

Received 6 December 2015.

Accepted 2 April 2016.

GIS MAPPING OF DIALECT VARIATIONS IN NORTH PERAK¹

Zaharani AHMAD, Nor HASHIMAH JALALUDDIN & Yusmaniza MOHD YUSOFF

National University of Malaysia*

zaharani@ukm.edu.my / shima@ukm.edu.my / yusmaniza_mohdyusoff@ymail.com

Abstract

The aim of this paper is to describe the variations of Malay dialects spoken in north Perak, and accordingly displaying their distribution in a geolinguistic map by employing a Geographical Information System (GIS) analytical tool. The analysis on dialect variations is based on phonological modifications of liquid segments word final. The data, which were collected in 24 different villages in the region, are manually transcribed and classified into variants based on their phonological realisations. These variants, which were initially recorded in the Excel Program, are then transferred into the GIS programme where the linguistic data are processed systematically. The GIS analytical tool produces geolinguistic maps demonstrating the distribution of dialect variations and the topographic profiling of the region. It is apparent that there is a correlation between the two where the latter plays a significant role in determining the former.

Keywords

dialect, geolinguistics, phonological modification, variants

CARTOGRAFÍA GEOREFERENCIADA SOBRE LAS VARIACIONES DIALECTALES EN EL NORTE DE PERAK

Resumen

El objetivo de este artículo es describir las variaciones de los dialectos malayos que se hablan en el norte de Perak y, en consecuencia, mostrar su distribución en un mapa geolingüístico empleando Sistemas de Información Geográfica (SIG). El análisis de las variaciones dialectales se basa en modificaciones fonológicas de los segmentos líquidos de final de palabra. Los datos, que fueron recogidos en 24 aldeas diferentes de la región, se han transcrito manualmente y se han clasificado en variantes

¹ This work has been developed under AP-2012-002 Project, funded by Universiti Kebangsaan Malaysia.

* Linguistic Program, Faculty of Social Sciences and Humanities, Universiti Kebangsaan Malaysia.

basadas en sus realizaciones fonológicas. Estas variantes, que inicialmente se registraron en el programa Excel, se transfirieron luego al programa SIG, donde los datos lingüísticos se procesaron sistemáticamente. La herramienta de análisis SIG produce mapas geolingüísticos que demuestran la distribución de las variaciones dialectales y el perfil topográfico de la región. Es evidente que existe una correlación entre ambos, donde este último juega un papel importante en la determinación de los primeros.

Palabras clave

dialecto, geolingüística, modificaciones fonológicas, variantes

1. Introduction

The advent of modern Information and Communications Technology (ICT) has inspired the development of many disciplines of study. The utilization of high-end computer technology in Geographical Information System (GIS) application enhances its analytical performance and capabilities. Modern GIS and its advanced spatial analysis tools allow sophisticated and efficient analysis of spatial data in many fields. One of them is geolinguistics, an interdisciplinary field that often incorporates language maps depicting spatial patterns of language or dialect location in a region.

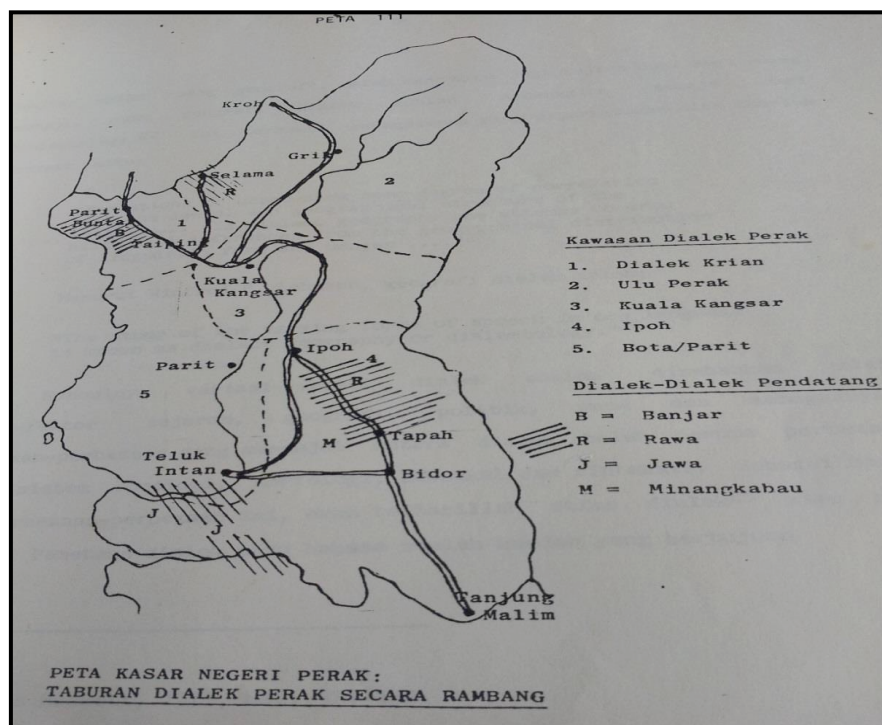
Accordingly, this paper attempts to describe and map the distribution of Malay dialects spoken in north Perak that covers the district of Larut Matang Selama (LMS) and Kuala Kangsar (KK) by using the GIS² tool. Choropleth maps produced by GIS will display the spatial pattern of dialect distribution, as well as exhibiting the topography of the area, such as hills, mountains, and rivers. Based on these maps, it is apparent that there is a correlation between the two where the latter plays a significant role in determining the former.

2. Literature review

As recorded in the literature, there were some studies on geodialect of Malay that have been conducted by previous scholars (Harun 1983, Ajid 1985, Asmah 1985, and

² This research is supported by the Universiti Kebangsaan Malaysia grant- UKM AP-2012-002.

Rohani 1986). However, the maps that they produced merely depicted the distribution of Malay dialects spoken in that particular area without displaying its topographical features, such as rivers, hills, and mountains. In addition, since the maps were manually sketched impressionistically, their validity and accuracy are dubious and unreliable. Examples of such maps are shown in Maps 1 and 2 below.

