Forensic Linguistic Profiling of Defendant 'DRL' in the Helwa Beautycare Case

[1] Nur Anzlina*, [2] Liu Zuxi, [3] Irwan Suswandi, [4] Tsurayya Yusra Zahirah Zahra

[1,3] Indonesian Literature, Faculty of Literature, Culture, and Communication, Universitas Ahmad Dahlan, 55166, Indonesia
[2] Faculty of Asian Languages and Cultures, Guangdong University of Foreign Studies, 510420, China
[4] Faculty of Cultural Sciences, Universitas Gadjah Mada, 55281, Indonesia
[1] 2200025038@webmail.uad.ac.id; [2] 20230602084@mail.gdufs.edu.cn; [3] irwan.suswandi@idlitera.uad.ac.id;
[4] tsurayyayusrazahirahzahra2098@mail.ugm.ac.id
*Correspondent email author: 2200025038@webmail.uad.ac.id

ARTICLE INFO

ABSTRACT

Linguistic profile analysis was used to analyze the dialect, idiolect, register, and language variation of the defendant "DRL" in this study. This study is a forensic linguistic study using a sociolinguistic approach. This type of study is descriptive and qualitative. The research data consisted of spoken language transcripts taken from three videos uploaded by "DRL" using the listen and note method. Phonetic analysis used IPA tables and Praline software for acoustic linguistic analysis. Based on the results of the study, the defendant "DRL" showed the use of the Palembang and Betawi dialects in his speech. The defendant also exhibited idiolects such as word repetition, random syntactic repetition, and phoneme changes in certain words. The defendant's register was identified as a YouTube content creator based on the choice of vocabulary in his speech. The defendant also appeared to use language variation in the form of code-mixing in his speech. This research can be one of the insights that can strengthen the contribution of sociolinguistics in the field of forensic linguistics by emphasizing that linguistic analysis is not only descriptive but can also be applied in law enforcement and become one of the developments in linguistic profiling methods that utilize linguistic theory to prove the speaker's intent, attitude and identity in the courtroom.

1. Introduction

The language a person uses can reflect the characteristics of the speaker and serve as relevant linguistic evidence in legal proceedings. Forensic linguistics is the study of language applied in a legal context to assist in the process of enforcing justice (Kusno et al., 2022; Sholihatin, 2024; Wicaksana et al., 2023). According to Coulthard & Johnson (2010), forensic linguistics organizes material into two parts: language in legal proceedings and language as evidence. One of the roles of forensic linguistics is to identify language (Anggreni et al., 2019; Casalan, 2015; Fitri et al., 2021; Jahara et al., 2022). This is in line with the opinion that forensic linguistics is the application of linguistic science principles and methods in legal issues and law enforcement (Kristianto, 2015; Rusdiansyah, 2020; Ubaidah et al., 2024).

The role of forensic linguistics becomes very important when language is used as evidence in legal cases, such as hate speech, threats, defamation, and fraud. Subyantoro (2019) also reveals that

forensic linguistics also deals with the identification of speakers based on dialect, style of speech, or accent, and sometimes even analyzes the suspect's handwriting to obtain their profile. All analysis results identified by forensic linguists will be involved in the courtroom as a consideration. The information obtained from forensic linguist expert witnesses is information that can be considered by the judge. The efforts that must be made by language witnesses in the court process are to produce accurate evidence so that it can be supporting evidence for reporting a case of defamation against someone who feels insulted and has filed a report as a form of dissatisfaction due to defamation (Akbar et al., 2023; Gea, 2024; Ode & Huda, 2022) .

Thus, the knowledge of judges in court requires the important role of forensic linguistic expert witnesses (Fitria, 2024; Naznin et al., 2022; Rosaniati et al., 2021). Mahsun (2018) states in his book that forensic linguists can use the results of several linguistic studies, such as linguistic research in pure linguistics and applied linguistics. The fields of linguistics that can be applied in forensic linguistics include phonology, morphology, semantics, syntax, pragmatics, sociolinguistics, psycholinguistics, and discourse analysis. From this explanation, it can be concluded that linguistics or language studies play a significant role in resolving legal issues.

The application of forensic linguistics in linguistic cases can be applied, one of which is in the case of hate speech committed by the defendant "DRL". The case was reported by the owner of Helwa Beautycare because the company owner claimed that the defendant "DRL" had caused significant damage to Helwa Beautycare products after the defendant reviewed the products on social media. Thus, the case can be analyzed from a forensic linguistic perspective because it contains a lot of linguistic evidence on the defendant's social media. Based on this explanation, this study aims to analyze the linguistic profile of the defendant with the initials "DRL" in the hate speech case through a sociolinguistic approach, specifically by highlighting the elements of dialect, idiolect and language variation that appear in the content of his YouTube video upload.

The researcher is interested in analyzing this case because the case has linguistic evidence in civil law cases. In addition, there are seven aspects that are part of sociolinguistic issues, namely (i) the social identity of the speaker, (ii) the social identity of the listener or speech partner, (iii) the social environment, (iv) speech events that occur in a place, (v) synchronic and diachronic analysis through social dialects, (vi) different social assessments by speakers based on behavior in speech forms, (vii) the level of linguistic variation and diversity, and the application of sociolinguistics (Anggara, 2024; Chaer, 2004) . Sociolinguistics also studies how cultural backgrounds shape the linguistic identity of speakers. Bucholzt & Hall (2005) argue that in sociocultural linguistics, the linguistic means used to signify and interpret speech acts vary according to cultural background.

