



## Derivasi Verba dalam Bahasa Lampung melalui Pendekatan Linguistik Komputasional

Dewi Ratnaningsih<sup>1\*</sup>, Ifan Iskandar<sup>2</sup>, Miftahulhairah Anwar<sup>3</sup>

<sup>1</sup>Universitas Muhammadiyah Kotabumi, Kotabumi, Lampung, Indonesia, 34517

<sup>2,3</sup>Universitas Negeri Jakarta, Jakarta Timur, Daerah Khusus Jakarta, Indonesia, 13320

<sup>1</sup> [dewi.ratnaningsih@umko.ac.id](mailto:dewi.ratnaningsih@umko.ac.id), <sup>2</sup> [ifaniskandar@unj.ac.id](mailto:ifaniskandar@unj.ac.id), <sup>3</sup> [miftahulhairah@unj.ac.id](mailto:miftahulhairah@unj.ac.id)

\*Corresponding Author: [dewi.ratnaningsih@umko.ac.id](mailto:dewi.ratnaningsih@umko.ac.id)

### Abstrak

*This study analyzes verbal derivation in Lampung Language Dialects A and O using a computational linguistic approach to identify affixation patterns, morphophonemic processes, and the semantic and syntactic roles of affixes. Verbal data were classified into prefixes, infixes, suffixes, and circumfixes. The findings reveal that both dialects share highly consistent derivational systems. The nasal prefix (N-) is the most productive, marking active transitivity and causativity, with parallel nasal assimilation patterns in both dialects (b/p → m-, d/t → n-, k/g → ng-, s/c/j → ny-, vowels → nge-/ng-). The prefix di- marks passivization, while be-/ber- and te-/ti- indicate reciprocal, stative, or unintentional passive meanings. Major suffixes (-ken, -ko, -an, -kon) function as causative, applicative, or result markers, whereas infixes (-el-, -em-, -en-, -er-) mainly affect phonological structure or add semantic nuance. Dialect comparison confirms strong structural similarity with minor phonetic variation, such as object suffixes (-ni/-na in Dialect A vs. -no in Dialect O) and limited use of bu- (Dialect A) and per- (Dialect O). The formalization of derivation rules in Python successfully predicted most derived forms, although refinement is required for vowel-initial nasal assimilation and vowel deletion in suffixation. Overall, the study demonstrates the fundamental similarity of Lampung verbal morphology and supports computational applications for regional language preservation, while contributing to broader Austronesian morphological typology.*

### Status Artikel:

Diterima: 25-11-2025

Direvisi: 26-11-2025

Diterima: 30-12-2025

### Kata Kunci:

Computational Linguistics, Derivation, Lampung Language, Verbs.



© 2025 Dewi Ratnaningsih, Ifan Iskandar, Miftahulhairah Anwar

This work is licensed under a

[Creative Commons Attribution-ShareAlike 4.0 International License](https://creativecommons.org/licenses/by-sa/4.0/).

## PENDAHULUAN

Bahasa Lampung, sebagai salah satu bahasa daerah di Indonesia yang kaya akan warisan budaya, menunjukkan kompleksitas morfologis yang signifikan, terutama dalam sistem derivasi verbanya. Kompleksitas ini dapat dijelaskan melalui teori morfologi generatif yang menyoroti peran aturan produktif dalam pembentukan kata baru dan ekspansi makna. Dalam kerangka morfologi derivatif, verba terbentuk melalui proses afiksasi yang tidak hanya memperluas leksikon, tetapi juga memungkinkan penyesuaian makna gramatikal seperti perubahan transitivitas, valensi, dan aspek. Hal ini sejalan dengan pandangan Afrianto et al yang menegaskan bahwa morfologi berperan penting dalam dinamika struktur kata dan fleksibilitas sintaksis bahasa (Afrianto et al., 2020), (Yan, 2021), (Nijs et al., 2025). Oleh sebab itu,

memahami mekanisme derivasi verba dalam Bahasa Lampung menjadi sangat esensial, tidak hanya untuk kajian linguistik teoretis, tetapi juga dalam upaya pelestarian bahasa dan pengembangan teknologi bahasa, khususnya pada era digital saat ini (Ajani et al., 2024). Studi menunjukkan bahwa instruksi morfologi berbasis teknologi dapat meningkatkan literasi dan pemahaman morfologis secara signifikan (Bratlie et al., 2022), (Colenbrander et al., 2024), (Wood et al., 2025).

Bahasa Lampung mengenal dua tipe utama verba: verba infleksional yang diambil dari kata dasar verba dan verba derivasional yang dihasilkan dari dasar non-verba seperti nomina, adjektiva, adverbial, pronomina, numeralia. Proses derivasi melibatkan penambahan afiks (prefiks, sufiks, infiks, konfiks) yang tidak hanya memperluas leksikon, tetapi juga mengubah makna gramatikal seperti transitivitas, valensi, dan aspek (Abidin et al., 2021). Proses derivasi melibatkan penambahan afiks (prefiks, sufiks, infiks, konfiks) yang tidak hanya memperluas leksikon, tetapi juga mengubah makna gramatikal seperti transitivitas, valensi, dan aspek. Penambahan prefiks serta sufiks pada verba dasar dapat mengubah transitivitas (dari intransitif menjadi transitif atau sebaliknya) dan menambah makna kausatif, resiprokal, atau aspek tertentu (Nugraha, 2024b). Infiks dan konfiks juga berperan dalam membentuk verba baru dengan perubahan kelas kata dan fungsi gramatikal, misalnya mengubah nomina/adjektiva menjadi verba, atau menambah argumen dalam struktur kalimat (Agus et al., 2024a).

Linguistik komputasional, sebagai cabang interdisipliner yang memadukan linguistik dan ilmu komputer, menawarkan pendekatan yang kuat untuk menganalisis dan memodelkan fenomena bahasa secara sistematis (Solovyev et al., 2022), (Bernard & Winterstein, 2022), (Alduais et al., 2025). Pendekatan ini memungkinkan peneliti tidak hanya mendeskripsikan aturan derivasi secara eksplisit, tetapi juga memvalidasi serta mengidentifikasi variasi antar-dialek dengan presisi tinggi melalui pemrosesan data digital. Dengan metode komputasi, analisis dapat dilakukan secara kuantitatif pada korpus besar, sehingga memperkuat validitas hasil dan membuka peluang aplikasi praktis seperti pengembangan alat penerjemah otomatis, sistem pembelajaran bahasa, dan dokumentasi digital yang mendukung pelestarian Bahasa Lampung (Luz, 2022). Model komputasional telah digunakan untuk mengidentifikasi dan memetakan variasi dialek secara otomatis, bahkan pada data besar seperti media sosial dan korpus digital (Strycharczuk et al., 2020).

Fokus penelitian pada analisis derivasi verba dalam dua dialek utama Bahasa Lampung, yakni Dialek A dan Dialek O, memperkaya kajian tipologi morfologi Austronesia. Berdasarkan studi dari berbagai dialek (Sasak, karo, Selulung, dan Kinamano), perbedaan fonologis dan leksikal antar dialek seringkali berdampak pada variasi morfologis, termasuk dalam proses afiksasi dan morfofonemik (Syarifurrahman et al., 2021; Wijayanti et al., 2025). Analisis komputasional yang diterapkan dalam penelitian ini memungkinkan pemetaan pola afiksasi (prefiks, infiks, sufiks, dan konfiks), identifikasi proses morfofonemik yang mendasari pembentukan verba, serta analisis kontribusi semantik dan fungsi sintaksis dari masing-masing afiks. Pendekatan ini juga mendukung prinsip modularitas dalam linguistik, di mana setiap komponen morfologi dapat dianalisis secara terpisah dan kemudian disintesis untuk memahami sistem secara holistik (Ahda et al., 2025; Nugraha, 2024a), (Sudana et al., 2024), (Agus et al., 2024b).

Selain itu, penelitian ini mengintegrasikan prinsip-prinsip analisis linguistik modern dengan metode komputasional untuk mengoptimalkan akurasi dalam identifikasi dan

klasifikasi proses afiksasi. Pemanfaatan data digital memungkinkan peneliti untuk melakukan pelacakan perubahan morfofonemik secara lebih detail, termasuk fenomena asimilasi dan penghilangan vokal yang kerap terjadi dalam proses derivasi verba Lampung (Amrulloh & Humairo, 2020). Dengan pendekatan ini, analisis dapat dilakukan secara longitudinal, sehingga pola-pola evolusi morfologi dapat terdeteksi dari waktu ke waktu dan mendukung pelestarian bahasa secara berkelanjutan. Penelitian ini juga menyoroti pentingnya kolaborasi lintas disiplin, di mana ahli bahasa, pengembang perangkat lunak, serta komunitas lokal dapat saling mendukung dalam proses dokumentasi dan revitalisasi Bahasa Lampung (Miaschi et al., 2021; Patria & Merdeka, 2023).

Tidak banyak penelitian terkini yang secara khusus membahas derivasi verba. Penelitian (Astawan & Sadwika, 2020) mendeskripsikan afiks derivasional verba Bahasa Lampung Dialek Abung (BLDA) menggunakan teori morfologi generatif. Penelitian (Septianingtiyas et al., 2025) menganalisis secara leksikal dan fonologis variasi dialektal kosakata 'kelingking' dalam Bahasa Lampung di Kabupaten Pringsewu, Provinsi Lampung. Namun, penelitian-penelitian tersebut masih terbatas pada pembahasan afiksasi di satu dialek atau pada satu kosakata spesifik, sehingga belum mengulas secara menyeluruh sistem derivasi verba lintas dialek utama Bahasa Lampung. Selain itu, belum banyak kajian yang mengintegrasikan pendekatan linguistik komputasional untuk memetakan pola morfofonemik dan kontribusi sintaksis afiks secara sistematis. Dengan demikian, masih terdapat kekosongan dalam analisis komparatif dan pemodelan digital derivasi verba yang dapat mendukung pelestarian dan pengembangan teknologi bahasa daerah.

Dengan membandingkan temuan antara kedua dialek, penelitian ini diharapkan dapat memberikan kontribusi pada pemahaman tentang konsistensi dan divergensi dalam sistem derivasi verba Bahasa Lampung. Teori variasi dialektal (Yue, n.d.) menegaskan pentingnya analisis komparatif untuk mengidentifikasi pola universal dan lokal dalam bahasa. Pendekatan ini menggunakan analisis statistik dan pemodelan matematis untuk mengukur jarak linguistik antar dialek, menggabungkan faktor geografis dan sosial untuk memprediksi variasi pengucapan dan leksikal (Huisman et al., 2021; Wieling et al., 2011). Formalisasi aturan derivasi melalui pendekatan komputasional diharapkan menjadi fondasi bagi pengembangan aplikasi linguistik yang lebih canggih, seperti penganalisis morfologi dan sistem pembelajaran bahasa berbasis AI, sehingga mendukung revitalisasi dan dokumentasi Bahasa Lampung secara berkelanjutan (Wood et al., 2025).

