

BAB III

METODOLOGI PENELITIAN

3.1 Jenis dan Desain Penelitian

Jenis penelitian yang digunakan dalam studi ini adalah *mixed methods* atau metode campuran, dengan desain *Explanatory Sequential Mixed Methods Design*. Pemilihan metode campuran didasarkan pada kebutuhan untuk memperoleh gambaran yang komprehensif mengenai variasi fonologis dan leksikal bahasa Jawa dialek Cirebon. Menurut Creswell (2022), *mixed methods* merupakan pendekatan penelitian yang mengintegrasikan prosedur kuantitatif dan kualitatif dalam satu rangkaian studi, sehingga data yang diperoleh dapat saling melengkapi.

Dalam penelitian ini, tahap kuantitatif digunakan untuk menghitung jumlah desa yang memperlihatkan variasi fonologis dan leksikal berdasarkan distribusi geografis serta distribusi sosial antargenerasi penutur. Hasil kuantitatif tersebut memberikan peta sebaran numerik yang bersifat objektif. Selanjutnya, tahap kualitatif dilakukan untuk mendeskripsikan dan menginterpretasikan data numerik tersebut, sehingga diperoleh pemahaman yang lebih mendalam mengenai faktor-faktor sosial, budaya, dan historis yang melatarbelakangi pola variasi kebahasaan.

Desain *explanatory sequential* dipandang paling sesuai karena penelitian dimulai dengan analisis kuantitatif, lalu dilanjutkan dengan analisis kualitatif untuk menjelaskan hasil temuan numerik. Dengan cara ini, integrasi kedua metode memungkinkan penelitian ini tidak hanya memetakan distribusi variasi kebahasaan secara statistik, tetapi juga mengungkap makna sosial-budaya di balik data tersebut. Oleh karena itu, pendekatan ini diyakini mampu memberikan gambaran yang utuh tentang dinamika variasi fonologis dan leksikal bahasa Jawa dialek Cirebon, sekaligus mengungkap indikasi pergeseran bahasa yang terjadi di masyarakat penuturnya.

Dengan demikian, penggunaan metode campuran dengan desain *Explanatory Sequential* memberikan kerangka penelitian yang terstruktur dan komprehensif. Pendekatan ini sejalan dengan tradisi penelitian geolinguistik yang menekankan pentingnya integrasi antara analisis kuantitatif dan kualitatif dalam kajian variasi bahasa. Kretzschmar et al. (1996) menegaskan bahwa penelitian geolinguistik tidak cukup hanya memetakan bentuk bahasa, tetapi juga perlu menggunakan kuantifikasi data survei linguistik untuk memahami distribusi variasi, sekaligus memadukannya dengan penafsiran kualitatif terhadap faktor sosial dan historis yang melatarbelakangi perbedaan tersebut. Oleh karena itu, penerapan metode campuran dalam penelitian ini bukan hanya pilihan teknis, melainkan juga bagian dari tradisi metodologis dalam studi geolinguistik modern.

Lebih jauh, penerapan metode ini memungkinkan penelitian tidak hanya berhenti pada deskripsi bentuk variasi fonologis dan leksikal, tetapi juga dapat mengungkap pola distribusi berdasarkan faktor geografis dan sosial yang berimplikasi pada gejala *language shift*. Artinya, metode campuran membantu menjawab inti permasalahan penelitian, yakni bagaimana variasi bahasa Jawa dialek Cirebon tidak hanya terpetakan secara spasial, tetapi juga dapat diinterpretasikan sebagai indikator perubahan dan pergeseran bahasa yang sedang berlangsung. Dengan demikian, metode ini diarahkan langsung untuk menjawab pertanyaan penelitian secara menyeluruh, baik dari aspek bentuk variasi, distribusi, maupun makna sosial-budayanya.

Setelah menetapkan jenis dan desain penelitian, langkah berikutnya adalah menentukan populasi dan sampel yang menjadi dasar pengumpulan data, agar hasil penelitian mampu merepresentasikan kondisi kebahasaan di daerah pengamatan secara akurat.