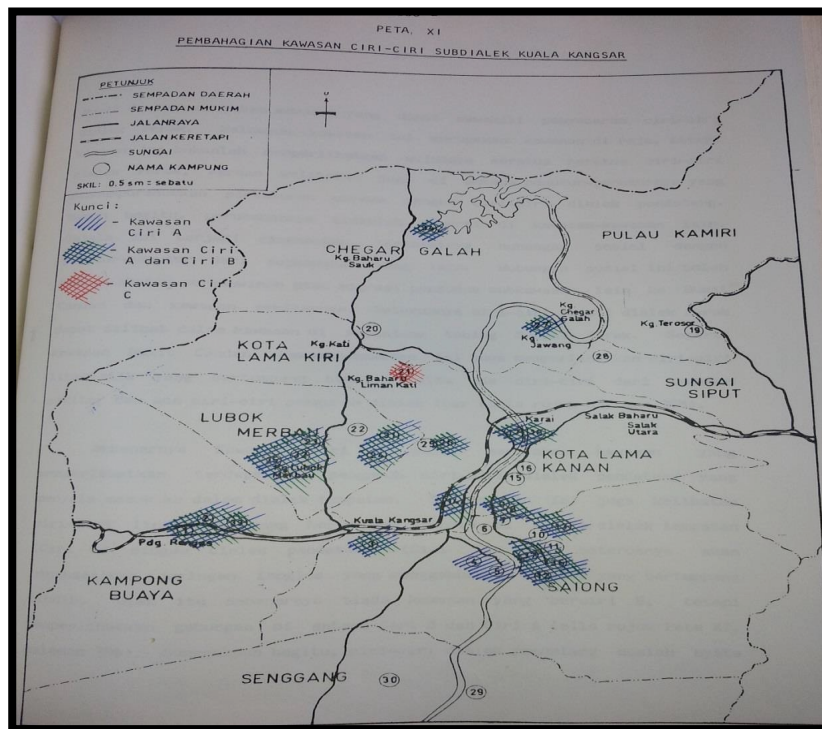


Map 1. The distribution of Perak Dialects (Harun 1983)

The above map displays the distribution of five main dialects of Perak Malay which are categorized randomly based on phonological variations, and they are dubbed according to their geo-political names and locations. The dialect spoken in Krian is called Krian dialect, the dialect spoken in Kuala Kangsar is identified as Kuala Kangsar dialect, and so forth. In addition, there are also immigrant dialects spoken in the state, namely Banjar, Rawa, Jawa and Minangkabau. It is apparent that Harun's classification on Malay dialects spoken in Perak is purely based on geopolitical boundaries rather than linguistic boundaries.

Rohani's (1986) study, however, does not cover the entire state of Perak, but solely focuses on the district of Kuala Kangsar. She claims that this area is a dialectal

transition region. Her study is also based on phonological variations, particularly those that involved vowel and consonant modifications. Areas with similar phonological behaviors are grouped together and their distribution in the vicinity is displayed in an imprecise map as shown in Map 2.



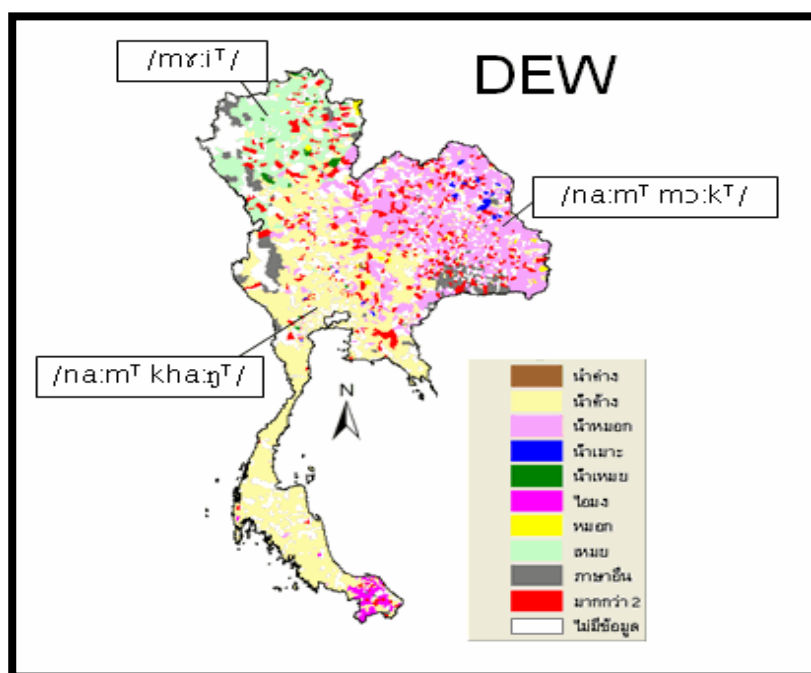
Map 2. The distribution of Perak subdialect in Kuala Kangsar (Rohani 1986)

In comparison, Rohani's map (1986) seems to be better than Harun (1983) in the sense that the former includes some topographical features in her map, such as rivers, roads, railway tracks and the name of the villages. Nevertheless, additional features such as hills and mountains are not reflected here.

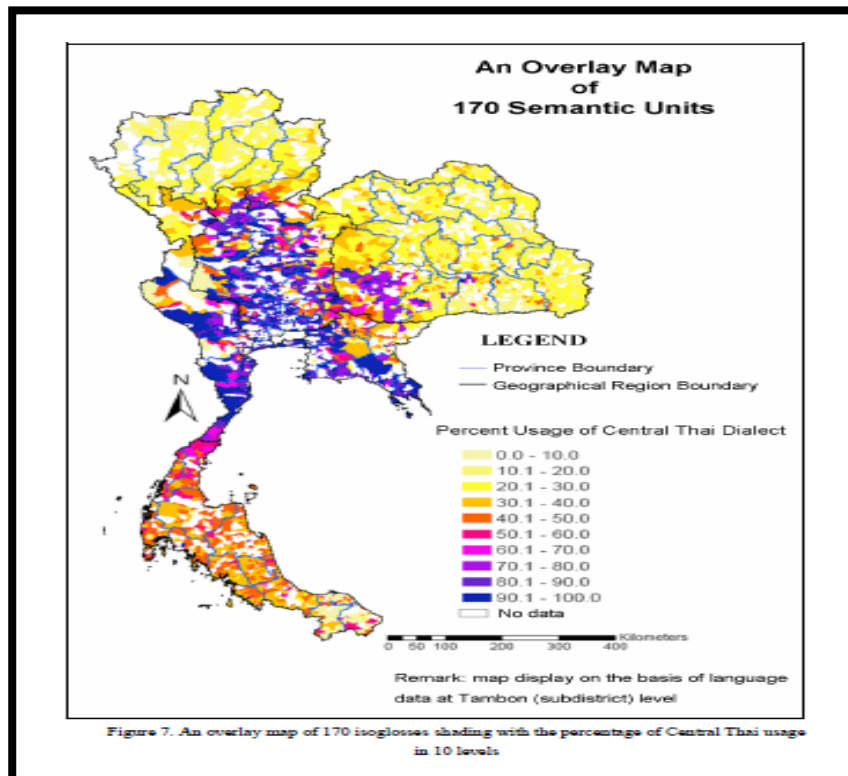
Unlike the situation in Malaysia, the use of modern technology in producing a reliable geodialectal map is flourishing in foreign countries, such as Thailand (Teerarojanarat & Tingsabadh 2011b), Japan (Onishi 2010) and USA (Hoch and Hayes 2010) where modern GIS analysis tool is used in defining dialect or language boundaries in the region. Teerarojanarat & Tingsabadh (2011a) state that the process of mapping dialect boundaries is not an easy job because it needs to take into account two important factors, namely language behavior and mapping technique. The use of GIS is aimed to improve the quality of the maps in representing dialect boundaries in Thailand.

