***Dialectologia 6*** *(2011), 55-75.*
*ISSN: 2013-2247*

*Received 24 February 2010.*
*Accepted 1 July 2010.*

# A GIS-BASED APPROACH FOR DIALECT BOUNDARY STUDIES

Sirivilai TEERAROJANARAT[1] and Kalaya TINGSABADH[2]
[1]Department of Geography, Faculty of Arts, Chulalongkorn University
Sirivilai.t@chula.ac.th
[2]Department of Linguistics, Faculty of Arts, Chulalongkorn University
kalaya.t@chula.ac.th

**Abstract**

For linguists, up to now drawing language boundaries has usually been problematic. This paper aims to present a way by which GIS and linguistic methodology are integrated, especially in the area of dialect geography, to help improve the quality of boundary drawing. Applied methodological flow with the exemplified application to the whole of Thailand is illustrated and discussed. Resultant maps are promising and so is the role of GIS as an essential tool for assisting linguists to better demarcate and interpret the language borders with more confidence, particularly in ambiguous areas.

**Keywords**

Dialect boundary, lexical study, GIS, Thailand

**Resumen**

Hasta ahora para los lingüistas trazar los límites lingüísticos ha sido una tarea generalmente problemática. Este trabajo tiene como objetivo presentar una vía a través de la cual los SIG y la metodología lingüística se integran, sobre todo en el ámbito de la geografía dialectal, para ayudar a mejorar la calidad del trazado de fronteras. Se ilustra y discute la aplicación ejemplificada del método al conjunto de Tailandia. Los mapas resultantes son prometedores y también lo es el papel de los SIG como una herramienta esencial para ayudar a los lingüistas a delimitar e interpretar más correctamente las fronteras del idioma, especialmente en las zonas lingüísticamente ambiguas.

**Palabras clave**

Fronteras dialectales, estudio léxico, SIG, Tailandia

©Universitat de Barcelona

## 1. Introduction

Drawing dialect boundaries is not an easy task. Generally, its complexity is caused by two key factors, the nature of language and traditional mapping techniques. The former can lead to many problems. Firstly, the transition from one dialect to another is often gradual rather than abrupt. It is hard to completely identify the location of the beginning or the end of one dialect (Langacker, 1968; Mackey, 1988). Secondly, the mixture of dialects spoken in a certain area, so-called the zone of transition, makes it difficult to decide which dialect is dominant and where the boundaries should be drawn (Chambers and Trudgill, 1980). Besides, the dialect spoken in one area may be subdivided further into varieties on the basis of phonological or lexical differentiation. Past and present-day research works have suggested different theoretical or methodological approaches for conducting the study of dialect boundaries e.g. the work of Weijnen (1946) in (Kremer, 1999) proposing the "little arrow method" as a way to conduct a subjective boundary study; the work of Williams and Ambrose (1988) investigating several boundary techniques for measuring language borders in western France; the work of Mase (1999) using the thickness of boundaries on a map to show the frequencies of responses; and the work of Long (1999) creating the dialect boundaries on the basis of the geographical perception of respondents. Such a variety of methods and techniques can result in different drawing of language boundaries. The second factor that creates complexity in dialect boundary studies involves the conventional mapping techniques. Unavoidably, working with dialect data, especially in the area of dialect geography, directly involves the spatial issues including field survey, data collection and recording, and displaying map results. In most previous dialect boundary studies, spatial issues in data collection and map presentation have been mostly problematic. Map making and cartographic display were, in all cases, carried out manually. Usually, overlay technique has been conducted by manually superimposing a bundle of paper maps, and drawing the boundaries between dialects has been roughly defined by hand. Superimposing of criss-crossing isoglosses also generates the uncertainty of borderlines drawing. These resultant boundaries, in many parts, have been unclear and ambiguous. Further different scales of study, e.g. one province or one region, may result in different fine and coarse outcomes. Distortion can thus be exaggerated due to the overlay of different scales of map sources, the use of



©Universitat de Barcelona

copied paper map source and so on. A boundary map created in that way, as a result, is inaccurate, imprecise, and possibly unreliable.

The advancement of computer technology in the last few decades has brought in big improvement of geographical technologies. Geographical tools and softwares used for spatial data measurement, collection and analysis have been developed to be easy-to-use and more powerful. Among these, Geographic Information System (GIS) is ideal for storing and managing big volumes of spatial data, performing analysis and producing cartographic displays of map results. Nowadays GIS has been widely exploited in various fields e.g. population and environmental studies. However, their applications to linguistics and dialect studies are limited and still in their infancy.

In this work, the key objective is to demonstrate a methodology by which the traditional linguistic approach in dialectology is integrated with GIS to improve the quality of dialect boundaries, and to apply this methodology to the whole of Thailand for creating dialect boundaries between Central and Non-Central Thai dialects. The benefits of GIS integration are also discussed.

## 2. Dialect data and its relation to Geographic Information System (GIS)

In the geographic perspective, by its nature dialect comprises two types of data: spatial and attribute. Spatial data refers to the locations where language data is collected while attribute data refers to language information associated to those locations e.g. collected language data, names of informants (see Figure 1). Their locations are usually specified to be geographically referenced, using the spatial concept of namely x and y coordinates, such as longitude and latitude. Representing spatial data is usually symbolized in three forms of features; point, line, and polygon. On a spatial basis, paper maps and GIS are key tools to handle the language data attempting to represent or model the real world. Unlike the traditional paper maps, GIS transforms those maps into digital form, stores and links both data types in the geographical database. As many themes of data may be involved, GIS uses the concept of layers to organize these themes e.g. one layer for roads, one layer for rivers, one layer for topography, and one layer for the lexical items representing the same meaning or other linguistic features.



©Universitat de Barcelona

Such concept allows researchers to choose only one or some of the layers for displaying, retrieving or analyzing information.