3.2 Wilayah Penelitian

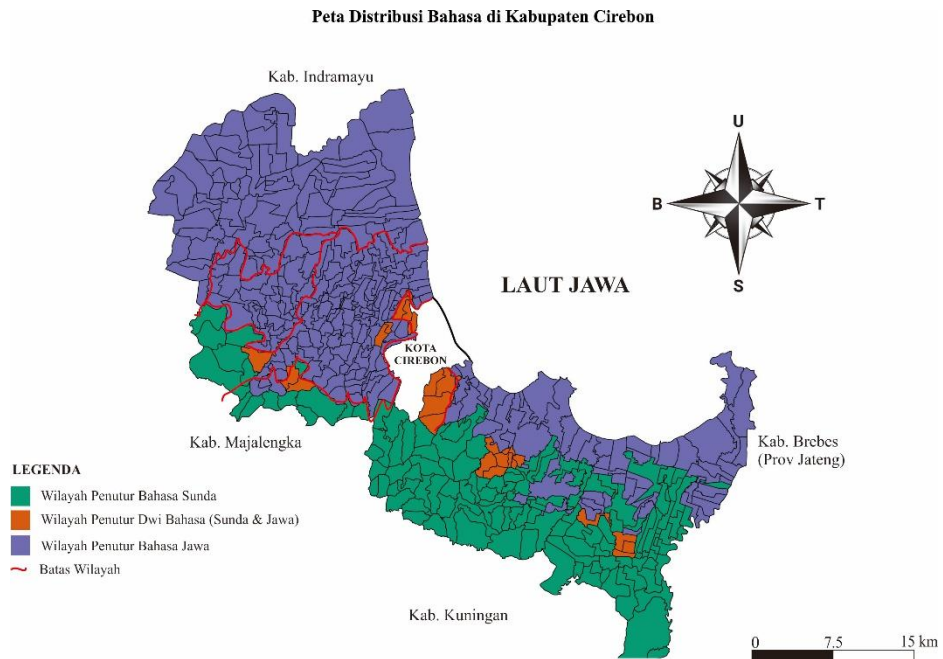
Penelitian ini dilaksanakan di wilayah Kabupaten Cirebon, Provinsi Jawa Barat, yang secara geografis terletak pada koordinat $108^{\circ}40'-108^{\circ}48'$ Bujur Timur dan $6^{\circ}30'-7^{\circ}00'$ Lintang Selatan. Secara administratif, Kabupaten Cirebon berbatasan dengan:

1. Sebelah utara: Kabupaten Indramayu.
2. Sebelah barat laut: Kabupaten Majalengka.
3. Sebelah selatan: Kabupaten Kuningan.
4. Sebelah timur: Kota Cirebon dan Kabupaten Brebes (Provinsi Jawa Tengah).

Berdasarkan data Badan Pusat Statistik (BPS) Kabupaten Cirebon (Ubaid et al., 2025), wilayah ini terdiri atas 40 kecamatan dan 412 desa. Dari jumlah tersebut, hasil observasi menunjukkan distribusi bahasa sebagai berikut:

1. Bahasa Jawa digunakan di 267 desa.
2. Bahasa Sunda digunakan di 122 desa.
3. Bahasa campuran (Jawa–Sunda) ditemukan di 23 desa.

Distribusi bahasa tersebut sejalan dengan kondisi geografis serta sejarah interaksi antarkelompok masyarakat. Peta distribusi bahasa di Kabupaten Cirebon dapat dilihat pada Gambar 1 berikut.



Gambar 1 Peta Distribusi Bahasa di Kabupaten Cirebon

Berdasarkan Gambar 3.1, bahasa Jawa (ditunjukkan dengan warna ungu pada peta) tersebar di wilayah utara, barat laut, tengah, dan timur Kabupaten Cirebon. Bahasa Sunda (warna hijau) mendominasi wilayah selatan dan barat yang berbatasan dengan Kabupaten Majalengka dan Kuningan. Adapun bahasa campuran (warna oranye) muncul secara sporadis di wilayah transisi, terutama di daerah perbatasan antara penutur bahasa Jawa dan Sunda, baik di bagian tengah, selatan, maupun barat.

Kondisi tersebut menjadikan Kabupaten Cirebon sebagai wilayah multibahasa yang menarik untuk dikaji secara geolinguistik karena memperlihatkan fenomena distribusi, kontak bahasa, dan potensi pergeseran (*language shift*) antargenerasi penutur. Pemilihan lokasi penelitian ini sekaligus didasarkan pada relevansinya dengan tujuan penelitian, yakni memetakan variasi fonologis dan leksikal bahasa Jawa dialek Cirebon yang dominan digunakan di wilayah penelitian.

