

**A CORPUS-BASED STUDY OF PRONOUNS IN BAHASA INDONESIA:
A CASE STUDY OF AKU/SAYA AND KITA/KAMI**
(*Pronomina Berbasis Korpus dalam Bahasa Indonesia:
Studi Kasus Aku/Saya dan Kita/Kami*)

Eva Tuckyta Sari Sujatna^a, Kasno Pamungkas^b, Meita Lukitawati Sujatna^c

^{abc}Universitas Padjadjaran

Jl. Raya Bandung Sumedang KM.21, Kabupaten Sumedang, Jawa Barat 45363

Pos-el: eva.tuckyta@unpad.ac.id, kasno.pamungkas@unpad.ac.id,
sujatna.meita@widyatama.ac.id

Naskah Diterima Tanggal 30 Desember 2023; Direvisi Akhir Tanggal 6 Juni 2024;

Diterbitkan Tanggal 17 juni 2024

DOI: <https://doi.org/10.26499/sawer.v30i1.1307>

Abstract

The first-person pronouns in Bahasa Indonesia have two forms: singular and plural. The singular pronoun forms are saya and aku for the singular pronoun forms, while the plural pronoun forms for the first-person are kita and kami. This paper aims to figure out the highest number of both singular and plural forms in Bahasa Indonesia, the number of the first-person singular pronouns form aku and saya, and the number of the plural pronouns for first-person forms kami and kita. The corpus linguistic method is applied in this research. Using corpus software, vis. AntConc 3.5.8 tool, the pronouns aku, saya, kita, and kami were identified. The 8,432,978 tokens were obtained from four different sources: newspapers, magazines, short stories, and novels. From the number of collected tokens, 224,796 corpora were classified as singular and plural pronouns. The result of the research shows that 1) the amount of the singular pronoun for the first-person is higher than the plural form; 2) the quantity of the singular pronoun form for the first-person aku is higher than saya; and 3) the usage of the plural pronoun of the first-person forms kami is higher than kita established in the four sources of the corpus.

Keywords: *pronouns, bahasa Indonesia, a corpus-based*

Abstrak

Pronomina orang pertama di dalam bahasa Indonesia memiliki dua bentuk: tunggal dan jamak. Bentuk pronomina tunggal adalah saya dan aku, sedangkan bentuk jamaknya adalah kita dan kami. Tujuan penelitian ini mendeskripsikan pronomina yang memiliki jumlah terbanyak yang berbentuk tunggal dan jamak dalam bahasa Indonesia, baik pronomina orang pertama berbentuk tunggal saya dan aku serta pronomina orang pertama yang berbentuk jamak kami dan kita. Penelitian ini menggunakan metode linguistik korpus. Dengan menggunakan perangkat lunak korpus, vis. aplikasi AntConc 3.5.8, kata ganti aku, saya, kita, dan kami diidentifikasi. Sebanyak 8.432.978 token yang berasal dari koran, majalah, cerita pendek, dan novel digunakan sebagai data penelitian ini. Dari sejumlah token tersebut, 224.796 korpus diklasifikasikan. Hasil penelitian menunjukkan bahwa 1) jumlah pronomina orang pertama tunggal lebih banyak dari pronomina orang pertama jamak; 2) penggunaan pronomina aku lebih banyak dari saya; dan 3) dari empat sumber korpus yang digunakan dilaporkan bahwa penggunaan pronomina orang pertama jamak kami lebih banyak dari kita.

Kata-kata kunci: *pronominal, bahasa Indonesia, berbasis korpus*

INTRODUCTION

Pronouns serve a significant aspect in a language. Pronouns are the replacement for nouns. According to Muryasov et al., the controversy surrounding the group of words known as 'pronouns' in linguistic leads to endless discussions. Some linguists view them as independent parts of speech, while others view them as functional words, which brings up the question of their status in the system of parts of speech (Muryasov et al., 2024).

A pronoun is a word that refers to participants in the discourse, in English; pronouns are *I, we, you, they, he, she, and it*. Pronouns, in Bahasa Indonesia, can be categorized into three different types: the first, the second, and the third person with singular and plural forms of each. This paper focuses on the first-person singular and plural pronoun forms. In Bahasa Indonesia, the singular form could be *saya* or *aku* and the plural form could be *kita* or *kami* while the singular is "I" and the plural is "we" in English. Some questions arise from the four sources: Which first-person pronoun is more dominant, the singular or plural form? From the four sources, which one is commonly used as the first-person singular pronoun form, *aku* or *saya*? Or from the four types of sources, which one is more commonly used as the first-person plural pronoun form, *kita* or *kami*?

The corpus-based terminology in this paper refers to a method using corpora of texts (Scott & Tribble, 2006). As a database of language, the corpus is also known as a depository of samples (McEnery & Hardie, 2012), by Jones & Waller who claimed that to ease the readers to search the text, it was stored electronically (Jones & Waller, 2015) while Viana et al. (eds) in 2011 argued that corpus as a textual population is fully representative (Viana et al., 2011). Related to the approaches mentioned, this study is based on the corpus collected from four sources: newspapers, magazines, short stories, and novels. The corpus collected from the four sources are 8,432,978 tokens. By using corpus software, vis. AntConc

3.5.8 tool, the pronouns *aku, saya, kita, and kami* were identified and analyzed.