Penelitian ini menawarkan kebaruan konseptual dan metodologis dalam kajian morfologi Bahasa Lampung. Secara konseptual, penelitian ini tidak hanya mendeskripsikan pola derivasi verba secara struktural, tetapi juga memetakan produktivitas dan distribusi afiks secara kuantitatif berbasis korpus leksikal skala besar. Secara metodologis, penelitian ini merupakan studi pertama yang mengintegrasikan analisis morfologi derivatif Bahasa Lampung dengan pendekatan linguistik komputasional berbasis rule-based yang terstruktur dan replikatif. Berbeda dengan penelitian sebelumnya yang umumnya bersifat deskriptif-kualitatif dan terbatas pada sampel kecil, penelitian ini memproses 30.255 entri leksikal dari berbagai sumber kamus dan memformulasikan model algoritmik yang mampu mengidentifikasi pola derivasi secara sistematis. Dengan demikian, penelitian ini berkontribusi pada penguatan dokumentasi digital bahasa daerah serta membuka peluang pengembangan perangkat analisis morfologi otomatis untuk Bahasa Lampung dan bahasa daerah lainnya di Indonesia. Dengan demikian,

penelitian ini tidak hanya memperkuat dasar teoretis kajian morfologi bahasa daerah, tetapi juga menawarkan solusi praktis untuk pelestarian dan pengembangan teknologi bahasa lokal di Indonesia. Berikut ini adalah *pseude-code* berbasis aturan yang digunakan penelitian.

```
CLASS MorfologiLampungEkstraktor:
    FUNCTION __init__():
        // Inisialisasi daftar prefiks bahasa
        Lampung (terurut dari terpanjang)
        awalan_lampung = OrderedDict([
            ('peN', 1), // peN- dengan peluluhan
            ('puN', 2), // puN- (hanya Dialek A)
            ('memper', 3), // memper-
            ('diper', 4), // diper-
            ('penge', 5), // penge-
            ('meng', 6), // meng-
            ('meny', 7), // meny-
            ('peng', 8), // peng-
            ('peny', 9), // peny-
            ('pem', 10), // pem-
            ('pen', 11), // pen-
            ('mem', 12), // mem-
            ('men', 13), // men-
            ('per', 14), // per-
            ('ter', 15), // ter-
            ('ber', 16), // ber-
            ('be', 17), // be- (Dialek A dan O)
            ('bu', 18), // bu- (hanya Dialek A)
            ('di', 19), // di-
            ('ti', 20), // ti- (hanya Dialek A)
            ('pe', 21), // pe-
            ('te', 22), // te-
            ('ke', 23), // ke-
            ('se', 24) // se-
        ])

        // Prefiks N- dengan alomorf
        awalan_nasal = OrderedDict([
            ('nge', 1), // nge-
            ('ng', 2), // ng-
            ('ny', 3), // ny-
            ('m', 4), // m-
            ('n', 5) // n-
        ])

        // Infiks (sisipan)
        sisipan = OrderedDict([
            ('el', 1), // -el-
            ('em', 2), // -em-
            ('en', 3), // -en-
            ('er', 4) // -er-
        ])

        // Akhiran umum (Dialek A dan O)
        akhiran_umum = OrderedDict([
            ('kon', 1), // -kon (Dialek O)
            ('ken', 2), // -ken (Dialek O)
            ('an', 3), // -an (Dialek A)
            ('on', 4), // -on (Dialek O)
            ('ko', 5) // -ko
        ])

        // Akhiran Dialek A
        akhiran_dialek_a = OrderedDict([
            ('ni', 1), // -ni
            ('ne', 2), // -ne
            ('na', 3) // -na
        ])

        // Akhiran Dialek O
        akhiran_dialek_o = OrderedDict([
            ('nou', 1), // -nou
            ('no', 2) // -no
        ])
    ])
```

```
// Konfiks (awalan + akhiran)
konfiks = [
    ('ke', 'an'), // ke-...-an
    ('ke', 'on'), // ke-...-on (Dialek O)
    ('se', 'an'), // se-...-an
    ('se', 'ni'), // se-...-ni (Dialek A)
    ('se', 'no'), // se-...-no (Dialek O)
    ('pe', 'an'), // pe-...-an
    ('pe', 'on'), // pe-...-on (Dialek O)
    ('te', 'an'), // te-...-an
    ('be', 'an'), // be-...-an
    ('di', 'an') // di-...-an
]
END FUNCTION
END CLASS
```

### Pseudo-code 1. Inisialisasi Ekstraktor

```
FUNCTION ekstrak_semua(kata, dialek='A'):
    awalan = ekstrak_awalan(kata, dialek)
    sisipan = ekstrak_sisipan(kata)
    akhiran = ekstrak_akhiran(kata, dialek)
    gabungan = ekstrak_gabungan(awalan,
    akhiran, dialek)

    RETURN {
        'awalan': awalan,
        'sisipan': sisipan,
        'akhiran': akhiran,
        'gabungan': gabungan
    }
END FUNCTION
```

### Pseudo-code 2. Ekstraksi Elemen Morfologi

```
FUNCTION proses_ekstraksi_morfologi_lampung():
    // Inisialisasi
    ekstraktor = NEW
    MorfologiLampungEkstraktor()

    // Load dataset
    df = READ_CSV('data/dataset-5-seleksi-
    verba.csv')

    // Tambahkan kolom baru untuk Dialek A
    df['lampung_A_awalan'] = None
    df['lampung_A_sisipan'] = None
    df['lampung_A_akhiran'] = None
    df['lampung_A_gabungan'] = None

    // Tambahkan kolom baru untuk Dialek O
    df['lampung_O_awalan'] = None
    df['lampung_O_sisipan'] = None
    df['lampung_O_akhiran'] = None
    df['lampung_O_gabungan'] = None

    // Inisialisasi counter statistik
    total_A_awalan = 0
    total_A_sisipan = 0
    total_A_akhiran = 0
    total_A_gabungan = 0
    total_O_awalan = 0
    total_O_sisipan = 0
    total_O_akhiran = 0
    total_O_gabungan = 0

    // Proses setiap baris
    FOR EACH row IN df:
        lampung_kata = row['lampung_kata']
        dialek = row['dialek']
```

```
// Skip jika kata kosong
IF lampung_kata is NULL OR empty THEN
  CONTINUE
END IF

// Ekstrak berdasarkan dialek
IF dialek == 'A' THEN
  hasil =
ekstraktor.ekstrak_semua(lampung_kata,
dialek='A')
  row['lampung_A_awalan'] =
hasil['awalan']
  row['lampung_A_sisipan'] =
hasil['sisipan']
  row['lampung_A_akhiran'] =
hasil['akhiran']
  row['lampung_A_gabungan'] =
hasil['gabungan']

  // Update statistik
  IF hasil['awalan'] THEN
total_A_awalan++
  IF hasil['sisipan'] THEN
total_A_sisipan++
  IF hasil['akhiran'] THEN
total_A_akhiran++
  IF hasil['gabungan'] THEN
total_A_gabungan++

  ELSE IF dialek == 'O' THEN
  hasil =
ekstraktor.ekstrak_semua(lampung_kata,
dialek='O')
  row['lampung_O_awalan'] =
hasil['awalan']
  row['lampung_O_sisipan'] =
hasil['sisipan']
  row['lampung_O_akhiran'] =
hasil['akhiran']
  row['lampung_O_gabungan'] =
hasil['gabungan']

  // Update statistik
  IF hasil['awalan'] THEN
total_O_awalan++
  IF hasil['sisipan'] THEN
total_O_sisipan++
  IF hasil['akhiran'] THEN
total_O_akhiran++
  IF hasil['gabungan'] THEN
total_O_gabungan++
  END IF
END FOR

// Simpan hasil
SAVE_CSV(df, 'data/dataset-5A-ekstraksi-
morfologi-lampung.csv')

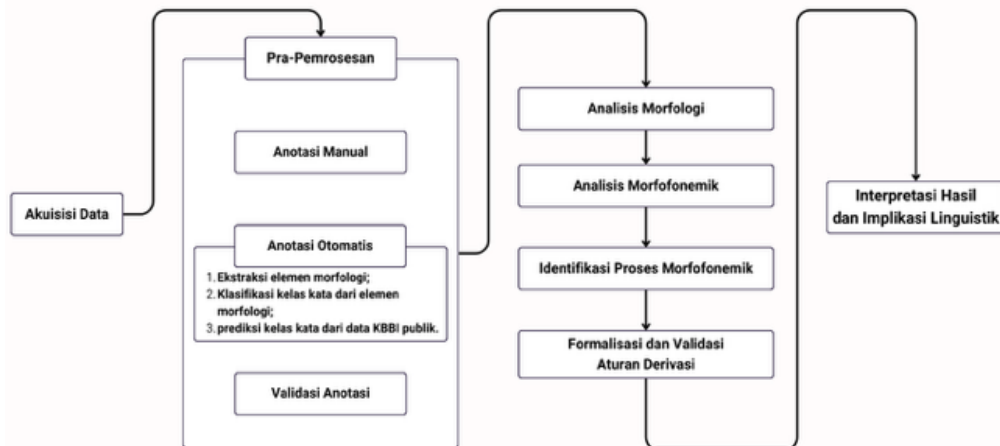
RETURN df
END FUNCTION
```

**Pseudo-code 13 Proses Utama Ekstraktor**

## **METODE PENELITIAN**

Penelitian ini mengadopsi pendekatan linguistik komputasional untuk menganalisis derivasi verba dalam Bahasa Lampung Dialek A dan Dialek O. Metodologi yang digunakan dalam penelitian ini terdiri dari enam tahapan, yaitu (1) akuisisi dan pra-pemrosesan data; (2) rekayasa fitur; (3) analisis morfologi; (4) identifikasi proses morfofonemik; (5) formalisasi dan validasi aturan derivasi; (6) Interpretasi hasil dan implikasi linguistik. Dalam kaitannya dengan linguistik komputasional, peneliti menggunakan Python dan Google Colab untuk melakukan

dua aktivitas utama yaitu (1) pra-pemrosesan data dan (2) analisis derivasi verba. Gambar 1 berikut ini adalah penjelasan masing-masing tahapan.



