Alih Kode dan Campur Kode dalam Siniar 'Malaysians speaking `Arabic?': Kajian Sosiolinguistik pada Penutur Multilingual

Naila Rahmaridha[1], Nasikul Mustofa Efendi[2], Muhammad Bachrul Ulum[3]
Prodi Bahasa dan Sastra Arab, Fakultas Adab dan Humaniora,
Universitas Islam Negeri Sunan Ampel Surabaya[123]
naeeworkspace@gmail.com nasikul.mustofa@uinsa.ac.id bachrul.ulum@uinsa.ac.id

Abstract:

This study aims to identify, classify, and describe the forms of code-switching and code-mixing, and to discuss the factors underlying their use in the podcast "Malaysians Speaking Arabic?". The podcast features a conversation in Saudi-dialect Arabic as the main language between two multilingual speakers of Thai and Malaysian backgrounds who once lived in Saudi Arabia. This study was written using a qualitative descriptive method with the sociolinguistic theoretical framework of Chaer and Agustina, which distinguishes integration and interference from code-mixing. The data source consists of words or sentences of utterances taken from the transcript of a 26-minute 51-second podcast video, collected through the non-participatory observation technique (simak bebas libat cakap) and note-taking, then analyzed with intralingual and extralingual matching methods (padan). Each finding of a linguistic phenomenon counts as one datum. The results identify 147 data, classified into 113 code-mixing, 12 code-switching, 16 integration, and 6 morphological interference. Code-mixing is the most frequent category, in the forms of 70 words, 23 phrases, 16 clauses, 3 idioms, and 1 reduplication. Of the 113 code-mixing data, 107 involve English elements, while Malay and Thai elements are found in smaller numbers. In general, the two speakers' shared repertoire, the presence of the audience as a third party, topic shifts, and discourse markers explain this pattern. This study also considers the commonness of loanword use in informal conversation as a basis for distinguishing it from code-mixing phenomena.
Keywords: code-switching; code-mixing; interference; integration; podcast

Abstrak:

Penelitian ini bertujuan untuk mengidentifikasi, mengklasifikasi, dan mendeskripsikan bentuk-bentuk alih kode dan campur kode, serta membahas faktor yang melatarbelakangi penggunaannya dalam siniar "Malaysians Speaking Arabic?". Siniar tersebut menampilkan percakapan berbahasa Arab dialek Saudi sebagai bahasa utama antara dua penutur multilingual berlatar belakang Thailand dan Malaysia yang pernah tinggal di Arab Saudi. Penelitian ini ditulis menggunakan metode deskriptif kualitatif dengan pendekatan kerangka teori sosiolinguistik Chaer dan Agustina yang membedakan integrasi dan interferensi dari campur kode. Sumber data penelitian berupa kata atau kalimat tuturan yang diambil dari data transkrip video siniar berdurasi 26 menit 51 detik yang dikumpulkan melalui teknik simak bebas libat cakap dan catat, lalu

dianalisis dengan metode padan intralingual dan ekstralingual. Setiap temuan fenomena kebahasaan terhitung sebagai satu data. Hasil penelitian mengidentifikasi 147 data yang diklasifikasikan menjadi 113 campur kode, 12 alih kode, 16 integrasi, dan 6 interferensi morfologis. Campur kode merupakan kategori terbanyak, dengan bentuk berupa 70 kata, 23 frasa, 16 klausa, 3 idiom, dan 1 perulangan. Dari 113 data campur kode tersebut, 107 data melibatkan unsur bahasa Inggris, sedangkan unsur bahasa Melayu dan Thai ditemukan dalam jumlah lebih sedikit. Secara umum, repertoar bersama kedua penutur, kehadiran penonton sebagai pihak ketiga, pergantian topik, dan penanda wacana menjelaskan pola tersebut. Penelitian ini juga mempertimbangkan kelaziman penggunaan kata pinjaman dalam percakapan informal sebagai dasar untuk membedakannya dari gejala campur kode.
Kata kunci: alih kode; campur kode; interferensi; integrasi; siniar

PENDAHULUAN

Globalisasi dan perkembangan teknologi komunikasi telah memperluas interaksi antarpengguna bahasa dengan latar belakang linguistik yang beragam. Perkembangan media digital turut menyediakan ruang bagi praktik kebahasaan tersebut untuk hadir dalam bentuk percakapan yang spontan, informal, dan dapat diakses oleh khalayak luas. YouTube, misalnya, tidak hanya menjadi media penyebaran informasi dan hiburan, tetapi juga menghadirkan representasi penggunaan bahasa dalam interaksi sehari-hari (Amelia et al., 2026).

Siniar di YouTube mempertemukan obrolan santai dengan penonton yang tidak terlihat. Dua orang berbincang di depan kamera, sementara ribuan orang menyimak dan menulis komentar di bawah video. Dalam ruang seperti itu, penutur multibahasa jarang bertahan pada satu bahasa. Mereka menyelipkan satu kata, memotong kalimat dengan bahasa lain, atau berpindah penuh ke bahasa kedua karena berbagai faktor. Dalam kajian sosiolinguistik, fenomena tersebut disebut dengan alih kode dan campur kode. Kedua fenomena tersebut memperlihatkan bahwa praktik berbahasa tidak selalu berlangsung dalam satu kode yang tetap, tetapi dapat melibatkan pergantian maupun penyisipan unsur bahasa lain sesuai dengan kebutuhan komunikasi.

Zahara dan Perangin-angin (2025) mencatat Fenomena ini pada video seorang kreator Indonesia di Korea Selatan. Dari 58 kasus yang mereka kumpulkan, sekitar dua pertiga berupa sisipan di dalam kalimat, dan sisanya berupa peralihan antarkalimat dan penanda singkat. Sosiolinguistik menyediakan istilah untuk Fenomena tersebut. Menurut Chaer dan Agustina, alih kode merupakan peralihan penggunaan suatu bahasa ke bahasa lain dalam peristiwa tutur, sedangkan campur kode adalah penyisipan unsur bahasa lain ke dalam bahasa yang sedang digunakan. Keduanya berbeda dari integrasi, yaitu unsur bahasa asing yang telah diserap ke dalam bahasa penerima, serta interferensi, yakni penyimpangan kaidah bahasa akibat pengaruh bahasa lain. Pembedaan ini penting dalam penelitian percakapan informal karena terdapat banyak kata yang telah lazim digunakan dan dianggap mapan dalam suatu bahasa. Pembedaan itu menentukan cara data dihitung.

