina jamak maskulin *أنتم* mengonstruksi identitas audiens secara eksplisit. Kalimat deklaratif ini bersifat afirmatif dan memotivasi. Visualnya berupa anggukan kepala yang menandakan persetujuan dan penguatan pesan. Nada afirmatif dalam mode audio memperkuat fungsi motivasionalnya. Berbeda dari dua data sebelumnya yang bersifat imperatif, ujaran ini lebih bersifat empowerment. Namun tetap dalam bingkai relasi guru–murid yang asimetris.

4. هذا مهم جدًا (27x)

Makna: “Ini sangat penting.”

Ini adalah data dengan frekuensi tertinggi (27 kali), yang menunjukkan fokus utama pada penegasan urgensi materi. Secara verbal, struktur deklaratif otoritatif membangun klaim kebenaran yang tidak diperdebatkan. Visual berupa tangan mengepal ringan memberi efek penegasan dan keseriusan. Dalam mode audio, tekanan kuat pada kata *مهم* menandakan intensifikasi makna. Multimodalitas di sini berfungsi untuk membangun legitimasi dan urgensi, sekaligus mengarahkan perhatian audiens pada poin tertentu.

5. اسمعوا النطق الصحيح (12x)

Makna: “Dengarkan pelafalan yang benar.”

Bentuk imperatif edukatif ini secara langsung berkaitan dengan fungsi fonetik atau pelafalan. Secara visual, mendekat ke kamera menciptakan efek intimasi dan fokus. Volume suara yang diturunkan (dramatic pause) dalam mode audio membangun ketegangan dan perhatian selektif. Strategi ini menunjukkan teknik pedagogis berbasis kontrol atensi melalui manipulasi jarak visual dan dinamika suara.

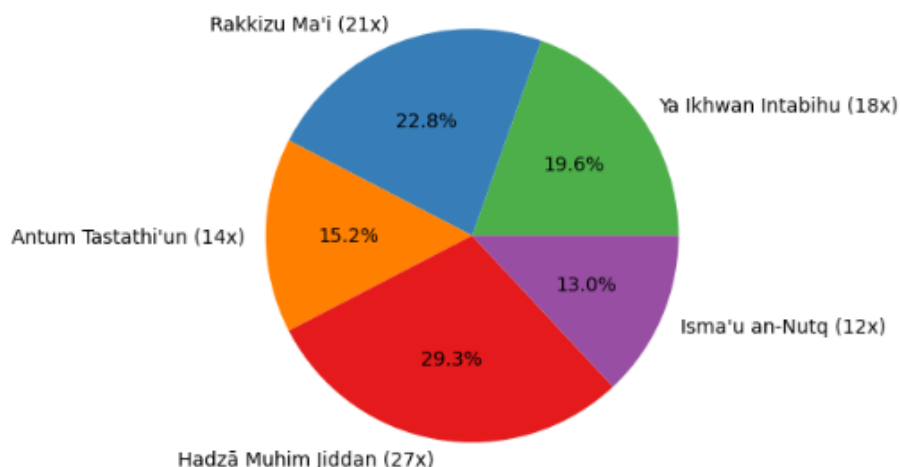


Diagram 4.1 Distribusi lima ujaran dominan

Diagram kedua memperlihatkan distribusi frekuensi lima ujaran utama yang paling sering diulang dalam video. Ujaran dengan frekuensi tertinggi adalah “هذا مهم جدًا” (27x), diikuti “ركزوا معي” (21x), “يا إخوان انتبهوا” (18x), “أنتم تستطيعون” (14x), dan “اسمعوا النطق الصحيح” (12x).

Dominasi “هذا مهم جدًا” (sekitar 29,3% dari lima data utama) menunjukkan bahwa strategi pedagogis utama yang digunakan adalah penekanan urgensi materi. Kalimat deklaratif ini membangun klaim kebenaran yang tidak diperdebatkan. Ketika suatu pernyataan diulang paling sering dan diberi tekanan visual serta audio, hal tersebut berfungsi untuk mengarahkan fokus kognitif audiens pada poin-poin tertentu yang dianggap krusial.

Frekuensi tinggi pada bentuk imperatif seperti “ركزوا معي” dan “يا إخوان انتبهوا جيدًا” memperlihatkan dominasi struktur komando dalam relasi guru–murid. Imperatif bukan hanya alat instruksi, tetapi juga

penanda relasi kuasa. Sementara itu, “أنتم تستطيعون” berfungsi sebagai ujaran motivasional, namun tetap dalam bingkai asimetris karena guru tetap menjadi sumber validasi kemampuan audiens.