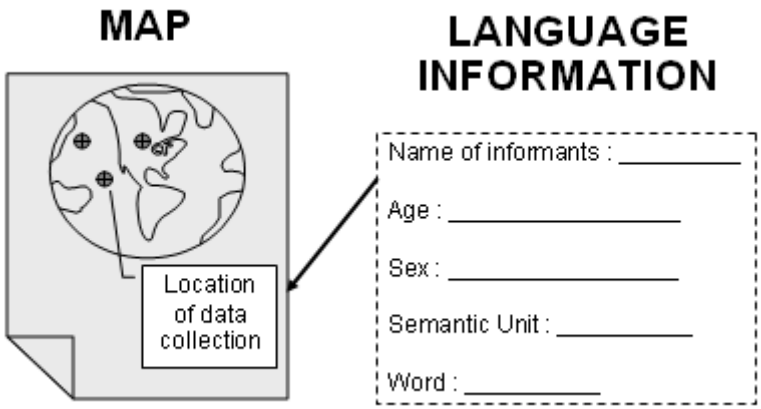



Figure 1. Map and language data as a relationship between spatial and attribute data in GIS

Apart from its capabilities of data storage, database management, and cartographic display, GIS has the ability to perform spatial analysis. Spatial analysis involves both simple and complex spatial processes to transform raw data into useful information (Longley et al, 2005). Before the development of GIS, some traditional methods of spatial analysis were done by hand or by the use of measuring devices such as ruler. Once GIS has been integrated, these maps are operated by computation - greatly facilitating complicated operations e.g. the overlay of multiple maps. In addition, repetitive manners of operations can be conducted more easily. Some spatial analysis functions of GIS are the overlay analysis, the feature selection and query analysis, the measurement and proximity analysis, etc. Details of spatial analysis functions are classified and described extensively elsewhere e.g. the books of Tomlin (1990), Maguire et al. (1991), Longley et al. (1999), or Longley et al. (2005).

Up to now, the applications of GIS to linguistics and dialect studies have been relatively few. The following examples in which there are a variety of study areas and methodological contents, success of GIS integration to linguistic applications has been conclusively affirmed. Most of them used GIS for data storage and cartographic assistance. Lee and Kretschmar (1993) created a geographical database under a GIS environment to store and display the linguistic data obtained from the database of the Linguistic Atlas of the Middle and South Atlantic States (LAMSAS). Another example is the research work of Luo et al. (2000) who used GIS to help visualize the settlement



©Universitat de Barcelona

pattern of Tai minority groups in southern China while Masron et al. (2005) used the geographical technologies of GPS (global positioning System) and GIS to capture the coordinates of the locations of respondents, then stored and represented the spatial patterns of dialectal variation spoken by Melanau speakers in Sarawak, Malaysia. In Thailand, Premsrirat et al. (2004) reported the use of GIS as a geographical database to store and map the distribution of ethnolinguistic groups for the whole country while Teerarojanarat and Tingsabadh (2008) developed the geographical database for storing 170 semantic units in a study of lexical variation covering the whole of Thailand and using the benefit of map display to reveal the word distribution.

Within this decade, there have been quite a few efforts to incorporate the analysis of GIS to help derive useful information. The research study of Wang et al. (2006) is a pioneer that used GIS mapping and spatial analysis through the spatial interpolation functions (e.g. spatial smoothing and trend surface modelling) to help reconstruct the linguistic and historical past settlement of the Tai from place names in southern China and Southeast Asia. Based on the Western Pennsylvania dialect Survey conducted by the English department at Clarion university, the research work of Ayad and Luthin (2009) exploited the advantageous GIS functions of cartographic display to be incorporated with social factors e.g. age, gender, occupation, level of education for mapping dialect distribution in the study area. None of these studies, however, have attempted to integrate GIS for dialect boundary studies. This paper is therefore different from the above examples in that GIS was integrated to help improve the quality of boundary drawing which will be demonstrated in the next section.

## 3. Exemplified GIS application to Thai dialectology: an exploratory study of methodology

In this paper the whole of Thailand was chosen as a study area with the key aim at demonstrating the integration of GIS research methodology to Thai dialectology. Geographically Thailand is situated in the heart of Southeast Asia, sharing its boundary with Myanmar in the west and north, Laos in the northeast, Cambodia in the east and Malaysia in the south. Reported by Premsrirat et al. (2004), over 60 ethnolinguistic groups with a diversity of languages were found in Thailand. In spite of the language



©Universitat de Barcelona

diversity, Thai is spoken in every part of the country. It consists of four main dialects - Northern Thai, North-eastern Thai, Central Thai, and Southern Thai. Their names traditionally represent parts of the country where these dialects are spoken. Studies in Thai dialectology using the manually mapping approach have been carried out on all of these dialects and many of their sub-dialects.

Up to now, drawing boundaries between dialects in many areas of Thailand has been problematic as these boundaries are unclear. In this study we integrated the advantageous GIS tool with the linguistic methodology to create a language boundary map based on lexical variation. The division of two Thai dialect regions – Central Thai and Non-Central Thai is the main focus. It should be noted that the term Non-Central Thai used in this study refers to Northern Thai, Northeastern Thai and Southern Thai as well as various non-Thai languages spoken in the country. The areas where more than one lexical item from different dialects is used are also included into the group of Non-Central Thai).

### *3.1. Data source and scope*

Data used in this study is primarily from the geographical database of the Word Geography of Thailand project (Teerarojanarat and Tingsabadh, 2008). The purpose of that project was to collect Thai dialect vocabulary. The study covers the whole of Thailand at subdistrict level except the Bangkok Metropolis. Briefly, its geographical database comprises the linking between two main data sources; the administrative boundary map and the language data.

#### 3.1.1. Thailand's administrative boundary map

Thailand's administrative boundary map is available at the scale of 1:250,000. The map was obtained from the Thailand Environment Institute (TEI), Thailand. It is a shapefile (vector data model). Features are stored as polygons having the detail of subdistrict level.



©Universitat de Barcelona

### 3.1.2. Language data

Language data, collected in the form of lexical items, was from a questionnaire of 170 questions – each representing a semantic unit. The questionnaires were sent in 2003 to all subdistricts, so-called Tambon in Thai, via the Ministry of Culture network in the first round and by post in the second round. Data collection was completed within 1 year with the total return of valid questionnaires from 6,379 Tambons or 88% of the study locations.

With the linkage of these two main data sources, the geographical database consists of 170 lexical variation maps - one map per one semantic unit. Each map contains a set of lexical items or words covering the whole country at Tambon level. Designation for data storage in the GIS database was based on the assumption that each Tambon has a uniform lexical usage. Thus, each single Tambon, in the form of a polygon, was used to be the smallest representative unit for storing and displaying language data. Figure 2 shows a map of the semantic unit "DEW" containing a set of lexical items and shading their variations on the basis of each Tambon's polygon. From the figure, for example, /mɤ:i$^{T}$ / clustering in the north, /na:m$^{T}$ mɔ:k$^{T}$/ in the northeast, and /na:m$^{T}$ kha:ŋ$^{T}$/ in the central region and south. These 170 maps are used as the primary data for the GIS analysis of this study. The other lexical variation maps of the project may be visited at the website of Word Geography Maps of Thailand (http://www.arts.chula.ac.th/~ling/geoling/WGT/).

## *3.2. Methodology and results*

Figure 3 depicts the overall methodological diagram of this study. It also highlights three GIS processes by which traditional linguistic practices are replaced. To begin with, the linguistic approach still plays a key role in the analysis of lexical classification and grouping. This step involves analysis and classification of the lexical items or words of each semantic unit into two groups; Central and Non-Central Thai. Lexical analysis involves the process of measuring the relative degrees of lexical similarity.