Penelitian terdahulu pada media berbahasa Arab hanya memakai dua kategori pertama (Karima et al., 2025; Nurfaiqa et al., 2025), sehingga kata pinjaman yang sudah

mapan dapat ikut terhitung sebagai sisipan. Mubasyiroh (2020) beralih ke wacana tulis. Ia menelaah percakapan WhatsApp mahasiswa pembelajar bahasa Arab di UIN Maulana Malik Ibrahim Malang dan menemukan campur kode berbentuk kata dan frasa dari beberapa kategori kata. Faktor yang ia catat meliputi hubungan penutur dengan mitra tutur, perubahan situasi dari formal ke informal, perubahan topik, serta latar belakang pendidikan dan tingkat kemampuan. Bahasa Arab menambahkan satu catatan. Penuturnya memakai ragam kolokial untuk percakapan sehari-hari dan ragam standar untuk situasi resmi (Ni'mah, 2024). Seluruh percakapan dalam objek penelitian ini berlangsung dalam ragam kolokial, yang tampak antara lain pada kata دحين 'sekarang' dan مره 'sangat'. Penelitian ini tidak memakai pembagian ragam sebagai alat analisis. Ragam kolokial hanya ditetapkan sebagai kode dasar tempat unsur bahasa lain dihitung.

Objek penelitian ini adalah siniar YouTube berjudul "Malaysians speaking Arabic?" yang tayang di kanal Omni Malaysia dan dapat diakses melalui https://youtu.be/7GKFIlIaHb8. Dua penutur, Nono dan Kadi, membahas perbedaan hidup di Arab Saudi dan di Malaysia. Topiknya bergerak dari perundungan dan media sosial ke pengalaman sebagai orang asing, makanan, perlakuan terhadap perempuan, cara mereka belajar bahasa Arab, dan akhirnya sebuah permainan melafalkan kata bahasa Inggris.

Kedua penutur tumbuh di Arab Saudi dan pindah ke Malaysia untuk kuliah. Pada awal episode, Kadi menyebut dirinya orang Thailand yang lahir dan besar di Arab Saudi dan datang ke Malaysia pada 2019 (01:59). Nono mengatakan bahwa ia sempat tinggal di Arab Saudi tetapi lahir di Malaysia, dan keduanya sudah enam tahun tinggal di Malaysia (02:16). Repertoar mereka karena itu mencakup Arab kolokial sebagai bahasa percakapan, Inggris, Melayu, dan sedikit Thai. Transkrip memperlihatkan keempatnya: kalimat seperti سو تيل مي ايش رايك عن هذا الشيء 'so tell me, what do you think about this' (00:18), kata Melayu *warung* dan *nasi campur* (10:16), dan salam Thai *sawadee khap* (21:25).

Tiga celah muncul dari telaah tersebut. Pertama, penelitian alih kode dan campur kode bahasa Arab pada media sebagian besar memakai dialog yang ditulis penulis naskah (Karima et al., 2025; Nurfaiqa et al., 2025) atau tulisan pembelajar (Mubasyiroh, 2020). Percakapan pada objek ini berlangsung secara spontan tanpa naskah percakapan dan tidak menggunakan bahasa asal penutur sebagai kode bahasa utama. Kedua, penelitian-penelitian itu tidak memisahkan alih kode dan campur kode dari integrasi dan interferensi, padahal tuturan penutur multibahasa lazim memuat kata pinjaman yang sudah mapan. Ketiga, bahasa Arab hampir selalu dipasangkan dengan satu bahasa lain, Inggris atau Indonesia, sedangkan penutur yang tumbuh di Arab Saudi dan kuliah di Malaysia membawa empat bahasa yakni Arab, Inggris, Thai, dan Melayu.

Maka dari itu, penelitian ini bertujuan (1) mengklasifikasikan tuturan siniar ke dalam alih kode, campur kode, integrasi, dan interferensi berdasarkan kerangka di atas, (2) mendeskripsikan bentuk dan bahasa sumber unsur yang ditemukan, dan (3) merumuskan secara umum faktor yang melatarbelakangi pemakaian alih kode dan campur kode oleh penutur dan lawan tutur.

Penelitian ini berkontribusi pada tiga hal. Pertama, penelitian ini menghitung integrasi dan interferensi secara terpisah sehingga jumlah sisipan tidak membengkak oleh kata pinjaman. Kedua, subjeknya penutur multilingual yang menggunakan bahasa Arab kolokial sebagai kode dasar dan tiga bahasa lainnya sebagai sisipan atau peralihan. Ketiga, penelitian ini membatasi pembahasannya pada identifikasi dan klasifikasi bentuk kebahasaan, sehingga hasilnya diharapkan dapat menjadi dokumentasi empiris bagi kajian kontak bahasa dalam komunikasi digital.

METODE

Penelitian ini berjenis kualitatif-deskriptif dengan pendekatan sosiolinguistik. Data berupa tuturan yang dimaknai dalam konteks sosialnya, sehingga pendekatan kualitatif sesuai. Metode deskriptif menggambarkan objek sebagaimana adanya dan tidak bertujuan menggeneralisasi (Karima et al., 2025; Nurfaiqa et al., 2025). Angka dan persentase pada bagian Hasil meringkas sebaran data dan tidak dipakai untuk uji statistik.

Data pada penelitian ini adalah kata atau kalimat yang teridentifikasi alih kode dan campur kode berdasarkan kerangka teori sosiolinguistik. Data penelitian yang dimaksud adalah kata, frasa, klausa, dan kalimat dalam transkrip yang memuat unsur bahasa selain Arab. Sumber data adalah siniar YouTube "Malaysians speaking Arabic?" yang diakses melalui https://youtu.be/7GKFIlIaHb8. Peneliti memakai transkrip takarir episode itu, yang berdurasi 26 menit 51 detik dan terbagi dalam 204 baris berpenanda waktu. Kemudian diverifikasi menggunakan teknik simak bebas libat cakap. Dua penutur muncul dalam episode, tetapi transkrip tidak menandai pembicara, sehingga analisis dilakukan pada tingkat tuturan dan bukan pada tingkat individu. Video bersifat publik dan peneliti hanya menganalisis isi yang disiarkan, tanpa mengumpulkan informasi pribadi di luar episode.