Dialect is a variation of language based on the place or region of origin of its speakers (Chaer, 2004). This idiolect is directly related to distinctive features of voice, word choice, style of language, sentence structure, and so on (Chaer, 2004). Wagner et al. (2014) mention that linguistic register is a different style of speech that reflects the identity, relationship, social status, and environment of the speaker and listener; that is, they signify the social identity and social situation of the speaker. Muysken (2000) mentions that *code-mixing* has three forms. First, *insertion*, which is the insertion of words or phrases into the main sentence structure that originate from another language. Second, *alternation*, which is the alternating change of grammatical structures from two languages in one discourse. Third, *congruent lexicalization*, which is the mixing of two languages in the same lexical structure. Sociolinguistic studies provide limitations to understanding the context of conversation and articulation in order to understand the meaning of language. This is in line with Gumprez's opinion (Gordon & Tannen, 2023), the insight that speech in conversation can only be interpreted by referring to speech activities identified by what he calls 'contextualization cues'. These theories

are used in this study to form the linguistic profile of the speaker, both the reported linguistic profile. The linguistic profile that is formed makes it easier for researchers to find out the background or identity of the reported person in this case.

Linguistic profile analysis was also conducted by Zahra & Wei (2023) entitled "Excavation of the Suspect's Identity by Analysis of the Speaker Profiling Case Evidence of Fake Voice Recordings". This study used a forensic linguistic approach with a linguistic profiling method based on *aural perception*. The research data consisted of recordings obtained from YouTube video uploads and original recordings broadcast on television. The results of the study show differences between the data obtained through YouTube and the original recordings broadcast on television in terms of phonetics and phonology. These two aspects show differences in pronunciation, stress, frequency, and tone in the two research data, revealing that the data obtained from YouTube videos had been falsified.

This study also utilized Praaline and Praat software to examine the conformity of the original recordings with the comparative recordings. Although the focus of the study was on linguistic profiles using Praline software, this study also combined dialect analysis, idiolect, language variation, and register analysis using a sociolinguistic analysis approach. Thus, the findings in this study are expected to strengthen the contribution of sociolinguistics in the field of forensic linguistics by emphasizing that linguistic analysis is not only descriptive but can also be implemented in law enforcement. Thus, this study is one of the developments in linguistic profiling methods that utilize linguistic theories in proving the intentions, attitudes and identities of speakers in the courtroom.

2. Method

This type of research is classified as descriptive qualitative research. Qualitative research is often called naturalistic research because it is conducted in actual conditions (Sobry & Hadisaputra, 2020; Sugiyono, 2013). Meanwhile, descriptive research is data presented in the form of words and images, rather than numbers (Abdussamad, 2021; Wekke, 2019). Thus, in this study, the data produced is in the form of speech. This study is included in forensic linguistics through a sociolinguistic approach in the form of linguistic profile analysis. YouTube social media was the source of data in this study. The researcher took three videos from the defendant's YouTube channel "DRL" containing the defendant's reviews of a beauty product called Helwa Beautycare.

The first video was uploaded on November 3, 2019, with the title "Parah Bethol!! Hasil Laboratorium Helwa Beauty Night Cream 2019!!!". Then, the second video was uploaded on February 1, 2020, with the title "Tercyduk Lagi! Bodylotion Abal-Abal, Ngaku Sudah Bpom? Mau Putih Jadi Hancur! Kejam-Nya Helwa". Next, The Third Video Was Uploaded on August 7, 2020, With The Title "Review Helwa Baru Nih!! Apakah Dia Sudah Tobat? Hasil Lab Agustus 2020". These three videos were utilized in this study because they pertain to the issues involving the defendant "DRL" and Helwa Beautycare. The defendant "DRL" is generally known as one of Indonesia's cosmetic doctors. The defendant also has several social media accounts, such as Instagram, TikTok, and YouTube. The defendant usually uploads content reviewing the quality of other people's beauty products. In addition, the defendant also has a main business, namely the Athena Group beauty clinic. Helwa Beautycare is a beauty product founded by Nabella Tamim, often referred to as Bella.

During data collection, researchers used observation methods and recording techniques. Mahsun (2019) states that observation is a method used in research to obtain data by listening to the speech of language users. In this study, researchers listened to the speech of the defendant "DRL" in three videos uploaded by the defendant. The defendant "DRL" is a doctor in Indonesia.

Meanwhile, the reporter or victim is Nabila Abdurrahman, who is also known as the owner and founder of Helwa Beautycare. The two parties were in conflict in the digital media realm, so the case was classified as *a cybercrime* with verbal statements as evidence. To facilitate data analysis, the researcher converted the statements in the videos into transcripts. Therefore, the researcher recorded the defendant's speech, which showed distinctive characteristics of dialect, idiolect, register, and language variation in the video, which will then be used as research data. The recording technique is an advanced technique used when someone applies the listening method in collecting data (Mahsun, 2019) .

The phonetic analysis in this study utilizes the *International Phonetic Alphabet* (IPA) table. Meanwhile, the acoustic linguistic analysis, which includes the frequency, tone and rhythm of the defendant's speech in his YouTube video, compares it with other YouTube videos that have the same dialectal speech by the defendant, utilizing the Praaline corpus software to validate the data analysis. In the idiolect analysis related to the defendant's grammatical patterns, the researcher refers to the opinion of (Anand & Korotkova, 2022) which states that the core structure of written Indonesian sentences is actually very simple, consisting only of a subject and predicate (S-P).