Many studies have referenced pronouns, including Bahasa Indonesia. In 2001, Djenar discussed locative pronouns in Bahasa Indonesia (*sini* 'here', *situ* 'there', and *sana* 'over there') (Djenar, 2001); in 2008, Djenar examined the choice of the two singular first-person pronouns *aku* and *saya* (Djenar, 2008). Besides the two articles by Djenar, Manns in 2012 described the variation of the first-person relates to stance and identity in Indonesia (Manns, 2012) while Tang & John (1999) described the first person "I" in exploring the writers' identity (Tang & John, 1999), in 2014, Wehrli & Awal discussed the second-person reference in Indonesian Christian prayer (Wehrli & Awal, 2014). The other studies discussed the pronouns in local languages in Indonesia, such as Fricke in 2017 illustrated the nouns and pronouns in Central Lembata Lamalohot (Fricke, 2017), Ratu in 2019 discussed pronouns in Mangondow language as one of the Indonesian local languages (Ratu, 2019). Flowerdew & Wang (2015), Walková (2018), Işık-Taş (2018), He (2020), Firdaus (2021), Li (2021), Yu (2021), Azar (2022), and Hardjanto (2022) discussed pronouns as authorial identity in academic writing and all the studies were inspired by Hyland (2001; 2002a; 2002b) (Azar & Hashim, 2022; Firdaus et al., 2021; Flowerdew & Wang, 2015; Hardjanto, 2022; He, 2020; Hyland, 2001, 2002b, 2002a; Işık-Taş, 2018; Li, 2021; Walková, 2018; Yu, 2021). Yusuf, et al. (2019) described the pronoun of address *ke* in Indonesian by Acehese speakers as reflecting the power and solidarity (Yusuf et al., 2019). The research on pronouns based on corpus has been done by Wijayanti et al. (2022), they discussed authorial pronouns *we* in EFL academic writing (Wijayanti et al., 2022) and Womack who discussed *we, us, and our* as pronouns (Womack, 2022).

Not only concerning pronouns, some studies also involve the corpus, such as Sujatna et al. (2022) who depicted the word

coronavirus in the NOW corpus (Sujatna et al., 2022). In 2021 Sujatna & Sugianto collected from corpora collection of Leipzig University and described the deictic time words in Bahasa Indonesia (Sujatna & Sugianto, 2021). In 2019, Yuliawati et al. illustrated the local wisdom related to linguistic construction of gender through the word “*awewe*” and “*lalaki*” in Mangle magazine (Yuliawati et al., 2019).

Since it was still limited information concerning pronouns related corpus, this study tries to determine the first-person pronouns in Bahasa Indonesia, based on the corpus. The corpora were taken from four different types of sources, they are newspapers, magazines, short stories, and novel which relates to first-person pronouns in Bahasa Indonesia.

THEORETICAL FRAMEWORK

Pronouns could be either a word or a group of words. Pronouns function as the replacement of nouns and they could refer to a person or thing (Alwi et al., 2003; Moeliono et al., 2017). Pronouns are classified into a closed class since nouns, verbs, adjectives, and adverbs are classified into open class. Pronouns in Bahasa Indonesia do not specify gender indication. Evans (1980) described pronouns into three categories: reference, coreference, and antecedent (Evans, 1980). In Bahasa Indonesia, pronouns could be demonstrative, such as *ini* ‘this’ *itu* ‘that’, and *situ* ‘there’ and the first, the second, and the third-person pronouns

First-person pronouns are categorized into singular and plural forms. The singular pronoun for the first-person form is “I” in English and ‘*saya*’, ‘*aku*’, or ‘*daku*’ are in Bahasa Indonesia while the plural pronoun form for the first-person “we” in English and ‘*kita*’ ‘*kami*’ in Bahasa Indonesia. The authors exclude the word *daku* since the pronoun is only found in literary works as argued (Moeliono et al., 2017). In Bahasa Indonesia, the choice of pronouns by the speakers relates to many variables, one of them is formality, while English does not

(Yusuf et al., 2019). In Bahasa Indonesia, the pronoun is used to distinguish the formality, and the pronoun ‘I’ is related to enacting stances (Manns, 2012). The following are the illustrations of the singular pronoun for the first-person form *saya* and *aku* in Bahasa Indonesia that refer to the pronoun ‘I’ in English.

- (1) *Saya sudah menunggunya selama dua jam saat itu.*
‘I have been waiting for her/ him for two hours at that time.’
- (2) *Mungkin aku kesal padanya.*
‘I was upset on him, maybe.’

The singular pronoun for the first-person form *saya* in (1) describes the singular pronoun for the first-person form *saya* is applied by the speaker in a formal situation. It aligns with Moeliono et al. & Alwi et al. that the first-person singular pronoun form *saya* is usually applied in a written and an oral conversation in a formal situation (Alwi et al., 2003; Moeliono et al., 2017). It differs from example (2), which illustrates the first-person singular pronoun form *aku* in (2) is described in an informal situation.

- (3) *Kami akan berangkat pukul enam pagi.*
‘We (excluding the listener) will go at six o’clock.’
- (4) *Kita akan berangkat pukul enam pagi.*
‘We (including the listener) will go at six o’clock.’

Examples (3) and (4) portray the usage of the pronoun *kami* and *kita* as the first-person plural pronoun forms. Both pronouns refer to the pronoun ‘we’ in English. Although, the two pronouns refer to the pronoun ‘we’, they have different meanings of each. As Moeliono (1988) cited by Sudartini (2009) that *kami* as an exclusive and *kita* as an inclusive first person pronoun in Bahasa Indonesia (Sudartini, 2009). The pronoun *kami* in (3) has an exclusive meaning; the listeners are assumed not to be involved in the conversation, while the pronoun *kita* has an inclusive meaning that the speakers and the