Pengumpulan data melalui tiga langkah. Pertama, peneliti menyimak episode secara berulang tanpa terlibat dalam percakapan, yaitu teknik simak bebas libat cakap (Sudaryanto, 2015, dikutip dalam Nurfaiqa et al., 2025). Kedua, peneliti membaca transkrip takarir yang dihasilkan otomatis. Transkrip itu menuliskan kata Inggris dan Melayu dengan huruf Arab, misalnya بيوتي أند ستوف ليك دات untuk *beauty and stuff like that* (04:37), sehingga setiap butir dipulihkan ke ejaan Latin dan dicocokkan dengan rekaman. Ketiga, peneliti mencatat butir yang relevan beserta penanda waktu, kutipan transkrip, bentuk pulihannya, dan konteks kalimatnya.

Peneliti menganalisis memakai metode padan. Padan intralingual membandingkan unsur dalam bahasa untuk menentukan bentuk dan kedudukan sisipan dalam kalimat. Padan ekstralingual menghubungkan tuturan dengan penutur, lawan tutur, topik, dan situasi untuk merumuskan faktor pemicu (Sudaryanto, 2015, dikutip dalam Nurfaiqa et al., 2025). Penelitian ini memakai satu kerangka teori dari sosiolinguistik Indonesia yang dirumuskan Chaer dan Agustina (2010). Kerangka itu memuat empat Fenomena kontak bahasa. Dua di antaranya, alih kode dan campur kode, menjadi sasaran utama. Dua lainnya, interferensi dan integrasi, menjadi penyaring: butir yang memenuhi rumusannya dikeluarkan dari alih kode dan campur kode.

Appel (1976, dikutip dalam Chaer & Agustina, 2010) memandang alih kode sebagai Fenomena peralihan pemakaian bahasa akibat perubahan situasi. Hymes (1974, dikutip

dalam Chaer & Agustina, 2010) melebarkan cakupannya: peralihan dapat terjadi antara dua bahasa atau lebih, antarragam dalam satu bahasa, bahkan antargaya tutur. Suwito (1985, dikutip dalam Mubasyiroh, 2020) menambahkan dua ciri, yaitu setiap bahasa masih menjalankan fungsinya sendiri dan fungsi itu menyesuaikan konteks yang berubah. Berdasarkan arahnya, Soewito (dikutip dalam Nurfaiqa et al., 2025) membagi alih kode menjadi intern, yaitu antarragam dalam satu bahasa, dan ekstern, yaitu antara bahasa asli dan bahasa asing.

Campur kode berbeda pada otonomi unsur yang terlibat. Thelander (1976, dikutip dalam Chaer & Agustina, 2010) menyebut campur kode terjadi bila di dalam satu klausa atau kalimat terdapat serpihan klausa atau frasa dari satu kode yang dipakai dalam kode lain. Chaer dan Agustina (2010) merumuskannya sebagai pemakaian sebuah kode dasar yang berfungsi penuh bersama kode lain yang hanya hadir sebagai serpihan tanpa fungsi atau otonomi sebagai kode. Bentuk serpihan itu, menurut Suwito (1983, dikutip dalam Karima et al., 2025), berupa kata, frasa, klausa, ungkapan atau idiom, baster, dan perulangan kata.

Ukuran yang membedakan klausa campur kode dari alih kode adalah posisi unsur asing. Bila ia berdiri sebagai klausa atau kalimat mandiri dengan fungsinya sendiri, peristiwanya alih kode. Contoh: so Khadijah, I would like to challange you to a game (24:15). Kalimat tersebut berdiri secara mandiri sesuai fungsinya, pada hal ini yaitu sebagai pergantian topik dari pembahasan perbandingan hidup di Arab dan Malaysia beralih menuju topik permainan kata. Bila ia tertanam sebagai serpihan di dalam kalimat berkode dasar lain, peristiwanya campur kode, contoh: (08.45) بس ما في ولا مطعم دكت في السعودية ذت ديسبوينت مي. Klausa “that dissapointed me” tidak berdiri sendiri, ia bersambung dengan kode bahasa lain.

Pada penelitian ini terdapat dua fenomena yang tidak masuk dalam alih kode dan campur kode. Fenomena pertama adalah interferensi. Weinreich (1953, dikutip dalam Chaer & Agustina, 2010) memakai istilah itu untuk perubahan sistem suatu bahasa akibat persentuhannya dengan unsur bahasa lain yang dikuasai penutur dwibahasa. Chaer dan Agustina (2010) menggolongkannya menurut tataran bahasa, yakni fonologi, morfologi, sintaksis, dan leksikal. Interferensi morfologis terjadi bila penutur memakai afiks satu bahasa pada dasar kata dari bahasa lain. Yang dinilai pada Fenomena ini adalah cara kata dibentuk, bukan sekadar kehadiran unsur asing.

Fenomena kedua adalah integrasi. Mackey (1968, dikutip dalam Chaer & Agustina, 2010) mendefinisikan integrasi sebagai pemakaian unsur bahasa lain dalam suatu bahasa yang dianggap dan diperlakukan sebagai bagian dari bahasa itu, sehingga tidak lagi terasa sebagai pinjaman. Dalam penelitian ini, penanda integrasi pada data tertulis adalah butir yang lazim dalam percakapan Arab sehari-hari dan muncul tanpa jeda, tanpa penjelasan, dan tanpa terjemahan dari penuturnya. Tabel 1 menerjemahkan keempat rumusan menjadi kriteria yang dapat diperiksa pada transkrip.

