Sutasoma 13 (1) (2025)

Sutasoma:
Jurnal Sastra Jawa

http://journal.unnes.ac.id/sju/index.php/sutasoma

A Comparison of Cognate Dialect Features in Javanese, Sundanese, Cirebon, and Indramayu

Ridha Mashudi Wibowo
Universitas Gadjah Mada
Corresponding Author: ridha@ugm.ac.id

DOI: https://doi.org/10.15294/sjsj.v13i1.21878
Submitted: February 28th 2025 **Accepted:** June 12th 2025 **Published:** June 27th 2025

Abstract

This study aims to identify cognates based on form and meaning among Javanese, Sundanese, Cirebon, and Indramayu vocabularies and analyze their dialectal relationships. Using a list of one hundred Swadesh vocabulary items sourced from dictionaries of the four languages, the data were phonetically transcribed and compared to identify shared lexemes. Dialectometric analysis was then applied to quantify lexical distances and classify dialectal relationships. The findings show that the four languages share 28 cognates and exhibit varying degrees of linguistic proximity: Javanese and Cirebon show the closest relationship, while other pairs are classified as distinct dialects. While the study integrates qualitative and quantitative approaches and contributes insights into less-studied dialects such as Cirebon and Indramayu, it is limited by its reliance on secondary data and lack of fieldwork. Future research should incorporate direct elicitation from native speakers, phonological correspondence analysis, the etymological distinction between cognates and loanwords, and sociolinguistic and geographical mapping to deepen the understanding of dialectal variation in Java.

Keywords: *dialectology, cognate, dialectometry, Javanese, Sundanese, Cirebon, Indramayu*

p-ISSN 2252-6307
e-ISSN 2686-5408

INTRODUCTION

Javanese and Sundanese, while classified as distinct languages, share similarities and differences across phonological, lexical, morphological, and syntactic levels. However, as part of the Austronesian language family (Adelaar & Himmelmann, 2005), both also have several linguistic similarities—the cultural and historical backgrounds of each influence this. In addition to having linguistic differences, Javanese and Sundanese have several linguistic similarities, both at the level of phonology, lexicon, morphology, and syntax. This is proven by the same/similar vocabulary in both languages. The similarity of linguistic words is called a *cognate*. Cognat */kog-nat/* means 1 identical; similar; 2 related because of blood relations (descent); 3 *Ling* related because it comes from the same language (about words from two or more languages) (https://kbbi.web.id/ kognat).

Campbell (2004: 45) explains that cognates provide important evidence for reconstructing proto-languages and forming language families. Through the analysis of cognates, linguists can reconstruct proto-languages and trace the kinship relationships between languages (Trask, 1996, pp.

112-115). Trask (1996) emphasizes that systematic comparison of cognates allows linguists to distinguish between words inherited from a common ancestor language and loanwords (Durie & Ross, 1996, pp. 58-63). Durie & Ross (1996), in *The Comparative Method Reviewed,* provide in-depth insights into how the comparative method, with a focus on cognates, is used to construct language family trees (Bellwood, 1997, pp. 88-92; Ewing, 2007: 12-15; Thomason & Kaufman, 1988: 15-22).

This research is important in the context of regional dialectology because the areas of West Java and Central Java are areas where several ethnolinguistic groups meet, which produce complex and interesting dialectal boundaries to study. In Sociolinguistics and Social Theory, Ewing (2007) highlights how social and geographical contexts shape dialectal variation in multilingual areas. Bellwood (1997) observes that interactions between language groups in this area have produced a diverse linguistic landscape with significant dialectal variation. In addition, a better understanding of the dynamics of language contact and language change in multilingual areas can be achieved through comparative dialectal studies (Thomason & Kaufman, 1988: 15-22). Thomason and Kaufman (1988) argue that intensive language contact can cause substantial linguistic change, including convergence and mixing language features."

"Cognates play an important role in the study of comparative historical linguistics. Cognates are defined as words in different languages derived from the same root language" (Campbell, 2004, p. 45). Campbell (2004: 45) explains that cognates provide important evidence for reconstructing proto-languages and forming language families. Through cognate analysis, linguists can reconstruct proto-languages and trace kinship relationships between languages (Trask, 1996, pp. 112-115). Trask (1996) emphasizes that systematic comparison of cognates allows linguists to distinguish between words inherited from the same ancestral language and loanwords. This study is important in the context of regional dialectology because the regions of West Java and Central Java are areas where several ethnolinguistic groups meet, which produce complex and interesting dialectal boundaries to study (Bellwood, 1997, pp. 88-92). Bellwood (1997) observes that interactions between language groups in this region have resulted in a diverse linguistic landscape with significant dialectal variation. In addition, a better understanding of the dynamics of language contact and language change in multilingual areas can be achieved through comparative dialectal studies (Thomason & Kaufman, 1988: 15-22). Thomason and Kaufman (1988) argue that intensive language contact can lead to substantial linguistic change, including convergence and blending of language features."