Distribusi ini menunjukkan bahwa pola komunikasi tidak seimbang secara netral, melainkan berorientasi pada kontrol perhatian, penekanan materi, dan reproduksi hierarki pedagogis melalui repetisi.

Pola Representatif dari Keseluruhan Data

Jika dilihat secara keseluruhan, terdapat beberapa pola penting:

1. Dominasi Bentuk Imperatif

Sebagian besar ujaran berbentuk perintah, menunjukkan relasi kuasa yang jelas antara pembicara dan audiens.

2. Konsistensi Maskulinitas Gramatikal

Semua pronomina dan bentuk kata kerja menggunakan bentuk maskulin jamak. Ini berimplikasi pada konstruksi audiens yang diasumsikan maskulin atau default maskulin dalam bahasa Arab.

3. Repetisi sebagai Strategi Pedagogis

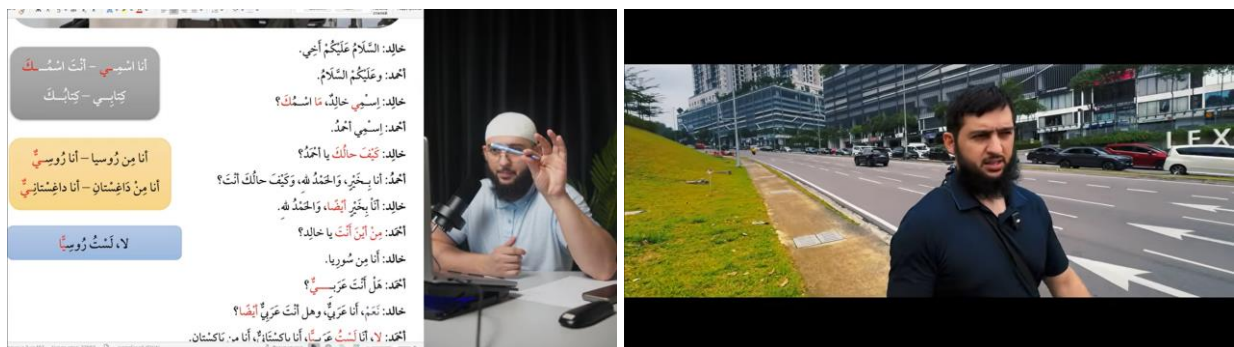
Frekuensi tinggi (terutama pada “هذا مهم جداً”) menunjukkan penggunaan repetisi untuk memperkuat internalisasi makna.

4. Koherensi Multimodal

Mode verbal, visual, dan audio saling menguatkan (reinforcing mode), bukan saling bertentangan. Setiap pesan linguistik diperkuat dengan gestur dan prosodi yang selaras.

Makna Interaktif: Negosiasi Dominasi dan Keintiman

Makna interaktif terlihat melalui tiga elemen utama: gaze, proximity, dan intonasi. Host menggunakan direct gaze dalam sebagian besar segmen, menciptakan efek “demand image” yang menuntut perhatian audiens. Jarak kamera medium close-up memperpendek jarak sosial sehingga membangun kedekatan semu, tetapi tetap mempertahankan hierarki melalui sudut kamera eye-level yang stabil.



Gambar 4.2 Tangkapan layar video *Learn Arabic with Khasu*

Menariknya, ketika menjelaskan konsep tata bahasa kompleks seperti fi'il amar (kata kerja perintah), host menurunkan volume suara dan memperlambat tempo bicara. Strategi audio ini berfungsi sebagai authenticity labor membangun kesan pedagogis yang sabar namun tetap otoritatif.