In their study, they have successfully collected 170 lexical of Thai language which represents one lexical semantic unit. Respondents were asked to fill up the questionnaires and subsequently sent them to the researcher. The collected data were analyzed and processed systematically by the GIS analytical tool where it finally produces a map that depicts linguistic boundaries of the lexical items.

Teerarojanarat & Tingsabadh's (2011a) study serves as a case in point how a geolinguistic study pertaining to dialect boundaries can be mapped by the GIS system. Based on the word 'dew', she has shown the distribution of dialect boundaries of central and non-central of Thailand. It must be noted that the overlay technique is also employed in the GIS in order to capture the overlapping between the lexical item and the area under studies. The GIS map produced is as follows.



Map 3. The distribution of the word 'dew' in Thailand (Teerarojanarat & Tingsabadh 2011a)



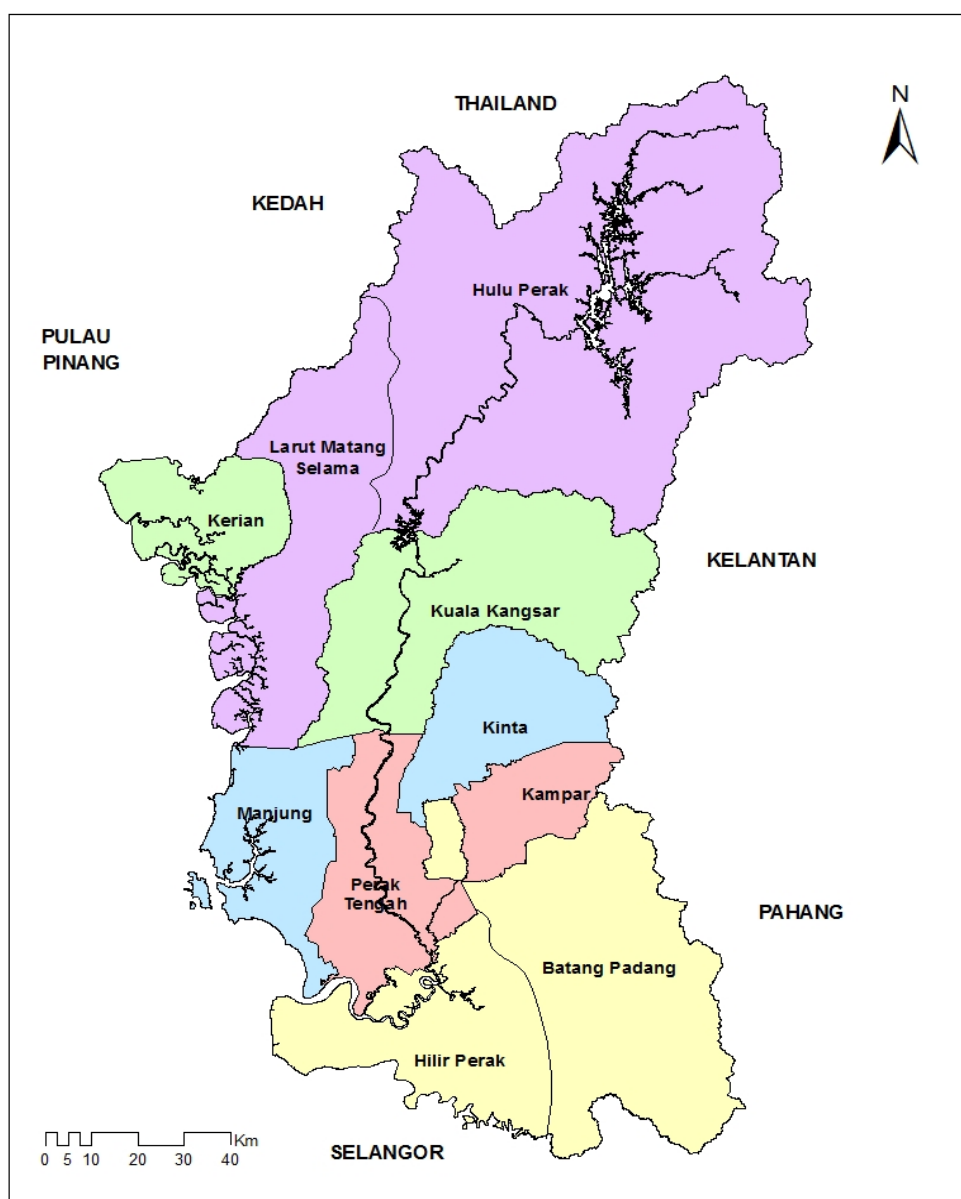
Map 4. An overlay map of 'dew' in Thailand (Teerarojanarat & Tingsabadh 2011b)