The conditions of the Cirebon and Indramayu languages are often compared to those of Javanese and Sundanese. Cirebon and Indramayu languages are often considered to be two different languages. Several times, confessions from cultural circles have emphasized this. Apart from these confessions, the Cirebon and Indramayu languages have many similarities in phonology, lexicon, morphology, and syntax. Both also absorb several vocabularies from Javanese and Sundanese. With the phenomenon of vocabulary absorption, the study of the comparison of dialectal vocabulary of Javanese, Sundanese, Cirebon, and Indramayu becomes interesting. Thus, the focus of the discussion in this article is aimed at what the cognates are based on their

form and meaning among the vocabulary of Javanese, Sundanese, Cirebon, and Indramayu, and what the dialectal relationships are between Javanese, Sundanese, Cirebon, and Indramayu? In line with the two research problems, this research aims to describe cognates based on their form and meaning among Javanese, Sundanese, Cirebon, and Indramayu vocabulary and to explain the dialectal relationships that exist between Javanese, Sundanese, Cirebon, and Indramayu languages.

Before this article was written, research was conducted on the relationship between the Javanese, Sundanese, Cirebon, and Indramayu languages. The research in question was a thesis that studied the proto-Austronesian language in Sundanese, Cirebon, Indramayu, and Javanese (Wiyanti, 2005). Research also studied the Sundanese language spoken in Dukupuntang District, Cirebon Regency (Winata, 2019). Another dialectological study was conducted by Itaristanti and Idah Faridah Laily (2023). They both studied the relic elements of the Cirebon language with a diachronic dialectological study. Another dialectological study compared Javanese spoken in Tanjung District, Brebes, with Javanese spoken in Losari District, Cirebon (Budiawan, 2019). Unlike the four studies, this study used a list of one hundred Swadesh vocabularies as a data provider instrument, used Javanese, Sundanese, Cirebon, and Indramayu dictionaries as data sources, and conducted dialectometric calculations at the lexicon level by ignoring language varieties.

METHOD

Dialectology, as a branch of linguistics, is concerned with studying language variation. Kridalaksana (2011:49) defines dialectology as the study of language varieties by treating them as unified structures. At the same time, Mahsun (1995:11) describes dialectology as the science of dialects that examines isolect differences as integrated systems. In this study, a dialectal comparison was conducted using a list of one hundred Swadesh vocabulary items collected from Javanese, Sundanese, Cirebon, and Indramayu. The Swadesh list is widely used in comparative linguistics to identify potential cognates (Lehmann, 1992).

The lexical data were obtained from published dictionaries of the four languages. While this method allows for systematic comparison, it is important to acknowledge that relying solely on dictionary sources may limit the representativeness of the data. Dictionaries often prioritize standardized or literary forms and may not fully capture the living language, dialectal nuances, or regional variants actively used by speakers. Thus, while this method provides a practical starting point, it is best viewed as preliminary.

Dialectometric analysis was used to calculate lexical distances between the language varieties using the following formula.

$$\frac{(S \times 100)}{n} = d\ \%$$

S = number of differences between observation points
n = number of items compared
d = percentage of lexical difference

The classification of dialectal relationships follows standard dialectometry (Séguy, 1971):

>81% → different language
51–80% → different dialect
31–50% → different subdialect
21–30% → speech variation

<20% → no difference

It should be noted that although some language pairs (e.g., Javanese and Cirebon) fall under the "<20%" threshold and are thus categorized as showing "no difference" by the metric, this does not necessarily imply absolute uniformity. Instead, it reflects high lexical similarity within the parameters of the Swadesh list. In dialectometry, "no difference" indicates a high degree of shared vocabulary in the measured items but does not rule out phonological, morpho-syntactic, or sociolinguistic differences. Therefore, terms such as "subdialect" or "speech variation" may still apply to capture subtle distinctions not detected by lexical metrics alone.

Furthermore, cognates were identified based on surface-level formal and semantic similarity. Given the extensive contact among the four language groups, some lexical similarities may be due to borrowing rather than common inheritance. Therefore, the cognates presented in this study should be regarded as *potential* cognates, pending further etymological validation.