©Universitat de Barcelona

The second process substitutes the way by which traditional manual drawing is performed. GIS was utilized by doing region grouping. That is, each semantic unit, classified Central Thai and Non-Central Thai from the previous step was coded and shaded as an isogloss map. A resultant example of this process is shown in Figure 4. From the figure, an isogloss map of the semantic unit "DEW" was produced by being coded in GIS as a binary variable identifying whether there is Central Thai (= 1) or Non-Central Thai (= 0) and then colored to symbolize two regions. By applying similar techniques, the 170 isogloss maps were created in this step.

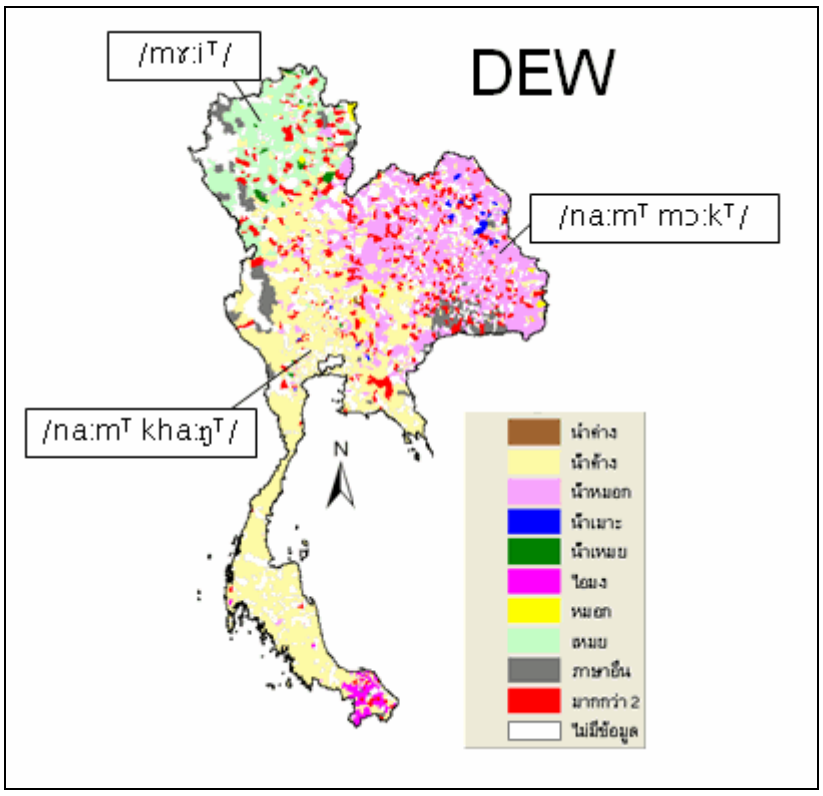



Figure 2. The "DEW" map showing the distribution of lexical items for the whole of Thailand at subdistrict or Tambon level



©Universitat de Barcelona

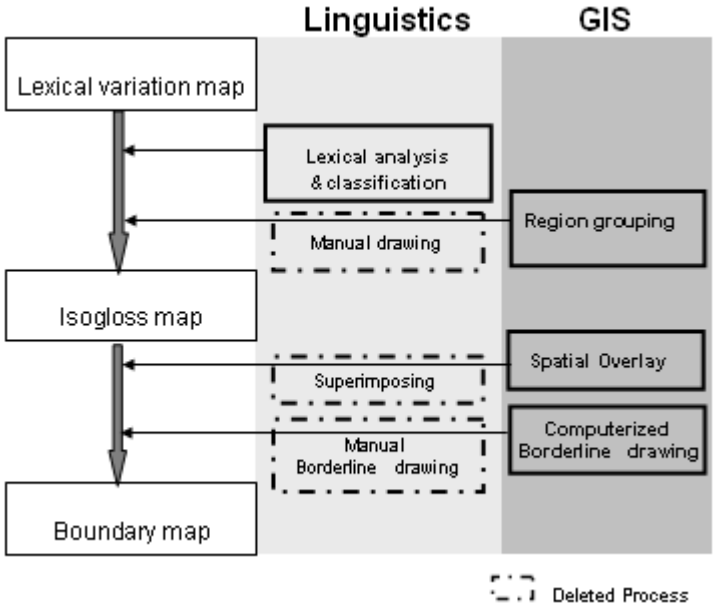



Figure 3. The conceptual methodology diagram showing the integrating of GIS tool and linguistic methodology for creation of a boundary map

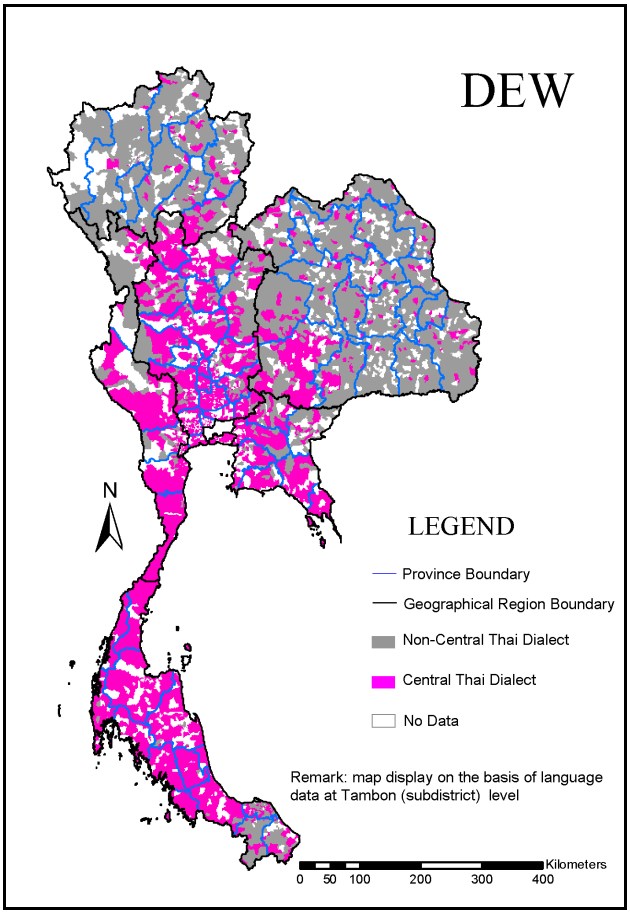



Figure 4. An isogloss map of a semantic unit "DEW" with the classification of two dialect regions - Central and Non-Central Thai



©Universitat de Barcelona

In the third step, the traditional linguistic methodology of manual superimposing of the bundle of isoglosses was also replaced by GIS functions. Under the GIS environment, each isogloss map is considered as a language layer. The spatial overlay analysis of the 170 isogloss layers in a vector mode on the basis of 'union' operation was then performed in order to create a composite map of dialect regional patterns. The 'union' operation can be considered as a Boolean logic 'OR'. For each union, it works in the way that two different layers are combined together to produce a new layer having both new spatial and attribute data (see Figure 5).