3. North Perak: a brief background

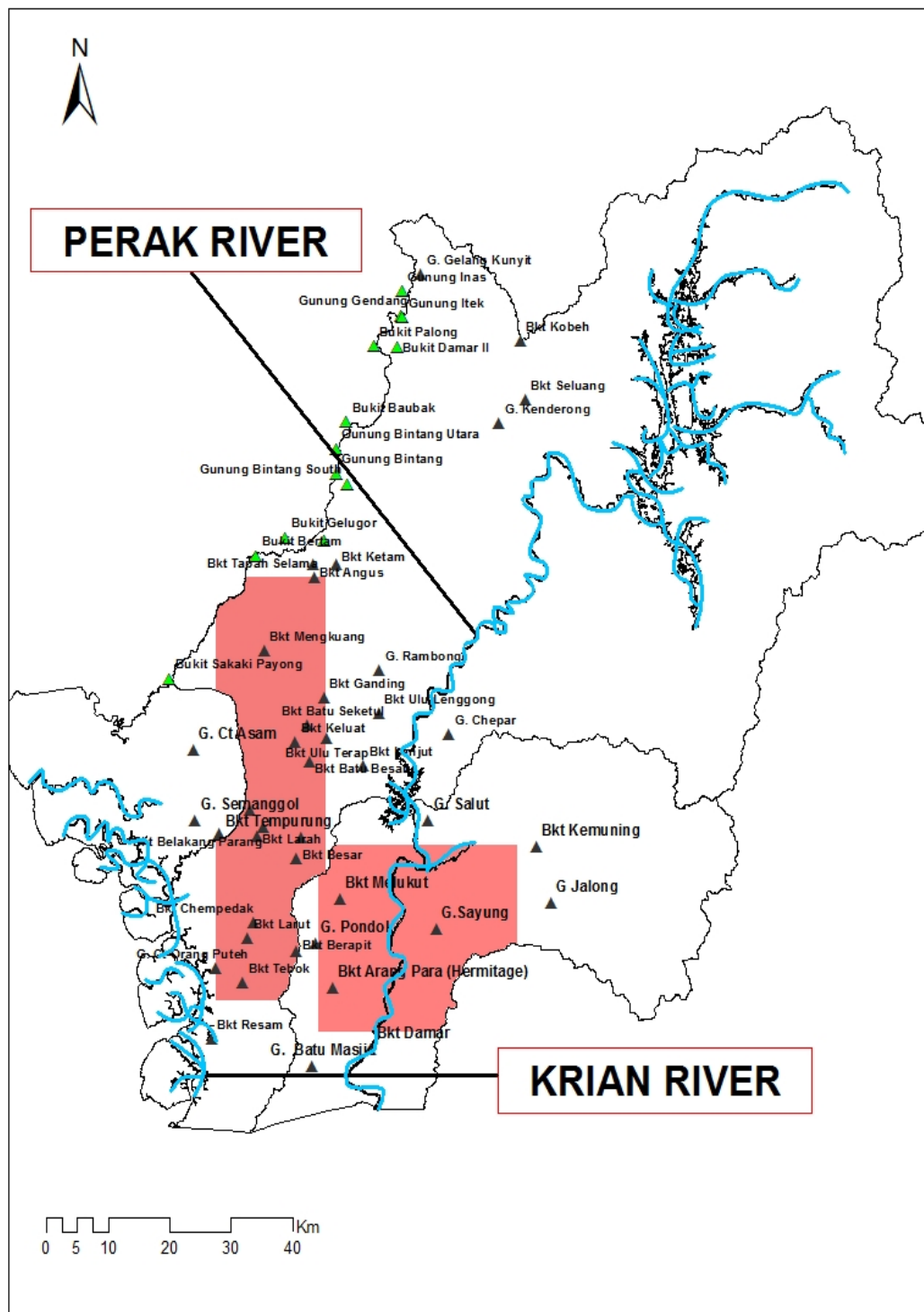
Perak is one the Malay states situated in Peninsula Malaysia. The state is bordered by Kedah in the west, Southern Thailand in the north, and Kelantan in the east. Administratively, the state of Perak is divided into ten districts, namely Hulu Perak, Larut Matang Selama, Kerian, Kuala Kangsar, Kinta, Manjung, Perak Tengah, Kampar, Batang Padang and Hilir Perak (see Map 5). For the purposes of this study, only two districts that will be examined that is Larut Matang Selama (SLM) and Kuala Kangsar (KK). The two districts were selected because as reported in the literature, the Malay dialects spoken here are pretty diverse. In addition, their topographic background is physically complex because the area is covered by mountains, hills, rivers and thick jungles (see Map 6). There are two major rivers that run across the state of Perak that is Krian river on the west side and the Perak river right in the interior. The settlements of traditional Malay communities were believed to be in the vicinity of these two rivers. As can be seen in Map 7, many Malay villages (i.e. 'kampung') are situated along the two

rivers. For the purpose of this study, a total of 24 villages were visited, as displayed on Map 7.

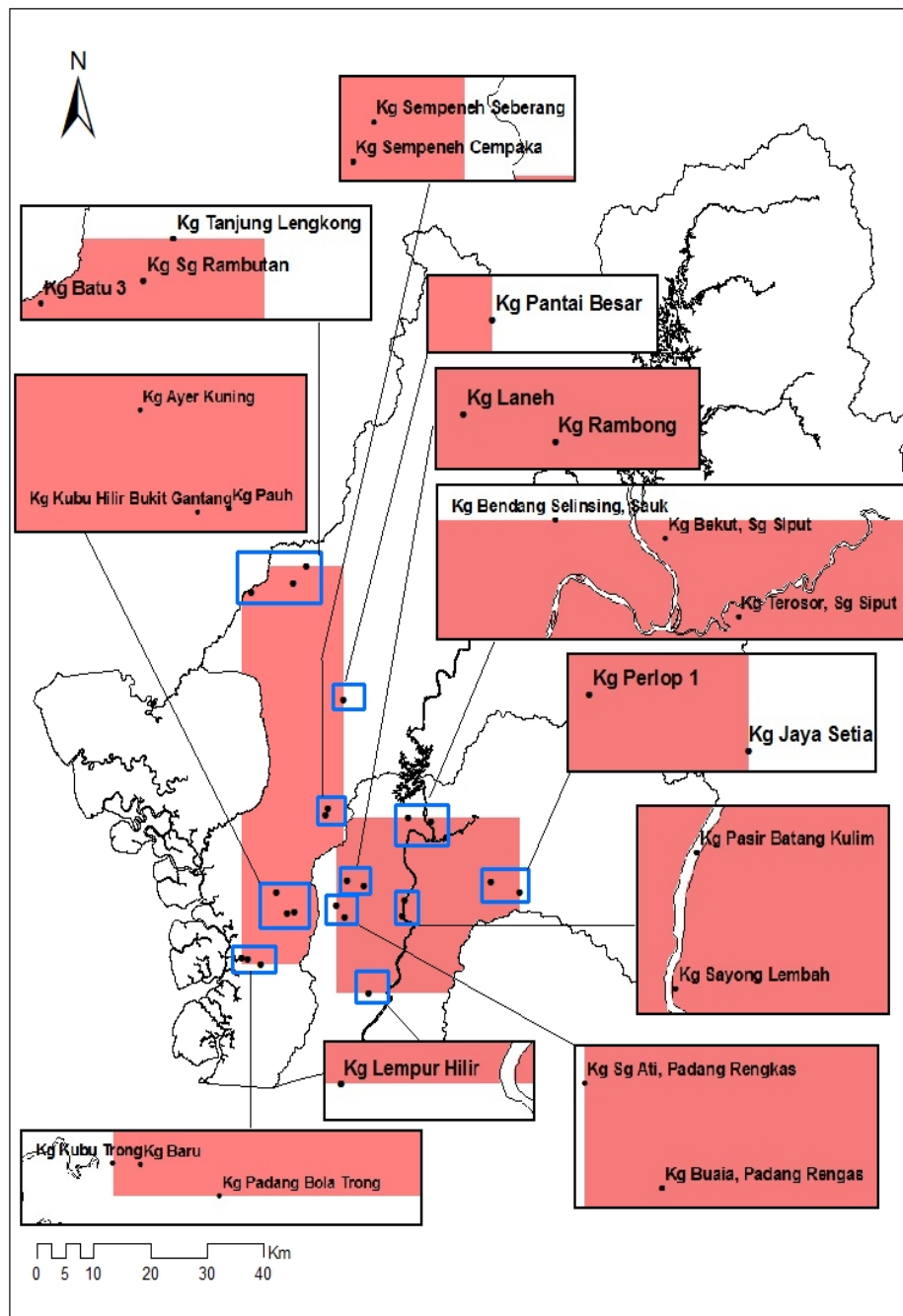
The diversity of the Malay dialects and the irregular nature of the surface shape and features of the region are two important aspects that will be scrutinized in this study. We attempt to account for whether the two attributes have a close relationship in the context of dialectal distribution in the north state of Perak.