It is important to note that this study's cognates identification was based on formal and semantic similarity as observed in dictionary data, without extensive etymological verification. Given the long-standing contact between Javanese, Sundanese, Cirebon, and Indramayu, some lexical similarities may result from borrowing rather than shared inheritance. Therefore, the cognates identified here should be regarded as potential cognates, pending confirmation through historical reconstruction and etymological analysis. This limitation is acknowledged and will be addressed in future studies incorporating comparative phono-logical methods and proto-language reconstruction.

RESULT AND DISCUSSION

Cognates without Form Difference

Table 1. Cognates Without Form Differences

Gloss	Java	Sunda	Cirebon	Indramayu
heart	*galih*	*galih*	*galih*	*galih*
blood	*gətih*	*gətih*	*gətih*	*gətih*
wind	*aŋin*	*aŋin*	*aŋin*	*aŋin*
bird	*manuʔ*	*manuʔ*	*manuʔ*	*manuʔ*
flower	*kəmbaŋ*	*kəmbaŋ*	*kəmbaŋ*	*kəmbaŋ*
mountain	*gunuŋ*	*gunuŋ*	*gunuŋ*	*gunuŋ*
hear	*pirəŋ*	*pirəŋ*	*pirəŋ*	*pirəŋ*

Table 1 shows that Javanese, Sundanese, Cirebon, and Indramayu have the same vocabulary for the glosses *hati, darah, angin, burung, bunga, gunung,* and *dengar*. For the gloss 'hati', all four languages have the same *galih* form in terms of form and meaning, but differ in language variety. *Galih* is included in the polite language variety in Javanese and Indramayu, while in Sundanese and Cirebon, it is not. *Manuʔ* in all four languages is the same in form, meaning, and variety. *Manuʔ* is not included in the polite language variety in all four languages. For the glosses *angin, bunga,* and *gunung*, all four languages have the same vocabulary in terms of form and meaning. However, *gunuŋ*, *aŋin*, and *kəmbaŋ* have different language varieties. In Javanese, Sundanese, and Cirebon, all three forms are not included in the polite language variety. The situation is different in Indramayu. In Indramayu, all three forms are included in the polite language variety and the non-refined language variety. The subtle language variety includes the *pirəŋ* form in the four languages.

Cognates with Different Shapes

1. Gloss 'Rain'

Table 2. Gloss 'Rain'

Gloss	Java	Sunda	Cirebon	Indramayu
rain	*udan*	*hujan*	*hudan*	*udan*

Javanese and Indramayu have the same form for the gloss 'rain', namely *udan*. In Sundanese, the gloss 'rain' has the form *hujan*. Meanwhile, in Cirebonese, the gloss 'rain' has the form *hudan*. The forms *udan*, *hujan*, and *hudan* are considered the same because the differences in their forms are not significant. The three forms only differ in their first and third phonemes. Phonetic differences like this are common in language change (Trask, 1996) and are often observed in cognate studies (Trudgill, 2004). The difference in the first phoneme of the three forms is θ:h:h, while the difference in the third phoneme is [d:j:d]. Regarding language variety, the four forms are included in the non-refined language variety in their respective languages.

2. Gloss 'Nose'

Table 3. Gloss 'Nose'

Gloss	Java	Sunda	Cirebon	Indramayu
nose	*iruŋ*	*iruŋ*	*iruŋ* *hiruŋ*	*iruŋ*

Javanese, Sundanese, Cirebon, and Indramayu have one form in common: *iruŋ*. However, in Cirebon, the nasal gloss has two similar forms: *iruŋ* and *hiruŋ*. The difference between the two forms does not mean that they are considered different because the difference is only in the first phoneme. The difference is [θ:h]. In the table above, the form of 'nose' in Javanese, Sundanese, and Cirebon is included in the non-refined language variety in their respective languages. Unlike the language varieties in Javanese, Sundanese, and Cirebon, the *iruŋ* form in Indramayu is included in both the polite and non-refined language varieties.

3. Gloss 'Teeth'

Table 4. Gloss 'Teeth'

Gloss	Java	Sunda	Cirebon	Indramayu
tooth	*untu*	*huntu*	*untu* *huntu*	*untu*

The pattern of shape differences for gloss 'tooth' in the four languages is somewhat similar to that for gloss 'nose'. The difference lies in the first phoneme, which is [θ:h:h:θ]. Cirebonese has the forms *untu* and *huntu*. The latter form is possibly a Sundanese loanword because it is the same. In terms of language variety, all forms of gloss gigi in Sundanese, Cirebonese, and Indramayu are non-standard language varieties. The form *untu* in Javanese is included in the *krama ngoko* variety.