3.3 Populasi dan Sampel

Dalam penelitian, populasi diartikan sebagai keseluruhan individu, objek, atau fenomena yang memiliki karakteristik tertentu dan menjadi fokus generalisasi hasil penelitian. Creswell (2022) mendefinisikan populasi sebagai kelompok besar yang menjadi sasaran penelitian. Namun, karena keterbatasan sumber daya, peneliti biasanya mengambil sebagian anggota populasi yang disebut sampel. Sampel dipilih sedemikian rupa agar dapat merepresentasikan populasi secara valid dan memungkinkan generalisasi hasil penelitian secara ilmiah.

Dalam kajian geolinguistik, pengertian populasi lebih diarahkan pada komunitas tutur di wilayah tertentu. Kretzschmar et al. (1996) menegaskan bahwa dalam survei linguistik, populasi mencakup seluruh penutur di suatu area penelitian, tetapi penelitian jarang sekali dapat menjangkau keseluruhan populasi tersebut. Oleh karena itu, digunakan sampel berupa titik pengamatan atau informan yang dipilih secara representatif untuk menggambarkan variasi bahasa. Pemilihan sampel menjadi sangat krusial karena analisis kuantitatif variasi linguistik sangat dipengaruhi oleh kualitas dan representasi data lapangan.

Berdasarkan kerangka tersebut, populasi dan sampel dalam penelitian ini dijelaskan sebagai berikut.

3.3.1 Populasi

Creswell (2022) mendefinisikan populasi sebagai keseluruhan kelompok individu atau objek yang memiliki karakteristik tertentu dan menjadi perhatian utama peneliti untuk ditarik generalisasi. Dalam konteks penelitian kebahasaan, *Pedoman Penelitian Pemetaan Bahasa* (Tim Pemetaan dan Pengembangan Bahasa Kemendikbud, 2018) menjelaskan bahwa populasi mencakup semua penutur dalam wilayah pengamatan yang secara linguistik relevan untuk didokumentasikan.

Berdasarkan pengertian tersebut, populasi penelitian ini adalah seluruh penutur bahasa Jawa dialek Cirebon yang berdomisili di Kabupaten Cirebon. Populasi ini bersifat heterogen, mencakup variasi usia, latar belakang sosial, serta distribusi wilayah (pesisir, pedalaman, hingga daerah perbatasan). Dengan demikian, populasi penelitian mencerminkan keseluruhan komunitas tutur bahasa Jawa dialek Cirebon di Kabupaten Cirebon.

3.3.2 Sampel

Menurut Creswell (2022), sampel merupakan bagian dari populasi yang dipilih untuk dianalisis lebih lanjut sehingga hasil penelitian dapat digeneralisasikan ke seluruh populasi. Sementara itu, *Pedoman Penelitian Pemetaan Bahasa* (2018) menegaskan bahwa sampel penelitian kebahasaan pada pemetaan bahasa ditentukan melalui titik pengamatan (daerah pengamatan) yang merepresentasikan keragaman bahasa di suatu wilayah.

Dalam penelitian ini, sampel berupa 267 desa di Kabupaten Cirebon yang dijadikan daerah pengamatan. Setiap desa dipilih untuk mewakili keragaman linguistik berdasarkan persebaran geografis (utara, tengah, selatan, timur, barat) serta kondisi sosial masyarakatnya. Selain itu, dari setiap desa diambil data bahasa dari dua kategori generasi, yaitu generasi tua (usia ≥ 40 tahun) dan generasi muda (usia < 40 tahun). Dengan strategi ini, sampel penelitian tidak hanya menggambarkan variasi bahasa lintas ruang, tetapi juga variasi antargenerasi penutur, sehingga mampu menunjukkan indikasi proses pergeseran bahasa (*language shift*) di Kabupaten Cirebon.

3.4 Pemilihan Informan

Pemilihan informan merupakan aspek penting dalam penelitian geolinguistik karena kualitas data sangat bergantung pada representativitas penutur yang dijadikan sumber. Menurut Kretzschmar et al. (1996), survei linguistik kuantitatif dalam penelitian atlas memerlukan informan yang mampu mencerminkan variasi bahasa, baik dari segi geografis maupun sosial. Hal ini

sejalan dengan *Pedoman Penelitian Pemetaan Bahasa* (Tim Pemetaan dan Pengembangan Bahasa Kemendikbud, 2018) yang menegaskan bahwa informan harus memenuhi syarat sebagai penutur asli, berdomisili tetap, serta masih aktif menggunakan bahasa daerah yang diteliti.