Gambar 1 Tahapan Implementasi Metode Penelitian

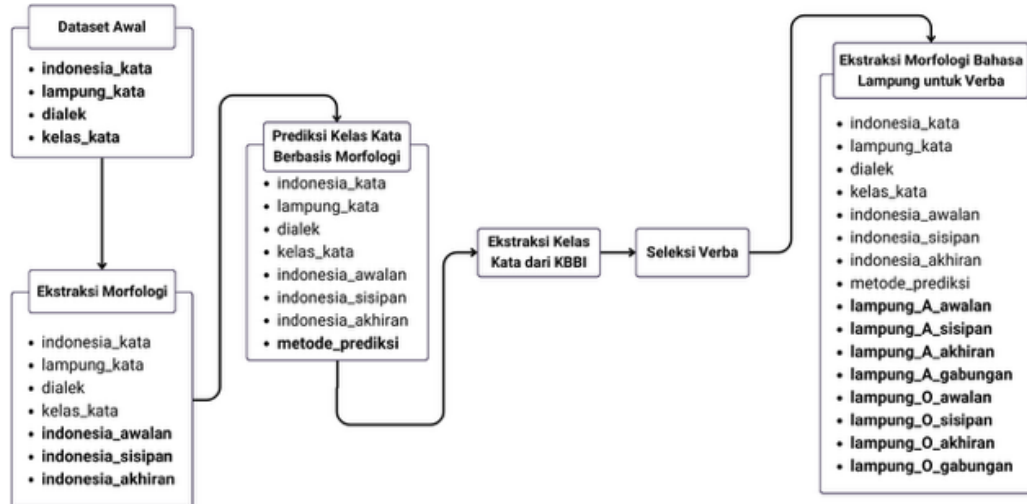
## 1. Sumber dan Akuisisi Data

Korpus bahasa Lampung diakuisisi dari empat kamus. Berdasarkan kamus-kamus tersebut, peneliti lalu menghimpun kata-kata dari dua dialek bahasa Lampung yaitu A dan O. Namun, tidak semua kamus mencantumkan keterangan kelas kata, sehingga mempersulit proses seleksi daftar kata verba. Berdasarkan permasalahan tersebut, peneliti melakukan dua teknik anotasi kelas kata sebelum dilakukan seleksi kata verba, yaitu (1) anotasi manual dan (2) anotasi otomatis berbasis linguistik komputasional. Anotasi manual dilakukan dengan cara memilih kelas tiap kata satu per satu secara manual. Sementara itu, anotasi berbasis linguistik komputasional dilakukan melalui tiga langkah yaitu (1) ekstraksi elemen morfologi; (2) klasifikasi kelas kata dari elemen morfologi; dan (3) klasifikasi kelas kata dari data KBBI. Ekstraksi elemen morfologi dilakukan untuk mengidentifikasi afiks, pola fonologis, struktur kata untuk mendapatkan ciri morfologis yang relevan. Klasifikasi kelas kata dari elemen morfologi dilakukan dengan menentukan kelas kata berdasarkan pola afiksasi atau ciri morfologis khas pada setiap bentuk kata. Sementara itu, klasifikasi kelas kata dari KBBI dilakukan dengan memanfaatkan padanan atau kemiripan bentuk Indonesia di KBBI untuk menebak kelas kata Lampung. Setelah semua mendapatkan kelas kata, peneliti melakukan validasi pada setiap kata teranotasi. Korpus verba lalu diseleksi dan dilakukan pemrosesan morfologi Bahasa Lampung.

## 2. Pra-pemrosesan dan Eksplorasi Data

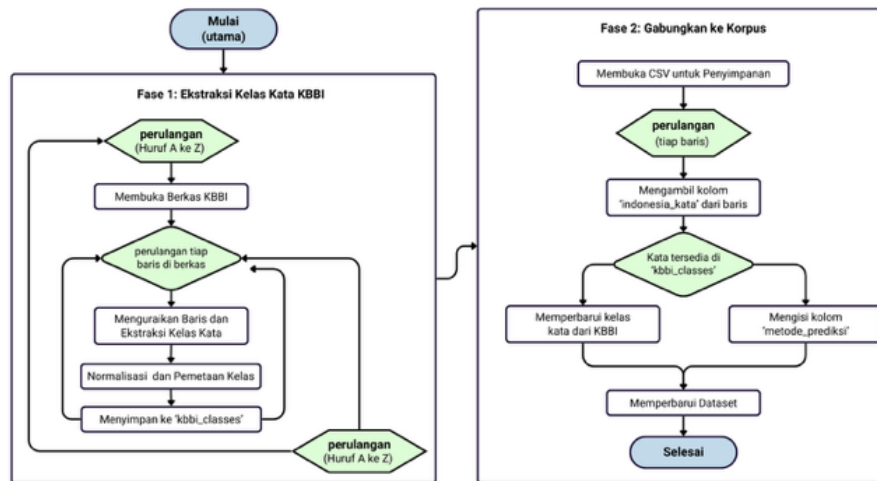
Transformasi data dilakukan melalui serangkaian tahapan pra-pemrosesan yang sistematis untuk menghasilkan korpus verba yang telah dianotasi. Setiap tahap menggunakan skrip khusus dan menghasilkan keluaran terstruktur yang menjadi masukan bagi tahap berikutnya. Pada tahap pertama, dataset awal terdiri atas 30.256 entri dengan delapan kolom utama, yakni id, indonesia\_kata, lampung\_kata, aksara, dialek, tanggal, sumber\_buku, dan kelas\_kata. Pada tahap ini belum dilakukan pra-pemrosesan; data digunakan sebagaimana adanya sebagai sumber masukan untuk proses selanjutnya. Pada tahap kedua, proses dimulai dengan validasi struktur dataset, termasuk pengecekan keberadaan kolom kelas\_kata. Kolom baru ditambahkan untuk menampung informasi morfologi. Oleh karena itu, pada tahap ini ditambahkan kolom id\_awalan, id\_sisipan, dan id\_akhiran. Kolom-kolom tersebut digunakan untuk mendapat

informasi kelas kata. Seperti yang telah dijelaskan sebelumnya bahwa informasi kelas kata tidak tersedia secara lengkap pada data kamus, sehingga perlu dilakukan anotasi kelas kata. Kolom kelas kata harus terisi sehingga daftar kata-kata verba dapat diseleksi dan digunakan dalam penelitian ini. Tahapan pra-pemrosesan yang dilakukan dapat dilihat pada gambar 2 berikut.



Gambar 2 Tahapan Implementasi Pra-pemrosesan Data

Ekstraksi morfologi dilakukan dengan mencocokkan awalan (21 jenis), sisipan (4 jenis), dan akhiran (14 jenis), dengan prioritas pada afiks terpanjang. Keluaran tahap ini menambahkan atribut morfologis bagi kata-kata yang memiliki afiks. Selanjutnya, peneliti melakukan klasifikasi kelas kata berbasis morfologi dengan cara memprediksi kelas kata berdasarkan pola afiksasi. Dataset divalidasi dan kolom *metode\_prediksi* ditambahkan. Baris yang telah memiliki anotasi kelas kata tidak diubah. Prediksi dilakukan melalui aturan berprioritas: pertama melalui konfiks, kemudian awalan tunggal, dan terakhir akhiran. Kata yang berhasil diprediksi diberi label kelas kata dan dicatat metode prediksinya sebagai berbasis “elemen morfologi”. Selain melakukan metode prediksi berdasarkan afiks, peneliti juga memanfaatkan korpus KBBI untuk menambah anotasi kelas kata, terutama kata-kata yang belum memiliki kelas kata. Pemanfaatan korpus KBBI untuk anotasi kelas kata disampaikan pada gambar 3. Tahapan ini diproses secara komputasional menggunakan Python seperti *regular expression* dan NLP. Setelah semua baris pada kolom kelas kata terisi, peneliti memilih entri kata verba.



Gambar 3 Anotasi Kelas Kata pada Korpus Bahasa Lampung Menggunakan KBBI

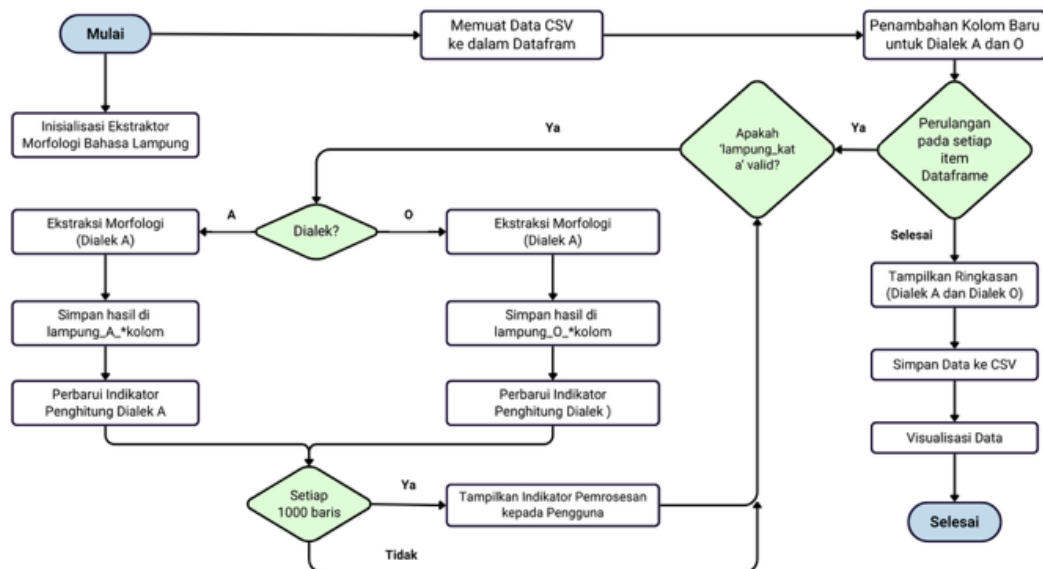
### 3. Rekayasa Fitur

Rekayasa fitur pada bahasa Lampung dalam penelitian ini dirancang untuk menghasilkan representasi terstruktur dari kata-kata yang dianalisis, baik dalam Dialek A maupun Dialek O. Prosesnya mencakup normalisasi teks, identifikasi afiks, serta pemetaan hubungan antara unsur morfologis tersebut. Setiap tahap mengikuti kaidah linguistik yang relevan dengan sistem morfologi bahasa Lampung. Terdapat lima langkah pada rekayasa fitur yaitu: (1) normalisasi dan konversi (konversi menjadi huruf kecil dan penghapusan data kosong); (2) penanganan frase; (3) ekstraksi awalan dengan mempertimbangkan alomorf; (4) ekstraksi akhiran spesifik dialek dan akhiran umum; dan (5) identifikasi konfiks. Seluruh ekstraksi dilakukan menggunakan aturan eksplisit mengenai afiksasi dan alomorf dalam bahasa Lampung. Sistem memisahkan proses berdasarkan Dialek A dan O untuk menangkap variasi morfologis. Penanganan frasa, huruf kapital, dan variasi teks memastikan hasil ekstraksi stabil dan konsisten. Prefiks nasal dievaluasi terlebih dahulu, diikuti prefiks umum, akhiran dialek, akhiran umum, dan akhirnya konfiks. Validasi panjang kata dan pembersihan format mencegah ekstraksi palsu (false positive).

### 4. Analisis Morfologi

Analisis dilakukan dengan pendekatan linguistik deskriptif dan linguistik komputasional berbasis aturan (*rule-based*), yang dirancang untuk menangkap afiksasi, alomorfisme, dan variasi dialektal secara sistematis. Analisis morfologi dalam penelitian ini berlandaskan pada prinsip bahwa kata merupakan satuan berstruktur yang tersusun atas morfem bebas dan morfem terikat. Kategori afiks yang dianalisis mencakup prefiks, infiks, sufiks, dan konfiks, termasuk variasi bentuk yang muncul akibat proses fonologis seperti peluluhan nasal. Penelitian ini mengadopsi paradigma *item-and-arrangement*, di mana struktur kata dipandang sebagai rangkaian morfem yang teratur dan dapat dipetakan secara linear. Inventaris afiks untuk bahasa Indonesia dan bahasa Lampung disusun berdasarkan literatur morfologi dan dikodifikasikan menjadi sejumlah kategori fungsional, yaitu afiks derivasional, infleksional, dan afiks netral. Afiks derivasional dianalisis berdasarkan kontribusinya terhadap pembentukan kelas kata baru, sedangkan afiks infleksional diperlakukan sebagai penanda kategori gramatikal seperti aspek, diatesis, atau valensi. Pada bahasa Lampung, analisis afiks disesuaikan karakteristik dialektal, yang menunjukkan perbedaan jelas dalam repertoar sufiks maupun distribusi prefiks tertentu.