4. Gloss 'You'

Table 5. Gloss 'You'

Gloss	Java	Sunda	Cirebon	Indramayu
you	*sira*	*siya*	*sira* *ira*	*sira* *ira*

The forms for the gloss *kau* in Javanese, Sundanese, Cirebon, and Indramayu have several variations. There is the form *sira* in Javanese, Cirebon, and Indramayu and the form *siya* in Sundanese. The forms *sira* and *siya* are considered the same because they only differ by one phoneme. The difference is r:y. In Cirebon and Indramayu, the form *sira* has a derivative form called *ira*. All forms for the gloss *kau* in Sundanese, Cirebon, and Indramayu in the table above are included in the non-refined language

variety. In Javanese, the form *sira* is included in the archaic variety.

5. Gloss 'That'

Table 6. Gloss 'That'

Gloss	Java	Sunda	Cirebon	Indramayu
that	*iku* *niku* *tiku*	*itu*	*niku*	*iku* *ikuh* *niku*

Niku , *tiku* , *iku* , *ikuh* , and *itu* forms for that gloss. All of these forms are considered the same because of the combination and order of their vowels. The forms *iku* (Javanese), *iku* (Indramayu), *ikuh (* Indramayu), *tiku* (Javanese), *itu* (Sundanese), and *niku* (Cirebon) are included in the non-refined language variety. The polite language variety includes *niku* (Javanese) and *niku* (Indramayu).

6. Gloss 'Correct'

Table 7. Gloss 'Correct'

Gloss	Java	Sunda	Cirebon	Indramayu
correct	*banər* *bənər*	*bənər*	*bənər*	*bənər*

All four languages have the form *bənər* for correct gloss. Javanese has one additional form, *banər*, which differs by only one phoneme from *bənər*. The difference between the two forms is only in the second phoneme. The difference is aːə. In the table above, all forms for correct gloss in Javanese, Sundanese, Cirebon, and Indramayu are non-refined language varieties.

7. Gloss 'Life'

Table 8. Gloss 'Life'

Gloss	Java	Sunda	Cirebon	Indramayu
life	*iḍup*	*hirup*	*hirup*	*hidup* *hirup*

The forms considered the same in all four languages for the gloss 'live' are *hidup*, *hirup*, and *iḍup*. The three forms' differences lie only in the first and third phonemes. The first phoneme difference is [h:h:θ], and the third phoneme difference is [ɗ:r:ɖ]. The form *hidup* in Indramayu is included in the polite and non-refined language variety. The other forms, namely *iḍup* (Javanese), *hirup* (Sundanese), *hirup* (Cirebon), and *hirup* (Indramayu), are included in the non-refined language variety in their respective languages.

8. Gloss 'Big'

Table 9. Gloss 'Large'

Gloss	Java	Sunda	Cirebon	Indramayu
big	*agəŋ* *agrəŋ* *aguŋ*	*agöŋ*	*agəŋ* *aguŋ*	*agəŋ*

In the 'big' glosses of the four languages, the forms *agrəŋ*, *agəŋ*, *aguŋ,* and *agöŋ are obtained*. The four forms differ in the third and fourth phonemes. The third phoneme difference is [r:θ:θ:θ], and the fourth phoneme difference is [ə:ə:u:ö]. The forms *agəŋ* (Javanese, Cirebon, and Indramayu) and *agöŋ* (Sundanese) are included in the polite language variety. The non-refined language variety includes *aguŋ* (Javanese and Cirebon) and *agrəŋ* (Javanese).

9. Gloss 'Four'

Table 10. Gloss 'Four'

Gloss	Java	Sunda	Cirebon	Indramayu
four	*papat* *pat*	*opat*	*papat* *pat*	*papat*

There are three forms for the gloss 'four', namely *papat* (Javanese, Cirebon, and

Indramayu), *opat* (Sundanese), and *pat* (Javanese and Cirebon). The three forms differ in the first and second phonemes. The first phoneme difference is [p:θ:θ]. The second phoneme difference is [a:o:θ]. All forms in the table above are non-refined language varieties in their respective languages.