Dalam penelitian ini, jumlah informan ditetapkan sebanyak 534 orang yang tersebar di 267 desa, dengan alokasi dua informan per desa. Penentuan jumlah tersebut didasarkan pada prinsip proporsionalitas agar setiap titik pengamatan terwakili secara merata, sehingga memungkinkan analisis variasi fonologis dan leksikal secara lebih komprehensif.

Adapun kriteria informan dalam penelitian ini adalah sebagai berikut:

1. Penutur asli bahasa Jawa dialek Cirebon, lahir dan besar di desa tempat tinggalnya, sehingga menjamin keaslian variasi linguistik yang diteliti.
2. Masih menggunakan bahasa Jawa dialek Cirebon secara aktif dalam komunikasi sehari-hari, baik dalam ranah keluarga maupun masyarakat.
3. Mewakili variasi sosial, yang dibedakan menurut:
 - a. **Usia:** terdiri atas dua kategori, yaitu generasi tua (40–65 tahun) dan generasi muda (20–40 tahun).
 - b. **Pendidikan:** terdiri atas tingkat pendidikan rendah dan tinggi untuk melihat pengaruh latar belakang pendidikan terhadap variasi bahasa.
4. Bersedia menjadi informan serta berpartisipasi secara terbuka dalam seluruh proses penelitian.

Dengan kriteria tersebut, pemilihan informan diharapkan dapat memberikan gambaran menyeluruh mengenai distribusi geografis dan sosial variasi fonologis serta leksikal bahasa Jawa dialek Cirebon, sekaligus menjadi dasar untuk menganalisis fenomena pergeseran bahasa (*language shift*) antargenerasi.

3.5 Instrumen Penelitian

Instrumen penelitian adalah alat yang digunakan peneliti untuk memperoleh, mencatat, dan menganalisis data secara sistematis. Dalam penelitian

geolinguistik, instrumen tidak hanya terbatas pada alat untuk mengumpulkan data di lapangan, tetapi juga mencakup perangkat untuk menganalisis data agar dapat menggambarkan variasi bahasa berdasarkan dimensi geografis maupun sosial. Oleh karena itu, instrumen penelitian ini dibedakan menjadi dua kategori, yaitu instrumen pengumpulan data dan instrumen analisis data.

3.5.1 Instrumen Pengumpulan Data

Instrumen pengumpulan data dalam penelitian ini menggunakan daftar kosakata 200 kata dasar Swadesh. Daftar ini berfungsi untuk mengukur variasi leksikal dan fonologis yang bersifat universal karena mencakup kosakata inti yang umum digunakan dalam berbagai bahasa. Menurut Kretzschmar & Schneider (1996), daftar Swadesh sangat efektif untuk penelitian dialektologi dan geolinguistik karena dapat digunakan secara konsisten pada banyak komunitas tutur sehingga memudahkan perbandingan distribusi leksikal antarwilayah.

Untuk memudahkan pengumpulan data di lapangan, daftar kosakata ini didigitalisasi dalam bentuk Google Form dengan tautan <https://forms.gle/GY9EPt5ttxQ3DQ7u6>. Formulir daring ini memungkinkan pencatatan langsung jawaban informan sekaligus meminimalisasi kesalahan transkripsi manual. Dengan demikian, instrumen pengumpulan data menggabungkan tradisi metode survei linguistik dengan teknologi digital sebagai sarana pencatatan (Creswell & Creswell, 2022).

3.5.2 Instrumen Analisis Data

Instrumen analisis data dalam penelitian ini mencakup perangkat lunak Microsoft Excel, CSV (*Comma-Separated Values*), *shapefile* dari Badan Informasi Geospasial (BIG), serta perangkat lunak QGIS. Excel digunakan untuk melakukan tabulasi awal data kosakata hasil pengumpulan, termasuk penghitungan distribusi kuantitatif antarwilayah dan kelompok sosial. Data yang telah dikategorisasikan kemudian diekspor ke dalam format CSV agar kompatibel dengan perangkat GIS.

Shapefile yang diperoleh dari BIG menjadi peta dasar yang digunakan sebagai acuan spasial, sementara QGIS dimanfaatkan sebagai instrumen utama untuk memvisualisasikan variasi fonologis dan leksikal dalam bentuk atlas linguistik digital. Menurut Kretzschmar & Schneider (1996), analisis berbasis atlas linguistik menuntut adanya keterpaduan antara data kuantitatif linguistik dengan pemetaan geospasial, sehingga pemilihan perangkat lunak GIS menjadi instrumen analisis yang relevan.