membiarkan empat belas panggilan	aku	abaikan. Kalau laki-laki itu
Wow, sejak kapan dia panggil	aku	aku Abang? Sejak kapan pula dia
aku pun menjulurkan tanganku	aku	abang etika, katanya. Aku
, dan 30etika30 sebagai istri birokrat keparat.	aku	, Abim, dan Amru juga tak
1 yang berarti tiga hari Minggu	aku	absen dari ritual ini? Bukan saja
yang fana adalah waktu, dan	kita	abadi.
Indoensia adalah sebuah suratn takdir bagi	kita	. Abai terhadap keragaman adalah
jauh ke depan, yang tidak dapat	kita	abaikan bagi kelangsungan
Hanya karena masalahnya terlalu kecil, lalu	kita	abaikan, begitu? Semua orang
diterapkan orangtuanya. Juga tak bisa	kita	abaikan faktor bawaan, ada yang
juga akan menyentuh banyak sisi yan	kita	abaikan selama ini. Tidak akan
atau berenang. Kau ingat, tidak? Teman	kita	Abe dengan polosnya berkata,
kepada Naila. Masih jauhkah kampung	kita	, abi? Sebentar lagi kita
sedang terhidang di halaman Tanah Air	kita	. Aceh pascatsunami adalah `pesta
tujuan besar kita. Inilah etika kenapa	kita	ada. Aku nggak akan secetek itu
Dongeng etikaku, ia ada sebab	kita	kita ada , atau sebaliknya: kita ada
terasa manis. bahkan supermanis di lidah	kita	. Ada Baklawa dan Sambousek Bil
dan pemulihan. Kedua hal yang	kami	abaikan selama mencari keadilan
atau tidak. Sejak dulu, orang tua	kami	acap menggunakan daun ganja atau
mengenal kami berdua pasti akan berpikir	kami	ada apa-apanya. Ini semua berkat
Perempatan yang superramai. Di depan	kami	ada bangunan berbentuk balok,
Paman Kay seperti mengerti ekspresi wajah	kami	. 'Ada banyak hal yang lebih
. Mungkin memeriksa etik di mobil	kami	ada bawaan yang mencurigakan.
sejarah. Anda tentu lihat, di tempat	kami	ada 4 benteng Spanyol dan satu
hari seberapa pun banyaknya ia makan.	kami	ada bersamanya etika ia dikunci
koloni tempat kalian menginduk. Untuk itulah	kami	ada. Bertahap, Zarah melihat titik-
keluar masuk tenda itu. Di antara	kami	, ada bocah laki-laki maupun
, kesadaran kami melejit ke ubunubun bahwa	kami	ada dalam datar pencarian orang.
hanya di Teluk Jakarta. "Beberapa crew	kami	ada di beberapa proyek mereka

To analyze more details, the authors classify them into two main categories (singular and plural) and then extracted them into four subcategories, *saya*, *aku*,

kita, and *kami* based on the four sources (newspapers, magazines, short stories, and novels) as described in the following Table 2.

Table 2
First-person Singular & Plural Pronoun Forms

Source	First-person pronouns			
	singular		plural	
	<i>saya</i>	<i>aku</i>	<i>kita</i>	<i>kami</i>
Newspaper	14,629	680	8,044	9,141
(%)	96	4	47	53
Magazine	9,405	1,047	9,784	4,804
(%)	90	10	67	33
Short story	9,014	28,918	4,435	7,649
(%)	24	76	37	63
Novel	6,055	77,439	12,334	21,418
(%)	7	93	37	63

Table 2 illustrates the number of four pronouns found in the data and the percentage of each pronoun to represent the complete understanding of the data. Table 2 depicts that related to the percentage of the utilization, the singular pronoun for the first-person *saya* in newspaper is the

highest among the pronouns. The singular pronoun for the first-person *aku* in newspaper is the lowest usage among the pronouns.

RESULT AND DISCUSSION

In this paper, the authors examine the two major problems: first-person singular and plural pronoun forms found in newspapers, magazines, short stories, and novels.

The Distribution of the Pronoun Saya and Aku

From 224,796 corpus extracted from 8,432,978 tokens, there are 147,187 corpora in the first-person singular pronoun form as portrayed in Figure 1.

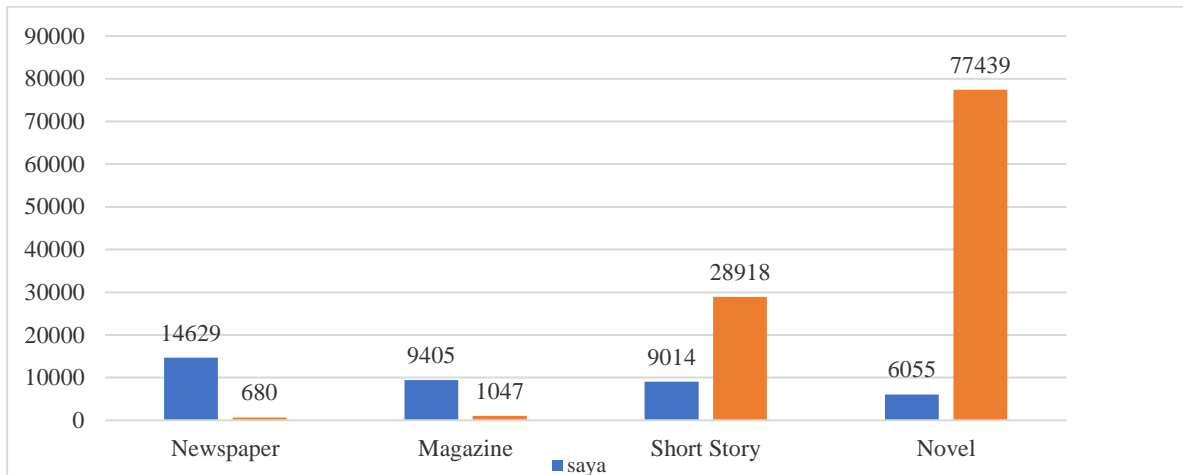


Figure 1.
Saya and Aku as First-person Singular Pronouns

Figure 1 portrays that sequentially, the first-person singular pronoun *aku* found in novels (77,439 from 83,494 corpora), short stories (28,918 from 37,932 corpora), magazines (1,047 from 10,452 corpora), and newspapers (680 from 15,309 corpora). Related to the first-person singular pronoun *saya* primarily found in newspapers (14,629 from 15,309 corpora), magazines (9,405 from 10,542 corpora), short stories (9,014 from 37,932 corpora), and novels (6,055 from 83,494 corpora). The followings are the samples of the corpus containing first-person singular pronoun forms *saya* and *aku* in Bahasa Indonesia, which were taken from newspapers, magazines, short stories, and novels.