Tabel 1
Kerangka Klasifikasi dan Penanda pada Data

Kategori	Rumusan	Rujukan	Penanda pada data
Alih kode	Peralihan ke bahasa lain pada tataran klausa atau kalimat; unsur asing berdiri mandiri dengan fungsinya sendiri.	Appel (1976); Hymes (1974); Suwito (1985); Soewito; Mubasyiroh (2020)	Klausa atau kalimat Inggris lengkap yang tidak bergantung pada struktur Arab di sekitarnya.
Campur kode	Serpihan kata, frasa, klausa, idiom, baster, atau perulangan dari bahasa lain di dalam kalimat berkode dasar Arab.	Thelander (1976); Chaer & Agustina (2010); Suwito (1983)	Unsur asing tertanam dalam kalimat Arab dan tidak memenuhi penanda integrasi atau interferensi.
Integrasi	Unsur pinjaman yang diperlakukan sebagai bagian dari bahasa penerima.	Mackey (1968)	Butir lazim dalam percakapan Arab sehari-hari; muncul tanpa jeda, penjelasan, atau terjemahan.
Interferensi (morfologis)	Pemakaian afiks atau pola satu bahasa pada unsur bahasa lain.	Weinreich (1953); Chaer & Agustina (2010)	Dasar kata asing yang diberi klitik Arab (ال, بـ) dan bukan butir mapan.

Sumber: disusun peneliti berdasarkan rujukan yang tercantum; Appel, Hymes, Thelander, Weinreich, dan Mackey dikutip melalui Chaer dan Agustina (2010).

Klasifikasi mengikuti empat pertanyaan berurutan yang bertumpu pada Tabel 1.

1. Apakah butir itu berasal dari bahasa selain Arab? Bila tidak, butir itu bukan data.
2. Apakah butir itu klausa atau kalimat mandiri yang tidak bergantung pada struktur Arab di sekitarnya? Bila ya, butir itu alih kode.
3. Apakah butir itu pinjaman mapan yang dipakai tanpa jeda, penjelasan, atau terjemahan? Bila ya, butir itu integrasi.
4. Apakah butir itu berupa dasar kata asing yang diberi klitik Arab dan bukan pinjaman mapan? Bila ya, butir itu interferensi morfologis. Butir lain yang tertanam dalam kalimat Arab adalah campur kode, dan bentuknya dirinci menurut Suwito (1983, dikutip dalam Karima et al., 2025).

Satu kemunculan dihitung satu data. Satu kalimat tuturan dapat memuat lebih dari satu data yang dihitung berdasarkan fenomena kebahasaan yang ditemukan. Rentang bahasa Inggris yang berlanjut tanpa sela bahasa Arab dihitung satu alih kode, sehingga

kalimat panjang pada 24:15 menjadi satu data. Tiga kelompok butir dikeluarkan dari hitungan: nama diri, merek, dan judul (misalnya TikTok, K-pop, dan *Boys Over Flowers*); empat kata stimulus pada permainan pelafalan di akhir episode (*colonel*, *Worcestershire*, *gnome*, dan *literally*), yang disebut sebagai bahan permainan dan bukan dipilih sebagai sarana bertutur. Faktor penggunaan alih kode dan campur kode dirumuskan secara umum pola seluruh episode dan tidak diklasifikasikan per data. Chaer dan Agustina (2010) mengidentifikasi adanya lima faktor yang melatarbelakainginya, yaitu penutur, mitra tutur, adanya pihak ketiga, pergantian topik, dan penanda wacana.

Data dalam penelitian ini disajikan menggunakan teknik penyajian informal dan formal. Penyajian informal dilakukan melalui uraian deskriptif untuk menjelaskan bentuk alih kode dan campur kode yang ditemukan dalam tuturan siniar *"Malaysians Speaking Arabic?"*, termasuk konteks percakapan yang melatarbelakangi kemunculannya. Sementara itu, penyajian formal dilakukan dalam bentuk tabel yang memuat kategori fenomena kebahasaan berdasarkan kerangka teori sosiolinguistik, kutipan tuturan, penanda waktu (*timestamp*), Bahasa yang digunakan, serta jumlah frekuensi data.

Keabsahan dijaga dengan tiga cara: menyimak ulang rekaman untuk setiap butir yang diragukan, mencantumkan penanda waktu pada semua data agar dapat ditelusuri. Butir pada perbatasan dua kategori dicatat secara terbuka pada bagian Pembahasan.

HASIL

Transkrip 204 baris menghasilkan 147 data. Tabel 2 memuat sebarannya. Campur kode mendominasi dengan 113 data, disusul integrasi 16 data, alih kode 12 data, dan interferensi morfologis 6 data. Penyajian di bawah mengikuti urutan data bahasa, klasifikasi, pola, dan temuan, dengan penanda waktu pada awal baris transkrip tempat butir muncul.

Tabel 2
Sebaran Data Menurut Kategori

Kategori	Jumlah data
Campur kode	**113**
Bentuk kata	70
Bentuk frasa	23
Bentuk klausa	16
Bentuk idiom	3
Bentuk perulangan kata	1
Integrasi	**16**
Alih kode (ekstern)	**12**
Interferensi (morfologis)	**6**
Total	**147**

Campur Kode

Campur kode berbentuk kata berjumlah 70 data dan menjadi bentuk terbanyak dari seluruh data campur kode. Hampir separuhnya, 33 data, adalah dua penanda wacana, yaitu *okay* (27 kali) dan *so* (6 kali), seperti pada اوكي هاي جايز (01:59) dan سو بالاخت كادي حنبدا (01:50). Dua belas data berikutnya adalah nama bahasa yang diucapkan dalam bentuk Inggris, *Malay* (8 kali) dan *English* (4 kali), misalnya انجلش ملاي انجلش ملاي (17:32). Sisa 25 data berupa kata isi seperti *option* (06:58), *experience* (14:14), *translate* (16:35), dan *accent* (26:00), ditambah empat kata Melayu: *warung* dan *nasi campur* (10:16) serta *bulan* dan *bulanku* (22:49).

Campur kode berbentuk frasa berjumlah 23 data. Frasa nominal dan preposisional paling banyak, misalnya *cyber bullying* dan *bullying in general* (00:10), *long term* (06:33), *universal language* (16:48), dan *extra vocabulary* (17:32). Dua frasa berbahasa Thai, سواديكاب *sawadee khap* dan كوبكونكاب *khop khun khap* (21:25), muncul ketika penutur menjelaskan keterbatasannya berbahasa Thai dengan saudara-saudaranya.