3. Result and Discussion

From November 2019 to August 2020, "DRL" uploaded a video on his YouTube channel containing a review of a *beauty care* product, namely Helwa Beautycare. "DRL" stated that the product was dangerous to use because it contained hazardous ingredients, such as *hydroquinone* and mercury, even though the product had passed BPOM testing. The video allegedly harmed Helwa Beauty care by violating laws and regulations and the *Kode Etik Kedokteran Indonesia (KODEKI)* by posting Helwa's cosmetic products. Furthermore, "DRL" is said to have misled the consumers with the aim of discouraging them from using Helwa's cosmetic products and instead using cosmetic products from the defendant "DRL"'s own clinic, which also produces cosmetic products similar to those of Helwa. Due to the widespread attention on social media, the defendant "DRL's" uploaded video containing reviews of the products has also been widely discussed. The abundance of linguistic evidence found in the defendant "DRL"'s uploaded video containing reviews of Helwa Beautycare products prompted researchers to conduct a linguistic profile analysis, including an analysis of the defendant "DRL"'s dialect, idiolect, register, and language variations.

Dialect Analysis of Defendant "DRL"

Dialect analysis was conducted on the phonetic aspects of the defendant "DRL" in the three videos uploaded to his YouTube channel. The analysis of the phonetic aspects was adjusted based on the International Phonetic Alphabet (IPA) Table 1. Based on the three YouTube videos uploaded by the defendant "DRL", it was found that he used a dialect spoken by the people of Sumatra, namely in the pronunciation of phonemes when the defendant "DRL" spoke.

Table 1. Spoken Vocabulary of DRL

Vocabulary	Phoneme Pronunciation
Kek; Merek; Ngeyel 'Like; Brand; Stubborn'.	Phoneme **[e]**, pronounced with *the vowel* **[ɛ]**, *Open-mid, near-front.*
Positif; Orang; Oke; Laptop 'Positive; Person; Ok; Laptop'.	Phoneme **[o]**, pronounced with the vowel **[ɔ]**, *mid-low, back.*
Manusia;Sama;Kedua;Buka 'Human; Same; Second; Open'.	Phoneme **[a]**, pronounced with vowel **[a]**, *open, mid.*

Based on the table, the pronunciation of the defendant's phonemes "DRL" is like [e] pronounced as [ɛ], [o] pronounced as [ɔ], and [a] pronounced as [a]. The pronunciation of these phonemes is widely used by speakers in the Sumatra region. This shows that the defendant "DRL" has a Sumatran dialect. In addition, in the defendant "DRL"'s YouTube channel video, the defendant uses intonation that begins with a high tone and sounds like shouting. The use of such intonation is one of the characteristics of the dialect spoken by people in the Sumatra region. Given the high intonation characteristic of the defendant "DRL" when speaking, it can be said that the defendant's intonation is influenced by the dialect spoken by the people of Sumatra.

In addition, investigators also found a video on the defendant's YouTube channel titled "Masak Besar di Palembang!!". In the video, the defendant's speech "DRL" indicates that the defendant "DRL" has a Palembang dialect. One of the characteristics of the Palembang dialect is the change of the vowel sound /a/ to /o/ in several words. As seen in Table 2, the words spoken by the defendant indicate the presence of the Palembang dialect, as follows.

Table 2. Words Uttered by DRL

Word Palembang Dialect	Translation Indonesian	English Translation
Semuanyo	*Semuanya*	Everything
Kito	*Kita*	We
Galo-Galo	*Semua*	All
Ado	*Ada*	There is/ There are
Biso	*Bisa*	Can
Baso	*Bahasa*	Language

The researchers also utilized one of *the* YouTube *reels* videos of public figures Dinda Hauw and Rey Mbayang, uploaded by the @reinfamily5678 account, in which they speak with a Palembang dialect. The researchers used the video to reinforce the data in the form of the defendant's speech using the Palembang dialect. In the YouTube *reel* uploaded by the account, Dinda Hauw and Rey Mbayang are seen speaking in the Palembang dialect, which involves changing the vowel /a/ to /o/ in words such as "*baso*" which means 'language'.

The defendant "DRL" also used the Betawi dialect in his statement. This was influenced by his place of residence, Jakarta. The defendant "DRL" was heard using the Betawi dialect in a YouTube video titled "Istana Dr. Richard Lee! Omset Sehari 41 Miliar, Pecahkan Rekor! #Grebekrumah" uploaded by the YouTube account "AH." The defendant "DRL" used elements of the Betawi dialect, particularly in the choice of pronouns and informal vocabulary, such as *"gue"* and *"gua"* which means 'me'. The use of such vocabulary is typical of the Betawi dialect and Jakarta slang, reflecting environmental influences.

Auditory Analysis

The auditory analysis includes a comparative analysis of several videos that use the Palembang dialect to reinforce the main data, which is the defendant's video that also uses the Palembang dialect. These videos are the defendant's main data video titled "Masak Besar Di Palembang!!". This is followed by two comparison videos, namely a YouTube *reels* video uploaded by the @lemakngonten account titled "Angkat Bicara! Ustad Felix Siauw Tentang Rendang Willy Salim Di

Palembang", and a video uploaded by the @Adiez Gilang account titled "Wong Kito Galo! Gilang Dirga & Rizal Amanda#AGvlog." The comparison data results from both videos, analyzed using the Corpus Praaline software, are as follows.