In our case, a new overlay map, for example, of two isoglosses is produced from a 'union' function. Since both polygon features have similar shapes, a new overlay map will differ in that new attributes (fields) will be added while the number of polygon features is still unchanged. One by one, a pair of isoglosses was overlaid by 'union' operation. Finally, an overlay map of 170 isoglosses was created. The following step involved counting and accumulating the frequency of language occurrence for each single Tambon from the attribute (tabular) data. Figure 6, as an example, shows the way to accumulate and calculate the percentage of Central/Non-Central Thai usage. From the figure, the table shows the overlay of 4 isoglosses of 6 Tambons. Each row or record represents one Tambon while the field 'cenn1', 'cenn2', 'cenn3, and 'cenn4' represent the overlaid isogloss map no.1, 2, 3 and 4 respectively. For each Tambon, the value of field 'cenn1', 'cenn2', 'cenn3, and 'cenn4'refers to the occurrence of Central Thai usage (coded as 1) or Non-Central Thai usage (coded as 0) of these isoglosses orderly. In the first row of the table representing Tambon no.1, usage of Central Thai (coded as 1) was found in 3 isoglosses regarding field 'cenn1', 'cenn3', and 'cenn4'. Thus it was occurred in the field 'TotalCen' equaling to 3. Based on the combination of a total of the 4 isoglosses, 75% was calculated and stored in the field 'PctCen'. This means 75% of the people in that Tambon speak Central Thai or 25% speak Non-Central Thai. The percentage values then can be plotted and classified on a map. Based on a similar concept, Figure 7 shows the resultant map of this process, shading with the percentage of Central Thai usage in 10 levels. The degree of 90 – 100% means the areas where Central Thai is spoken. In contrast the degree of 0 – 10% means the areas where people use Non-Central Thai.



©Universitat de Barcelona

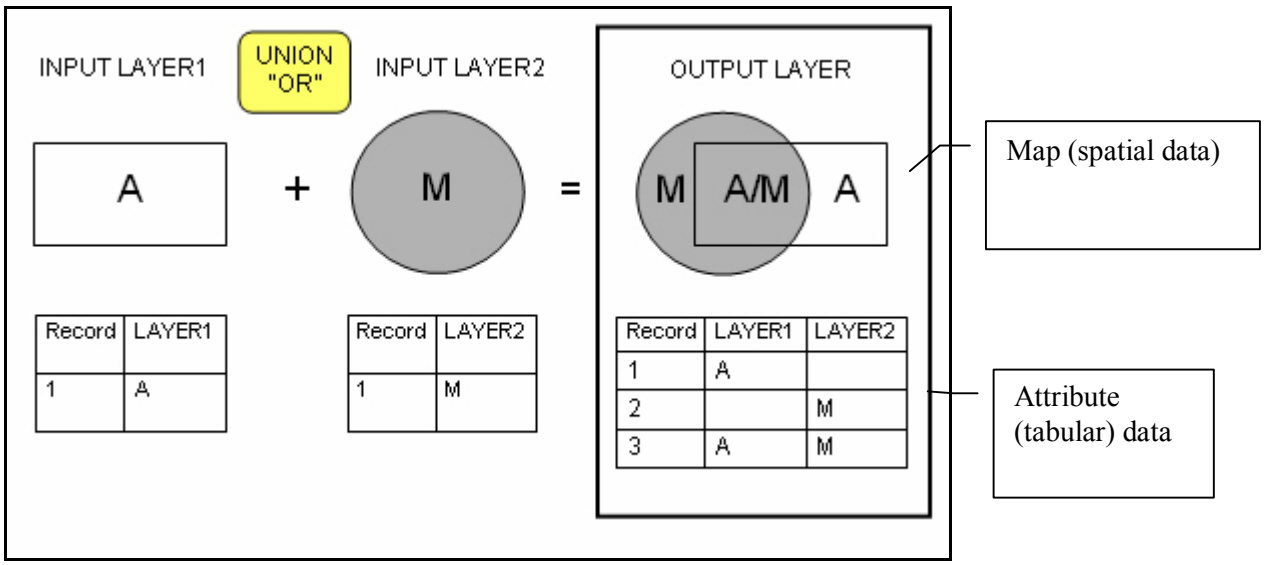



Figure 5. Principle overlay analysis of UNION function

| Record | Cenn1 | Cenn2 | Cenn3 | Cenn4 | TotalCen | PctCen |
|---|---|---|---|---|---|---|
| 1 | 1 | 0 | 1 | 1 | 3 | 75.0 |
| 2 | 1 | 1 | 1 | 1 | 4 | 100.0 |
| 3 | 1 | 0 | 0 | 0 | 1 | 50.0 |
| 4 | 0 | 0 | 0 | 0 | 0 | 0.0 |
| 5 | 0 | 0 | 1 | 0 | 2 | 25.0 |
| 6 | 0 | 1 | 1 | 1 | 3 | 75.0 |

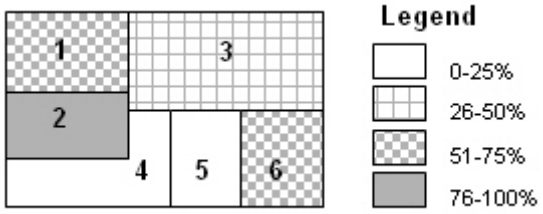



Figure 6. A simple form of a table illustrating the way to accumulate and calculate the percentage of Central/Non-Central Thai usage

The final step involves the way to draw the boundaries between the two Thai dialects. The final decision is still based on the judgment and the background knowledge of dialectologists. However, unlike the old technique in which the boundaries have been drawn manually and with uncertainty, here GIS utilized in the



©Universitat de Barcelona

earlier step provided different views of maps from an overlay map previously produced as shown in Figure 8, and thus enabled dialectologists to make better interpretations and judgments with more confidence. In this study, Central Thai usage at the degree of ≥ 50% was chosen as a key criterion to separate these two regions - Central and Non-Central Thai. This criterion was set from the fact that 50% usage was considered an unbiased figure. Once set, the final step involved the way to draw lines on the basis of political boundaries. At first, the political boundaries of Tambon (or subdistrict), the finest unit, were considered. However, as shown in Figure 8, it can be clearly seen that the appearance of no data (appearing in white area) scattering over the large areas in the map caused difficulty in drawing continuing lines. To overcome this problem, maps in Tambon or subdistrict level were aggregated using the GIS function into district, so-called Amphoe in Thai, as shown in Figure 9. Aggregation technique involves the way to accumulate the occurrence of language usage of Tambon level into Amphoe level. Having aggregated, the boundaries in the study were finally drawn based on the Amphoe boundaries as shown in Figure 10. Consequently, a boundary map of the two regions - Central and Non-Central Thai was created with sharply defined boundaries as shown in Figure 11. The map produced can thus be used as a point of reference for Thai dialectology on the basis of lexical study. The result was also published with an identical map but with a content adjusted to different readers by Teerarojanarat and Tingsabadh (2009) in Thai as a main purpose for discussing the map result and comparing it to another Thai boundary study conducted in the same area on the website (http://www.arts.chula.ac.th/~ling/geoling/BMTD/) in Thai as a summary of the boundary study.