(i) The first-person singular pronoun forms *saya*

- (5) *“Dalam sehari, saya mengisi acara di tiga tempat yang berbeda. Saya menjelaskan siapa Jokowi dan kenapa orang NU harus memilih dia,” kata dia, ketika ditemui Tempo di pondok pesantren yang ia pimpin, Selasa pekan lalu.*

‘In a day, I perform in three different occasions. I explained who Jokowi is and why NU people should vote for him,” he said, when met by Tempo at the islamic boarding school he leads on Tuesday last week.’

- (6) *“Bolak-balik saya ke Aceh atau mereka datang ke Jakarta,” kata Direktur Jenderal Otonomi Daerah Kementerian Dalam Negeri itu Kamis pekan lalu.*

‘I go back and forth or they visit to Jakarta, said Director General of Regional Autonomy of the Ministry of Internal Affairs last Thursday.’

- (7) *Berkali-kali saya mencoba untuk membaca. Tapi hasilnya selalu mengecewakan, setiap dicoba, bayang-bayang gelap itu menghalang pandang. Seperti menutup harakat dan ayat-ayat. Bayang-bayang itu selalu muncul setiap kali saya membaca, Ustaz.*

‘I have tried to read it many times. However, the result is always disappointing, every time it is tried, the dark shadows obstruct the view. Like closing the harakat and verses. Those

shadows always appear every time I read, Ustaz.’

- (8) *Terseok-seok, Danny berhasil pulang ke rumah. Dua hari ia di rumah, orang tuanya memberi tahu ada aparat membawa daftar 400 nama orang yang terlibat unjuk rasa berikut alamat rumah mereka. "Nama saya ada di dalam daftar itu," katanya. 'Shuffling around, Danny made it home. Two days after he was at home, his parent's told authorities were carrying a list of 400 names of people involved in the protests along with their home addresses. "My name is on that list," he said.'*

Data (5) illustrates the first-person singular pronoun, *saya* taken from a newspaper. The data (5) depicts that the speaker is in a formal situation; it is an interview. The speaker applies the first-person singular pronoun *saya* in that situation to show the formality.

The data (6) is obtained from a magazine. It describes the usage of the first-person singular pronoun form *saya*. The data (6) is also in an interview or a formal situation. The paragraph in data (6) portrays a formal conversation between the interviewer and the Director General of Regional Autonomy of the Ministry of Internal Affairs as the interviewee in a formal situation. It represents that the interviewee chooses the pronoun *saya* rather than *aku* to show a formality.

Data (7) is a fragment of a conversation taken from a short story. The data also illustrates the usage of the first-person singular pronoun form *saya* in a formal situation since the pronoun *saya* is chosen by the narrator as a student to show his respect to his Ustaz ('his teacher').

Data (8) depicts the first-person singular pronoun form *saya* which was found in a conversation in a novel. The narrator preferred to use the pronoun *saya* rather than *aku* to show his respect since he spoke to someone older than him.

- (ii) The first-person singular pronoun forms *aku*

- (9) *Sore harinya, aku mengunjungi A Ling. Kupakai baju terbaikku. A Ling tampak sangat anggun dibalut chong kiun berwarna biru laut, pakaian kebangsaannya khusus untuk hari istimewa.*

'In the afternoon, I visited A Ling. I wear my best clothes. A Ling looked very elegant wrapped in a navy blue chong kiun, her national dress was special for a special day.'

- (10) *Aku merasa tak nyaman. Mungkin karena aku tak terbiasa dengan suasana seperti ini. Bertemu dengan ratusan gadis cantik dan kaya. Pakaian dan kerudung mereka serasi dan elok dilihat.*

'I felt uncomfortable. Maybe it's because I'm not used to this kind of atmosphere. Met hundreds of beautiful and rich girls. Their clothes and veils are harmonious and beautifully seen.'

- (11) *Aku terpana. Begitu megahnya bangunan Masjid Agung sekarang. Lebih megah ketika dua puluh tahun lalu aku kemari bersama Ayah.*

'I was stunned. So magnificent is the building of the Great Mosque now. It was grander when twenty years ago I came here with Dad.'

- (12) *Namaku Pardanem, bukan nama asli. Pekerjaan tukang sortir kopi luak di lereng Batukaru yang dingin. Kemiskinan menyebabkan aku tak bisa meneruskan sekolah. Syukur aku lancar berbahasa Melayu dan badan gempal, aku sudah dua kali menjadi saksi partai politik saat pemilu.*

'My name is Pardanem, not my real name. The work of a badger coffee sorter on the cold slopes of Batukaru. Poverty caused me to be unable to continue my studies. Thankfully I speak Malay fluent and have been a witness twice by political parties during elections.'

The data (9) was captured from a novel while the data (10) was acquired from a short story. The two data describe the usage of the word *aku* in novels and short stories. Related to the meaning, the pronoun *aku* has a similar meaning to I in English. In Bahasa Indonesia, the singular pronoun for

the first-person form, *aku* is commonly employed in informal situations to show an intimate relationship among the interlocutors. The data (9) describes an informal situation; it is described that the speaker and the listener have an intimate relationship since the speaker chooses to use the word *aku* to refer to the pronoun I in English. It is displayed that the data illustrates an informal situation that was taken from a conversation in a novel. It also happens to the data (10); the speaker has spoken to the listener in an informal situation, and the speaker has chosen to use the pronoun *aku* refers to I in English. The data (10) is obtained from a conversation as a fragment of a short story.

Data (11) was fragmented from a story of a narrator about Mesjid Agung in a magazine. The narrator has chosen the first-

person singular pronoun form *aku* to describe an intimate relationship between the narrator and the readers/ listeners. Data (12) was adopted from a paragraph of a story published in a newspaper. The narrator chose to use the pronoun the first-person singular pronoun form *aku* to show a close relationship between the narrator and the readers.