The main data, namely the voice of the Defendant "DRL" has a frequency of 129.199–258.398 Hz (average fr. 193.7985), a pitch of C3-21c–C4-21c with decibels of -22 to -12. This data is supported by two other comparative data sets. The first is voice data from a video uploaded by @lemakngonten titled "Angkat Bicara! Ustad Felix Siauw Tentang Rendang Willy Salim Di Palembang" with a frequency of 129.199–258.398 (average fr. 193.7985), pitch C3-21c–C4-21c, decibels -31 - -20. The second video, uploaded by the account @Adiez Gilang with the title "WONG KITO GALO! Gilang Dirga & Rizal Amanda#AGvlog" has a frequency of 129.199–258.398 Hz (average fr. 193.7985), pitch C3-21c–C4-21c, decibels -18 to -15. These results are the analysis obtained when mentioning the word *"kito"* in the Palembang language, which means 'we'. The researcher chose the word *"kito"* because all three videos had one thing in common: they often mentioned the word *"kito"*.

Analysis of the Defendant's Idiolect "DRL"

Based on the video uploaded by the defendant "DRL", there are several idiolects that are often used by the defendant "DRL" in his speech, which are characteristic of the defendant "DRL". The idiolects that appear in the defendant's speech are heard repeatedly when the defendant delivers his speech. As seen in Table 3, the data in the form of a transcription of the speech taken from the defendant's YouTube channel is as follows.

The defendant exhibited an idiolect in the form of repetitive word patterns, such as "*abal-abal*, *ya*, and *share share share*" 'fake, yes, and share, share, share' which were uttered repeatedly. In addition to being a form of linguistic identity, the defendant's use of repetitive words was also used to express the speaker's emotions and to give a firm appeal to his YouTube audience. Thus, the use of repetitive words that appeared in the defendant's speech certainly reflected the characteristics of idiolect.

In addition, the defendant "DRL" also has a distinctive idiolect in the form of *scrambled syntax* or the habit of speaking with random syntax. Syntactically, the speaker does not seem to follow the standard sentence structure that has a systematic Indonesian grammar in the form of "S–P". As seen in Table 4, the use of random syntax by the defendant "DRL" can be proven in the following data.

As seen in Table 4, based on this data, it can be said that the grammatical patterns used by the defendant "DRL" are irregular, namely "K-S-P-O-P, K-S-Pel-P, Ket. –S–P–Ket.–S–O–P, P–K–P, K–S–P–Pel.–P". In addition, the irregular grammatical patterns in the defendant's speech appear when the defendant uses the word *"punya"* 'have'. This phenomenon shows that the defendant does not refer to the standard Indonesian sentence structure when speaking spontaneously, but rather relies on a style of speech based on habit. Thus, the use of repeated words and grammatically incorrect structures can be considered a characteristic of the defendant's idiolect.

The defendant "DRL" in the speech heard through the video uploaded to his *YouTube* channel has a characteristic idiolect of changing the sounds [s], [h] to [ʂ], [a] to [ä] in the word *"share,"* so that the word sounds like [ʂär] in the defendant's speech. The word *"share"* is also pronounced with a swaying tone by the defendant. In addition, the defendant "DRL" also changes the sound [a] to [ä] in several other words, such as in the word "*teman"* 'friend' to [temän]*, "benar"* 'right' to [benär]*, "pedas"* 'spicy' to [pedäs]*, "cepat"* 'fast' to [cepät]*.

Table 3. Palembang Dialect Vocabulary of DRL

Speech Transcription	Repeated Words
"Jadi kita gak usah ributin hal yang seperti itu punya, mau merkuri mau hidrokuinon, jelas ini produk berbahaya jadi kita gak usah terlalu ributin hal seperti itu punya, mempertanyakan hal yang seperti itu punya."	Punya
'So we don't need to worry about things like that, whether it's mercury or hydroquinone, clearly these are dangerous products, so we don't need to worry too much about things like that, questioning things like that.'	
"Kenapa saya bilang seperti ini punya?"	
'Why am I saying this?'	
"Ini baru namanya BPOM ya, ini baru pake, bangga lah saya ngomong ini baru namanya BPOM ya. Terus kita liat lagi toner ya, tonernya kaya gini, sekarang gak ada toner ya, terus BB creamnya kek gini ya, cuma ada merek doang, gak ada ingredientnya juga, gak ada apa-apa abal-abal ya"	Ya
'This is what BPOM is all about, right? This is new, and I'm proud to say this is what BPOM is all about. Then let's look at the toner again. The toner is like this. Now there's no toner, right? Then the BB cream is like this. There's only the brand name, no ingredients listed, nothing at all, just fake stuff.	
"Karena ya, ternyata ya saya masukin ke lab seperti biasa, saya YouTuber bermodal gak pake tetes-tetes ya, saya masukin ke lab."	
'Because, you see, I sent it to the lab as usual. I'm a YouTuber with no budget, so I sent it to the lab.'	
"Karna video ini hanya terbatas tentang penggunaan bodylotin abal-abal"	Abal-abal
'Because this video is only about the use of fake body lotion.'	
"Krim handbody abal-abal seperti ini punya" 'Counterfeit handbody cream like this has'	
"Sekali lagi saya bilang share, share, share"	Share
'Once again, I say, share, share, share.'	
"Jangan lupa share, share, share, selalu saya garis bawahi jangan lupa share"	
'Don't forget to share, share, share, I always emphasize, don't forget to share'	

Analysis of Defendant "DRL"'s Register

In this study, the register found by the researcher aims to identify the characteristics of the language used by a particular group. In this case, the researcher found that the defendant "DRL" used certain words in his YouTube channel videos. The words used or spoken are often used by a particular group at. The data in the form of transcripts taken from the defendant's YouTube channel videos are as follows.