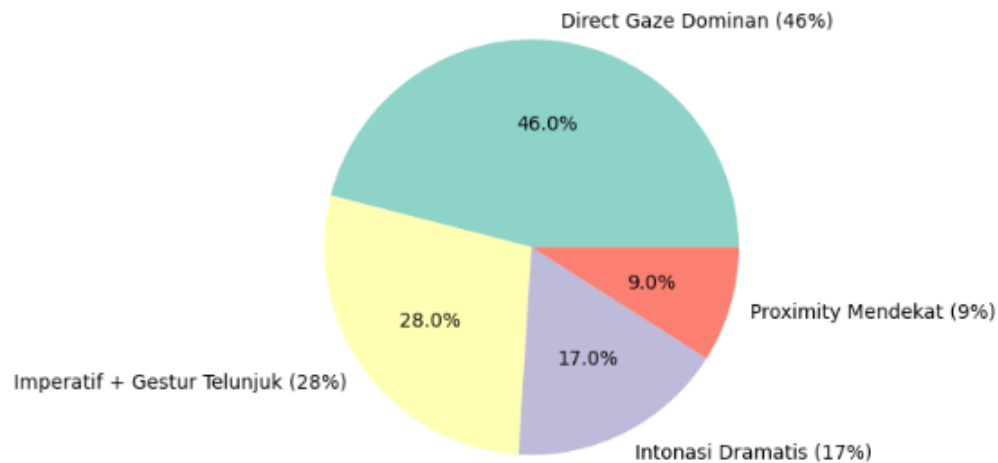


Diagram 4.3 Distribusi Mode Interaktif

Angka 46% pada direct gaze menegaskan kembali bahwa tatapan menjadi strategi utama dalam membangun relasi interaktif. Ini adalah fondasi visual dari konstruksi otoritas.

Sebesar 28% penggunaan imperatif yang diperkuat gestur telunjuk menunjukkan bahwa instruksi tidak hanya verbal, tetapi juga diperkuat secara kinestetik. Gestur menunjuk kamera menciptakan efek deiksis langsung, seolah-olah audiens ditargetkan secara personal. Hal ini mempertegas posisi pembicara sebagai pemberi komando.

Sebanyak 17% intonasi dramatis memperlihatkan pentingnya aspek prosodi dalam membangun intensitas makna. Tekanan suara dan perubahan volume menciptakan fokus selektif pada bagian tertentu materi. Sementara itu, 9% proximity mendekat ke kamera menunjukkan bahwa manipulasi jarak visual digunakan secara selektif untuk momen-momen penting, seperti penjelasan fonetik.

Distribusi Mode Interaktif:

- Direct Gaze Dominan: 46%
- Imperatif + Gestur Telunjuk: 28%
- Intonasi Penekanan Dramatis: 17%
- Proximity Mendekat ke Kamera: 9%

Makna Komposisional dan Mediasi Algoritma

Secara komposisional, seluruh video menggunakan latar bersih dengan pencahayaan terang, resolusi tinggi, serta thumbnail kontras tinggi dengan teks Arab besar berwarna kuning atau putih. Desain ini bukan hanya estetika, tetapi sinyal algoritmik yang meningkatkan click-through rate (CTR).

Thumbnail umumnya menampilkan ekspresi wajah serius atau fokus dengan teks seperti: تعلم العربية بسهولة (Belajar Bahasa Arab dengan Mudah).

Kombinasi warna kontras, framing wajah dominan, serta tipografi Arab tebal memperkuat nilai salience. Dalam perspektif makna komposisional, wajah host ditempatkan di pusat (centered composition) sehingga menjadi elemen informasi utama (nucleus of meaning).

Dari sisi mediasi algoritma, pola unggahan konsisten (2–3 video per minggu) serta durasi ideal 8–12 menit menunjukkan keselarasan dengan logika retensi penonton YouTube. Dengan demikian, otoritas maskulin yang dibangun tidak hanya bersifat kultural, tetapi juga terintegrasi dalam strategi visibilitas platform.

Secara keseluruhan, hasil penelitian ini menunjukkan bahwa konstruksi otoritas maskulin dalam saluran pembelajaran bahasa Arab tersebut merupakan hasil interaksi kompleks antara:

1. **Mode Verbal** → Dominasi imperatif maskulin dan sapaan kolektif pria.
2. **Mode Visual** → Gestur direktif, direct gaze, framing sentral.
3. **Mode Audio** → Intonasi tegas dengan penekanan selektif.
4. **Mode Komposisional** → Studio profesional dan thumbnail kontras tinggi.
5. **Mediasi Algoritmik** → Optimasi estetika dan durasi konten.

Maskulinitas di sini tampil sebagai performa pedagogis yang terencana dan dikenali secara algoritmik. Otoritas tidak hanya dihasilkan oleh kompetensi linguistik, tetapi juga oleh kemampuan mengelola orkestrasi multimodal yang selaras dengan tata kelola platform digital.

SIMPULAN

Berdasarkan hasil analisis terhadap 12 video pembelajaran bahasa Arab, dapat disimpulkan bahwa konstruksi otoritas maskulin dalam saluran tersebut terbentuk melalui orkestrasi multimodal yang konsisten dan sistematis. Otoritas tidak hanya dihasilkan oleh kompetensi linguistik pembicara, tetapi oleh interaksi kompleks antara mode verbal, visual, audio, komposisional, dan mediasi algoritmik.