The Distribution of the Pronoun Kita and Kami

In the previous explanation, it is introduced that the first-person pronouns could be singular and plural forms. The following figure describes the plural pronoun for the first-person forms *kami* and *kita* in Bahasa Indonesia found in the four sources (newspapers, magazines, short stories, and novels) described in Figure 2.

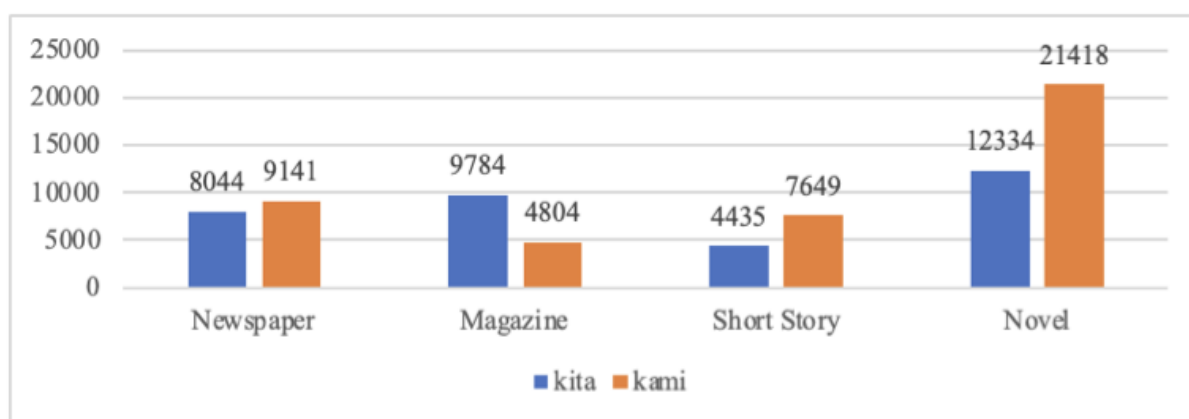


Figure 2
Kita and Kami as First-person Plural Pronoun

Figure 2 describes that the plural pronoun for the first-person *kami* sequentially found in novels, (21,418 of 33,752 corpora), short stories (7,649 of 12,084 corpora), newspapers (9,141 of 17,185 corpora), and magazines (4,804 of 14588 corpora). The data (5) – (12) capture the plural pronoun forms for the first-person singular, while the followings (13) – (20) are the data of the first-person plural pronoun forms in Bahasa Indonesia.

- (i) The first-person plural pronoun forms *kita*

- (13) *Pertamina mempunyai hak untuk mengevaluasi pada bulan ketiga soal kelanjutan kontrak dengan Sonangol. "Kalau happy kita lanjut, kalau tidak happy, ya kita stop."*

‘Pertamina has the right to evaluate in the third month the continuation of the contract with Sonangol. "If we are happy, if we are not happy, then we stop."’

- (14) *Kalau tidak dilindungi (keragaman hayati), kita hanya bisa mengenang mereka sebagai nama jalan, nama toko, atau merek sebuah barang.*

‘If they (biodiversity) were not protected, we (including the listeners or readers) just

can only remember them as the name of streets, stores, or brand of the products.’

- (15) *Kukernyitkan kening, “Sekar, bukannya Sabtu ini kita ada ujian di Sanggar?” Sekar tampak kaget, dia lupa bahwa anak-anak sanggar harus menghadapi ujian, dan kami adalah tim penilainya. “iya...” Keluar semacam bunyi-bunyian kecil dari mulut Sekar.*

‘Kukernyit frowned, "Sekar, isn't it this Saturday that we have an exam in sanggar?" Sekar looked shocked, he forgot that the children of the studio had to face the exam, and we were his assessment team. "yes..." There came a kind of small noises from Sekar's mouth.’

- (16) *Alfa mengambil kesempatan dan buru-buru menarik Kell ke luar. “Kecuali kalau kita ada di film The Flintstones, bagaimana bisa benda itu proyektor?” bisik Alfa sambil menunjuk batu-batu yang dipelototi Bodhi.*

‘Alpha seized the opportunity and hurriedly pulled Kell outside. "Unless we're in the movie The Flintstones, how can that thing be a projector?" the Alpha whispered, pointing to the stones Bodhi glared at.’

Data (13) illustrates that the first-person plural pronoun form *kita*, taken from newspapers, revealed twice. The data (13) describes both first-person plural pronouns *kita* refer to “we” in English. The pronoun *kita* in this context refers not only to the speaker who represents Pertamina but also to Senagol as the contractor. It is implied that if both sides (Pertamina and Senagol) are happy, they will continue the contract and if they are not happy, they will stop the contract after having the third month’s evaluation.

Data (14) portrays a part of a conversation captured from a magazine. In data (14), the speaker chooses the plural pronoun form for the first-person *kita*, to include the listeners or readers to be involved in the conversation. The speaker let the listeners and the readers be active in protecting biodiversity.

Data (15) depicts a conversation in a short story. The plural pronoun for the first-person *kita* refers to the speaker and the listener (Sekar) who have an exam in Sanggar. In this context, the speaker uses the pronoun *kita* rather than *kami*.

Data (16) is adopted from a fragment of conversation in a novel. It is depicted that the plural pronoun for the first-person *kita* that relates to the speaker, Kell, represents the listener. So, the speaker chooses *kita* rather than *kami*, showing that the listener is included in the situation.

- (ii) The first-person plural pronoun forms *kami*

- (17) *Bulan September nanti kami akan ada evaluasi dan laporan pelaksanaan pileg dan pilpres kemarin, kata Ferry.*

‘In the next September, we (excluding the listener) will evaluate and report the house of representative and president last elections, Ferry said.’