"Hai semuanya, welcome back to my channel. Jadi seperti yang kalian tahu, ni hari ini yang kalian tunggu-tunggu, kita akan mereview Helwa dengan kemasan baru. Jadi Helwa ini banyak banget direquest orang, katanya "sudah ada kemasan baru, dok", "sudah BPOM", "sudah lengkap", "sudah buka klinik, dok. Tolong review ulang." Oke deh saya review, khusus untuk Helwa saya kasih satu segmen khusus, jadi sebenernya saya ini paling sering ngendorsin Helwa, ya. Helwa ni temen lama

saya bahkan seterkenal itu hubungan saya dan Helwa sampai saya kalian search di YouTube "dokter Richard" muncul "dokter Richard Helwa". Padahal saya gak pernah kerja di Helwa, sumpah."

Hello everyone, welcome back to my channel. So, as you know, today is the day you've been waiting for. We will review Helwa with its new packaging. So, many people request for Helwa, saying, 'There's new packaging, doc,' 'It's been approved by BPOM,' 'It's complete,' 'You've opened a clinic, doc. Please review it again.'" Alright, I'll review it. Specifically for Helwa, I'll dedicate a special segment. Actually, I most often recommend Helwa. Helwa is an old friend of mine; in fact, my relationship with Helwa is so well-known that if you search for 'Dr. Richard' on YouTube, what appears is 'Dr. Richard Helwa.' Even though I've never worked at Helwa, I swear.'

Table 4. Words Repeated by DRL

Speech Transcription	Subject	Predicate	Object	Notes	Complement	Sentence Pattern
"Jadi kita gak usah ributin hal yang seperti itu punya…" 'So we don't need to bother with things like that…'	*Kita* 'we'	*- gak usah ributin* 'don't need to bother' *- punya* 'Like that'	*hal yang seperti itu* 'Things like that'	*Jadi* 'so'	-	K–S–P–O–P
"kenapa saya bilang seperti ini punya?" 'Why am I saying this?'	*Saya* 'me'	*- bilang* 'saying' *- punya* 'this'	*seperti ini* 'Like this'	*Kenapa* 'why'	*seperti ini* 'Like this'	K–S–Pel–P
"…jelas ini produk berbahaya jadi kita gak usah terlalu ributin hal seperti itu punya" '…clearly this is a dangerous product, so we don't need to make too much of an issue out of it.'	*- ini produk* 'This product' *- kita* 'we'	*-berbahaya* 'dangerous' *- gak usah terlalu ributin* ' don't need to bother too much *- punya* 'it'	*hal seperti itu* "things like that'	*- jelas* 'clearly' *- jadi* 'so'	-	Ket.–S–P–Ket.–S–O–P
"hati-hati dalam penggunaan hidrokuinon seperti ini punya" 'Be careful when using hydroquinone like this'	-	*- hati-hati* 'Be careful' *- punya* 'Like this'	*dalam penggunaan hidrokuinon seperti ini* 'Using hydroquinone like this'	-	-	P–K–P
"Kenapa saya bilang seperti ini punya?" 'Why am I saying this?'	*Saya* 'me'	*- bilang* 'saying' *- punya* 'this'	-	*Kenapa* 'why'	*seperti itu* 'Like this'	K–S–P–Pel.–P

Based on the above speech data, it can be identified that the speaker has a distinctive register and shows their identity as a YouTube content creator based on their choice of vocabulary. This register is reflected in the choice of vocabulary used when creating content on their social media. Several terms, such as "*welcome back to my* channel", "request", "review*", and *"share"* in the transcription above are part of a register that is closely related to the way content creators communicate, especially on the YouTube platform.

The phrase *"welcome back to my channel"* is an opening line that is very often used by YouTube creators when starting their videos. This phrase is not only a greeting to viewers or subscribers, but also usually serves as a marker of identity on their own channel and strengthens the relationship between content creators and *subscribers.* The use of this language shows that the speaker has an understanding of content creation in the YouTube community, which is characterized by the structure of video content openings on the platform that has been formulated in this way and aims to build audience engagement.

The use of the word "*request*" by speakers also indicates the existence of two-way communication dynamics between content creators and their audience. In the context of YouTube, the word *"request"* not only refers to a request, but also plays a role in shaping the interaction between content creators and their followers. The content created is often the result of suggestions, comments or feedback from viewers. The use of this word reflects the interactive and responsive style between content creators and their followers, which is a characteristic of content creators who rely on audience engagement.

Furthermore, the word *"review"* indicates that content creators produce evaluative or review-type content. In the YouTube community, *reviews* are a very common content idea, especially those related to products that are currently being talked about. The use of this word is not only a substitute for "review", but also reflects the speaker's tendency to use terms that are familiar in the realm of social media, as well as indicating an orientation towards a wider audience.

Finally, the use of the word *"share"* as an invitation to spread the video is part of a digital communication strategy oriented towards increasing audience reach. On the YouTube platform, this phrase usually appears at the end of a video and is part of a fixed formula that aims to increase interaction and the number of views of the content. The explicit use of this word shows that the speaker understands and uses language that serves to invite and promote in the context of social media.

Overall, the choice of vocabulary used by the speaker not only demonstrates their proficiency in using terms that are often heard on the YouTube platform, but also reflects the distinctive register of a content creator. This pattern indicates a strong competence in building relationships with audiences, responding to social media expectations and adapting language forms to communicative purposes on social media. Therefore, in forensic linguistic analysis, these words can be used as strong indicators in profiling the speaker's social background and digital activities in a more specific and targeted manner. The following are some examples of words that are often used by content creators.