### *3.3 Discussion*

While the complexity due to the nature of language still exists, this paper presents an alternative, a GIS-based approach, by which the mapping technique was enhanced to lessen the error and distortion during the process of boundaries drawing. The paper also demonstrated the applied concept of GIS overlay analysis and statistical means to be integrated for producing an overlay map of 170 semantic units, and finally creating a boundary map. The applied methodology benefits linguists and researchers in many ways.



©Universitat de Barcelona

Firstly, GIS can produce a more accurate overlay map than the traditional one. Manually superimposing maps can easily produce unexpected and incorrect output. Under a GIS environment, however, distortion due to the process of spatial overlay can be minimized. As a result, the quality of the map can be controlled throughout the whole process. Secondly, overlaying multiple maps, in our case, 170 semantic units is possible and convenient. It has been proven that using a GIS process saves a lot of time. The repetitive overlay tasks can be overcome by writing programming scripts and making it easy to operate. Thirdly, GIS can produce a quantified overlay map (see Figure 7). In our case GIS facilitates the way to compute the percentage of Central/Non-Central Thai. Degrees of mixing between these dialects can be easily observed and examined. The overlay map produced in this way aids in identifying where the heartland of Central Thai is located as well as pinpointing its spreading direction. It, thus, helps researchers understand spatial variation and distribution of dialects as well as dialect density over the space. Fourthly, GIS facilitates cartographic display. In this study GIS can produce different views of maps (see Figure 8) from the quantified overlay map. Researchers can then learn and get better knowledge on lexical study.

There are possibly some other cases in which data is collected and stored in a GIS database as point features. In these cases, transformation of point features to Thiessen polygon features, also known as 'Voronoi networks' and 'Delaunay triangulations', have been suggested to be performed first. By using the Thiessen polygon technique, the area of each point is proportionally divided and distributed into regions on the basis of the Delaunay criterion (ESRI, 2004). Figure 12 illustrates the transformation of points to Thiessen polygons. Clearly seen, the area of each Thiessen polygon is closer to the point on which the polygon is based than to any other point. Once Thiessen polygons are created, the approach developed in this study can be easily adapted and applied to other regions.



©Universitat de Barcelona

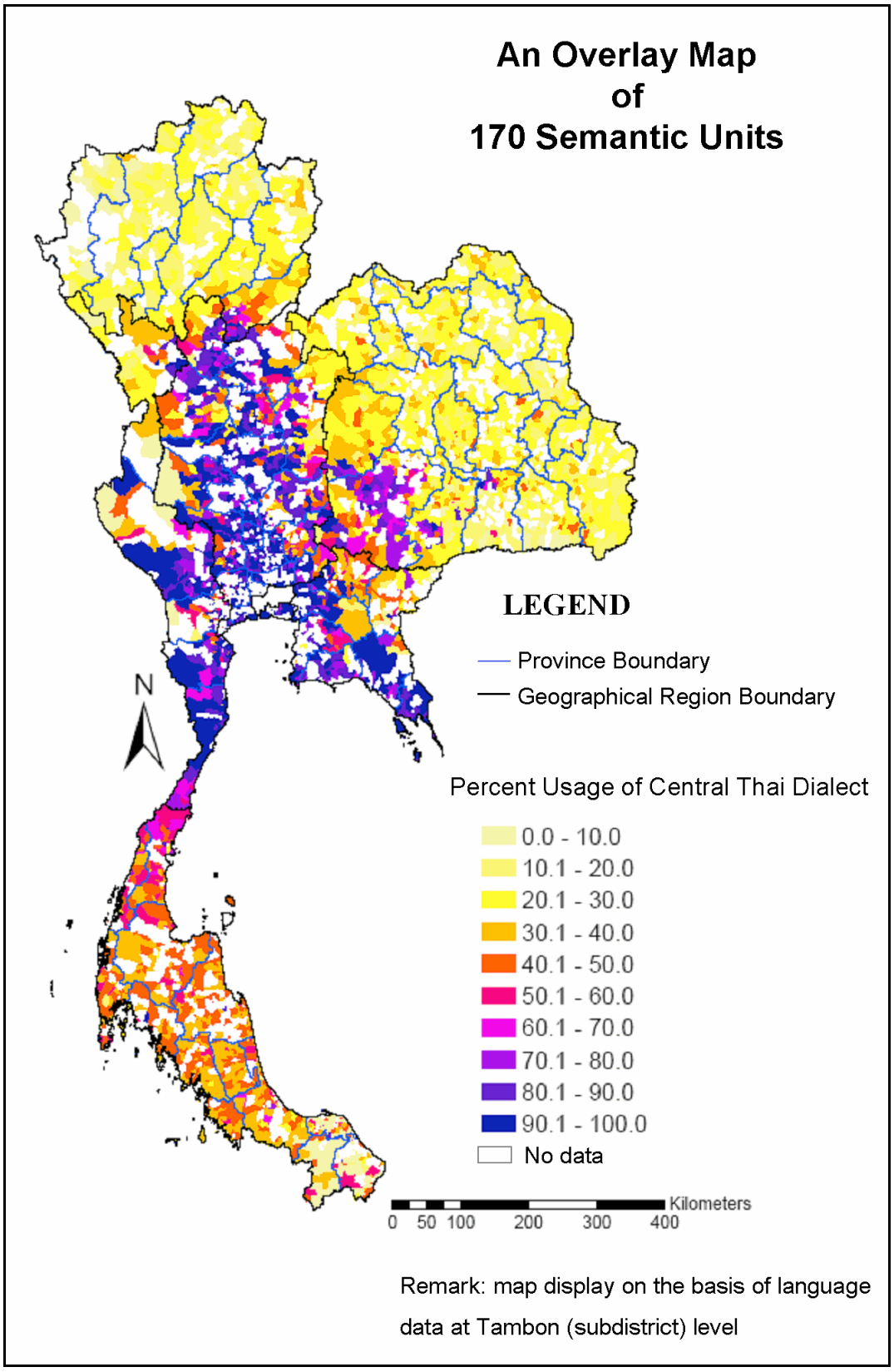



Figure 7. An overlay map of 170 isoglosses shading with the percentage of Central Thai usage in 10 levels