10. Gloss 'Green'

Table 11. Gloss 'Green'

Gloss	Java	Sunda	Cirebon	Indramayu
green	*ijo* *iju* *jo*	*hejo*	*hijo* *ijo*	*ijo*

In the green gloss of the four languages, the forms *hejo* (Sundanese), *hijo* (Cirebon), *hijo* (Cirebon), *ijo* (Javanese, Cirebon, and Indramayu), *iju* (Javanese), and *jo* (Javanese) are obtained. The six forms differ in the first, second, and fourth phonemes. The first phoneme difference is [h:h:h:θ:θ:θ]. The second phoneme difference is [e:i:i:i:i:θ]. The fourth phoneme difference is [o:o:o:o:u: o]. The six forms are included in the non-refined language variety in their respective languages.

11. Gloss 'Black'

Table 12. Gloss 'Black'

Gloss	Java	Sunda	Cirebon	Indramayu
black	*irəŋ* *iṭəŋ* *liṭəŋ*	*hidöŋ*	*ahirəŋ* *hirəŋ* *irəŋ*	*irəŋ*

From the table above, the forms *ahirəŋ* (Cirebon), *hirəŋ* (Cirebon), *irəŋ* (Javanese, Cirebon, and Indramayu), *hidöŋ* (Sundanese), *iṭəŋ* (Javanese), and *liṭəŋ* (Javanese) are obtained. The six forms differ in the first, second, fourth, and fifth phonemes. The first phoneme difference is [a:θ:θ:θ:θ:θ]. The second phoneme difference is [h:h:θ:h:θ:l]. The fourth phoneme difference is [r:r:r:d:ṭ:ṭ]. The fifth phoneme difference is [ə:ə:ə:ö:ə:ə]. The six forms are included in the non-refined language variety in their respective languages.

12. Gloss 'Stone'

Table 13. Gloss 'Stone'

Gloss	Java	Sunda	Cirebon	Indramayu
rock	*watu*	*batu*	*watu*	*watu*

In the table of 'batu' above, the forms of *batu* and *watu* are obtained from the four languages. The difference between the two forms is only the first phoneme. The difference is [b:w]. Both forms are included in the non-refined language variety in their respective languages.

13. Gloss 'Road'

Table 14. Gloss 'Road'

Gloss	Java	Sunda	Cirebon	Indramayu
road	*dalan* *jalan*	*dalan* *jalan*	*dalan*	*dalan* *jalan*

Dalan and *jalan* for the gloss *jalan* are obtained from the four languages. The forms of *dalan* and *jalan* differ only in their first phonemes. The difference in the first phoneme is d:j. Regarding language variety, the forms of *dalan* and *jalan* are included in the non-refined language variety in their respective languages.

14. Gloss 'Salt'

Table 15. Gloss 'Salt'

Gloss	Java	Sunda	Cirebon	Indramayu
salt	*uyah*	*uyah*	*uyah* *wuyah*	*uyah*

As for the forms obtained for road glosses, the four languages differ only in their first phonemes: uyah and *wuyah*. The difference in the first phoneme of the *uyah* and *wuyah* forms is [θ:w]. Both forms in their respective languages are included in the non-refined language variety.

15. Gloss 'Sit'

Table 16. Gloss 'Sitting'

Gloss	Java	Sunda	Cirebon	Indramayu
sit	*liŋgih* *liŋgɛh*	*liŋgih*	*aliŋgih*	*liŋgih*

In the sitting gloss, the forms *aliŋgih*, *liŋgih*, and *liŋgɛh* are obtained from the four languages. The three forms differ in their first and sixth phonemes. The difference in the first phoneme is a:θ:θ. The difference in the sixth phoneme is [i:i:ɛ]. The polite language variety includes the forms liŋgih in Sundanese and Indramayu. The forms *liŋgih* (Jawa-nese), *liŋgɛh* (Javanese), and *aliŋgih* (Cire-bon) are included in the non-refined language variety in their respective languages.

16. Gloss 'Count'

Table 17. Gloss 'Count'

Gloss	Java	Sunda	Cirebon	Indramayu
count	*etaŋ* *petaŋ*	*etaŋ*	*ketaŋ*	*etaŋ*

The forms *ketaŋ* (Cirebon), *petaŋ* (Javanese), and *etaŋ* (Javanese, Sundanese, and Indramayu) are obtained for the counting glosses of Javanese, Sundanese, Cirebon, and Indramayu. The difference between these forms is in the first phoneme. The difference in the first phoneme is [k:p:θ]. The forms *etaŋ* (Javanese and Indramayu) and *petaŋ* (Javanese) are included in the polite language variety in their respective languages. Meanwhile, the forms *etaŋ* (Sundanese) and *ketaŋ* (Cirebon) are included in the non-refined language variety in their respective languages.