1. "hai semua ketemu lagi sama aku, Tasyi. *Welcome to my channel*" Tasyi Athasyia.
2. "hai guys, *welcome back to my channel*" Jessica Jane, in a *YouTube* video on her channel titled "Cobain Makanan Tiktokers Viral".
3. "hari ini, Jessica mau *review* cireng ndutsnya Markibok." on Jessica Effendy's *YouTube reel* titled "Cobain Cireng Nduts".
4. "aku tuh lupa, waktu itu aku pengen *ngereview* ini." On Fuji's *YouTube* channel titled "Mukbang Tteokbokki Ter-Viral di *TikTok*, Beli Harus War Dulu".

5. "Terima kasih kepada kalian semua yang sudah me-*request* untuk gua mainin game *Creepy Shift: Roadside Diner* yang satu ini." On MiawAug's *YouTube* channel titled "Ini Game Horror Simulator – Creepy Shift: Roadside Diner Indonesia".
6. "Dan hari ini, sesuai dengan *request* kalian semua aku bakal borong semua peralatan bayi, *guys*" on the *YouTube* channel Frost Diamond entitled "Kita Borong Semua, Buat Surprise Lahiran Bayi Kembar Frost Diamond Junior Nanti".
7. "Jangan lupa *share* video ini ke teman-teman kamu yang membutuhkan" on Juliarimbaa's *YouTube* channel titled "Cara ampuh agar dia Takut kehilangan kamu.
8. "Jangan lupa *share* video ini ke temen-teme kamu yang , wah, pendapatannya banyak yang pengen tau tentang ini dengan baik, oke?" on Jang Hansol's *YouTube* channel Korea Reomit titled "Burning Sun.. Fakta akan terungkap".

As explained in the introduction, forensic linguists analyze linguistic evidence in a crime. Instead of determining the register of the defendant "DRL" as a doctor, this study found that the defendant's register was that of a content creator on YouTube based on the languages used by the defendant.

Analysis of the Language Variation of Defendant "DRL"

In this analysis, we will examine the language variation in the form of code-mixing used by the defendant "DRL". Code-mixing is the simultaneous use of two languages in a single utterance. In the videos uploaded by the defendant on his YouTube channel, it can be seen that the defendant often exhibits the phenomenon of code-switching in his speech. The following is some data in the form of transcripts of speech from the defendant's YouTube channel that can be presented in this study.

1. "Karna loh gak ada *ingredient*, gak ada label, kapan pun gak ada segel, kapan pun mau dia ganti isinya, *we never know*"
2. "Jadi, *better*, jangan dipake"
3. "sekali lagi saya bilang *share, share, share*"
4. "ini merek yang *TOP* bgt"
5. "Hai semuanya, *welcome back to my channel*"

In the first piece of evidence, the defendant "DRL" was providing information to viewers about the product being reviewed. The defendant said that the packaging of the product being reviewed was sold with incomplete information, such as no information about the ingredients used to make the product, no label and no seal on the bottle cap. Then, the defendant makes an assumption by saying *"we never know"* if the contents of the product can be replaced at any time because the product is sold with packaging that has minimal information about the product. The phrase *"we never know"* also gives the impression of ambiguity to the audience's understanding because the phrase is not based on clear facts that can raise doubts among the audience about the contents of the product reviewed by the defendant. Thus, this sentence can be categorized as opinion manipulation as a wild statement that has no responsibility for its meaning. The word *"ingredient"* and the phrase *"we never know"* in this statement are only fragments involved in the first data statement.

Furthermore, in the second piece of data, the defendant appears to be giving advice or a warning through his statement. This is evidenced by the defendant's statement warning that it is "better" not to use the product reviewed by the defendant on his YouTube channel because the product contains harmful ingredients after being tested in a laboratory by the defendant. The word *"better"* in this data clearly appears to be only a fragment of the utterance, so it is said that code-mixing has occurred.

In the third piece of data, the researcher found a phenomenon of code-mixing in the defendant's speech. The word that became a fragment in the speech was the English word "*share.*" In the third piece of data, the defendant appears to repeat the word "*share*" several times. The defendant's intention in repeating the word is to emphasize to viewers that they should share the video (Nurdiawan et al., 2025). The aim is to inform them that the product being reviewed is dangerous to use and needs to be disseminated.

In the fourth piece of data, it can be seen from the defendant's speech in the video he uploaded that he gave a review in the form of a rating of the product he reviewed using the word *"top"* which is taken from English. It can be seen that the word *"top"* in this piece of data is only a fragment of a word. In Indonesian, the word *"top"* means 'atas', meaning that the defendant's statement implies that the product being reviewed is ranked at the top.

The last piece of data, namely the fifth piece of data, shows that the defendant uttered this sentence when he gave his opening remarks in a video uploaded to his YouTube channel. The defendant opened the video by greeting his audience or YouTube subscribers with two words, namely *"hai semuanya*" 'hi everyone', followed by the sentence *"welcome back to my channel"*.

Based on this explanation, the researcher found code-mixing data in the defendant's YouTube upload video. In the video uploaded to his YouTube channel, the defendant used a lot of language variation in the form of code-mixing between Indonesian and English. In addition, the defendant also used language variation in the form of code-mixing to elicit reactions from the audience in the form of verbal responses during the video screening.