Salah satu komponen penting dalam analisis adalah pemodelan alomorf, khususnya prefiks nasal dalam bahasa Lampung yang menampilkan variasi bentuk bergantung pada fonem awal kata dasar. Proses analisis morfologi dilakukan secara komputasional seperti gambar 4 berikut ini.



Gambar 4 Alur Analisis Morfologi Secara Komputasional pada Verba Bahasa Lampung

Analisis juga memperhatikan keberadaan konfiks sebagai morfem diskontinu yang bekerja secara terintegrasi. Konfiks pada kedua bahasa diklasifikasikan berdasarkan fungsi semantis dan gramatikalnya, seperti pembentuk nomina abstrak atau penanda verbalitas. Dalam kerangka analisis, konfiks diberi prioritas interpretatif dibandingkan prefiks atau sufiks yang muncul secara terpisah, karena kedua komponennya membentuk satu unit makna yang tidak dapat ditafsirkan secara independen. Karena bahasa Lampung memiliki dua dialek besar yang berbeda secara morfologis, analisis dilaksanakan dengan mempertimbangkan distribusi afiks yang bersifat dialek-spesifik. Sufiks tertentu yang hanya muncul pada Dialek A atau Dialek O diperlakukan sebagai indikator morfologis yang relevan, dan aturan analisis dibedakan antara kedua dialek. Pendekatan ini memungkinkan representasi yang akurat terhadap kekhasan sistem morfologi Lampung yang bersifat geografis dan historis. Proses identifikasi morfologis dilakukan melalui seperangkat aturan deterministik yang mengatur prioritas analisis. Aturan tersebut disusun secara hierarkis dengan urutan: (1) pemeriksaan konfiks, (2) identifikasi prefiks, dan (3) identifikasi sufiks. Setiap aturan disertai batasan struktural, seperti panjang minimal kata, posisi infiks, serta ketentuan dialek.

## 5. Identifikasi Proses Morfofonemik

Proses morfofonemik utama dalam penelitian ini diidentifikasi melalui analisis terhadap sejumlah contoh representatif. Setiap proses dievaluasi untuk mengkaji pola perubahan bunyi yang muncul ketika suatu morfem berinteraksi dengan bentuk dasar. Secara khusus, dua proses morfofonemik berikut dianalisis secara sistematis: yaitu (1) **asimilasi nasal pada awalan N-** dan (2) **penyisipan (infiksasi)**. Asimilasi nasal mencakup variasi alomorfis seperti *m-*, *n-*, *ng-*, *nge-*, dan *ny-* yang muncul bergantung pada fonem awal kata dasar. Analisis difokuskan pada identifikasi aturan asimilasi yang mengatur perubahan konsonan awal tersebut, sehingga dapat dijelaskan kondisi fonologis yang memicu realisasi alomorf tertentu. Sementara dalam

penyisipan, infiks seperti *-el-*, *-em-*, *-en-*, dan *-er-* dikaji untuk memahami pola penyisipan morfem ke dalam kata dasar serta implikasi fonologis dan morfologis yang ditimbulkannya. Pemeriksaan ini bertujuan untuk menelaah bagaimana infiksasi memodifikasi struktur internal kata dan mempengaruhi bentuk permukaan yang dihasilkan. Proses morfofonemik kunci diidentifikasi melalui pemeriksaan contoh-contoh spesifik seperti pada tabel 1 berikut.

**Tabel 1 Proses, Contoh, dan Tujuan Analisis Morfofonemik**

Proses	Contoh Awalan/Sisipan	Tujuan Analisis
Asimilasi Nasal (Awalan N-)	N-(m), N-(n), N-(ng), N-(nge), N-(ny)	Identifikasi aturan asimilasi perubahan bunyi konsonan awal kata dasar saat dilekati awalan N-
Penyisipan (Infiksasi)	-el-, -em-, -en-, -er-	Memahami pola penyisipan dan dampaknya terhadap bentuk kata

## 6. Formalisasi Aturan Derivasi

Berdasarkan analisis morfofonemik, proses derivasi dalam bahasa Lampung diidentifikasi dan diformalisasi ke dalam bentuk fungsi komputasional. Formalisasi ini memungkinkan prediksi bentuk kata terderivasi dengan mempertimbangkan perubahan fonologis yang sistematis, termasuk asimilasi, infiksasi, sufiksasi, dan konfiksasi. Tabel 2 berikut merangkum aturan derivasi yang diimplementasikan.

**Tabel 2 Rincian Formalisasi Aturan Derivasi**

Nama Aturan/Fungsi	Deskripsi	Contoh
apply_nasal_prefix_refined	Mengaplikasikan aturan asimilasi nasal berdasarkan konsonan awal kata dasar (bilabial, alveolar, palatal, velar, vokal) untuk memprediksi bentuk awalan N-	N-
apply_suffix	Melekatkan berbagai akhiran (-ken, -ko, -an, -on, -kon) serta akhiran objek spesifik dialek (-ni, -na untuk Dialek A; -no untuk Dialek O) pada kata dasar	-ken, -ko, -an, -on, -kon, -ni, -na, -no
apply_circumfix	Mengintegrasikan logika awalan (termasuk asimilasi nasal) dan akhiran untuk memprediksi bentuk terderivasi dari konfiks	N-...-ken, be-...-an, di-...-an
apply_infix	Menyisipkan infiks setelah konsonan awal kata dasar, mencerminkan mekanisme infiksasi yang diamati	

## 7. Interpretasi Hasil dan Implikasi Linguistik

Analisis morfofonemik dalam proyek ini mengungkap pola derivasi sistematis pada bahasa Lampung. Prefiks N- mengalami asimilasi nasal sesuai konsonan awal kata dasar, menghasilkan variasi alomorf seperti n-, m-, ng-, nge-, dan ny-, yang mencerminkan distribusi

komplementer alomorf. Infiksasi (-el-, -em-, -en-, -er-) dan sufiksasi, termasuk akhiran spesifik dialek seperti -ni, -na (Dialek A) dan -no (Dialek O), menunjukkan mekanisme morfologis yang memengaruhi bentuk dan makna kata. Konfiks yang menggabungkan awalan dan akhiran menegaskan peran morfem diskontinu dalam pembentukan kata derivatif. Pendekatan *rule-based* yang memprioritaskan konfiks memastikan prediksi bentuk kata yang konsisten dan dapat direproduksi. Temuan ini memiliki implikasi linguistik signifikan, antara lain: dokumentasi afiks dan variasi dialektal bahasa Lampung, validasi prinsip fonologis melalui alomorfisme N-, dan pemahaman struktur kata serta mekanisme derivasi. Integrasi dengan sumber eksternal seperti KBBI memperkuat validitas metodologis, menjadikan kerangka ini berguna untuk penelitian lebih lanjut dalam morfologi dan fonologi bahasa daerah.

## HASIL DAN PEMBAHASAN

### 1. HASIL

Penelitian ini menghasilkan serangkaian temuan komprehensif mengenai derivasi verba dalam Bahasa Lampung dialek A dan dialek O. Temuan-temuan ini mencakup identifikasi pola afiksasi, analisis proses morfofonemik yang mendasari, kontribusi semantik dan fungsi sintaksis dari setiap afiks, perbandingan karakteristik antar-dialek, serta evaluasi terhadap aturan derivasi yang telah diformalisasi. Pemrosesan komputasional tidak hanya dilakukan pada analisis data, tetapi juga pada pra-pemrosesan data. Hasil dari penelitian ini adalah sebagai berikut.

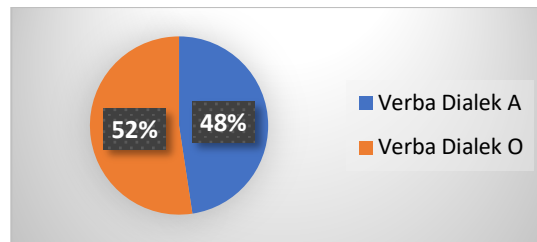
#### a. Karakteristik Data

Korpus verba dalam penelitian ini berjumlah 7.934 entri dengan rincian 3.786 dialek A dan 4.148 dialek O. Sebelum korpus verba diseleksi, jumlah data awal adalah 30.255 entri yang terdiri atas 14.080 dialek A dan 16.175 dialek O. Komponen afiksasi verba mencakup delapan kategori pada masing-masing dialek, disusun berdasarkan awalan, sisipan, akhiran, dan gabungan pada masing-masing dialek. Rekapitulasi data dapat dilihat pada tabel 3 berikut.

**Tabel 3 Rekapitulasi Karakteristik Data**

No	Komponen	Total Data
1	Dataset utama (berbagai kelas kata) <ul style="list-style-type: none"> <li>▪ Dialek A</li> <li>▪ Dialek O</li> </ul>	30.255 entri 14.080 entri (46.5%) 16.175 entri (53.5%)
2	Data verba <ul style="list-style-type: none"> <li>▪ Dialek A</li> <li>▪ Dialek O</li> </ul>	3.786 entri 4.148 entri
3	Afiksasi Verba Lampung <ul style="list-style-type: none"> <li>▪ Awalan Dialek A</li> <li>▪ Sisipan Dialek A</li> <li>▪ Akhiran Dialek A</li> <li>▪ Gabungan Dialek A</li> <li>▪ Awalan Dialek O</li> <li>▪ Sisipan Dialek O</li> <li>▪ Akhiran Dialek O</li> <li>▪ Gabungan Dialek O</li> </ul>	2.508 523 609 511 2.637 650 588 502

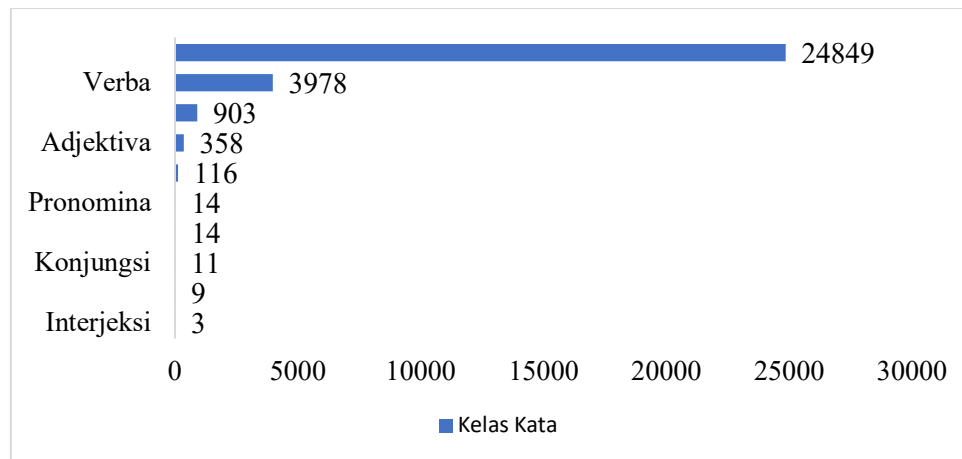
Berdasarkan hasil seleksi, jumlah verba dialek O lebih banyak daripada verba dialek A. Persentase verba tiap dialek divisualisasikan pada gambar 5 berikut.