Dengan pemisahan ini, instrumen penelitian tidak hanya mencakup sarana untuk mengumpulkan data, tetapi juga perangkat untuk menganalisis dan memvisualisasikan data, sehingga hasil penelitian dapat memberikan gambaran yang lebih menyeluruh tentang variasi bahasa Jawa dialek Cirebon berdasarkan distribusi geografis dan sosial.

Dengan demikian, penetapan populasi dan sampel dalam penelitian ini memberikan kerangka dasar yang jelas untuk memperoleh data yang representatif mengenai variasi fonologis dan leksikal bahasa Jawa dialek Cirebon. Populasi yang mencakup seluruh penutur di Kabupaten Cirebon memastikan cakupan penelitian yang luas, sementara pemilihan sampel berupa 267 desa dengan mempertimbangkan perbedaan geografis dan generasi penutur memungkinkan analisis yang lebih terarah terhadap pola variasi bahasa. Strategi ini sejalan dengan prinsip penelitian geolinguistik yang menekankan keterwakilan spasial dan sosial, sehingga data yang diperoleh dapat dianalisis secara komprehensif pada tahap pengumpulan data berikutnya.

3.6 Teknik Pengumpulan Data

Pengumpulan data merupakan tahap penting dalam penelitian karena kualitas analisis sangat ditentukan oleh bagaimana data diperoleh dari lapangan. Creswell & Creswell (2022) menegaskan bahwa dalam penelitian *mixed methods*, data dapat dikumpulkan melalui prosedur kualitatif maupun kuantitatif, baik

secara berurutan maupun simultan, untuk menjawab pertanyaan penelitian secara komprehensif.

Selain itu, pendekatan yang dikemukakan Kretzschmar et al. (1996) dalam penyusunan *linguistic atlas* digunakan untuk memastikan setiap data lapangan mencerminkan variasi nyata yang diujarkan penutur asli dalam konteks sosialnya. Adapun tahap pengumpulan data dalam penelitian ini adalah sebagai berikut.

1. Penentuan Wilayah Penelitian

Penentuan wilayah penelitian dilakukan dengan mengacu pada batas administratif Kabupaten Cirebon yang terdiri atas 40 kecamatan, 12 kelurahan, dan 412 desa. Namun, penelitian ini difokuskan pada desa-desa yang mayoritas penduduknya menggunakan bahasa Jawa dialek Cirebon dalam interaksi sehari-hari. Pendekatan ini sejalan dengan prinsip geolinguistik yang menekankan keterwakilan daerah pengamatan untuk menggambarkan keragaman bahasa secara spasial (Kretzschmar et al., 1996).

2. Observasi Lapangan

Observasi langsung dilakukan di lapangan untuk memastikan vitalitas penggunaan bahasa Jawa dialek Cirebon di setiap desa yang menjadi sampel penelitian. Menurut *Pedoman Pemetaan Bahasa* (Tim Pemetaan dan Pengembangan Bahasa Kemendikbud, 2018), observasi ini berfungsi untuk memverifikasi bahwa informan benar-benar menggunakan bahasa daerah dalam kehidupan sehari-hari, sehingga data yang diperoleh autentik dan merepresentasikan kondisi kebahasaan setempat.

3. Wawancara Langsung

Wawancara dilakukan dengan mengacu pada daftar kosakata dasar Swadesh yang berjumlah 200 kata. Daftar ini digunakan untuk mengungkap variasi fonologis dan leksikal secara konsisten di setiap titik pengamatan. Creswell (2022) menekankan bahwa wawancara terstruktur dengan instrumen yang jelas membantu menjaga keseragaman data, sehingga hasil penelitian dapat dibandingkan antardaerah maupun antargenerasi.

4. Pencatatan Digital

Hasil wawancara dicatat secara digital menggunakan Google Form melalui tautan <https://forms.gle/GY9EPt5ttxQ3DQ7u6>. Pencatatan digital dipilih untuk meningkatkan akurasi, memudahkan verifikasi data, serta mengurangi risiko kehilangan informasi. Praktik ini sejalan dengan perkembangan metode dokumentasi linguistik modern yang mengintegrasikan teknologi digital dalam pengumpulan data lapangan.