4. Conclusion

This study demonstrates that linguistic profiling provides a robust and systematic framework for identifying speaker identity in forensic contexts, particularly within digitally mediated discourse. By integrating dialectal, idiolectal, register, and language variation analyses, the findings reveal that the defendant "DRL" exhibits a distinctive linguistic profile shaped by regional affiliation, habitual speech patterns, and professional engagement as a digital content creator. Phonetic and acoustic analyses indicate the presence of Sumatran speech characteristics, most notably the realization of /e/ as [ɛ] and /o/ as [ɔ], accompanied by high pitch and rapid speech tempo. More specifically, consistent vowel shifts from /a/ to /o/ across multiple lexical items, together with Praaline-based acoustic measurements of the word "kito", align closely with Palembang dialectal patterns. The convergence of perceptual and instrumental evidence strengthens the reliability of dialect identification. In addition, the use of Betawi informal pronouns ("gue," "gua") reflects contact-induced variation influenced by the speaker's current sociolinguistic environment. At the idiolectal level, the defendant's speech is characterized by recurrent lexical repetition, non-canonical syntactic structures, and systematic phonetic deviations, including the realization of /s/ as [ʂ] and /a/ as [ä]. These features function as stable markers of individual linguistic behavior rather than incidental performance effects. Register analysis further identifies the defendant's communicative role as a YouTube content creator, evidenced by the repeated use of platform-specific discourse formulas such as "welcome back to my channel," "request," "review," and "share." Frequent Indonesian–English code-mixing serves pragmatic functions, including emphasis, evaluative framing, and audience engagement.

Notwithstanding these findings, this study is limited by its reliance on a restricted corpus of publicly available video data, which may not capture the full range of the speaker's linguistic variability across contexts. The absence of controlled elicitation also constrains the generalizability of certain phonetic and syntactic patterns. Despite these limitations, the study contributes to

forensic linguistics by demonstrating how multi-level linguistic analysis, combining sociolinguistic theory with acoustic evidence, can enhance the evidentiary value of language in legal settings. The findings underscore the applicability of linguistic profiling to digital media discourse and support its potential use in speaker identification and expert testimony. Future research should expand the corpus size, incorporate cross-contextual data, and develop standardized analytic protocols to improve reliability, replicability, and cross-case comparability in forensic linguistic practice..

Acknowledgment

References

Abdussamad, Dr. H. Zuchri S.I.K., M. S. (2021). *Metode Penelitian Kualitatif*. CV. syakir Media Press.

Akbar, A., Razak, N. K., & Anzar. (2023). Conversational Implicature pada Berita Acara Pemeriksaan Kasus Dugaan Pembunuhan Berencana di Makassar Tinjauan Linguistik Forensik. *Lingue : Jurnal Bahasa, Budaya, Dan Sastra*, *5*(1), 41–51. https://doi.org/10.33477/lingue.v5i1.5335

Anand, P., & Korotkova, N. (2022). How to theorize about subjective language: a lesson from "de re." In *Linguistics and Philosophy*. Springer. https://doi.org/10.1007/s10988-021-09331-0

Anggara, P. P. (2024). *Variasi Bahasa Idiolek Pada Tuturan Konten Kreator Youtube Ciduk Warrior: Kajian Sosiolinguistik*. *17*(April), 25–30. https://ejournal.unibba.ac.id/index.php/metamorfosis/article/view/1264

Anggreni, L. S., Nugroho, R. A., Luthfi, H. S., Kresna, I. M., & Santoso, T. B. (2019). Penggunaan kata umpatan di Twitter berdasarkan gender di pilkada Sumatera Utara 2018. *Jurnal Kajian Komunikasi*, *7*(1), 121. https://doi.org/10.24198/jkk.v7i1.18447

Bucholzt, M., & Hall, K. (2005). Identity and interaction: A sociocultural linguistic approach. *Discourse Studies*. https://doi.org/10.1177/1461445605054407

Casalan, M. C. (2015). *The language of police reports : A forensic linguistic analysis*.

Chaer, A. (2004). *Sosiolinguistik Perkenalan Awal*. Pt Rineka Cipta.

Coulthard, M., & Johnson, A. (2010). The Routledge handbook of forensic linguistics. In *The Routledge Handbook of Forensic Linguistics*. https://doi.org/10.4324/9780203855607

Fitri, N., Artawa, K., Satyawati, M. S., & Sawirman, S. (2021). Transitivitas dalam Teks Peradilan Indonesia: Kajian Linguistik Fungsional Sistemik. *Diglosia: Jurnal Kajian Bahasa, Sastra, Dan Pengajarannya*, *4*(2), 139–148. https://doi.org/10.30872/diglosia.v4i2.116

Fitria, T. N. (2024). Forensic Linguistics: Contribution of Linguistics in Legal Context. *PRASASTI: Journal of Linguistics*, *9*(1), 117. https://doi.org/10.20961/prasasti.v9i1.71527

Gea, B. H. (2024). Peran Ahli Bahasa Dalam Mengungkap Tindak Pidana Penghinaan Dan

Pencemaran Pada Media Elektronik. *Jurnal Surya Kencana Satu: Dinamika Masalah Hukum dan Keadilan*, *15*(1), 59–71. https://doi.org/10.20961/jv.v11i2.71624

Gharamah, F. M. A. (2025). Analysis Of Student Competence in Writing Scientific Papers Based on Information Technology (It). *Lingue: Jurnal Bahasa, Budaya, Dan Sastra Pusat Pengembangan Bahasa IAIN Ambon.* https://doi.org/10.33477/lingue.v7i1.9777