Campur kode berbentuk klausa berjumlah 16 data. Tujuh di antaranya adalah *I think*, baik sendirian maupun diikuti klausa pendek seperti *I think it's just* (12:50). Sembilan lainnya klausa yang tertanam dalam kalimat Arab: *that disappointed me* (08:45), *we give each other* (14:07), *even though I speak Arabic* (15:24), dan *it's really a blessing* (19:53). Tiga data berbentuk idiom, semuanya *oh my god* (04:14, 09:08, 13:50), dan satu berbentuk perulangan kata, *bye bye* (26:30). Dari 113 campur kode, 107 berbahasa Inggris, 4 berbahasa Melayu, dan 2 berbahasa Thai. Tabel 3 menyajikan contoh dari setiap bentuk.

Tabel 3
Contoh Campur Kode menurut Bentuk

Bentuk	Butir	Waktu	Bahasa	Jumlah
Kata	*okay*	sepanjang episode (pertama 01:59)	Inggris	27
Kata	*so*	01:50, 06:01, 15:46, 16:12, 25:43	Inggris	6
Kata	*option*	06:58	Inggris	1
Kata	*warung*, *nasi campur*	10:16	Melayu	2
Kata	*bulan*, *bulanku*	22:49	Melayu	2
Frasa	*bullying in general*	00:10	Inggris	1
Frasa	*long term*	06:33	Inggris	1
Frasa	*sawadee khap*, *khop khun khap*	21:25	Thai	2
Klausa	*I think*	04:14, 12:50, 15:07, 16:56, 21:40 (2), 22:25	Inggris	7

Bentuk	Butir	Waktu	Bahasa	Jumlah
Klausa	*we give each other*	14:07	Inggris	1
Klausa	*even though I speak Arabic*	15:24	Inggris	1
Idiom	*oh my god*	04:14, 09:08, 13:50	Inggris	3
Perulangan	*bye bye*	26:30	Inggris	1

Catatan: waktu mengikuti penanda baris transkrip. Sumber: data penelitian.

Alih Kode

Alih kode berjumlah 12 data. Seluruhnya ekstern, yaitu peralihan dari Arab ke Inggris. Tidak ada peralihan klausa ke bahasa Melayu atau Thai, dan tidak ada alih kode intern. Lima data berada dalam satu rentang di akhir episode (24:15 sampai 25:52), ketika penutur memainkan permainan melafalkan kata Inggris. Tujuh data lain tersebar sepanjang episode (Tabel 4). Rentang terpanjang dimulai pada 24:15 dan memuat beberapa kalimat Inggris berturut-turut, yang dihitung satu alih kode karena tidak disela bahasa Arab.

Tabel 4
Seluruh Data Alih Kode

No.	Waktu	Kutipan transkrip	Alih aksara
1	00:18	سو تيل مي	*So, tell me*
2	06:18	فور يو وش وند دي يو بريفير	*for you, which one do you prefer*
3	09:08	دس اش غن بي جود	*this is gonna be good*
4	09:08	ات ديسبوينت يو سم تيمز	*it disappoints you sometimes*
5	12:34	إس دات ا ورد	*is that a word*
6	20:24	يو لوز دات جانس	*you lose that chance*
7	20:47	اتس سو شيمفل اوف يو، يو دونت سبيك يور اون لنغويج	*it's so shameful of you, you don't speak your own language*
8	24:15	سو ذي نيو لنغويج فور هير از انغلش	*so the new language for her is English ... (rentang berbahasa Inggris sampai "...to pronounce")*
9	24:47	ايل غيف يو ذس وورد، بروناونس ذي فيرست وورد اند ذين	*I'll give you this word, pronounce the first word and then*
10	24:55	ذا فيست ورد از	*the first word is*

No.	Waktu	Kutipan transkrip	Alih aksara
11	25:26	اي ثينك ذي رايت برونانسييشن از	*I think the right pronunciation is*
12	25:52	نوم ازذا سمول تنز	*gnome is the small ...*

Sumber: data penelitian.

Integrasi

Integrasi berjumlah 16 data dari tujuh butir leksikal (Tabel 5). Butir terbanyak adalah *social media*, yang muncul enam kali dengan artikel Arab, misalnya بالسوشيال ميديا (26:30). Berikutnya *trend* (tiga kali, 00:00), *elevator* dalam bentuk اسانسير dan الاسانصير (12:34), serta *comment* dalam bentuk كومنتس (00:58) dan بالكومنت (26:23). Penutur memakai butir-butir ini tanpa jeda, tanpa penjelasan, dan tanpa terjemahan. Satu butir, بزنسات (10:32), menerima akhiran jamak Arab *-āt*.

Tabel 5
Data Integrasi

Butir	Jumlah	Bentuk pada transkrip	Waktu
social media	6	السوشيال ميديا, بالسوشيال ميديا	00:00, 00:10, 00:25, 00:40, 14:22, 26:30
trend	3	ترند	00:00
elevator	2	اسانسير, الاسانصير	12:34
comment	2	كومنتس, بالكومنت	00:58, 26:23
ice cream	1	ايس كريم	09:17
cake	1	هالكيك	13:50
business	1	بزنسات	10:32

Sumber: data penelitian.

Interferensi Morfologis

Interferensi morfologis berjumlah enam data (Tabel 6). Setiap butir berupa dasar kata asing yang diberi klitik Arab dan bukan pinjaman mapan: البلاتفورم *platform* (00:40), بالتوبك *topic* (05:53), الار&ار *R&R* (11:28), عبارات السويت *sweet* (22:25), اللف لانجوج *love language* (22:57), dan الاكسن *accent* (26:14).

Tabel 6
Data Interferensi Morfologis

Bentuk pada transkrip	Dasar asing	Klitik Arab	Waktu
البلاتفورم	*platform*	ال	00:40
بالتوبك	*topic*	بـ + ال	05:53

Bentuk pada transkrip	Dasar asing	Klitik Arab	Waktu
الار&ار	*R&R*	ال	11:28
السويت	*sweet*	ال	22:25
اللف لانجوج	*love language*	ال	22:57
الاكسن	*accent*	ال	26:14

Sumber: data penelitian.