©Universitat de Barcelona

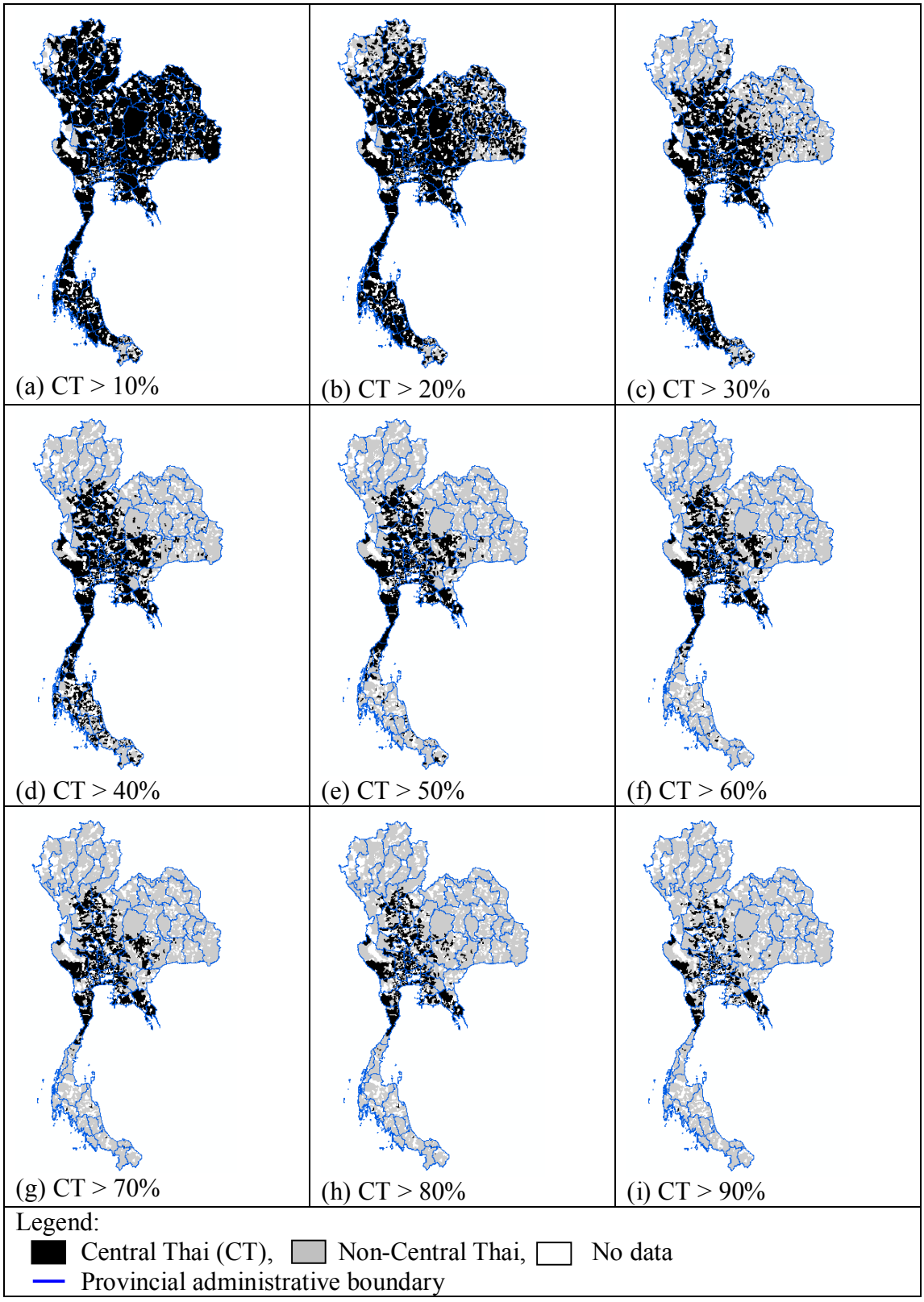



Figure 8. Different degrees of the areas where people speaking Central Thai (CT) in Thailand from (a) > 10% to (i) > 90%



©Universitat de Barcelona

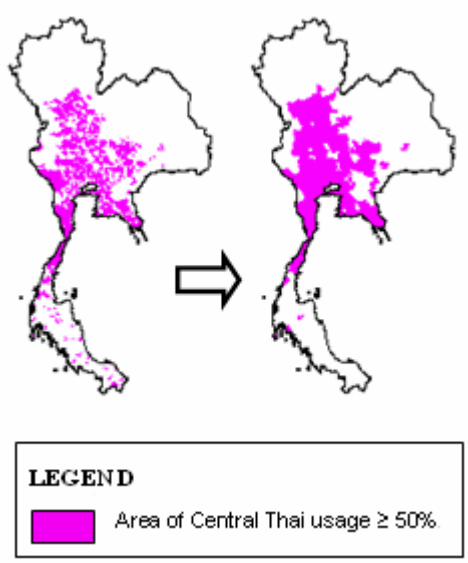



Figure 9. The selected Central Thai usage ≥ 50% of Tambon boundary (a), and at the more aggregated level of Amphoe boundary (b) (Remark: the remaining unshaded region refers to the Non-Central Thai)

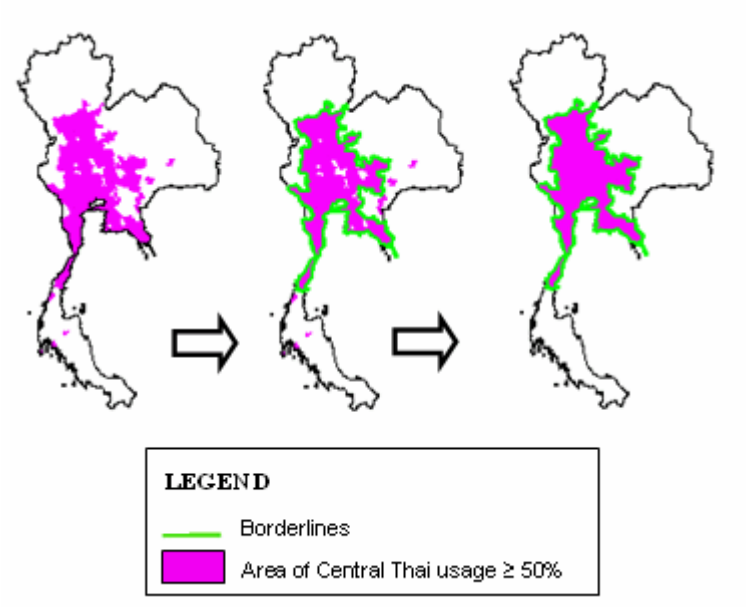



Figure 10. Drawing the lines on the basis of majority of adjacent Amphoe (district) boundaries