Gambar 5 Komposisi Verba Tiap Dialek di Bahasa Lampung

### b. Akuisisi dan Pra-pemrosesan Data

Sebelum diseleksi, korpus verba diproses dalam lima tahapan yaitu (1) akuisisi; (2) anotasi manual; (3) anotasi berbasis linguistik komputasional (4) validasi; dan (5) seleksi korpus verba. Proses akuisisi telah dijelaskan sebelumnya pada karakteristik data. Pada tahap selanjutnya, anotasi manual dan linguistik komputasional dilakukan untuk menentukan kelas pada masing-masing kata. Hal ini dilakukan karena data awal penelitian tidak memiliki kelas kata. Jumlah kata yang dianotasi secara manual, seperti pada gambar 6, adalah 5.406 kata (17%).



Gambar 6 Grafik Kelas Kata Setelah Dianotasi Secara Manual

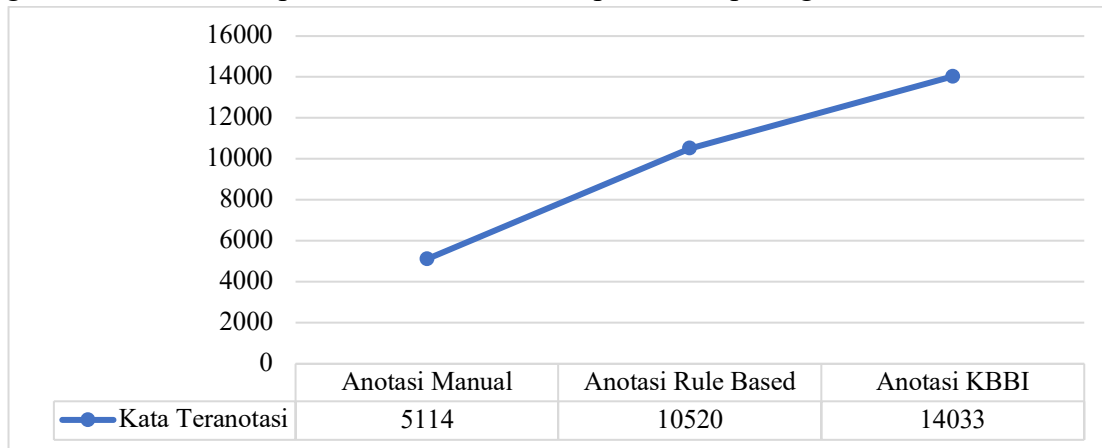
Setelah anotasi manual selesai, peneliti menggunakan teknik linguistik komputasional untuk menambah isian kelas kata. Pemrosesan dilakukan melalui ekstraksi morfologis yang kemudian menjadi dasar prediksi kelas kata secara *rule-based*. Hasil dari ekstraksi morfologi dapat dilihat pada tabel 4 berikut ini.

Tabel 4 Hasil Ekstraksi Morfologi

Statistik	Nilai
Total entri terproses	30.255
Baris dengan awalan teridentifikasi	8.830 (29,2%)
Baris dengan sisipan teridentifikasi	6.159 (20,4%)
Baris dengan akhiran teridentifikasi	5.101 (16,9%)

Aturan yang digunakan mengacu pada kaidah morfologi bahasa Indonesia, terutama yang berkaitan dengan kategori kata, karena bahasa Indonesia memiliki pola yang lebih terdefinisi dan kategori kata bersifat relatif universal. Kolom yang digunakan adalah *lampung\_kata*. Proses ini berhasil menambah 5.114 kelas kata, sehingga total kata dengan kelas kata pada

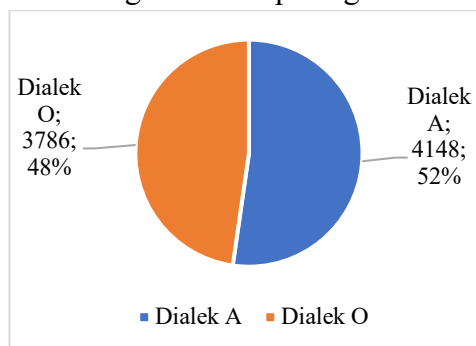
tahap ini menjadi 10.520 kata. Lalu, karena masih tersisa 19.735 tanpa kelas kata, peneliti memanfaatkan data KBBI dari publik untuk membantu anotasi kelas kata. Peneliti menggunakan Python untuk membantu dalam menentukan kelas kata berdasarkan data KBBI. Prediksi kelas kata berdasarkan KBBI menghasilkan 14.033 kelas kata, sehingga total kata dengan kelas kata mencapai 22.127. Proses ini dapat dilihat pada gambar X berikut ini.



Gambar 6 Perkembangan Hasil Anotasi Secara Manual, *Rule-Based*, dan KBBI

### c. Seleksi Verba dan Rekayasa Fitur

Seleksi verba merupakan langkah untuk memilah kata kerja dari kumpulan data yang berisi berbagai jenis kata. Dari total 30.255 entri pada data awal, proses seleksi ini menghasilkan 7.934 entri berupa verba, yang berarti sekitar 26.2% dari keseluruhan data. Distribusi verba berdasarkan dialek digambarkan pada gambar X berikut.



Gambar 7 Komposisi Verba Bahasa Lampung pada Dialek A dan O

Karakteristik morfologi verba bahasa Lampung adalah sebagai berikut.

**Tabel 6 Karakteristik Morfologi Verba Bahasa Lampung**

Karakteristik Morfologi	Total
Dengan awalan	4860
Dengan akhiran	2345
Dengan sisipan	2536

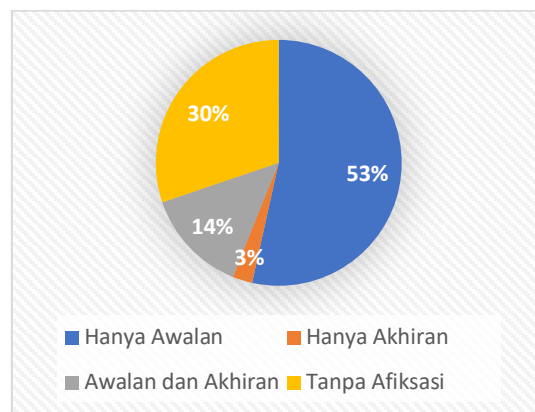
Rekayasa fitur dimulai dengan ekstraksi fitur morfologi bahasa Lampung dari kata-kata verba. Proses ini menambahkan kolom-kolom baru yang berisi informasi struktural tentang awalan, sisipan, akhiran, dan konfiks untuk kedua dialek (A dan O).

**Tabel 7 Fitur Morfologi yang Diekstraksi per Dialek**

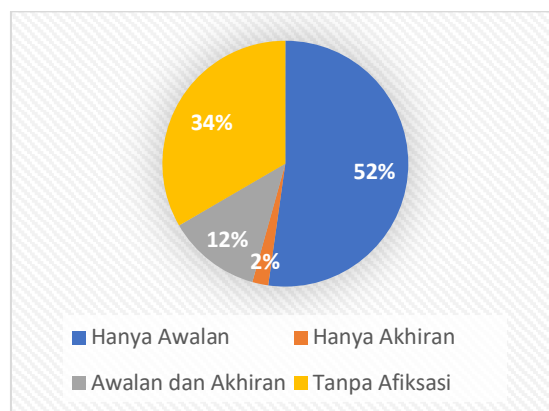
Dialek	Nama Fitur	Jenis Afiks	Contoh Bentuk / Pola
A	lampung_A_awalan	Prefiks	be-, bu-, ti-, peN-, puN-, te-, ke-, se-
A	lampung_A_sisipan	Infiks	-el-, -em-, -en-, -er-
A	lampung_A_akhiran	Sufiks	-ni, -ne, -na, -an
A	lampung_A_gabungan	Konfiks (awalan + akhiran)	ke-...-an, se-...-ni, pe-...-an
O	lampung_O_awalan	Prefiks	be-, peN-, te-, ke-, se-
O	lampung_O_sisipan	Infiks	-el-, -em-, -en-, -er-
O	lampung_O_akhiran	Sufiks	-nou, -no, -kon, -ken, -on, -ko
O	lampung_O_gabungan	Konfiks (awalan + akhiran)	se-...-no, be-...-an

#### d. Analisis Morfologi

Analisis mencakup distribusi awalan, akhiran, sisipan, dan konfiks pada kedua dialek (A dan O), serta pola-pola morfologi yang dominan dalam bahasa Lampung. Dari **7.934 verba** yang dianalisis, ditemukan berbagai pola morfologi yang menunjukkan keragaman dan kompleksitas struktur kata dalam bahasa Lampung. Analisis ini penting untuk memahami sistem morfologi pada bahasa Lampung. Berikut ini adalah distribusi struktur morfologi berdasarkan dialek.