- (18) *“Mana kuncinya?” katanya lagi. Aku menyerahkan kunci yang diminta dan dia membuka pintu rumah. Kemudian kami masuk bersama-sama ke dalam rumah.*

“Where's the key?” he said again. I handed over the requested key and he opened the door of the house. Then we went in together into the house.’

- (19) *Kami berkumpul di Stasiun Kota Baru, Malang. Teman saya mengajak tiga rekannya. Tepat pukul 07.00 WIB, kami bersepeda motor menuju arah Kecamatan Sumbermanjing Wetan, Kabupaten Malang.*

‘We met at Kota Baru Station, Malang. We gathered at Kota Baru Station, Malang. My friend took three of his colleagues. At exactly 07.00 WIB, we rode a motorbike towards the direction of Sumbermanjing Wetan District, Malang Regency.’

(20) *Ini kasus sensitif, kami mesti hati-hati," katanya. Diam-diam polisi sebenarnya bergerak mengumpulkan foto yang mereka perlukan. Antara lain, mencari foto siswa dan pegawai JIS ke sejumlah studio foto langganan sekolah tersebut.*

‘This is a sensitive case, we must be careful, he said. Secretly the police actually moved to collect the photos they needed. Among other things, looking for photos of JIS students and employees to several photo studios subscribed to the school.’

Data (17) as a fragment of a conversation in a newspaper illustrates that the first-person plural pronoun form *kami* refers to “we” (excluding the speaker). The utterance in data (17) describes only the speaker (Ferry and team) who will evaluate and report the House of Representatives and President’s last elections (excluding the listeners or the readers). Data (18) as a data which was taken from and (19) narrate the usage of the first-person plural pronoun form *kami* that refers to “we” (excluding the listeners/ readers). Data (18) was seized from a short story while the data (19) was occupied from a novel. Both the first-person plural pronoun *kami* in (18) and (19) refer to the “we” in English but the plural pronoun for the first-person *kami* in that context excludes the listeners or the readers; it just only expresses the speakers. Data (20) is a fragment from a conversation in a magazine. The first-person plural pronoun form *kami* excludes the listeners or the readers. *Kami* refers to the pronoun (the

narrator representing the Indonesian Child Protection Commission (KPAI)).

From the four sources of the corpus, the author try to decipher the 224,796 tokens found related to the percentage of the pronoun usage, as explained in Figure 3. Figure 3 outlines that the singular pronoun for the first-person form, *aku*, has the highest number among the four types of the first-person pronoun. It reaches 108,084 times or 48%, in Bahasa Indonesia found in the 224,796 tokens from the four sources (newspapers, magazines, short stories, and novels). The second position in the first-person pronoun is *kami*. This pronoun achieves 43,102 times (19%) from 224,796 tokens, while the next, as the third position, is the pronoun *saya*. This pronoun attains 39,103 times from 224,796 tokens or 17% of the four pronouns. The last position is the plural pronoun for the first-person form *kita*, which has 34,597 times or 15% of the total tokens.