Sebaran Fenomena Kebahasaan Sepanjang Episode dan Data yang Tidak Diklasifikasikan

Per lima menit, sebarannya 23 data (00:00 sampai 04:59), 18 (05:00 sampai 09:59), 23 (10:00 sampai 14:59), 34 (15:00 sampai 19:59), 34 (20:00 sampai 24:59), dan 15 (25:00 sampai 26:40). Seluruh 34 data pada menit 15 sampai 19 adalah campur kode, dan segmen itu berisi cerita penutur tentang cara mereka belajar bahasa. Tujuh dari 12 alih kode jatuh setelah menit ke-20. Integrasi paling padat pada lima menit pertama (8 data), saat topik perundungan dan media sosial dibahas.

Butir yang tidak dihitung berupa dua kategori. Pertama, kategori nama diri, merek, dan judul: TikTok, K-pop, *Boys Over Flowers*, dan Al-Baik. Kedua, empat kata stimulus pada permainan pelafalan, yaitu *colonel*, *Worcestershire*, *gnome*, dan *literally*.

PEMBAHASAN

Bagian ini menafsirkan hasil melalui kerangka pada Tabel 1. Urutannya mengikuti besar kategori: campur kode sebagai pola utama, alih kode sebagai pengecualian, batas keduanya dengan integrasi dan interferensi, lalu bahasa sumber dan faktor pemakaian. Perbandingan dengan penelitian terdahulu dan sumbangan penelitian ditutup pada subbagian terakhir.

Campur Kode sebagai Pola Utama

Dari 125 data yang berstatus alih kode atau campur kode, 113 adalah campur kode. Angka itu menunjukkan bahwa percakapan siniar berjalan dengan satu kode dasar, yakni Bahasa Arab dialek Saudi, dan bahasa lain masuk sebagai serpihan. Rumusan Thelander (1976, dikutip dalam Chaer & Agustina, 2010) dan rumusan Chaer dan Agustina (2010) cocok dengan data ini. Bahkan butir sepanjang klausa, seperti نحاول انه وي غيف إيتش أذر ‘kami berusaha agar *we give each other*’ (14:07), tidak berdiri sendiri. Klausa Inggris itu menjadi pelengkap bagi kerangka Arab *nuhāwil annahu*, sehingga ia tetap serpihan dan bukan peralihan.

Komposisi serpihan itu menjelaskan mengapa angkanya besar. Hampir separuh campur kode berbentuk kata (33 dari 70) adalah *okay* dan *so*, dua butir yang hampir tidak membawa isi rujukan. Fungsinya mengatur arah bicara: *so* membuka agenda (01:50), *okay* menanggapi permintaan lawan tutur (01:59). Fraser (1999, dikutip dalam Zahara & Perangin-angin, 2025) menyebut unsur semacam ini penanda wacana yang menjaga koherensi percakapan. Dalam data berbentuk kata pada film *Honeymoonish*, *okay* juga muncul paling sering (Karima et al., 2025). Dengan demikian, sebagian besar

volume campur kode di media berbahasa Arab tampaknya ditopang oleh penanda wacana, bukan oleh kosakata baru.

Kata isi dan frasa menunjukkan jenis alasan yang lain. Penutur memilih *option*, *experience*, *hospitality*, dan *long term* walaupun padanan Arab tersedia. Kerangka Chaer dan Agustina (2010) tidak mensyaratkan padanan itu hilang, karena campur kode hanya menuntut adanya kode dasar dan serpihan. Pilihan atas istilah Inggris lebih mungkin mencerminkan repertoar penutur daripada kekosongan kosakata Arab. Dugaan ini tidak diuji pada penelitian ini dan hanya berlaku bagi penutur dalam siniar.

Tabel 7 menempatkan proporsi itu di samping penelitian terdahulu. Porsi campur kode terhadap gabungan alih kode dan campur kode dalam *Honeymoonish* menunjukkan perbandingan 75,0% lebih banyak penggunaan campur kode dan 55,6% lebih banyak penggunaan campur kode dalam *Bayt Thahir*. Sedangkan persentase perbandingan penggunaan campur kode pada penelitian ini mencapai 90,4%.

Tabel 7
Porsi Campur Kode terhadap Gabungan Alih Kode dan Campur Kode

Penelitian	Objek	Alih kode	Campur kode	Porsi campur kode
Karima et al. (2025)	Film *Honeymoonish*	7	21	75,0%
Nurfaiqa et al. (2025)	Serial *Bayt Thahir*	4	5	55,6%
Penelitian ini	Siniar tidak berskrip	12	113	90,4%

Catatan: angka Karima et al. memakai rincian 15 kata dan 6 frasa; Zahara dan Perangin-angin (2025) tidak disertakan karena tipologinya berbeda. Sumber: diolah dari penelitian yang dirujuk.

Alih Kode: Pengalihan bahasa

Jika campur kode adalah kebiasaan, alih kode adalah pengecualian yang punya sebab. Appel (1976, dikutip dalam Chaer & Agustina, 2010) mengaitkan alih kode dengan perubahan situasi, dan Suwito (1985, dikutip dalam Mubasyiroh, 2020) menambahkan bahwa tiap bahasa menjalankan fungsinya sendiri. Kedua ciri itu tampak jelas pada permainan di akhir episode. Pada 24:15, penutur menyatakan bahwa bahasa Inggris adalah bahasa baru bagi lawan tutur dan menantangnya melafalkan kata. Sejak titik itu, bahasa Inggris menjadi bahasa kegiatan: instruksi (24:47), pengumuman kata (24:55), dan penilaian (25:26) memakai bahasa Inggris. Ketika penutur perlu menanyakan arti kata, ia kembali ke bahasa Arab, ايش معناها (25:43). Peralihan dan kembalinya sama-sama mengikuti perubahan kegiatan.

Lima alih kode lain muncul ketika penutur menirukan suara orang lain atau pikirannya sendiri: pertanyaan yang sering diajukan orang kepada mereka (06:18), pikiran saat melihat harga mahal (09:08, dua data), komentar orang yang mempermalukan penutur yang tidak berbahasa ibu (20:47), dan pertanyaan pada diri sendiri tentang sebuah kata (12:34). Dua data sisanya adalah pembuka giliran bicara *so, tell me* (00:18) dan penilaian *you lose that chance* (20:24).