Map 5. Geopolitical boundaries of state districts in Perak



Map 6. Topographic background of LMS and KK



Map 7. List of villages in LMS and KK

4. Methodology

The data used for the study were collected in the field by interviewing the respondents based on a questionnaire sets with 96 lexical items. Respondents were asked to pronounce each lexical item in their own native dialect either directly or

indirectly. The selected respondents must be non-mobile and categorized as adults (22-49 years) and senior citizens (50 years and over). In total, there are about 200 respondents participated in this study. The conversation was then recorded by using a smart phone, MP3 and digital camera.

After data collection, the next step is transcribing the raw data. All lexical items were transcribed in a broad phonetic transcription. Each lexical and its variants are identified and categorized accordingly based on their phonological realizations. For example, the word /bantal/ 'pillow' surfaces as four variants, namely [bantaj], [bata], [bante] and [banta]. These variants are then labeled as L1, L2, L3 and L4 consecutively. List of respondents with their various pronunciations are keyed into Excel program. Overall data in Excel Program is subsequently transferred into the GIS software for analysis. Spatial maps were then produced automatically.

LEXICAL		BANTAL				ULAR			AIR			
VILLAGES	CODE	V1	V2	V3	V4	V1	V2	V3	V1	V2	V3	V4
		bante	bantaj	bata	banta	ulo	ula	ulaŋ	ajo	ae	ε	ajaŋ
DAERAH LARUT MATANG SELAMA (LMS)												
MUKIM BATU KURAU												
KG. SEMPENEH SEBERANG	1WT	bante				ulo					ε	
	2WT	bante				ulo					ε	
	3WT		bantaj				ula				ε	
	4WT	bante				ulo					ε	
	5WT	bante					ula				ε	
	6WT	bante				ulo					ε	
	7WT		bantaj				ula				ε	
	1D		bantaj				ula				ε	
	2D	bante					ula				ε	
KG. SEMPENEH CEMPAKA	3D	bante				ulo					ε	
	1WT		bantaj				ula				ε	
	2WT	bante				ulo					ε	
	1D	bante					ula				ε	
	2D	bante					ula				ε	
	3D		bantaj			ulo					ε	
MUKIM BUKIT GANTANG												
KG. AIR KUNING	1WT		bantaj				ula			ae		
	1D		bantaj				ula			ae		
	2WT	bante						ulaŋ				ajaŋ
	2D		bantaj			ulo			ajo			
	3WT		bantaj			ulo			ajo			
	4WT		bantaj				ula			ae		
	3D		bantaj				ula			ae		
	5WT		bantaj				ula			ae		
	6WT		bantaj					ulaŋ		ae		
	7WT		bantaj					ulaŋ		ae		ajaŋ

Table 1. List of respondents and variants tabled in Excel Program

5. Phonological analysis

Generally, all lexical items, specifically those with word final liquid consonants are pronounced differently by the respondents which are dubbed as variants. For example, the word /bantal/ 'pillow' surfaces as four variants, namely (i) [bantaɲ], (ii) [bata], (iii) [bante] and (iv) [banta]. These variants are commonly identified as representing certain dialects of Malay spoken in the Peninsula Malaysia. For instance, [bante] is associated to Perak Malay (Zaharani 1991, Raja Mukhtaruddin 1986, Harun 1983), [bata] is Pattani Malay (Ruslan 2011), and [bantaɲ] is Kedah Malay (Asmah 1985). In a linguistic standpoint, these variants have a lot of similarities as well as differences. Similarities suggest that the dialects are preserving the shared linguistic heritage inherited from the same source, and the differences signify that the dialects have undergone self-development and modifications as time passed by. According to Chambers & Trudgill (1972), the existence of various dialects in a language is a result of changes in the history of a language which was originally uniform. In the contexts of input and output relationship, it can be assumed that the source language acts as an input, while dialectal variations surfaces as variants are the output. In relation to the above assumptions, this study presupposes that the lexical /bantal/ as the input and the four surface variants as the outputs, namely (i) [bantaɲ], (ii) [bata], (iii) [bante] and (iv) [banta]. Linguistically, in the rule-based approach the change from input to output is driven by phonological rules. In general, there are three rules involved in the modifications, namely (i) deletion rule, (ii) feature changing rule, and (iii) coalescence rule. Variants with deletion rule represent Pattani Malay, while variants with feature changing rule characterize Kedah Malay, and finally variants with coalescence rule correspond to Perak Malay. The above description can be summarized as follows.

1. Pattani Malay

(i) Deletion rule - liquid consonants are dropped word finally

Input : /bantal/ Output : [bata] [banta]

Input : /ular/ Output: [ula]

Input : /air/ Output: [ɛ] [ae]

2. Kedah Malay

(i) Glide formation rule - /l/ consonant becomes glide [j] word finally

Input : /bantal/ Output : [bantaɟ]

(ii) Pharyngealization rule - /r/ becomes pharyngeal consonant [ʕ] word finally

Input : /ular/ Output: [ulaʕ]

Input : /air/ Output: [aɟaʕ]

3. Perak Malay

(iii) Coalescence rule - /al/ becomes [ɛ] and /ar/ becomes [ɔ].

Input : /bantal/ Output : [bantɛ]

Input : /ular/ Output : [ulɔ]

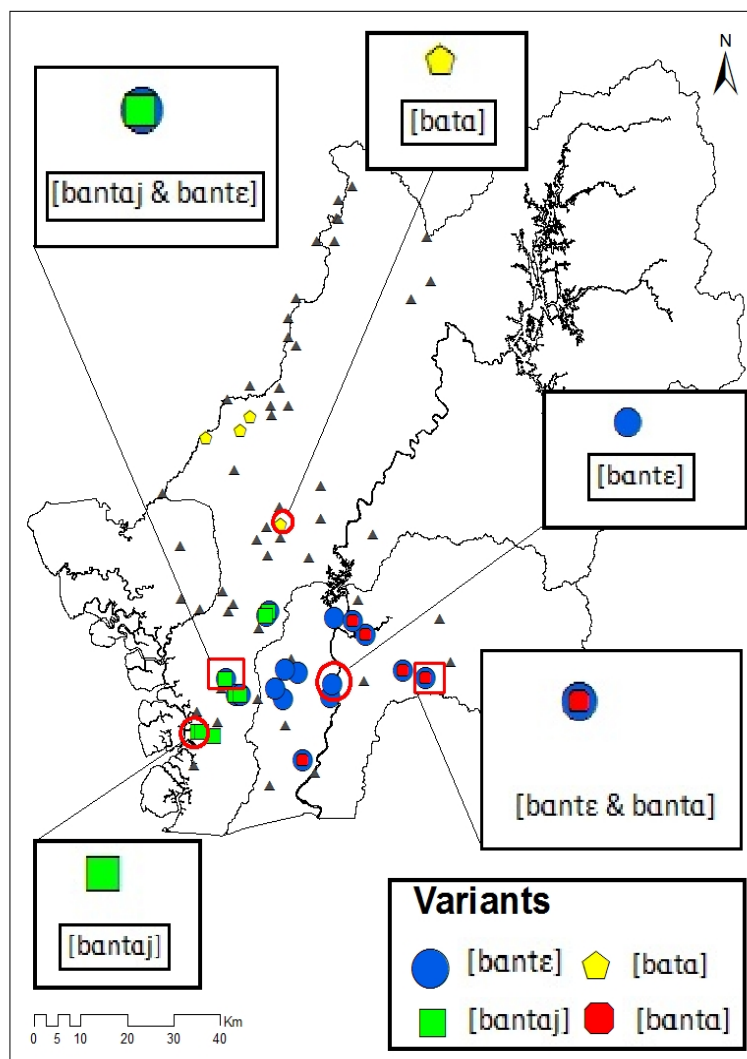
Input : /air/ Output : [aɟɔ]