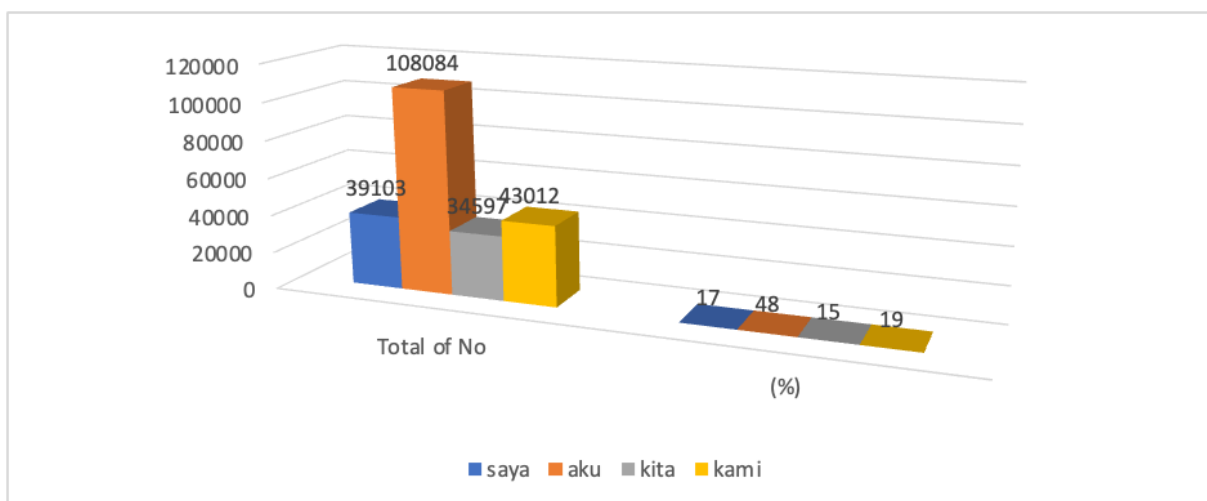


Figure 3

Saya, Aku, Kita, and Kami in the Sources of Corpus

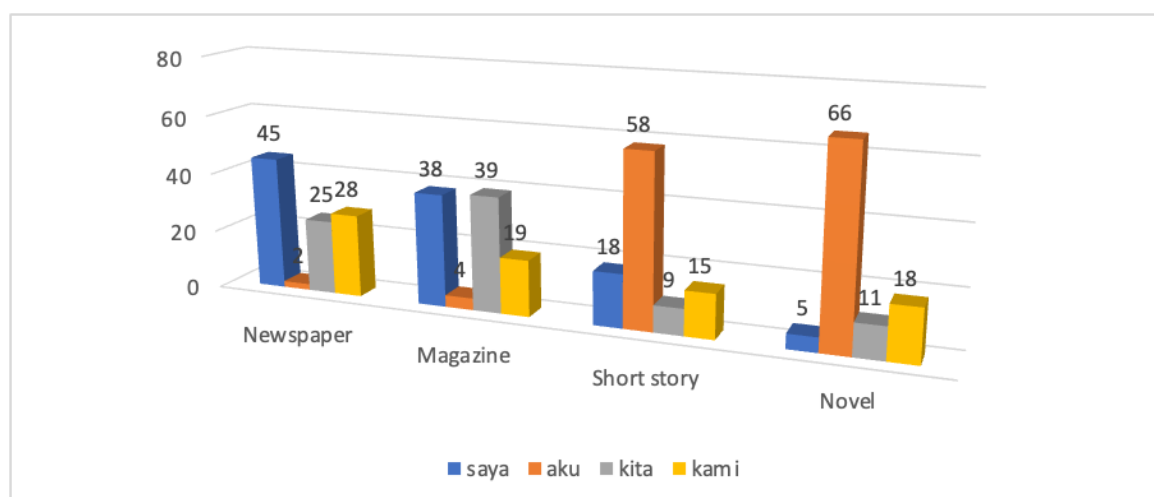


Figure 4
First-Person Pronouns: Singulars and Plurals

Figure 4 compares the pronouns *saya*, *aku*, *kita*, and *kami*. In the newspaper, sequentially, the dominant pronouns are *saya* (45%), *kami* (28%), *kita* (25%), and *aku* (2%). Related to the source, in the newspapers, the first-person pronoun *saya* is more frequently used than *aku* since the first-person pronoun *saya* is more formal than *aku*. In the magazine, the dominant pronouns found are *kita* (39%), *saya* (38%), *kami* (19%), and *aku* (4%). The dominant pronoun applied in newspapers is *saya* as a singular, and the dominant pronoun in magazines is *kita* as a plural form. It could be assumed that the writers or the speakers employ the first-person pronoun *kita* in magazines to involve the readers' or listeners' roles in their conversation.

Figure 4 also identifies that the singular pronoun for the first-person *aku* is the dominant pronoun found in short stories and novels. In novels, it was reported that *aku* (66%), *kami* (18%), *kita* (11%), and *saya* (5%) while in the short stories, *aku* (58%), *saya* (18%), *kami* (15%), and *kita* (9%). It implied that the singular pronoun for the first-person *aku* is more frequently used in novels and short stories to show intimate relationships in informal situations than the singular pronoun for the first-person form *saya*.

CONCLUSION

This research focused on describing the highest number of both singular and plural forms in Bahasa Indonesia, the number of the first-person singular pronouns form *aku* and *saya*, and the number of the plural pronouns for first-person forms *kami* and *kita*. Related to the data taken from four corpus sources (newspapers, magazines, short stories, novels), it was stated that the singular pronoun for the first-person form is more frequently used or the number of singular is higher than the plural form. Secondly, the number of the singular pronoun form for the first-person *aku* is higher than *saya*, and the pronoun *aku* is commonly applied in novels and short stories. The speaker chose singular pronoun for the first-person form *aku* is to show an intimate relationship between the interlocutors. Thirdly, the number of the plural pronouns for the first-person form *kami* is higher than *kita* found in the four sources of corpus. Based on the corpus it could be believed that the writers or the speakers retain the plural pronoun for the first-person *kita* in magazines to engage the readers' or listeners' roles in the conversation.

As stated earlier, there are the first, second, and third person pronouns in

Bahasa Indonesia. Since this study is limited to the first person pronouns, it is suggested for the further research to investigate the second and the third first

person pronouns in Bahasa Indonesia based on the corpus.

REFERENCES

- Alwi, H., Dardjowidjojo, S., Lapoliwa, H., & Moeliono, A. M. (2003). *Tata Bahasa Baku Bahasa Indonesia*. Balai Pustaka.
- Anthony, L. (2019). *AntConc 3.5.8 (MACintosh OS X)*.
- Azar, A. S., & Hashim, A. (2022). Analysing Authorial Identity Construction in the Review Article Genre in Applied Linguistics. *Studies in English Language and Education*, 9(1), 94–114. <https://doi.org/10.24815/siele.v9i1.21898>
- Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches*. SAGE Publications Ltd.
- Djenar, D. N. (2001). Indonesian “Locative” Pronouns: Deictic or Anaphoric? *Australian Journal of Linguistics*, 21(1), 49–71. <https://doi.org/10.1080/07268600124060>
- Djenar, D. N. (2008). Which self? Pronominal choice, modernity, and self-categorizations. *International Journal of Sociology of Language*, 189, 31–54. <https://doi.org/10.1515/IJSL.2008.002>
- Evans, G. (1980). Pronouns. *Linguistic Inquiry*, 11(2), 337–362.
- Firdaus, S. F., Soemantri, Y. S., & Yuliawati, S. (2021). A corpus-based study of self-mention markers in English research articles. *International Journal of Language and Teaching Education*, 37–46. <https://doi.org/10.22437/ijolte.v5i2.15695>
- Flowerdew, J., & Wang, S. H. (2015). Identity in Academic Discourse. *Annual Review of Applied Linguistics*, 35, 81–99. <https://doi.org/10.1017/S026719051400021X>
- Fricke, H. (2017). Nouns and pronouns in Central Lembata Lamaholot (Austronesian, Indonesia). *Wacana*, 18(3), 746–771. <https://doi.org/10.17510/wacana.v18i3.635>
- Hardjanto, T. D. (2022). “However, we argue that ...”: The Construction of Authorial Identities in English Research Articles. *Theory and Practice in Language Studies*, 12(8), 1525–1536. <https://doi.org/10.17507/tpls.1208.08>
- He, F. (2020). Identity Construction in Academic Writing of Student Writers Who Use English as an Additional Language: A Literature Review. *Chinese Journal of Applied Linguistics*, 43(4), 506–524. <https://doi.org/10.1515/CJAL-2020-0033>
- Hyland, K. (2001). Humble servants of the discipline? Self-mention in research articles. *English for Specific Purposes*, 20, 207–226. [https://doi.org/10.1016/S0889-4906\(00\)00012-0](https://doi.org/10.1016/S0889-4906(00)00012-0)
- Hyland, K. (2002a). Authority and invisibility: *Journal of Pragmatics*, 34, 1091–1112. [https://doi.org/10.1016/S0378-2166\(02\)00035-8](https://doi.org/10.1016/S0378-2166(02)00035-8)
- Hyland, K. (2002b). Options of identity in academic writing. *ELT Journal*, 56(4), 351–358. <https://doi.org/10.1093/elt/56.4.351>
- Işık-Taş, E. E. (2018). Authorial identity in Turkish language and English language research articles in Sociology: The role of publication context in academic writers’ discourse choices. *English for Specific Purposes*, 49, 26–38. <https://doi.org/10.1016/j.esp.2017.10.003>
- Jones, C., & Waller, D. (2015). *Corpus Linguistics for Grammar*. Routledge. <https://doi.org/10.4324/9781315713779>
- Li, Z. (2021). Authorial presence in research article abstracts: A diachronic investigation of the use of first person pronouns. *Journal of English for Academic Purposes*, 51, 100977. <https://doi.org/10.1016/j.jeap.2021.100977>
- Manns, H. (2012). First-Person Pronominal Variation, Stance and Identity in Indonesia. *Australian Journal of Linguistics*, 32(4), 435–456. <https://doi.org/10.1080/07268602.2012.744265>