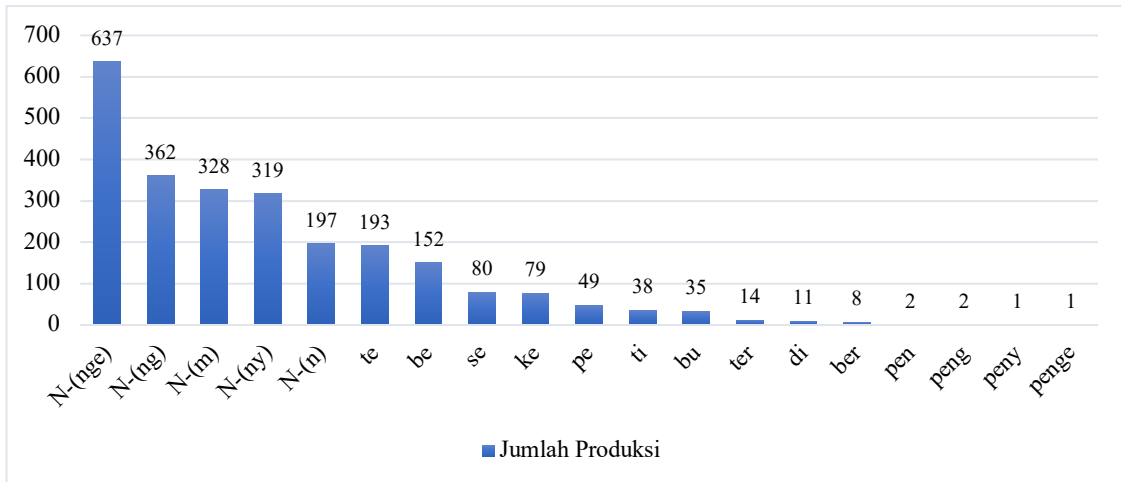


Gambar 8 Komposisi Afiks Dialek A

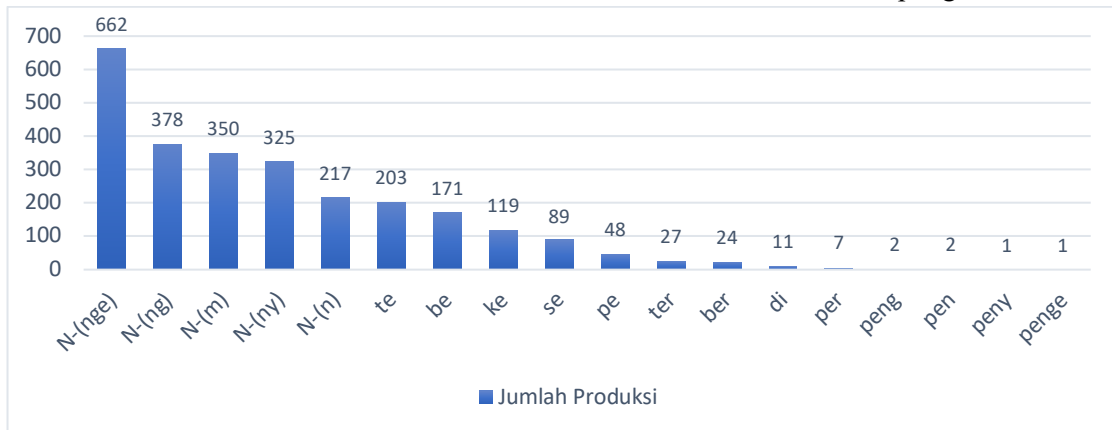


Gambar 9 Komposisi Afiks Dialek O

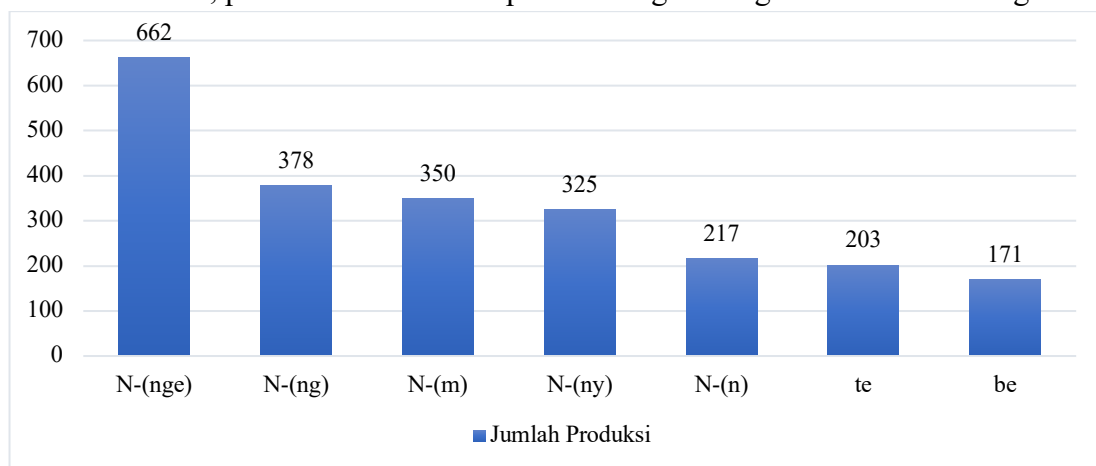
Selain dilakukan visualisasi pada komposisi afiks pada tiap dialek, berikut ini ditampilkan daftar 15 teratas awalan dialek A dan O yang menunjukkan produktivitas tiap awalan.



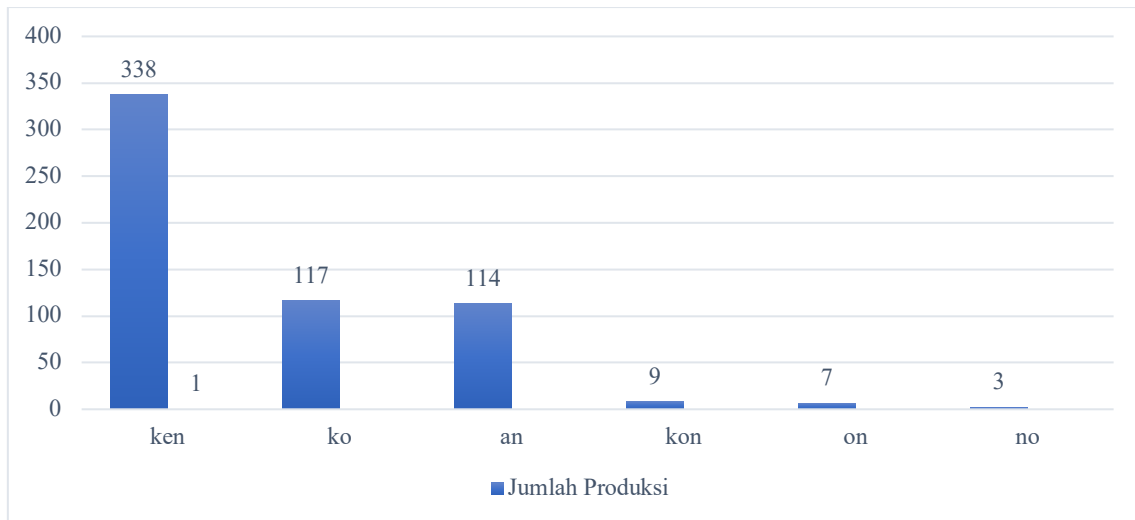
Gambar 10 Daftar 15 Teratas Produktivitas Awalan Verba Bahasa Lampung Dialek A



Gambar 11 Daftar 15 Teratas Produktivitas Awalan Verba Bahasa Lampung Dialek O  
 Sementara itu, produktivitas akhiran pada masing-masing dialek adalah sebagai berikut.



Gambar 12 Daftar Produktivitas Akhiran Verba Bahasa Lampung Dialek A



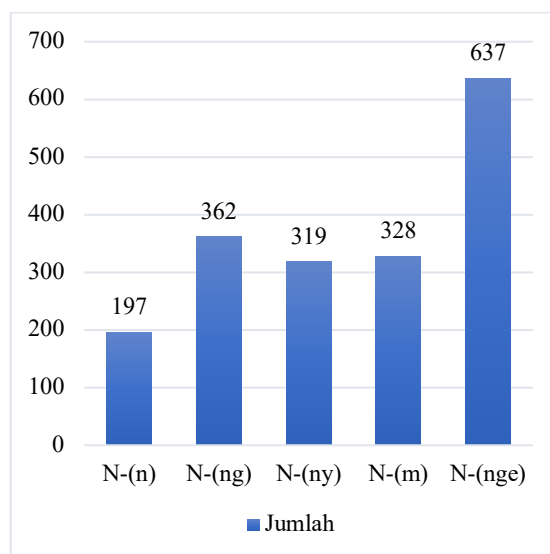
Gambar 13 Daftar Produktivitas Akhiran Verba Bahasa Lampung Dialek O

**e. Analisis Morfofonemik**

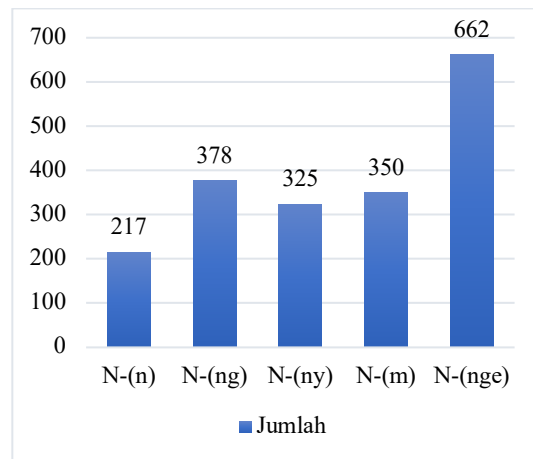
Analisis ini mengkaji perubahan-perubahan fonologis yang terjadi dalam verba bahasa Lampung, termasuk alomorf, peluluhan konsonan, dan variasi dialek. Penelitian ini menganalisis berbagai fenomena morfofonemik dalam verba bahasa Lampung, seperti yang tertuang pada tabel 8 berikut ini.

**Tabel 8 Analisis Morfofonemik**

Istilah	Penjelasan	Contoh
Alomorf	Variasi bentuk morfem yang sama	N- menjadi n-, ng-, ny-, m-, nge-
Peluluhan	Penghilangan konsonan awal kata dasar saat afiks ditambahkan	-
Asimilasi	Perubahan fonem untuk menyesuaikan dengan lingkungan fonologis	-
Variasi Dialek	Perbedaan morfofonemik antara dialek A dan O	-



Gambar 14 Distribusi Alomorf dalam Dialek A



Gambar 15 Distubusi Alomorf dalam Dialek O

Verba bahasa Lampung menunjukkan kompleksitas sistem morfologi yang tinggi. Prefiks N- memiliki 5 alomorf yang muncul berdasarkan konsonan awal kata dasar, menunjukkan proses asimilasi nasal yang sistematis. Peluluhan konsonan pada prefiks **peN-** juga menunjukkan pola yang konsisten, dengan variasi *pen-*, *peng-*, *peny-*, *pem-*, dan *penge-* yang dapat diprediksi berdasarkan konsonan awal kata dasar. Perbedaan morfofonemik dialek A dan O menunjukkan variasi linguistik yang signifikan, dengan beberapa awalan dan akhiran yang spesifik untuk masing-masing dialek.

#### f. Identifikasi Pola Afiksasi dan Jenis Derivasi Verba

Analisis mengidentifikasi empat kategori utama afiks: prefiks, infiks, sufiks, dan konfiks. Setiap kategori memiliki peran spesifik dalam pembentukan verba dalam bahasa Lampung.