6. GIS Mapping

The following discussion attempts to illustrate the distribution of Malay dialects based on the three lexical items discussed above by presenting choropleth maps generated by the GIS analytical tool.

6.1 The distribution of /bantal/variants

As already been mentioned, the word /bantal/ 'pillow' has four variants, namely (i) [bantaɟ], (ii) [bata], (iii) [bantɛ], and (iv) [banta]. The first variant represents Kedah Malay, the second variant is Pattani Malay and the third variant is Perak Malay. The fourth variant is quite interesting because its existence is a result of merging between two dialects that is Pattani Malay and Perak Malay. Phonological feature of Pattani Malay is signified by /l/ deletion, whereas for Perak Malay it involves the preservation of homorganic nasal word internally. The distribution of these variants is shown in the map below.



Map 8. The distribution of /bantal/ 'pillow' variants in LMS and KK

The above map demonstrates that the [bantaj] variant identified as Kedah Malay spreads in the southern part of LMS district. All the three villages located on the banks of Krian River (i.e. kampung Kubu Trong, kampung Baru, and kampung Padang Bola Trong) entirely speak the dialect. It then spreads to the north towards the next villages. However, villages in the northwestern part of LMS that are located adjacently to the state of Kedah do not use Kedah Malay. These villages are speaking Pattani Malay or Kelantan Malay instead. The state of Kelantan is located far away from the village.

Based on the map above, it can be predicted that Kedah Malay comes into the state of Perak via the Krian River. The dialect subsequently spreads to the north because the topography of the area is relatively flat and low land, and the presence of

some scattered small hills does not hinder the spreading process. The dialect, however, does not spread to the northwestern part of LMS because the area is covered by mountains and hills which accordingly impede the movement. The same explanation goes to the region in the north where the two states are bordered by hilly and mountainous highlands.

Remote villages in the middle of the mountain range (i.e. kampung tanjung Lengkong, kampung Sungai Rambutan dan kampung Batu 3) adversely use [bata], a variant designating the Pattani Malay. Apparently, the dialect must be originating from Pattani/Kelantan in the northeast and progresses along the Perak River before it moves inland towards the west region.

In KK district, almost all the villages are using the [bantɛ] variant. This variant is a typical pronunciation of Perak Malay, and claimed to be a distinctive feature of the dialect (Zaharani 1991, Asmah 1985, Harun 1983, Raja Mukhtaruddin 1986). As can be seen, Perak river runs across most parts of KK district, and all the inhabited villages are located along the river basin. Since most of the area is flat and lowlands, the dialect is widely disseminated at all corners.

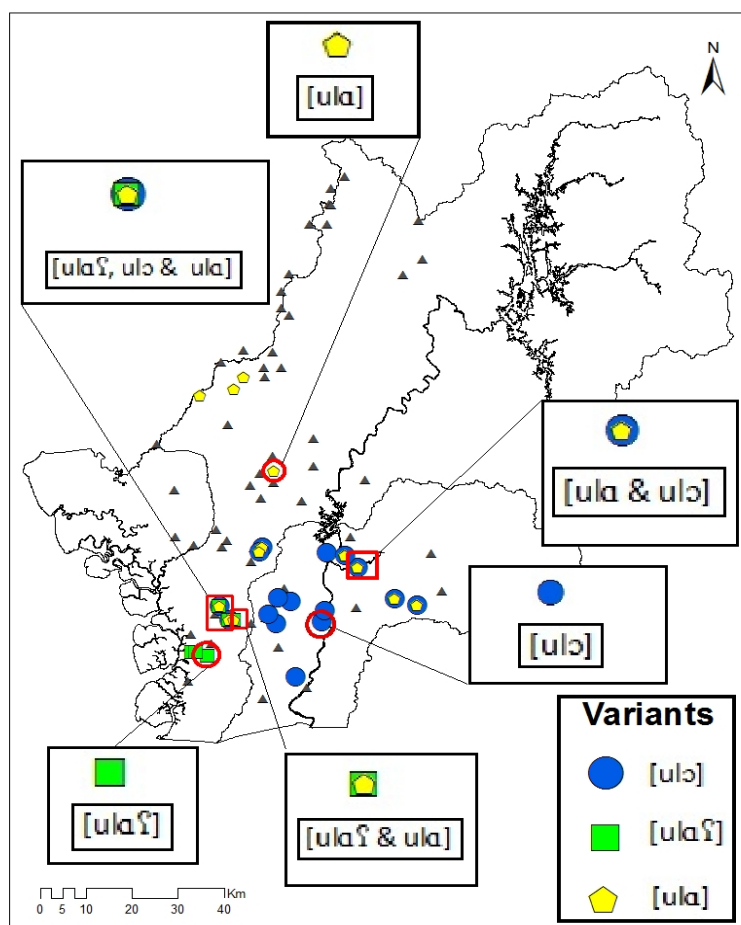
For villages located closely at the district boundaries, a mixture of dialects is often used. Although there are some small hills here and there, their presence does not hamper interaction and communication. The villagers in kampung Sempeneh and kampung Sempeneh Cempaka speak a mixture of two dialects that is Kedah Malay and Perak Malay.

An interesting observation and worth to be noted is that the occurrence of [banta] which is spoken in the north and east of KK. This variant has its own characteristic which entails a merger of two dominant dialects discussed thus far, that is Pattani Malay and Perak Malay. The phonological feature of the former involved deletion of liquid word finally, and in the latter it retains the homorganic nasal word internally. The amalgamation of these two phonological features gives rise to a new Malay subdialect in the state of Perak. Concomitantly, Perak Malay is also spoken here, and as a result dialect mixing is observable in certain areas. And again, mixing of dialects possibly takes place because there is no topographical barrier that can stop the distribution.