- McEnery, T., & Hardie, A. (2012). *Corpus Linguistics*. Cambridge University Press. <https://doi.org/10.1093/oxfordhb/9780199276349.013.0024>
- Moeliono, A. M., Lapoliwa, H., Hasan Alwi, H., Sasangka, S., & Sugiyono. (2017). *Tata Bahasa Buku Bahasa Indonesia* (Keempat). Kementerian Pendidikan dan Kebudayaan.
- Muryasov, R. Z., Zheltukhina, M. R., & Sergeeva, O. V. (2024). Personal/impersonal and subject/null subject phenomena in the multi-structural languages. *XLinguae*, 17(2), 54–66. <https://doi.org/10.18355/XL.2024.17.02.04>
- Ratu, D. M. (2019). Mongondow Pronoun: The Teaching of Indonesian Local Language in the Context of National Language. *International Journal of Recent Technology and Engineering (IJRTE)*, 7(6S5), 1091–1098.
- Scott, M., & Tribble, C. (Eds.). (2006). *Textual Patterns Key words and corpus analysis in language education*. John Benjamins Publishing Company. <https://doi.org/10.1075/sci.22>
- Sudartini, S. (2009). The Question of Grammatical Equivalence in Translation. *Journal of English and Education (JEE)*, 96–108. <https://doi.org/10.20885/jee.v3i1.6485>
- Sujatna, E. T. S., & Sugianto, L. P. M. (2021). Picturing Deictic Time Words of Bahasa Indonesia Corpora Collection Leipzig University. *Lexicography and Language Documentation*, 185–194.
- Sujatna, E. T. S., Sujatna, M. L., Sugianto, L. P. M., & Mulyono, P. M. S. (2022). Picturing “Coronavirus” in News on the Web (NOW) Coprus: A Corpus Linguistic Study. *Sosiohumaniora: Jurnal Ilmu-Ilmu Sosial dan Humaniora*, 24(1), 114–121. <https://doi.org/10.24198/sosiohumaniora.v24i1.34926>
- Tang, R., & John, S. (1999). The “I” in identity: Exploring writer identity in student academic writing through the first person pronoun. *English for Specific Purposes*, 18(SUPPL. 1). [https://doi.org/10.1016/S0889-4906\(99\)00009-5](https://doi.org/10.1016/S0889-4906(99)00009-5)
- Viana, V., Zyngier, S., & Barnbrook, G. (Eds.). (2011). *Perspectives on Corpus Linguistics*. John Benjamins Publishing Company. <https://doi.org/10.1075/sci.48>
- Walková, M. (2018). Author’s self-representation in research articles by anglophone and Slovak linguists. *Discourse and Interaction*, 11(1), 86–105. <https://doi.org/10.5817/DI2018-1-86>
- Wehrli, A., & Awal, N. M. (2014). Second Person Reference in Indonesian Christian Prayer: A Sociolinguistic Analysis. *Journal of Language Studies*, 14(3), 175–189. <https://doi.org/10.17576/GEMA-2014-1403-11>
- Wijayanti, L. T., Sholikhati, N. I., & Verrysaputro, E. A. (2022). Authorial Pronoun We: Capturing Politeness in EFL Academic Writing. *J-Lalite: Journal of English Studies*, 3(1), 76–84. <https://doi.org/10.20884/1.jes.2022.3.1.5934>
- Womack, P. (2022). Pronouns: we/us/our. *Critical Quarterly*, 64(2), 93–99. <https://doi.org/10.1111/criq.12655>
- Yu, C. (2021). Authorial Identity Behind First Person Pronouns in English Research Articles of Native and Non-Native Scholars Winter and Spring. *Journal of New Advances in English Language Teaching and Applied Linguistics*, 3(1), 462–477. <https://doi.org/10.22034/JELTAL.2021.129414>
- Yuliawati, S., Dienaputra, R. D., Sujatna, E. T. S., Suryadimulya, A. S., & Lukman, F. (2019). Looking into “Awewe” and “lalaki” in the Sundanese Magazine Mangle: Local Wisdom and a Corpus Analysis of the Linguistic Construction of Gender. *International Journal of Advanced Science and Technology*, 28(8s), 549–559.
- Yusuf, Y. Q., Nasir, C., & Andib, N. (2019). Power and solidarity: The pronoun of address ke [ke] used in Indonesian by Acehese speakers. *International Journal of Language Studies*, 13(1), 77–98.

ACKNOWLEDGEMENT

Big gratitude is conveyed to the Rector of Universitas Padjadjaran, Indonesia who has facilitated the process of the research until this paper is published.