Tabel 9 Afiksasi dan Bentuk Derivasi Verba

Afiks	Bentuk	Fungsi	Keterangan/Dialek
Prefiks Nasal (N-)	N-(m), N-(n), N-(ng), N-(nge), N-(ny)	Membentuk verba aktif transitif, makna kausatif atau agen	Afiks paling produktif
Prefiks di-	di-	Menandai bentuk pasif, subjek penerima tindakan	-
Prefiks be-/ber-	be-, ber-	Membentuk verba intransitif, keadaan, tindakan resiprokal, refleksif	-
Prefiks te-/ti-	te-, ti-	Verba pasif tak sengaja, keadaan statif, kemampuan	te- (Dialek O), ti- (Dialek A)
Prefiks ke-	ke-	Keadaan tidak sengaja atau potensi	Serupa dengan te-/ti-
Prefiks pe-/pen-/peng-/peny-/penge-	pe-, pen-, peng-, peny-, penge-	Menandai tindakan atau kausatif	Dalam konteks verba
Infiks	-el-, -em-, -en-, -er-	Fonologis atau penambah nuansa makna halus	Bukan perubahan kategori gramatikal drastis
Sufiks	-an, -ken, -ko, -on, -kon, -ni, -na, -no	Kausatif/aplikatif (-ken, -ko, -kon), hasil/resiprokal (-an), penanda objek (-ni, -na, -no, -on)	-ni, -na (Dialek A); -no, -on (Dialek O)

Afiks	Bentuk	Fungsi	Keterangan/Dialek
Konfiks	N-...-ken, be-...-an	Menggabungkan makna dan fungsi awalan-akhiran, derivasi verba kompleks	N-...-ken (kausatif aktif), be-...-an (resiprokal/iteratif)

### g. Perbandingan Dialek Lampung dan Validasi Aturan Derivasi Terformalisasi

Perbandingan kedua dialek menunjukkan **kesamaan morfologis yang sangat kuat** dalam sistem derivasi verba. Sebagian besar afiks dan aturan derivasi berfungsi secara identik. Perbedaan utama bersifat minor dan cenderung fonetis atau kuantitatif:

**Tabel 10 Perbandingan Aturan Derivasi Dialek A dan O**

Fitur	Dialek A	Dialek O	Keterangan
Variasi Fonetis Sufiks Objek	-ni, -na	-no, -on	Perbedaan vokal akhir dialek
Frekuensi Afiks	Sedikit variasi awalan nasal	Sedikit variasi awalan nasal	Dipengaruhi preferensi leksikal/lingkungan fonologis
Afiks Unik Minor	bu- lebih menonjol	per- lebih menonjol	Diferensiasi inventaris afiks kurang produktif

Fungsi Python yang memodelkan aturan derivasi berhasil memprediksi sebagian besar bentuk turunan verba. Namun, beberapa area membutuhkan penyempurnaan lebih lanjut. Secara keseluruhan, temuan ini mengkonfirmasi konsistensi tinggi dalam morfologi verba Bahasa Lampung dan memberikan dasar yang kuat untuk pengembangan model komputasi yang lebih akurat. Tabel 10 berikut menunjukkan fenomena dan kasus yang ditemukan

**Tabel 11 Model Aturan Devivasi menggunakan Phyton**

Fenomena	Contoh Kasus	Catatan/Permasalahan
Asimilasi Nasal Kata Dasar Berawalan Vokal/h	ambil (ngambil vs. ngeambil), hitung (ngitungken vs. N-hitungken)	Kompleksitas aturan asimilasi nasal untuk kata dasar berawalan vokal/konsonan h belum sepenuhnya tertangkap
Penghilangan Vokal pada Sufiksasi	kumpul + -on → kumplon (bukan kumpulon)	Penghilangan vokal pada suku kata terakhir sebelum sufiksasi belum diimplementasikan
Infixasi Kompleks	saut → semuat	Transformasi fonologis/leksikalisasi lebih dalam daripada penyisipan langsung

## 2. PEMBAHASAN

### a. Ringkasan Tujuan Analisis

Analisis ini bertujuan untuk mendalami derivasi verba dalam Bahasa Lampung, khususnya pada Dialek A dan Dialek O, dengan memanfaatkan pendekatan linguistik komputasi. Melalui ekstraksi, kategorisasi, dan formalisasi aturan afiksasi, kami berupaya mengidentifikasi pola morfologis, memahami proses morfofonemik, serta membandingkan karakteristik derivasi verba antara kedua dialek. Tujuan utamanya adalah untuk membangun pemahaman yang lebih terstruktur tentang bagaimana verba dibentuk dan dimodifikasi dalam Bahasa Lampung, yang dapat menjadi dasar untuk aplikasi komputasi di masa depan.

## b. Integrasi Temuan-Temuan Kunci

### 1) Pola Afiksasi dan Jenis Derivasi

Terdapat empat jenis afiks utama: awalan (prefiks), sisipan (infiks), akhiran (sufiks), dan gabungan (konfiks). Masing-masing memiliki peran spesifik dalam derivasi verba:

**Tabel 12 Afiksasi dan Fungsi Verba**

Bentuk Afiks	Fungsi	Contoh
Awalan N- (Nasal)	Membentuk verba aktif transitif, kausatif, pelaku tindakan	minjak (bangkit), nangani (mengatasi), ngusung (membawa), nyani (membuat)
Awalan di-	Penanda pasif, mengubah fokus ke penerima	dibang (digantung)
Awalan be- / ber-	Menandai verba resiprokal, refleksif, statif	bekawan (berteman), berdikir (berzikir)
Awalan te- / ti-	Tindakan tidak disengaja, kemampuan, keadaan statis	tetok (terbacok), tisalok (terayom)
Awalan pe- / pen- / peng- / peny- / penge-	Nomina agen, derivasi verba kausatif, verba tindakan/proses	pendiko (memberitahu)
Sisipan (-el-, -em-, -en-, -er-)	Memengaruhi nuansa fonologis/aspek verba	belajagh (b-ajar), gemul (g-mul/akar beruang)
Akhiran (-an, -ken, -ko, -ni, -na, -no, -on, -kon)	Kausatif, aplikatif, hasil, resiprokal/iteratif, sufiks objek/kepemilikan	ngabadiken (mengabadikan), megenalko (mempromosikan), nyani (membuatnya/membuat)
Konfiks (Awalan + Akhiran)	Makna derivasional kompleks, gabungan fungsi konstituen	N-...-ken (aktif-kausatif), be-...-an (resiprokal-iteratif), di-...-an (pasif-lokatif)

### 2) Frekuensi Afiks

Analisis frekuensi menunjukkan bahwa awalan nasal (N-(nge), N-(ng), N-(m), N-(ny), N-(n)) adalah yang paling produktif di kedua dialek, mendominasi pembentukan verba aktif. Untuk sufiks, -ken, -ko, dan -an merupakan yang paling sering digunakan, menyoroti pentingnya kategori kausatif, aplikatif, dan hasil dalam verba Lampung. Sisipan cenderung memiliki frekuensi yang lebih rendah dibandingkan awalan dan akhiran.

### 3) Proses Morfonemik

**Tabel 13 Morfonemik dalam Bahasa Lampung**

Proses	Deskripsi	Pola/Contoh	Fungsi
Asimilasi Nasal	Proses morfonemik paling menonjol, awalan N- berasimilasi dengan konsonan awal kata dasar	b/p menjadi m-, d/t menjadi n-, s/c/j menjadi ny-, k/g menjadi ng-, vokal/kasus tertentu menjadi nge-	Memfasilitasi kelancaran pengucapan, menunjukkan kekhasan fonologi Bahasa Lampung
Penyisipan (Infixation)	Sisipan seperti -el-, -em-, -en-, -er- disisipkan setelah konsonan awal kata dasar	-el-, -em-, -en-, -er-	Bersifat fonologis atau menambah nuansa makna yang halus, kurang mengubah kategori semantis secara eksplisit

#### 4) Perbandingan Dialek (Dialek A vs. Dialek O)

**Tabel 14 Perbandingan Afiksasi Dialek A dan O**

Aspek	Dialek A	Dialek O	Keterangan
Konsistensi Morfologis	Berlaku	Berlaku	Afiks, aturan derivasi, proses morfofonemik identik
Frekuensi Awalan Nasal	N-(m), N-(ny) lebih sering	N-(nge), N-(ng) lebih dominan	Perbedaan kuantitatif/fonetis
Sufiks Objek	-ni, -na	-no	Perbedaan vokal akhir sufiks
Afiks Unik	bu- lebih menonjol	per- lebih menonjol	Penggunaan tidak seproduktif afiks inti
Vokal Akhir Sufiks	'a', 'i'	'o'	Dialek O cenderung vokal 'o'

#### c. Kontribusi Semantik dan Fungsi Sintaksis Afik

**Tabel 15 Verba dan Penanda Semantik**

Jenis Verba	Penanda	Deskripsi	Contoh
Transitif Aktif	N-	Subjek melakukan tindakan langsung pada objek	menulis dari tulis, menggambar dari gambar
Pasif	di-	Subjek adalah penerima tindakan	ditulis dari tulis, digambar dari gambar
Kausatif	-ken, -ko, -kon	Subjek menyebabkan sesuatu terjadi atau membuat objek melakukan tindakan	membacakan dari baca, mempromosikan dari kenal
Resiprokal/Refleksif	be-/ber-, -an	Tindakan saling atau tindakan pada diri sendiri	berteman dari teman, bercanda dari canda
Statif/Tidak Sengaja	te-/ti-	Keadaan atau tindakan yang terjadi tanpa disengaja oleh subjek	terjatuh dari jatuh
Aplikatif	-ni, -na, -no, -on	Menandai penerima, benefaktif, atau instrumen dari suatu tindakan	tujuni (menujui sesuatu), dapatno (mendapatkannya)
Sisipan	-	Mengubah aspek atau modalitas verba, seringkali dalam konteks historis atau fonologis	-

Temuan penelitian menunjukkan bahwa prefiks memiliki tingkat produktivitas lebih tinggi dibandingkan sufiks dan konfiks dalam pembentukan verba derivatif Bahasa Lampung. Pola ini serupa dengan bahasa Melayu, namun distribusi dan frekuensi afiks tertentu menunjukkan karakteristik morfologis yang lebih spesifik dan tidak sepenuhnya paralel dengan bahasa serumpun. Berbeda dari penelitian sebelumnya yang bersifat deskriptif, studi ini mengkuantifikasi produktivitas afiks berbasis korpus besar melalui pendekatan komputasional, sehingga menghasilkan deskripsi morfologi yang lebih objektif, terukur, dan replikatif. Perbedaan produktivitas antarafiks juga mengindikasikan bahwa sistem morfologi Bahasa

Lampung bersifat selektif, dipengaruhi oleh distribusi leksikal dan frekuensi penggunaan dalam korpus.