©Universitat de Barcelona

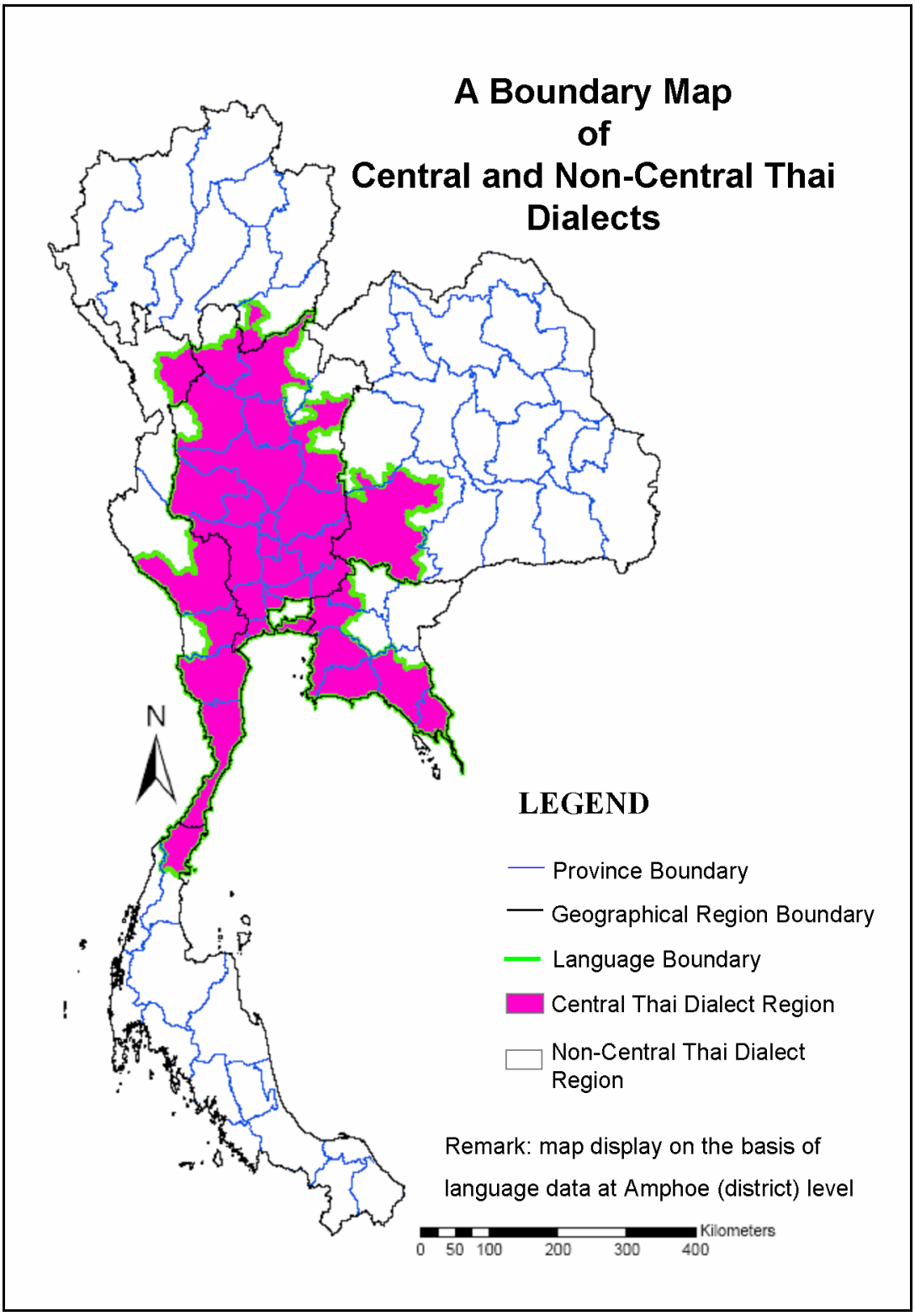



Figure 11. A boundary map showing two Thai dialect regions - Central and Non-Central Thai



©Universitat de Barcelona

## 4. Conclusion and future research direction

The GIS-based approach was applied in this study with the exemplified application to Thai dialectology. The by-product outcome, namely the overlay of 170 semantic units, facilitates better interpretation of results with more confidence, particularly in the unclear areas. The final boundary map created in this research, as a result, was promising as it has higher degree of accuracy, higher quality and more reliability than the traditional one. Furthermore, since a resultant map is stored in a digital format, other types of layers such as political boundaries, topography, settlement patterns, social features i.e. cultural affiliation, ethnicity can be incorporated easily.