Secara verbal, dominasi bentuk imperatif maskulin dan penggunaan pronomina jamak maskulin membangun relasi kuasa yang hierarkis antara guru dan audiens. Maskulinitas gramatikal berfungsi sebagai mekanisme linguistik yang secara implisit mengonstruksi audiens dalam kerangka default maskulin. Repetisi ujaran dengan frekuensi tinggi memperkuat internalisasi makna serta mempertegas posisi pembicara sebagai sumber legitimasi pengetahuan.

Secara visual, penggunaan *direct gaze*, gestur direktif, dan framing sentral menciptakan kombinasi antara dominasi dan kedekatan. Strategi ini memungkinkan pembicara mempertahankan otoritas sekaligus membangun keintiman pedagogis. Dari sisi audio, variasi intonasi, penekanan selektif, dan kontrol volume menjadi instrumen pengelolaan atensi yang efektif.

Pada tingkat komposisional dan algoritmik, profesionalitas visual serta konsistensi format konten menunjukkan bahwa otoritas maskulin juga diproduksi dalam keselarasan dengan logika platform digital. Dengan demikian, maskulinitas dalam konteks ini tampil sebagai performa pedagogis yang terstruktur dan dikenali secara algoritmik.

Penelitian ini menegaskan bahwa dalam ruang digital, identitas dan otoritas tidak bersifat natural, melainkan dikonstruksi melalui praktik multimodal yang terencana. Temuan ini membuka peluang penelitian lanjutan mengenai relasi gender, pedagogi, dan mediasi algoritma dalam konteks pembelajaran daring.

DAFTAR PUSTAKA

- Atma, S., & Awad, Z. (2023). Gender stereotypes in TV commercials: A multimodal analysis approach. *Theory and Practice in Language Studies*, 13(9), 2365-2374. <https://doi.org/10.17507/tpls.1309.24>
- Abdullah, N. I. binti, Sarudin, A. binti, Osman, Z., Mohamed, H. F., Redzwan, R., Muhammad, M. M., & Idris, M. (2021). Elements of multimodality in cartoon through linguistic mode and visual mode. *Theory and Practice in Language Studies*, 11(2), 191-201. <http://ejournal.unp.ac.id/index.php/humanus/index>
- Al Atoom, M., Alzoubi, Z., Al-Momani, H., & Al-Khawaldeh, R. (2024). Development of gender identity during adolescence. *Journal of Adolescence*, 96(2), 215-229. <https://doi.org/10.1080/02673843.2024.2312853>
- Andersson, M. (2021). The climate of climate change: Impoliteness as a hallmark of homophily in YouTube comment threads on Greta Thunberg's environmental activism. *Journal of Pragmatics*, 178, 93-107. <https://doi.org/10.1016/j.pragma.2021.03.003>
- Andersson, M. (2024). Multimodal expression of impoliteness in YouTube reaction videos to transgender activism. *Discourse, Context & Media*, 58, 100760. <https://doi.org/10.1016/j.dcm.2024.100760>
- Asif, S., et al. (2025). A multimodal discourse analysis: Representation of women in feminine hygiene product. *ResearchGate*. <https://doi.org/10.71281/jals.v3i4.471>
- Balestri, R. (2024). *Examining multimodal gender and content bias in ChatGPT-4o*. arXiv. <https://arxiv.org/abs/2411.19140>
- Bateman, J. A., & Tseng, C.-I. (2023). Multimodal discourse analysis as a method for revealing narrative strategies in news videos. *Multimodal Communication*, 12(3), 261-285. <https://doi.org/10.1515/mc-2023-0029>
- Belgrimet, S., & Rabab'ah, G. (2021). A multimodal discourse analysis of English posters in violence awareness campaigns against women. *Theory and Practice in Language Studies* <http://dx.doi.org/10.17507/tpls.1102.12>