17. Gloss 'Dog'

Table 18. Gloss 'Dog'

Gloss	Java	Sunda	Cirebon	Indramayu
dog	*asu*	*asu* *basu*	*asu*	*asu*

The basic forms are obtained from the four languages in the gloss' dog'. The difference between the basic (Sundanese) and basic (Javanese, Sundanese, Cirebon, and Indramayu) forms is only in the first phoneme. The difference in the first phoneme is [b:θ]. The basic forms are included in the non-refined language variety in their respective languages.

Two Cognates without Difference in Form

Table 19. Two Cognates Without Difference in Form

Gloss	Java	Sunda	Cirebon	Indramayu
child	*anaʔ* *putra*	*anaʔ* *putra*	*anaʔ* *putra*	*anaʔ* *putra*

The table above shows that the *anaʔ* and *putra* forms are obtained from the four languages. Both forms exist in all four languages. The non-refined language variety includes the *anaʔ* form in all four languages. The *putra* form is included in the polite language variety in Javanese and Indramayu and the non-refined language variety in Sundanese and Cirebon.

Two Cognates with Different Forms

Table 20. Two Cognates with Different Forms

Gloss	Java	Sunda	Cirebon	Indramayu
wife	*ɛstri* *garwa* *gərwa*	*istri* *garwa*	*estri* *stri* *garwa*	*istri* *garwa*

Two forms are obtained in the gloss 'wife' of the four languages, along with several variations. The variant forms are *istri* (Sundanese and Indramayu), *estri* (Cirebon), *ɛstri* (Javanese), and *stri* (Cirebon). The second forms are *garwa* (Javanese, Sundanese, Cirebon, and Indramayu) and *gərwa* (Javanese). The forms *istri*, *estri*, *ɛstri*, and *stri* differ only in their first phoneme. The difference is [i:e:ɛ:θ]. The form *ɛstri* is included in the polite language variety in Javanese. The forms *istri* (Sundanese and Indramayu), *estri* (Cirebon), and *stri* (Cirebon) are included in the non-refined language variety in their respective languages. The forms *garwa* and *gərwa* differ only in their second phoneme. The difference in the second phoneme referred to is [a:ə]. The forms *garwa* (Javanese and Indramayu) and *gərwa* (Javanese) are included in the polite language variety in their respective languages. Meanwhile, the form *garwa* (Sundanese and Cirebon) is included in the non-refined language variety in their respective languages.

Two Cognates, One of Which is Different in Shape

Table 21. Two Cognates, One of Which Has a Different Shape

Gloss	Java	Sunda	Cirebon	Indramayu
I	*kula* *kawula*	*kula* *kawla*	*kula* *kawla*	*kula* *kawula*
if	*bilih* *kalamun* *kəlamun* *lamon* *lamun*	*bilih* *lamun* *mun*	*bilih* *lamun* *namun* *mun*	*bilih* *lamon* *lamun*

Two forms are obtained in the gloss 'aku', one of which varies. The form kula is obtained from the four languages for the gloss 'aku'. The form *kula* exists in all four languages without any differences in form. The form *kula* is included in the polite language variety in Javanese and Indramayu. In Sundanese, the form *kula* is included in the non-refined language variety. In Cirebon, the form *kula* is included in the polite and non-refined language varieties. In addition, the forms *kawula* (Javanese and Indramayu) and *kawla* (Sundanese and Cirebon) are obtained. The two forms differ only in their fourth phoneme, [u:θ]. The form *kawula* is included in the polite language variety in Indramayu and the non-refined language variety in Javanese. The form *kawla* is included in the polite language variety in Cirebon and the non-refined language variety in Sundanese.

Furthermore, forms are obtained in the gloss 'if' from the four languages, one of which is present in all four languages without any phonological differences. This form is *bilih*. The *bilih form* is included in the polite language variety in the four languages. In addition to bilih, there are also the forms *kalamun* (Javanese), *kəlamun* (Javanese), *lamun* (Javanese, Sundanese, Cirebon, and Indra-mayu), *lamon* (Javanese and Indramayu), *namun* (Cirebon), and *mun* (Sundanese and Cirebon). The six forms differ in the first, second, third, fourth, and sixth phonemes. The first phoneme difference is [k:k:θ:θ:θ:θ]. The second phoneme difference is [a:a:θ:θ:θ]. The third phoneme difference is

[l:l:l:l:n:θ]. The fourth phoneme difference is [a:a:a:a:a:θ]. The sixth phoneme difference is u:u:u:o:u:u. The forms *kalamun* (Javanese), *kəlamun* (Javanese), *lamun* (Javanese, Sundanese, Cirebon, and Indramayu), *lamon* (Javanese and Indramayu), *namun* (Cirebon), and *mun* (Sundanese and Cirebon) are included in the non-refined language varieties in their respective languages.