Based on the GIS map displayed thus far, it can be concluded that the distribution of Malay dialects in the state of Perak has a very close relationship with the topography of the area concerned. The central focal point of Malay dialect dissemination in Perak state is along the river basin. Kedah Malay spreads along Krian river, while Perak Malay disperses along Perak River basin. For mountainous and hilly regions, dialect distribution is obstructed, even though the position of the villages is closely located.

6.2 The Distribution of /ular/ variants

The word /ular/ 'snake' has three variants representing the three main dialect of Malay discussed earlier, that is [ulaʔ] for Kedah Malay, [ulɔ] represents Perak Malay, and [ula] symbolizes Pattani Malay. The distribution of these variants is shown in the Map below.



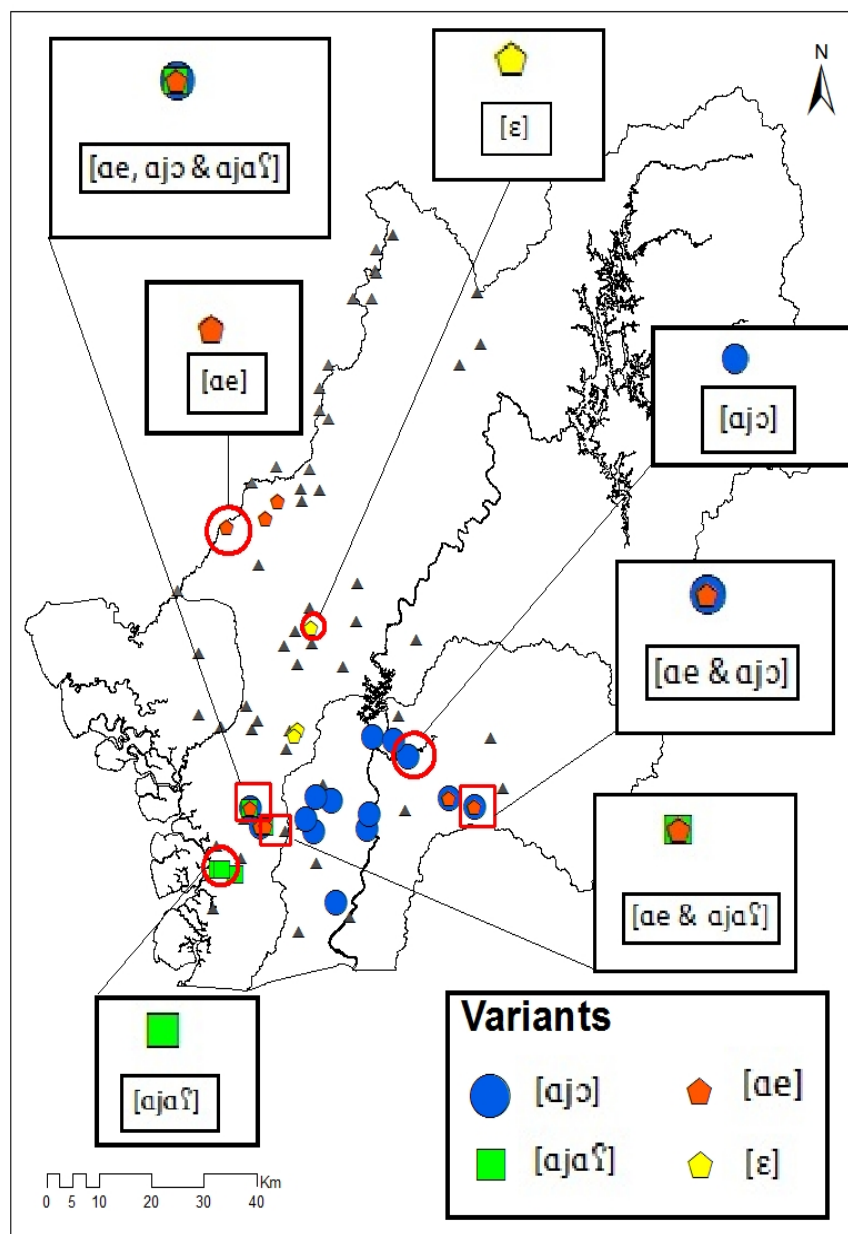
Map 9. The distribution of /ular/ 'snake' variants in LMS and KK

In general, the distribution of /ular/ variants has the same dispersal pattern as in the previous case. The Kedah Malay variant identified as [ulaʕ] spreads along the Krian River, whereas [ula] classified as Pattani Malay is widely used in the north part of LMS. Despite the fact that those villages are bordered by the state of Kedah, Kedah Malay does not disseminate into the region because the area is covered by mountain ranges. It is deemed that Pattani Malay is spoken there because the Pattani people come from the east via the Perak River and later settle down there. The Perak Malay variant signified as [ulo] is spoken in all the villages located along the Perak River. Since the area is a flat river basin, interaction and communication remain intact without any hindrance.

One significant difference between the two lexical items that can be observed is that their distributional pattern at the district boundary is not the same. In the previous case, all the villages at the proximity use a mixture of Kedah Malay and Perak Malay. Whereas in the present case, the villagers in the north speak a mixture of Pattani and Perak Malay, while the villagers further south use a mixture of Pattani, Kedah and Perak Malay.

6.3 The distribution of /air/ variants

Similarly to the case of /bantal/ 'pillow', the word /air/ 'water' also has four different variants, namely [aɔ], [aɔʕ], [ɛ] and [œ] which are representing the Perak Malay, Kedah Malay and Pattani Malay consecutively. The distribution of these variants is as follows.



Map 10. The distribution of /air/ 'water' variants in LMS and KK

In the present case, the Pattani Malay has two variants, namely [ae] and [ε]. As can be seen, the latter is used in the north part of LMS (i.e kampung Tanjung Lengkong, kampung Sg. Rambutan, kampung batu 3), whereas the latter is spoken in the interior part of LMS (i.e kampung Pantai Besar, kampung Sempeneh Seberang, kampung Sempeneh Cempaka). Although the latter is located closely to the KK border, unlike the previous cases, it is not influenced by the Perak Malay dialect. In the south, Kedah Malay is still dominant along the Krian River. For the southern villagers that are located at the

KK border, they speak a mixture of three different dialects – Kedah, Perak and Pattani Malays.

Along the Perak River, as usual the most prominent dialect spoken here is the Perak Malay which is represented by [ɑjɔ] variant. Just as in the previous case, a mixing of dialects also occurs in the east part of KK, especially at the border where the areas are flat and lowland region. Despite the fact that there are some small hills scattered all over the place, its presence does not prevent the interaction between the local populations.