- Bhatia, A. (2022). Analysing online videos. In C. Vasquez (Ed.) *Research Methods for Digital Discourse Analysis*.
<https://www.bloomsbury.com/us/research-methods-for-digital-discourse-analysis-9781350166837/>
- Bugshan, O. A. S. (2025). *Similarities and Differences in Two Adult Learners' Communicative Repertoires in Arabic as an Additional Language: A Study of Sociolinguistic and Contextual Alignment in the Saudi Hijazi Context* (Doctoral dissertation, La Trobe).
<https://doi.org/10.26181/29937047>
- Dewi, N. P., & Tawami, T. (2022). Representation of Writing Skill Importance In a Video Advertisement From The Perspective of Multimodal Discourse Analysis. *International Journal of Education, Information Technology, and Others*, 5(4), 95-109.
<https://jurnal.peneliti.net/index.php/IJEIT/sitemap>
- Discourse Analyzer. (2024, August 16). *Kress and van Leeuwen in discourse analysis*.
<https://discourseanalyzer.com/kress-and-van-leeuwen-in-discourse-analysis/>
- Dontcheva-Navratilova, O. (2023). Persuasion in multimodal digital genres: Building credibility in video abstracts. *ESP Today*, 11(2), 213-236. <https://doi.org/10.18485/esptoday.2023.11.2.2>
- Firmansyah, M. B., Suchaina, S., & Ridwan, M. (2025). *Multimodal discourse analysis in learning media: A case study of promotional content in Pasuruan*. AL-ISHLAH: Jurnal Pendidikan, 17(2), 2991–3004. <https://doi.org/10.35445/alishlah.v17i2.7334>
- Fitriansyah, N., & Lubis, T. (2023). Mediating Alghorithm Mediating Da'wa: The New Preacher and Optimization of Social Media for Da'wa in the Case of Habib Ja'far. *Jurnal Penelitian*, 1-12.
<https://doi.org/10.28918/jupe.v20i1.1092>
- Gezici, G., & Saygin, Y. (2022). Measuring gender bias in educational videos: A case study on youtube. *arXiv preprint arXiv:2206.09987*.
<https://doi.org/10.48550/arXiv.2206.09987>
- Ghiglini, C., & Tabasso, N. (2024). Endogenous identity in a social network. *arXiv Preprint arXiv:2406.10972*. <https://doi.org/10.48550/arXiv.2406.10972>
- Hoffmann, C. (2021). *How multimodal cohesion secures narrative orientation in film trailers*. *Discourse, Context & Media*, 42, 100515. <https://doi.org/10.1016/j.dcm.2021.100539>
- Kilby, L. (2021). *When words are not enough: Combined textual and visual analysis in multimodal communication*. *Methods in Psychology*, Volume 5, December 2021, 100071.
<https://doi.org/10.1016/j.metip.2021.100071>
- Kress, G., & van Leeuwen, T. (2006). *Reading Images: The Grammar of Visual Design* (2nd ed.). London: Routledge. <https://doi.org/10.4324/9781003099857>
- Le, T. K. C. A. (2024). *The construction of national identity in a state without nation* (Master's thesis, University of Groningen). University of Groningen Research Database.
<https://campus-fryslan.studenttheses.ub.rug.nl/id/eprint/407>
- Leeder, G. F. (2024). Masculine, independent, and “not acting”: Hegemonic masculinity, femmephobia, and queer men. *Sociological Forum*. <https://doi.org/10.1111/socf.12991>
- Lennon, H. W., & Kilby, L. (2021). A multimodal discourse analysis of ‘Brexit’: Flagging the nation in political cartoons. In M. Demasi, S. Burke, & C. Tileagă (Eds.), *Political communication: Discursive perspectives* (pp. 115–146). Palgrave Macmillan. https://doi.org/10.1007/978-3-030-60223-9_5
- Lim, F. V., & Ang, J. (2024). Commodifying the self: a multimodal analysis of college YouTubers' first day videos. *The Journal of International Communication*, 30(1), 111-134.
<https://doi.org/10.1080/13216597.2023.2245384>
- Lindenberg, D. (2023). *An SF-informed multimodal discourse analysis*. *English for Specific Purposes*, Volume 69, Pages 67-79. <https://doi.org/10.1016/j.esp.2022.10.002>
- Lisnarti, L. (2025). Uses and Gratifications dalam Media Sosial Algoritmik: Tantangan dan Kesenjangan antara Gratifications Sought dan Gratifications Obtained. *Liwaul Dakwah: Jurnal Kajian Dakwah dan Masyarakat Islam*, 15(2), 162-173. <https://doi.org/10.47766/liwauldakwah.v15i2.7095>
- Meirandani, A., Ulmasitoh, C., Aprilia, L., & Nargis, N. (2025). Analysis Gender and Language in Online Communication (Case Study: Tiktok and Instagram). *Variable Research Journal*, 2(01), 116-121.
<https://variablejournal.my.id/index.php/VRJ/article/view/168>
- Moya-Guijarro, A. J. (2024). *A multimodal analysis of character–character interaction in visual narratives*. *Linguistics and Education*, Volume 82, 101312.
<https://doi.org/10.1016/j.linged.2024.101312>