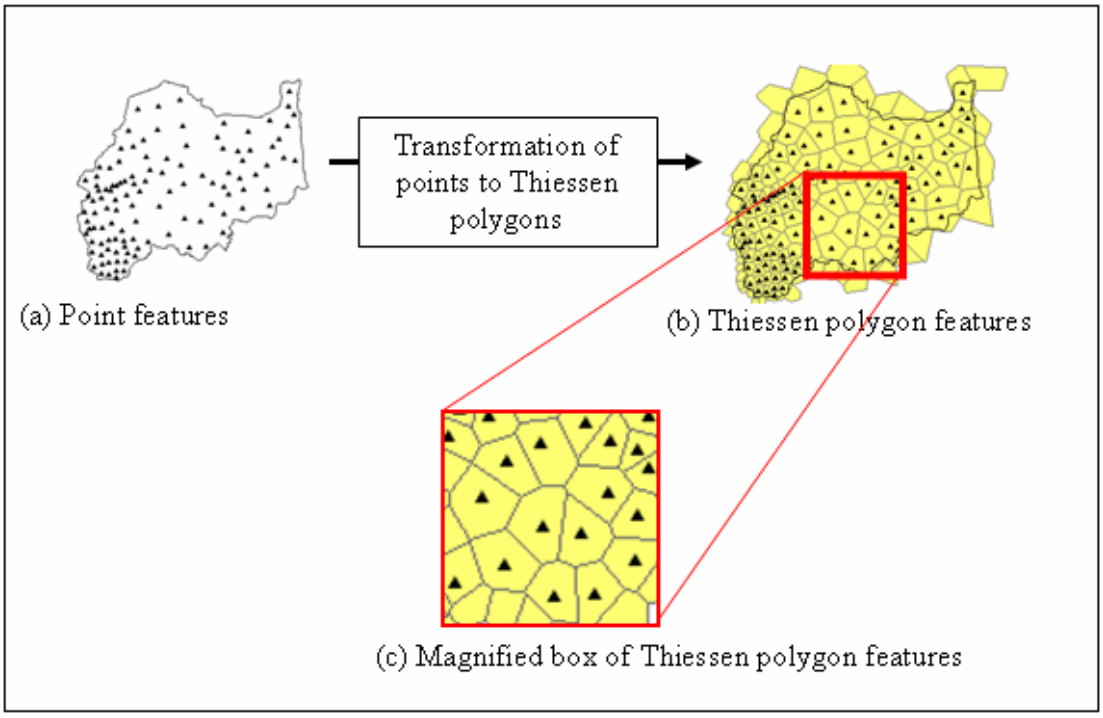



Figure 12. Transformation of point features to Thiessen polygon features

Further work will adapt the GIS-based approach to produce four Thai dialects – Northern Thai, Northeastern Thai, Central Thai, and Southern Thai. We are also interested in exploring the relation between language data and settlement patterns including topography, transportation, and population settlement in the near future.



©Universitat de Barcelona

## References

AYAD, Y. and LUTHIN, H. (2009) "Mapping Dialect: GIS Applications in Studying Dialect of Western Pennsylvania", in *The 2009 ESRI International User Conference Proceedings - San Diego, California, USA. 13 - 17 July 2009*. [http://proceedings.esri.com/library/userconf/proc09/index.html]. Accessed: 5 January 2010.

CHAMBERS, J. K. and P. TRUDGILL (1980) *Dialectology*, Cambridge: Cambridge University Press.

ESRI. (2004) *ArcGIS Desktop Help*, Redlands: ESRI Press.

KREMER, L. (1999) "The Netherlands-German National Border as a Subjective Boundary", in Dennis R. Preston (ed.), *Handbook of Perceptual Dialectology volume 1*, Amsterdam: John Benjamin Publishing Company, 31-36.

LANGACKER, R. W. (1968) *Language and its structure: some fundamental linguistic concepts*, New York: Harcourt, Brace & World.

LEE, J. and W. A., Jr. KRETZSCHMAR (1993) "Spatial analysis of linguistic data with GIS functions", *International Journal of Geographic Information Systems*, 7(6), 541-560.

LONGLEY, P.A., M.F. GOODCHILD, D.J. MAGUIRE, and D.W. RHIND (eds.) (1999) *Geographical Information Systems: Principles, Techniques, Applications and Management*, New York: John Wiley & Sons.

LONGLEY, P.A., M.F. GOODCHILD, D.J. MAGUIRE, and D.W. RHIND (2005) *Geographic Information Systems and Science*, West Sussex: John Wiley & Sons.

LONG, D. (1999) "Geographical Perceptions of Japanese Dialect Regions", in Dennis R. Preston (ed.), *Handbook of Perceptual Dialectology volume 1,* Amsterdam: John Benjamin Publishing Company, 177-198.

LUO, W., J. F. HARTMANN, J., LI, and V. SYSAMOUTH (2000) "GIS Mapping and Analysis of Tai Linguistic and Settlement Patterns in South China", *Geographic Information Sciences*, 6(2), 129-136.

MASE, Y. (1999) "Dialect Consciousness and Dialect Variations: Examples in the Nagano-Gifu Boundary Region", in Dennis R. Preston (ed.), *Handbook of Perceptual Dialectology volume 1,* Amsterdam: John Benjamin Publishing Company, 31-36.

MACKEY, W. F. (1988) "Geolinguistics: Its Scope and Principles", in Colin H. Williams (ed.), *Language in Geographic Context*, Clevedon: Multilingual Matters, 20-46.



©Universitat de Barcelona

MAGUIRE, D.J., M.F. GOODCHILD, D.W. and RHIND (1991) *Geographical Information Systems: Principles and Applications*, Harlow: Longman Scientific and Technical.

MASRON, T., R. RAINIS, S. GHAZALI, S.C. LAH, A. SAZALI, A. A. GHANI, and S. A. GHAFOR (2005) "Using GIS to grasp dialectal variation", in *Conference. Proceedings of Map Asia 2005 – Jakarta, Indonesia. 22 -25 August 2005.* [http://www.gisdevelopment.net/application/miscellaneous/ma0538.htm]. Accessed: 5 January 2010.

TEERAROJANARAT, S. and K. TINGSABADH (2008) "Development of Geographical Database of Lexical items in Thai Dialects", *Journal of Language and Linguistics*. Department of Linguistics, Faculty of Liberal Arts, Thammasat University, 26(2), 1–20. (in Thai).

TEERAROJANARAT, S. and K. TINGSABADH (2009) "Mapping Dialect Region of Thailand: a Case Study of the Borderlines of Central Thai Dialect" (*Journal of Arts*, Faculty of Arts, Chulalongkorn University, 38(1) June – December 2009, 166-202). (in Thai).

PREMSRIRAT, S. and OTHERS (2004) *Ethonolinguistics Maps of Thailand*, Bangkok: Office of the National Culture Commission in Ministry of Culture. (in Thai).

TOMLIN, C. D. (1990) *Geographic Information Systems and Cartographic Modelling*, Englewood Cliffs: Prentice Hall.

WANG, F., J. HARTMANN, W. LUO, and P. HUANG (2006) "GIS-Based Spatial Analysis of Tai Place Names in Southern China: An Exploratory Study of Methodology", *Geographic Information Sciences,* 12(1), 1-9.

WEIJNEN, A. A. (1946) "De grenzen tussen de Oost-Noordbrabantse dialecten onderling" [The borders between the dialects of eastern North Brabant], in Antonius A. Weijnen, J. M. Renders, and Jac. Van Ginnecken (eds.), *Oost-Noordbrabantse dialect problemen* [Eastern North Brabant dialect problems] Bijdragen en Mededelingen der Dialectencommissie van de koninklijke Nederlandse Akademie van Wetenschappen te Amsterdam, 8, 1-15.

WILLIAMS, C. H. and J. E. AMBROSE (1988) "On Measuring Language Border Areas", in Colin H. Williams (ed.), *Language in Geographic Context,* Clevedon: Multilingual Matters, 93-135.

## Acknowledgements

The authors wish to thank the Word Geography of Thailand Project which provided the data source for analysis. We are also very grateful to Chulalongkorn University Centenary



©Universitat de Barcelona

Academic Development Project and the Faculty of Arts, Chulalongkorn University for the financial support that made this study possible. Without their support this study would not have been completed.



©Universitat de Barcelona