Dialectometric Calculations

The dialectometric comparison among Javanese (Jw), Sundanese (Snd), Cirebon (Crb), and Indramayu (Imy) was conducted to quantify the degree of lexical differences between each language variety. The results are shown in Table 22.

Table 22. Dialectometric Comparison among Four Dialects

Language Pair	Number of Differences	Percentage Difference (%)	Classification
Jw: Snd	23	32.8%	Subdialect difference
Jw: Crb	13	18.5%	No difference
Jw: Imy	38	54.2%	Dialect difference
Snd: Crb	40	57.1%	Dialect difference
Snd: Imy	54	77.1%	Dialect difference
Crb: Imy	44	62.8%	Dialect difference

Based on the dialectometric results, the lexical distance between Javanese and Cirebon is the smallest (18.5%), suggesting they are part of the same variety without significant dialectal diffe-rence. Javanese and Sundanese show a moderate difference (32.8%), corresponding to a subdialectal relationship. All other pairs (Jw: Imy, Snd: Crb, Snd: Imy, Crb: Imy) exceed 50% lexical diver-gence, classifying them as different dialects under standard dialectometric interpretation (see Séguy, 1971).

While this study focuses on lexical comparisons and quantitative dialectometric calculations, it is important to recognize that sociolinguistic factors—such as speaker attitudes, language pres-tige, and identity—also significantly shape dialect boundaries. For instance, speakers may align themselves with certain linguistic forms not purely due to structural similarity, but also due to socio-cultural affiliations or prestige considerations. In regions where language contact is intense, such as in the border areas between Central and West Java, dialect convergence or divergence can be influenced by political identity, urbanization, educational policy, and inter-ethnic interaction. Therefore, this study's lexical similarities or differences should be interpreted with sociolinguistic realities. Future research should incorporate sociolinguistic field-work, including interviews and perceptual dialect-ology methods, to better understand how speakers conceptualize and negotiate their linguistic identi-ties.

CONCLUSION

This study has shown that Javanese, Sundanese, Cirebon, and Indramayu share 28 lexical items considered cognates based on form and meaning. These include glosses such as nose, tooth, heart, blood, I, you, child, wife, that, true, life, big, four, green, black, wind, rain, bird, dog, stone, flower, road, mountain, salt, sit, hear, count, and if. Dialectometric calculations reveal varying levels

of lexical distance: Javanese and Cirebon show no significant difference (18.5%), while Javanese and Sundanese show subdialectal variation (32.8%). Other pairs—Javanese–Indramayu (54.2%), Sundanese–Cirebon (57.1%), Sundanese–Indramayu (77.1%), and Cirebon–Indramayu (62.8%)—exhibit different dialects. These results suggest that while mutual intelligibility may still exist in dialect pairs with moderate divergence, communication is likely more seamless among varieties classified as having no difference or only subdialectal variation.

However, these lexical similarity and divergence patterns should not be viewed in isolation from their broader historical, cultural, and political contexts. West and Central Java regions have long been zones of interaction among ethnic groups, influenced by historical trade routes, colonial administrative divisions, migration, and cultural exchange. Such dynamics may have facilitated lexical borrowing, convergence of speech forms, or the maintenance of distinct identities through language. The overlapping vocabulary observed between Javanese, Cirebon, and Indramayu may reflect inherited linguistic structures and socio-political forces that have shaped language use over time. Future studies would benefit from incorpo-rating historical and ethnographic data to comple-ment linguistic findings and better understand how dialects in this region have evolved.

ACKNOWLEDGEMENT

This research article was realized with independent research funds outside the 2024 scheme provided by the Faculty of Cultural Sciences, Universitas Gadjah Mada, Yogyakarta. High appreciation is conveyed to Mohammad Farhan, who has become a reliable field research staff member.

REFERENCES

Adelaar, A., & Himmelmann, Nikolaus (Eds.). (2005). *The Austronesian Languages of Asia and Madagascar*. Canberra: Pacific Linguistics.

Bellwood, P. (1997). *Prehistory of the Indo-Malaysian Archipelago* (Revised ed.). Honolulu: University of Hawai'i Press.