## SIMPULAN

Analisis data menunjukkan bahwa proses formalisasi aturan derivasi verba dalam Bahasa Lampung melalui pendekatan linguistik komputasional berhasil dilakukan dengan tingkat akurasi yang kuat. Berbagai aturan morfologis—meliputi asimilasi nasal, sufiksasi, konfiksasi, dan infiksasi—telah diterjemahkan secara sistematis ke dalam fungsi-fungsi Python seperti *apply\_nasal\_prefix\_refined*, *apply\_suffix*, *apply\_circumfix*, dan *apply\_infix*. Pertama, aturan **asimilasi nasal** berhasil diformalkan, terutama terkait perubahan konsonan awal, misalnya ‘b’/‘p’ menjadi *m-*, dan ‘d’/‘t’ menjadi *n-*. Aturan untuk kata berawalan vokal juga disusun, termasuk penggunaan *nge-* atau *ng-*. Sementara itu, proses **sufiksasi** ditangani oleh fungsi *apply\_suffix* yang mampu mengolah berbagai sufiks seperti *-ken*, *-ko*, *-an*, *-on*, dan *-kon*, termasuk varian sufiks objek antar dialek, seperti *-ni / -na* pada Dialek A dan *-no* pada Dialek O. Pada **konfiksasi**, fungsi *apply\_circumfix* berhasil menggabungkan logika prefiks dan sufiks, misalnya pada bentuk *N-...-ken* atau *be-...-an*. **Infiksasi** juga berhasil dimodelkan dengan penyisipan infiks *-el-*, *-em-*, *-en-*, dan *-er-* setelah konsonan pertama dalam bentuk dasar. Validasi terhadap aturan tersebut dengan data uji baru memperlihatkan bahwa model bekerja cukup akurat, namun beberapa aspek memerlukan penyempurnaan lebih lanjut. Pada proses asimilasi nasal untuk kata berawalan vokal, prediksi seperti *ngeambil* atau *ngeisi* kurang sesuai dengan bentuk yang diharapkan (*ngambil*, *ngisi*), sehingga aturan prefiks *N-* untuk kondisi ini perlu diperhalus. Pada Dialek O, ditemukan pula fenomena **penghilangan vokal** dalam sufiksasi, misalnya *kumpul + -on* yang secara prediktif menghasilkan *kumpulon*, padahal bentuk yang benar adalah *kumplon*. Selain itu, kasus kata berawalan *h-* seperti *hitung* yang seharusnya menjadi *ngitungken* menunjukkan bahwa aturan asimilasi nasal untuk konsonan *h-* masih belum terakomodasi. Dari sisi fungsi semantis dan sintaktis afiks, prefiks seperti *N-*, *di-*, *be-/ber-*, *te-/ti-*, *ke-*, serta variasi *pe-*, *bu-* (Dialek A), dan *per-* (Dialek O) terbukti mempengaruhi *voice*, transitivitas, aspek, serta nuansa intensional. Prefiks *N-* merupakan yang paling produktif, terutama dalam membentuk verba aktif transitif dengan makna kausatif atau agen. Sufiks seperti *-ken*, *-ko*, *-kon*, *-an*, dan sufiks objek berperan dalam membentuk makna kausatif, aplikatif, iteratif, serta struktur valensi. Konfiks menghasilkan makna kompleks melalui kombinasi fungsi prefiks dan sufiks, sedangkan infiks memberi kontribusi yang lebih bersifat fonologis atau aspekual dibandingkan perubahan sintaktis yang besar. Perbandingan antara **Dialek A dan Dialek O** menunjukkan bahwa keduanya memiliki pola morfologis dan proses morfofonemik yang sangat serupa. Perbedaan terutama terletak pada aspek fonetik, seperti penggunaan sufiks objek (*-ni / -na* pada Dialek A dan *-no* pada Dialek O) serta variasi kecil dalam distribusi prefiks yang kurang produktif, seperti *bu-* di Dialek A dan *per-* di Dialek O.

## REFERENSI

- Abidin, Z., Wijaya, A., & Pasha, D. (2021). Aplikasi Stemming Kata Bahasa Lampung Dialek Api Menggunakan Pendekatan Brute-Force dan Pemograman C#. *JURNAL MEDIA INFORMATIKA BUDIDARMA*, 5(1), 1–8.  
<https://doi.org/10.30865/mib.v5i1.2483>

- Afrianto, Sujatna, E. T. S., Darmayanti, N., Ariyani, F., & Cooke-Plagwitz, J. (2020). Clause and predicative constituents in an Austronesian language: Lampung language. *Topics in Linguistics*, 21(2), 62–79. <https://doi.org/10.2478/topling-2020-0010>
- Agus, N., Jerniati, Garing, J., Herianah, Musayyedah, Hastianah, & Syamsulrijal. (2024a). Verbalization Affixes in the Limola Language: A Descriptive Analysis. *Theory and Practice in Language Studies*, 14(12), 3712–3723. <https://doi.org/10.17507/tpls.1412.05>
- Agus, N., Jerniati, Garing, J., Herianah, Musayyedah, Hastianah, & Syamsulrijal. (2024b). Verbalization Affixes in the Limola Language: A Descriptive Analysis. *Theory and Practice in Language Studies*, 14(12), 3712–3723. <https://doi.org/10.17507/tpls.1412.05>
- Ahda, F. A., Wibawa, A. P., Prasetya, D. D., Sulisty, D. A., & Nafalski, A. (2025). Minangkabau Language Stemming: A New Approach with Modified Enhanced Confix Stripping. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 9(3), 677–687. <https://doi.org/10.29207/resti.v9i3.6511>
- Ajani, Y. A., Oladokun, B. D., Olarongbe, S. A., Amaechi, M. N., Rabi, N., & Bashorun, M. T. (2024). Revitalizing Indigenous Knowledge Systems via Digital Media Technologies for Sustainability of Indigenous Languages. *Preservation, Digital Technology & Culture*, 53(1), 35–44. <https://doi.org/10.1515/pdte-2023-0051>
- Alduais, A., Yassin, A. A., & Allegretta, S. (2025). Computational linguistics: A scientometric review. *Quality & Quantity*, 59(5), 4097–4136. <https://doi.org/10.1007/s11135-025-02138-2>
- Amrulloh, M. A., & Humairo, L. (2020). MORPHO-PHONOLOGICAL OF LOANWORDS IN LAMPUNG LANGUAGE FROM ARABIC. *Arabiyat : Jurnal Pendidikan Bahasa Arab Dan Kebahasaaraban*, 7(1), 78–97. <https://doi.org/10.15408/a.v7i1.13174>
- Astawan, N., & Sadwika, I. N. (2020). AFIKS DERIVASIONAL VERBA BAHASA LAMPUNG: SEBUAH KAJIAN MORFOLOGI GENERATIF. *Stilistika : Jurnal Pendidikan Bahasa Dan Seni*, 8(2), 173–191. <https://doi.org/10.59672/stilistika.v8i2.766>
- Bernard, T., & Winterstein, G. (2022). Introduction to the special section on the interaction between formal and computational linguistics. *Journal of Language Modelling*, 10(1). <https://doi.org/10.15398/jlm.v10i1.325>
- Bratlie, S. S., Gustafsson, J., & Torkildsen, J. V. K. (2022). Effectiveness of a Classroom-Implemented, App-Based Morphology Program for Language-Minority Students: Examining Latent Language-Literacy Profiles and Contextual Factors as Moderators. *Reading Research Quarterly*, 57(3), 805–829. <https://doi.org/10.1002/rrq.447>
- Colenbrander, D., Von Hagen, A., Kohnen, S., Wegener, S., Ko, K., Beyersmann, E., Behzadnia, A., Parrila, R., & Castles, A. (2024). The Effects of Morphological Instruction on Literacy Outcomes for Children in English-Speaking Countries: A Systematic Review and Meta-Analysis. *Educational Psychology Review*, 36(4), 119. <https://doi.org/10.1007/s10648-024-09953-3>
- Huisman, J. L. A., Franco, K., & Van Hout, R. (2021). Linking Linguistic and Geographic Distance in Four Semantic Domains: Computational Geo-Analyses of Internal and External Factors in a Dialect Continuum. *Frontiers in Artificial Intelligence*, 4, 668035. <https://doi.org/10.3389/frai.2021.668035>
- Luz, S. (2022). *Computational linguistics and Natural Language Processing*. <https://doi.org/10.4324/9781315158945>
- Miaschi, A., Brunato, D., & Dell'Orletta, F. (2021). A NLP-based stylometric approach for tracking the evolution of L1 written language competence. *Journal of Writing Research*, 13(1), 71–105. <https://doi.org/10.17239/jowr-2021.13.01.03>

- Nijs, J., Van De Velde, F., & Cuyckens, H. (2025). Is Word Order Responsive to Morphology? Disentangling Cause and Effect in Morphosyntactic Change in Five Western European Languages. *Entropy*, 27(1), 53. <https://doi.org/10.3390/e27010053>
- Nugraha, D. S. (2024a). Analyzing Prefix /me(N)-/ in the Indonesian Affixation: A Corpus-Based Morphology. *Theory and Practice in Language Studies*, 14(6), 1697–1711. <https://doi.org/10.17507/tpls.1406.10>
- Nugraha, D. S. (2024b). Analyzing the Derivational Verb of Indonesian Based on the Derivational Morphology. *South Asian Research Journal of Arts, Language and Literature*, 6(01), 1–11. <https://doi.org/10.36346/sarjall.2024.v06i01.001>
- Patria, R., & Merdeka, P. H. (2023). Creative Strategies in the Recovery of Endangered Languages. *Journal of Literature Language and Academic Studies*, 2(02), 57–61. <https://doi.org/10.56855/jllans.v2i2.671>
- Septianingias, V., Nugraha, I., & Susilawati, L. (2025). VARIASI DIALEKTAL ‘KELINGKING’ PADA BAHASA LAMPUNG. *Kongres Internasional Masyarakat Linguistik Indonesia, 2025*. <http://kimli.mlindonesia.org/index.php/kimli/article/view/188>
- Solovyev, V. D., Solnyshkina, M. I., & McNamara, D. S. (2022). Computational linguistics and discourse complexology: Paradigms and research methods. *Russian Journal of Linguistics*, 26(2), 275–316. <https://doi.org/10.22363/2687-0088-31326>
- Strycharczuk, P., López-Ibáñez, M., Brown, G., & Leemann, A. (2020). General Northern English. Exploring Regional Variation in the North of England With Machine Learning. *Frontiers in Artificial Intelligence*, 3, 48. <https://doi.org/10.3389/frai.2020.00048>
- Sudana, D., Hardini, T. I., & Zifana, M. (2024). Teaching Implication of Meaning Construction at Morphological Level of Affixation in a Complex Indonesian Verb Formation. *International Journal of Language Education*, 8(2). <https://doi.org/10.26858/ijole.v8i2.64088>
- Syarifaturrahman, W. K., Sutarman, & Abdussamad, Z. (2021). A Morpho-phonemic Analysis on Sasak Affixation. *International Journal of Linguistics, Literature and Translation*, 4(1), 122–130. <https://doi.org/10.32996/ijllt.2021.4.1.13>
- Wieling, M., Nerbonne, J., & Baayen, R. H. (2011). Quantitative Social Dialectology: Explaining Linguistic Variation Geographically and Socially. *PLoS ONE*, 6(9), e23613. <https://doi.org/10.1371/journal.pone.0023613>
- Wijayanti, K. D., Rahadini, A. A., & Veronika, P. (2025). Konsistensi Perubahan Bunyi Morfem Dalam Bahasa Jawa. *Kamaya: Jurnal Ilmu Agama*, 8(2), 297–302. <https://doi.org/10.37329/kamaya.v8i2.3913>
- Wood, C., Torres-Chavarro, M., Tibi, S., Schatschneider, C., & Ke, F. (2025). Computer-Delivered Morphology-Focused Vocabulary Instruction: A Pilot Study. *American Journal of Speech-Language Pathology*, 34(2), 818–833. [https://doi.org/10.1044/2024\\_AJSLP-24-00319](https://doi.org/10.1044/2024_AJSLP-24-00319)
- Yan, J. (2021). Morphology and word order in Slavic languages: Insights from annotated corpora. *Voprosy Jazykoznanija*, (4), 131. <https://doi.org/10.31857/0373-658X.2021.4.131-159>
- Yue, Z. (n.d.). *Linguistic Variation in English: A Systematic Review*.