5. Pengelompokan Data

Setelah data terkumpul, tahap selanjutnya adalah pengelompokan berdasarkan faktor geografis (utara, tengah, selatan, timur, barat Kabupaten Cirebon) dan faktor sosial (generasi tua dan muda). Pengelompokan ini mengacu pada prinsip analisis dialektologi dan geolinguistik yang menekankan hubungan antara variasi bahasa dengan persebaran spasial dan demografi penutur (Kretzschmar & Schneider, 1996).

3.7 Teknik Analisis Data

Analisis data dalam penelitian ini dilakukan secara bertahap dengan mengacu pada prinsip penelitian geolinguistik yang memadukan pendekatan kuantitatif dan kualitatif. Pendekatan kuantitatif digunakan untuk mengklasifikasikan dan menghitung distribusi variasi fonologis maupun leksikal berdasarkan faktor geografis dan sosial. Sementara itu, pendekatan kualitatif dipakai untuk menafsirkan makna distribusi tersebut dalam konteks perubahan bahasa. Dengan demikian, analisis tidak hanya berfokus pada frekuensi kemunculan variasi, tetapi juga pada interpretasi geolinguistik yang mengungkap pola konservasi, transisi, dan inovasi dalam Bahasa Jawa dialek Cirebon.

1. Entri dan Penyusunan Data

Data yang telah terekam melalui Google Form diekstrak ke dalam format spreadsheet, kemudian diperiksa kembali untuk memastikan kelengkapan dan

konsistensi. Seluruh data selanjutnya dimasukkan ke dalam basis data komputer agar memudahkan pengelompokan dan analisis.

2. Klasifikasi Data

Klasifikasi data dilakukan dengan mengelompokkan hasil pengumpulan data fonologis dan leksikal berdasarkan distribusi geografis dan sosial. Distribusi geografis dibagi ke dalam empat wilayah utama di Kabupaten Cirebon, yaitu wilayah utara, barat laut, selatan, dan timur. Sementara itu, distribusi sosial dikategorikan berdasarkan rentang usia per sepuluh tahun (20–30 dan 31–40 sebagai kelompok generasi muda; 41–50 dan 51–65 sebagai kelompok generasi tua), serta latar belakang pendidikan (tidak sekolah, sekolah dasar, sekolah menengah, hingga perguruan tinggi). Pendekatan klasifikasi ini sejalan dengan pedoman penelitian variasi bahasa yang menekankan pentingnya distribusi sosial dan geografis dalam menganalisis fenomena kebahasaan (Creswell, 2022; Kretzschmar et al., 1996).

3. Analisis Pola Variasi

Analisis pola variasi dilakukan dengan mengidentifikasi gejala fonologis berupa variasi vokal, variasi konsonan, proses elisi (penghilangan bunyi), dan epentesis (penambahan bunyi). Selain itu, variasi leksikal dianalisis berdasarkan jumlah varian yang muncul, mulai dari satu hingga dua belas variasi pada kosakata tertentu.

4. Dokumentasi Digital

Data yang telah diklasifikasikan kemudian didokumentasikan dalam format digital menggunakan Microsoft Excel untuk mempermudah kategorisasi dan pengolahan. Selanjutnya, data diekspor ke dalam format CSV (*Comma Separated Values*) agar kompatibel dengan perangkat lunak pemetaan geospasial.

5. Visualisasi Data

Visualisasi data dilakukan dengan perangkat lunak Geographic Information System (QGIS). Data CSV yang telah diolah diproyeksikan pada peta shapefile

Kabupaten Cirebon, kemudian diberikan skema warna menggunakan ColorBrewer. Proses ini menghasilkan peta linguistik yang menampilkan zona konservatif, transisional, dan inovatif, sebagaimana dianjurkan dalam praktik pemetaan dialektologi modern (Tim Pemetaan dan Pengembangan Bahasa Kemendikbud, 2018).

6. Interpretasi Geolinguistik

Tahap akhir adalah interpretasi geolinguistik, yaitu membandingkan hasil pemetaan variasi fonologis dan leksikal dengan teori variasi bahasa serta model penyebaran inovasi linguistik. Interpretasi ini mengacu pada kerangka teori variasi bahasa (Chambers & Trudgill, 2004; Labov, 1994) untuk memahami dinamika pergeseran bahasa dan pembentukan zona linguistik di Cirebon