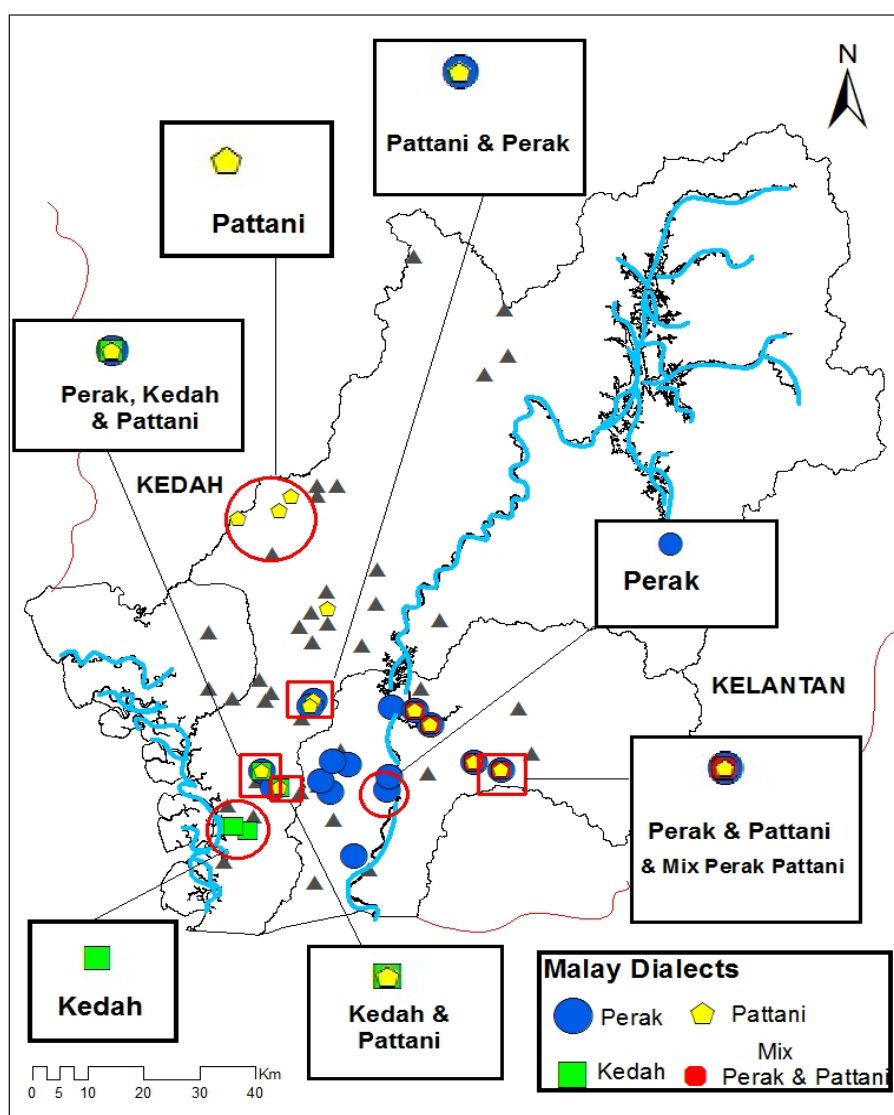
7. Summary

The finding of the study demonstrates that the distribution of the Malay dialects spoken in the districts of KK and LMS is correlated with the topographical features of the region. The maps produced by GIS analytical tool are reliable and authentic because the software employs the same coordinates used in Skewed Rectified Orthomosaic (RSO) and Cassini in Malaysia, and the topographical information provided by the agencies is correct and accurate. An overall summary of the findings is as follows:

i) The distribution of Perak Malay spreads across the villages in the district of KK, particularly along the Perak River. The villages here retain the original features of the dialect. This dialect also spans into the area of LMS, particularly in kampung Air Kuning, kampung Kubu Hilir and kampung Pauh in Bukit Gantang subdistrict because they are closely located to KK and the area is not separated by mountain ranges. Consequently, there is a mixture of Perak Malay and Kedah Malay spoken in the area. In the east part of KK, Perak Malay mixes with Pattani Malay which subsequently gives rise to a new subdialect with a merger of both. At the same time, the dominant dialect of Perak Malay is also spoken in the area.

ii) Pattani Malay spreads in the northern part of LMS, as it is in the proximity to Hulu Perak, Kelantan and Pattani. Perak Malay as well as the Kedah Malay fails to penetrate into the villages because the area is covered by mountain ranges and dense forests which complicate the process of communication.

iii) Kedah Malay spreads in the southern part of LMS, mainly along the Krian River. The dialect, then move northward as the land is flat and level. For areas that are adjacent to the KK district, there is a mixture between Kedah Malay and Perak Malay. Although there are small-scattered hills around, interaction and communication among the speakers are still taking place.



Map 11. The distribution Malay dialects in LMS and KK

References

- AJID, Che Kob (1985) *Dialek Geografi Pasir Mas*, Bangi: Penerbit Universiti Kebangsaan Malaysia.
- ASMAH, Omar (1985) *Susur Galur Bahasa Melayu*, Kuala Lumpur: Dewan Bahasa dan Pustaka.
- CHAMBERS, Jack K. & Peter TRUDGILL (1980) *Dialectology*, Cambridge: Cambridge University Press.
- HARUN, Mat Piah (1983) "Dialek negeri Perak", *Bengkel Kepemimpinan dan Pengajian Kebudayaan Negeri Perak*, 20-23, October 1983, Ipoh.
- HOCH, S. & J. J. HAYES (2010) "Geolinguistics: The incorporation of geographic information systems and science", *Gamma Theta Upsilon in the Geographical Bulletin*, Volume 51-1, 23-36.
- ONISHI, Takuichiro (2010) "Analyzing dialectological distributions of Japanese", *Dialectologia*, Special Issues, 1, 123-135. < <http://www.publicacions.ub.edu/revistes/dialectologiaSP2010/>>
- RAJA MUKHTARUDDIN, Raja Mohd Daim (1986) *Dialek Perak*, Ipoh: Yayasan Perak.
- ROHANI, Yusof (1986) *Dialek Geografi Kuala Kangsar: Suatu Kajian Perbandingan*, MA Thesis, Kuala Lumpur: Universiti Malaya.
- RUSLAN, Uthai (2011) *Keistimewaan Dialek Melayu Patani*, Bangi: Penerbit Universiti Kebangsaan Malaysia.
- TEERAROJANARAT, Sirivilai & Kalaya TINGSABADH (2011a) "A GIS-based approach for dialect boundary studies", *Dialectologia*, 6, 55-75. < <http://www.publicacions.ub.edu/revistes/dialectologia6/>>
- TEERAROJANARAT, Sirivilai & Kalaya TINGSABADH (2011b) "Using GIS for linguistic study: a case of dialect change in the northeastern region of Thailand", *Procedia Social and Behavioral Sciences*, 21, 362-371.
- ZAHARANI, Ahmad (1991) *The Phonology & Morphology of the Perak Dialect*, Kuala Lumpur: Dewan Bahasa dan Pustaka.