Semua alih kode bersifat ekstern, dan tidak ada yang beralih ke bahasa Melayu atau Thai. Hal itu sejalan dengan cerita penutur. Bahasa Melayu adalah bahasa rumah bagi salah satu dari mereka (19:35 sampai 19:53), tetapi di depan lawan tutur dan penonton yang berbahasa Arab dan Inggris, ia hanya muncul sebagai butir leksikal.

Batas dengan Integrasi dan Interferensi

Penyaringan integrasi dan interferensi mengubah angka secara nyata. Bila 16 integrasi dan 6 interferensi digabung ke campur kode, jumlah campur kode naik dari 113 menjadi 135. Dua penelitian terdahulu pada media Arab (Karima et al., 2025; Nurfaiqa et al., 2025) tidak melakukan penyaringan ini, sehingga tidak diketahui berapa banyak data mereka yang berupa pinjaman mapan. Temuan ini menjadi alasan untuk memisahkan keduanya pada penelitian berikutnya.

Integrasi di sini bertumpu pada rumusan Mackey (1968, dikutip dalam Chaer & Agustina, 2010): unsur dipakai dan diperlakukan sebagai bagian dari bahasa penerima. Butir seperti *social media*, *trend*, dan *elevator* memenuhi rumusan itu. Penutur memakainya enam, tiga, dan dua kali tanpa menjelaskan atau menerjemahkannya, dan sebagian membawa artikel atau awalan Arab. Bandingkan dengan *love language*, yang disusul penjelasan *in Arabic* dan karena itu diperlakukan sebagai istilah asing.

Interferensi morfologis adalah kategori yang paling mudah diperdebatkan. Bahasa Arab lazim memberi artikel *al-* pada pinjaman, sehingga sebagian peneliti akan menilai البلاتفورم sebagai adaptasi biasa, bukan penyimpangan. Penelitian ini menggolongkannya sebagai interferensi karena Weinreich (1953, dikutip dalam Chaer & Agustina, 2010) berbicara tentang pengaruh sistem satu bahasa pada bentuk bahasa lain, dan klitik Arab pada dasar asing yang belum mapan memenuhi rumusan itu. Konsekuensinya dapat dihitung. Bila enam butir itu dipindahkan ke campur kode, porsi campur kode menjadi 81,0%. Bila dipindahkan ke integrasi, porsi integrasi menjadi 15,0%. Kesimpulan utama, bahwa campur kode adalah pola dominan, tidak berubah pada kedua skenario.

Bahasa Sumber: Inggris, Melayu, dan Thai

Bahasa Inggris menyumbang 94,7% campur kode. Angka itu sejalan dengan dua penelitian terdahulu pada media Arab yang juga hanya menemukan sisipan Inggris (Karima et al., 2025; Nurfaiqa et al., 2025). Salah satu penutur menyebut bahasa Inggris sebagai bahasa yang ia pakai bersama teman-temannya (17:32), dan keduanya kuliah di Malaysia, sehingga bahasa itu menjadi sumber kosakata yang paling dekat. Yang berbeda dari penelitian terdahulu adalah kehadiran bahasa Melayu dan Thai, meskipun jumlahnya kecil.

Pada keempat butir Melayu dan dua butir Thai, sebab pemakaiannya lebih spesifik. *Warung* dan *nasi campur* (10:16) merujuk pada hal yang khas Malaysia, sedangkan *bulan* dan *bulanku* (22:49) muncul saat penutur menerjemahkan sapaan sayang bahasa Arab ke bahasa Melayu. Dua frasa Thai muncul ketika penutur menyebut saudara-saudaranya yang berbahasa Thai (21:25). Pola serupa ditemukan pada kreator Indonesia di Korea Selatan, yang menyisipkan istilah Korea untuk makanan dan institusi yang tidak punya padanan langsung (Zahara & Perangin-angin, 2025). Bahasa

lingkungan tempat tinggal dengan demikian masuk sebagai kosakata lokal, sedangkan bahasa global seperti Inggris masuk sebagai kosakata umum.

Faktor Pemakaian oleh Penutur dan Lawan Tutur

Chaer dan Agustina (2010, dikutip dalam Nurfaiqa et al., 2025) menyebut lima faktor alih kode: penutur, mitra tutur, kehadiran pihak ketiga, perubahan situasi, dan perubahan topik. Campur kode, menurut Nurfaiqa et al. (2025), biasanya dipicu identifikasi peran, identifikasi ragam, atau keinginan menjelaskan sesuatu dengan lebih mudah. Mubasyiroh (2020) memperoleh faktor yang bertumpang tindih pada percakapan pembelajar. Dalam siniar ini, pola umumnya dapat dirumuskan tanpa mengklasifikasikan data satu per satu.

Pertama, kedua penutur bergantian menjadi penutur dan lawan tutur, dan keduanya menguasai Arab, Inggris, dan Melayu. Karena repertoar mereka sama, sisipan Inggris hampir tidak memicu permintaan penjelasan. Satu-satunya pertanyaan arti pada kata Inggris muncul pada kata stimulus permainan (25:43). Kedua, penonton hadir sebagai pihak ketiga. Salam pembuka (01:33) dan penutup (26:14 sampai 26:35) ditujukan kepada mereka, dan keduanya memuat unsur Inggris seperti *hi guys* dan *bye bye*. Ketiga, topik berganti, dan setiap topik membawa kosakatanya. Perundungan dan media sosial membawa butir integrasi, cerita belajar bahasa membawa nama bahasa (*Malay* dan *English*), dan permainan membawa alih kode terpanjang. Keempat, penutur memakai sisipan untuk menandai sikap dan giliran bicara, seperti *I think*, *okay*, dan *so*.

Dibandingkan dengan Karima et al. (2025) dan Nurfaiqa et al. (2025), penelitian ini berbeda pada empat hal: jenis data, jumlah kategori, bahasa sumber, dan perlakuan terhadap kata pinjaman. Data mereka berupa dialog berskrip, kategori mereka dua, dan bahasa sumbernya Inggris. Penelitian ini memakai percakapan tidak berskrip, empat kategori, dan tiga bahasa sumber. Dibandingkan dengan Mubasyiroh (2020), objeknya lisan dan bukan tulisan pembelajar, dan penuturnya menguasai ragam kolokial sebagai bahasa pertama. Dibandingkan dengan Zahara dan Perangin-angin (2025), penelitian ini sama-sama berobjek video YouTube, tetapi melibatkan bahasa Arab dan memisahkan integrasi serta interferensi.