Gordon, C., & Tannen, D. (2023). Framing and related concepts in interactional sociolinguistics. *Discourse Studies*, *25*(2), 237–246. https://doi.org/10.1177/14614456231155073

Jahara, Munirah, & Andi Sukri Syamsuri. (2022). Application Statement Analysis Conversation of the Murder Case at the Makassar District Court (Forensic Linguistics). *RETORIKA: Jurnal Ilmu Bahasa*, *8*(2), 110–118. https://doi.org/10.55637/jr.8.2.5435.110-118

Kristianto, Y. (2015). Linguistik Forensik:Sebuah Tinjauan Bahasa Dalam Ranah Hukum. *LITERA Jurnal Bahasa Dan Sastra*, *1*(1), 97–102. https://doi.org/10.14421/ajbs.2012.11102

Kusno, A., Arifin, M. B., & Mulawarman, W. G. (2022). Pengungkapan Pemerasan dan Pengancaman pada Alat Bukti Kasus Pinjaman Online (Kajian Linguistik Forensik). *Diglosia: Jurnal Kajian Bahasa, Sastra, Dan Pengajarannya*, *5*(3), 555–570. https://doi.org/10.30872/diglosia.v5i3.423

Kusumawati, A. (2025). Indonesian and English sentence patterns: A contrastive syntactic perspective. *Linguistics Initiative.* https://pustaka.my.id/journals/live/article/view/341

Mahsun, Prof. Dr., M. S. (2019). *Metode Penelitian Bahasa Tahapan, Strategi, Metode, Dan Tekniknya* (Edisi Keti). Pt Rajagrafindo Persada.

Mahsun, P. D. M. . (2018). *Linguistik Forensik Memahami Forensik Berbasis Teks Dengan Analogi Dna.* Pt Rajagrafindo Persada.

Muysken, P. (2000). *Bilingual speech : a typology of code-mixing.* Cambridge University Press.

Naznin, S., Bashir, & Khan, F. (2022). The Importance of Forensic Language in Law. *Competitive Social Sciences Research Journal (CSSRJ)*, *3*(2), 801–809. https://cssrjournal.com/index.php/cssrjournal/article/view/450

Nurdiawan, I., Apriyani, T., & Li, H. (2025). Beyond the Ideal: Deconstructing Perfection in Aveus Har's 'Istri Sempurna.' *COMMICAST*, *6*(2), 364–372. https://doi.org/10.12928/commicast.v6i2.12854

Ode, M. I., & Huda, N. (2022). Peran Linguistik Forensik terhadap Penegakan Hukum Indonesia: Studi Analisis Ujaran Kebencian Suku Lampung. *Tadabbur: Jurnal Integrasi Keilmuan*, *1*(01). https://doi.org/10.15408/tadabbur.v1i1.27874

Rosaniati, Sanyoto, & Wasi, R. (2021). Penerapan Saksi Ahli Linguistik Forensik Sebagai Alat Bukti Dalam Tindak Pidana Turut Serta Menyiarkan Berita Bohong Dan Menimbulkan Keonaran Terkait Hasil Swab Test (Studi Putusan Nomor 225/Pid.Sus/2021/Pn.Jkt.Tim). *Soedirman Law Review*, *3*(4), 664–680. https://doi.org/10.20884/1.slr.2021.3.4.170

Rusdiansyah. (2020). Hukum Dan Linguistik Forensik. *Al-Amwal Journal of Islamic Economic Law* *5*(1), 21–31. https://doi.org/10.24256/alw.v5i1.1420

Sholihatin, D. E. (2024). *Linguistik Forensik Dan Kejahatan Berbahasa.* Pustaka Pelajar.

Sobry, M., & Hadisaputra, P. (2020). Penelitian kualitatif Penelitian kualitatif. In *Bandung: PT. Remaja Rosda Karya* (Issue c).

Subyantoro. (2019). Linguistik Forensik: Sumbangsih Kajian Bahasa dalam Penegakan Hukum. *Adil Indonesia Jurnal*, *1*. https://doi.org/10.35473/aij.v1i1.141

Sugiyono, P. D. (2013). *Metode Penelitian Kuantitatif, Kualitatif Dan R & D*. Alvabeta, Cv.

Ubaidah, R., Setianingrum, F. A., Novari, R. D., Bahasa, P., & Darussalam, S. (2024). Forensic linguistic study of the case of vicky prasetyo and angel lelga in the case of defamation. *Jurnal Skripta 1*0(1), 37–41. https://doi.org/10.31316/skripta.v10i1.5959

Wagner, L., Vega-Mendoza, M., & Van Horn, S. (2014). Social and linguistic cues facilitate children's register comprehension. *First Language*, *34*(4), 299–314. https://doi.org/10.1177/0142723714544413

Wekke, I. S. (2019). *Metode Penelitian Pendidikan Bahasa* (Yogyakarta). Penerbit Gawe Buku.

Wicaksana, L. S. D., Pastika, I. W., & Satyawati, M. S. (2023). Kajian Linguistik Forensik dalam Penyidikan Kasus Pembunuhan Engeline. *Humanis*, *27*(2), 217. https://doi.org/10.24843/jh.2023.v27.i02.p10

Zahra, T. Y. Z., & Wei, Z. (2023). Excavation of the Suspect's Identity by Analysis of the Speaker Profiling Case Evidence of Fake Voice Recordings. *NOTION: Journal of Linguistics, Literature, and Culture*, *5*(2), 169–178. https://doi.org/10.12928/notion.v5i2.8038