Bloomfield, L. (1933). *Language*. New York: Henry Holt and Company.

Chambers, J. K., & Trudgill, P. (1998). *Dialectology* (2nd ed.). Cambridge: Cambridge University Press. Campbell, Lyle. 2004. *Historical Linguistics: An Introduction* (2nd ed.). Edinburgh: Edinburgh University Press.

Danadibrata, R. (2015). *Sundanese Dictionary*. Bandung: Kiblat Buku Utama.

Durie, M. & Malcolm Ross (eds.). (1996). *The Comparative Method Reviewed: Regularity and irregularity in language change*. New York: Oxford University Press.

Dutta, Jhulan & Das, Swagata. 2013. Language Variation and Change: Dialectology. *International Journal of Humanities and Social Science Invention*, *2* (1), 28–34.

Ewing, M. J. (2007). *Sociolinguistics and Social Theory*. Edinburgh: Edinburgh University Press

Kridalaksana, Harimurti. 2011. *Linguistic Dictionary Fourth Edition*. Jakarta: Gramedia Pustaka Utama.

Lehmann, Winfred P. 1992. *Historical Linguistics: An Introduction*. London: Routledge.

Mahsun. (1995). *Diachronic Dialectology: An Introduction*. Yogyakarta: Gadjah Mada University Press.

Nerbonne, J., & Heeringa, W. 1997. *Computational Dialectology: Problems and prospects*. In Proceedings of the Ninth Conference of the European Chapter of the Association for Computational Linguistics (pp. 221-228).

Nothofer, B. (1981). *The classification of the Austronesian languages*. Canberra: Department of Linguistics, Research School of Pacific Studies, The Australian National University.

Saefullah, I. (2017). *Indramayu Language Dictionary*. Indramayu: Kainoe Books.

Séguy, J. (1971). La dialectométrie dans l'atlas linguistique de Gascogne. *Revue de linguistique romane*, 35(138), 1–24.

Sudaryanto. (1993). *Metode dan aneka teknik analisis bahasa: Pengantar penelitian wahana kebudayaan secara linguistis*. Yogyakarta: Duta Wacana University Press.

Sudjana, TD, Salana, Susilo, Atmaja Warsito, and Hadi Soerojo. 2015. *Cirebon Language Dictionary*. Bandung: Humanities.

Thomason, S. Gray, & Terrence Kaufman. (1988). *Language Contact, Creolization, and Genetic Linguistics*. Berkeley: University of California Press.

Trask, R.L. (1996). *Historical Linguistics*. London: Arnold.

Trudgill, P. (2004). *Sociolinguistic Typology: Social determinants of linguistic complexity*. Oxford: Oxford University Press.

Uhlenbeck, E. M. (1960). Some general points of Indonesian linguistics. *Lingua*, 9, 327–336.

Utomo, Imam Budi. (2021). *Javanese-Indonesian Dictionary*. Yogyakarta: Language Center of the Special Region of Yogyakarta Province.

Wiyanti, Sri. (2005). "Proto Austronesian Reflection on the Core Vocabulary of Sundanese, Cirebon, Indramayu, and Javanese in Comparative Historical Linguistics Studies". Thesis. Postgraduate Program, Indonesian University of Education. Bandung.

Winford, D. (2003). *An Introduction to Contact Linguistics*. Oxford: Blackwell Publishing

List of Pages

Budiawan, Raden Yusuf Sidiq., and Faidatun Mujawanah. 2019. "Comparison of Javanese Language Variations in Tanjung Brebes District with Losari District, Cirebon Regency". *Synthesis of Scientific Journal of Culture*. 13(2). 57—64. https://doi.org/10.24071/sin.v13i2.2282 accessed on April 2, 2024.

Itaristanti, I., and Laily, IF 2023. "Synchronous Dialectology Study of Cirebon Language: An Effort to Explore Relic Elements". *Satwika: Study of Cultural Sciences and Social Change.* 7(2), 606—616. https://doi.org/10/22210/satwika.v7i2. 32709 accessed on April 2, 2024.

Winata, Nana Triayana. 2019. "Cirebon Sundanese Language in Dukupuntangg District, Cirebon Regency: A Study of Dialect Geography". *Bahtera Indonesia: Journal of Research on Indonesian Language and Literature Education*. 4(1). 39—47. https://doi.org/10.31943/bi.v4i1.5 accessed on April 2, 2024.

https://kbbi.web.id/kognat on June 2, 2024.