- Muassomah, M., Abdullah, I., Wahyuni, I., Burhan, A., Yelfi, D., Wargadinata, W., & Billah, B. G. M. (2025). The linguistic construction of gender: Masculinity and femininity in Arabic grammar. *Theory and Practice in Language Studies*, 15(4), 1207-1216. <https://doi.org/10.17507/tpls.1504.19>
- Nugroho, A. A. P., & Seinsiani, I. G. (2025, August). The analysis of code switching and code mixing in the 'Historytelling' YouTube segment by Nessie Judge. In *ELT Forum: Journal of English Language Teaching* (Vol. 14, No. Special Issue, pp. 265-279). <https://doi.org/10.15294/elt.v14iSpecial%20Issue.28569>
- Rafi'atussyifa, G., & Nor, H. (2023). Gender Disparities in Utilizing Language Features on YouTube Vlogs. *Muadalah*, 11(2), 125-134. <https://doi.org/10.18592/muadalah.v11i2.11268>
- Rahmadani, L., & Sahiruddin, A. L. (2023). A Multimodal Discourse Analysis of Gender Representation in Three Indonesian Sanitary Pad TV Advertisements. In *Icon Laterals 2023: Proceedings of the 4th International Conference Entitled Language, Literary, And Cultural Studies, Icon Laterals 2023, 11-12 July 2023, Malang, Indonesia* (p. 185). European Alliance for Innovation. <https://doi.org/10.4108/eai.11-7-2023.2340594>
- Rahmadani, L., et al. (2024). *A multimodal discourse analysis of gender representation in advertisements*. Proceedings of the EAI International Conference. <https://doi.org/10.4108/eai.11-7-2023.2340594>
- Ramadhan, G. (2025, June). Wacana Visual Dan Verbal Dalam Video Pembelajaran BIPA di youtube: Analisis Multimodal. In *Prosiding Seminar Nasional Bahasa Indonesia Bagi Penutur Asing* (Vol. 1, No. 1, pp. 185-189). <https://proceeding.umsu.ac.id/index.php/sebipa/article/view/1094/pdf>
- Sari, N. (2024). A socio-semiotic study of memes among Indonesian Gen Z. *Communica: Journal of Communication Studies*, 5(2), 45–60. <https://doi.org/10.61978/communica.v2i3.775>
- Sari, G. G., Udasmoro, W., & Noviani, R. (2025). The politics of masculinity: male virility and social coercion in Indonesian cinema. *Jurnal Komunikasi*, 19(3). <https://doi.org/10.20885/komunikasi.vol19.iss3.art7>
- Sherwani, K. A., & Harchegani, M. K. (2022). The impact of multimodal discourse analysis on the improvement of Iraqi EFL learners' reading comprehension skill. *International Journal of Instruction*, 15(1), 681–696. <https://doi.org/10.25130/jtuh.29.12.2.2022.22>
- Silitonga, N., Sitompul, S., Pasaribu, S., Silitonga, S., & Natsir, M. (2025). Multimodality analysis in YouTube educational content (Case study on the realization of meaning in *SUARA BERKELAS* video “Kalau kamu mau punya critical thinking, jangan skip obrolan ini!”). *Jurnal Riset Rumpun Ilmu Bahasa*, 4(3), 283–291. <https://doi.org/10.55606/jurribah.v4i3.7028>
- Wood, R. (2021). ‘What I’m not gonna buy’: Algorithmic culture jamming and anti-consumer politics on YouTube. *New Media & Society*, 23(9), 2754-2772. <https://doi.org/10.1177/1461444820939446>
- Xia, Y. (2023). *A multimodal genre analysis of TED Talk videos*. *Journal of English for Academic Purposes*, 59, 101277. <https://doi.org/10.1016/j.esp.2022.11.007>
- Yuliani, S. Y., Al Farisi, M. Z., & Supriadi, R. (2025). Multimodal-Pragmatic Analysis of Arabic–Indonesian Subtitle Translation in the Animated Film Muhsin and Mahasin: The Episode “The Danger of Boasting”. *Al-Fathin: Jurnal Bahasa dan Sastra Arab*, 8(02), 234-252. <https://e-journal.metrouniv.ac.id/al-fathin/en/article/view/11868>