Sumbangan penelitian berupa kerangka empat kategori terbukti dapat dioperasikan pada transkrip berdurasi penuh dan menghasilkan angka dasar yang dapat dibandingkan oleh peneliti lain.

KESIMPULAN

Penelitian ini mengklasifikasikan tuturan dalam satu episode siniar “Malaysians speaking Arabic?” berdasarkan kerangka alih kode, campur kode, integrasi, dan interferensi menurut Chaer dan Agustina (2010). Dari 147 data, ditemukan 113 campur kode, 16 integrasi, 12 alih kode, dan 6 interferensi morfologis. Campur kode berbentuk kata sejumlah 70 data, frasa 23 data, klausa 16 data, idiom 3 data, dan perulangan kata sejumlah 1 data.

Bagi kajian bahasa Arab, temuan ini memperlihatkan dua hal. Pertama, 93 dari 113 data sisipan dalam bahasa Arab kolokial paling sering berbentuk kata dan frasa, sedangkan sisipan berbentuk klausa Inggris hanya 16 data. Pada campur kode, kerangka kalimat

tetap Arab dan bahasa asing masuk sebagai serpihan. Kedua, bahasa Arab menyediakan morfologinya sendiri bagi unsur asing, berupa artikel *al-*, awalan *bi-*, dan akhiran jamak *-āt*, seperti بالتوبك. Penutur yang memperoleh bahasa Arab di lingkungan Saudi tanpa berketurunan Arab menunjukkan pola yang sejalan dengan dialog Arab pada film dan serial yang telah diteliti (Karima et al., 2025; Nurfaiqa et al., 2025), yaitu sisipan Inggris dalam kerangka Arab.

Bahasa Inggris menyumbang 107 dari 113 campur kode, bahasa Melayu 4, dan bahasa Thai 2. Seluruh alih kode berupa peralihan dari Arab ke Inggris dan berkumpul pada tiga situasi: permainan pelafalan, penirukan suara orang lain, dan pembuka giliran atau komentar singkat. Secara umum, pemakaian alih kode dan campur kode oleh penutur dan lawan tutur dilatarbelakangi repertoar bersama yang meliputi Arab, Inggris, dan Melayu; kehadiran penonton sebagai pihak ketiga; pergantian topik; dan kebutuhan menandai sikap serta giliran bicara.

Kebaruan penelitian ini terletak pada pemisahan integrasi dan interferensi dari campur kode, pada subjek penutur yang tumbuh di Arab Saudi dan kuliah di Malaysia, dan pada objek percakapan siniar yang tidak berskrip. Pemisahan itu mengubah jumlah campur kode sebesar 19,5%, dan tidak mengubah kesimpulan bahwa campur kode adalah pola dominan.

Penelitian ini memiliki lima keterbatasan. Korpusnya satu episode. Transkripnya dihasilkan otomatis dan harus dicocokkan dengan rekaman. Transkrip tidak menandai pembicara, sehingga perbedaan antarpenutur tidak dapat dianalisis. Klasifikasi integrasi dan interferensi bergantung pada penilaian peneliti berdasarkan indikator kerangka teori dan perlu diuji oleh penilai kedua serta kamus Arab. Interferensi fonologis, sintaksis, dan leksikal belum dinilai. Penelitian lanjutan dapat menambah episode dan kanal, menandai pembicara dari rekaman, memakai penilai kedua dengan laporan kesepakatan, dan membandingkan penutur Arab di Malaysia dengan penutur Arab di Arab Saudi.

DAFTAR PUSTAKA

Amelia, T., Al Farisi, M. Z., & Nurmala, M. (2026). Alih kode dan campur kode Arab-Indonesia dalam vlog YouTube Ibrohim Elhaq. *JP-BSI (Jurnal Pendidikan Bahasa dan Sastra Indonesia), 11*(1), 1–15. https://doi.org/10.26737/jp-bsi.v11i1.8483

Chaer, A., & Agustina, L. (2010). *Sosiolinguistik: Perkenalan awal*. Rineka Cipta.

Karima, M. A., Rohanda, R., & Adriadi, I. (2025). Alih kode dan campur kode dalam film Arab Honeymoonish karya Elie El Semaan. *Lingua Franca: Jurnal Bahasa, Sastra, dan Pengajarannya*, *9*(1), 1–20.

Mubasyiroh. (2020). Alih kode dan campur kode bahasa Arab (Studi kasus percakapan WhatsApp mahasiswa Fakultas SAINTEK UIN MALIKI Malang). *Deskripsi Bahasa*, *3*(2), 182–193.

Ni'mah, R. A. (2024). Analisis penggunaan variasi bahasa Arab fusha dan 'amiyah dalam ruang lingkup sosial bahasa (sosiolinguistik). *Al-Fathin: Jurnal Bahasa dan Sastra Arab*, *7*(2), 45–57. https://doi.org/10.32332/1be6qq32

Nurfaiqa, Y., Khomisah, & Muslikah, S. (2025). Alih kode dan campur kode dalam serial Arab Bayt Thahir karya Sultan Al-Abdulmohsen. *Titian: Jurnal Ilmu Humaniora*, *9*(1), 259–280.

Omni Malaysia. (2025, August 28). *Malaysians speaking Arabic?* [Video]. YouTube. https://youtu.be/7GKFIlIaHb8

Suwito. (1983). *Pengantar awal sosiolinguistik: Teori dan problema*. Henary Offset.

Wulandari, R., Rohman, A., & Khoiroh, H. (2025). Analisis campur kode dan alih kode dalam film pendek Gontor Liayyin (kajian sosiolinguistik). *Konferensi Nasional Mahasiswa Bahasa dan Sastra Arab (KNM-BSA) 2025*, 577–590.

Zahara, R., & Perangin-angin, A. Br. (2025). Campur kode dalam YouTube Xaviera Putri analisis: Sosiolinguistik pada interaksi digital. EduGrows: Education and Learning Review, 1(1), 1–1